Author's Response

Date: 20 March 2019

We thank the Editor, reviewers and those who provided short comments on the manuscript for their inputs. The feedback has helped to improve both the clarity and content of the manuscript. We have provided responses to both the short comments and full reviews. We indicated by section, specified by paragraph, where revisions were made within the manuscript text. Since we include figures in response to both the short comments and the full reviews, there is a letter code to indicate if the figure is related to a short comment (e.g. SC1.1) or to a reviewer comment (e.g. RC3.1). The figures in the manuscript itself maintain normal numbering convention (e.g. Figure 1, Figure 2, etc.).

The following is a list of the major changes to the manuscript:

- We have revised the manuscript text for clarity and brevity. In particular we have shortened the Introduction and rearranged the text regarding the treatment of snow depth in each satellite data product.
- Based on input received during the review, we have revised the text of Section 2.1 and the information provided in Table 1 to clarify specific aspects of the processing chain and waveform retrackers used in each satellite thickness product.
- We have extended Table 1 so as to include the details of two additional satellite-derived sea ice thickness data sets, although these data are not included in the further analysis. This decision is a compromise between providing the pertinent details of publicly-available data products, while not overwhelming the reader with too much information in the figures and tables.
- We have replaced the CS2SMOS data set used in the original submission with an updated version of the data set, and revised all figures and tables containing CS2SMOS data.
- We have updated figures and tables wherever possible with new data that has become available since the original submission. In particular, we now include the BGEP ULS ice draft observations for the 2016-2017 season.
- We expanded our results to include winter growth rates, adding a new table (Table 5).
- The reviewers highlighted concerns regarding the original methods used to calculate the correlation between data products, and that using a near-neighbour interpolation with a search radius of 50 km could potentially artificially improve the correlation results. To address these concerns, we have revised the approach to calculate the correlation statistics between the satellite data products, as well as between the satellite and airborne observations. In the revised manuscript the thickness observations are placed onto a common grid (0.4° latitude by 4° longitude) before common grid cells are compared and correlation statistics calculated. This follows the approach originally taken in Laxon et al. (2013) as well as in subsequent studies and allows the reader to place our results in the context of the published literature. We note that this did not change the results of the correlation analyses in a substantial way.
- We changed the author order, switching the places of second and third author, to correspond with the actual amount of effort, which changed during the review process.

Within the manuscript text, all edits (additions/deletions) are indicated in red font. The manuscript version indicating track changes is posted as a separate author's comment in order to keep the response document concise.

Response to Review #1 (Jack Landy)

Assessment of Contemporary Satellite Sea Ice Thickness Products for Arctic Sea Ice
Sallila et al., 2018

This study provides the first genuine inter-comparison of pan-Arctic sea ice thickness data products obtained from three different satellite sensors. It is systematic in approach, equitably comparing the datasets to identify seasonal and interannual patterns of ice thickness variability, as well as assessing biases versus independent reference data. Results from the assessment will prove to be a valuable contribution to the field of Arctic climate science and potentially to operational stakeholders.

We thank the Reviewer for their encouraging comments on the usefulness of our study for the science community and stakeholders.

However, I have some concerns about the robustness of methodology used to process and compare datasets. These concerns focus on the procedure used to resample data, as described in the comments below. I have also made a few suggestions to improve the analysis and impact of the author's findings. I'd recommend this manuscript is published in The Cryosphere, following these revisions.

Please do get in contact if you have questions regarding these comments. Kind regards, Jack Landy

General comments:

RC 1.1. The method used to re-grid the satellite products may be introducing errors into the data inter-comparison and artificially improving the correlations. By gridding to 5 km with a search radius of 50 km you are introducing significant spatial autocorrelation between adjacent grid cells (through your method cells can theoretically only be treated as independent samples at a length scale 25 km). It is not statistically robust to be calculating the pixel-wise correlation between spatially-dependent samples, and this may be artificially improving the coefficient and significance. Perhaps you could try recalculating the correlation coefficient for a down-sampled 25-km version of your current grid? Take a look at this paper https://www.cambridge.org/core/journals/annals-of-glaciology/article/impact-of-spatial-aliasing-on-seaice-thickness-measurements/AB5DB924171B219D25E64A829E418840 to consider your gridding approach. (This is also worth considering for your OIB comparisons in Fig 9).

Agreed. Originally data were placed on a grid with 5 km resolution using a nearest neighbour interpolation function with a search radius of 50 km. As you and other reviewers have commented, this could lead to an autocorrelation between adjacent grid cells, and an overestimate of the correlation between independent data products. As suggested, we have revised our approach and now place the data sets onto a common 0.4° latitude by 4° longitude grid before calculating the correlations between products. We have revised the description of the gridding procedure in Section 3.1, updated figures 4, 9 and 10, and Tables 2 and 3. We have updated Sections 3, 4, and 5 to reflect the changes in the Figures and

Tables. Our revised method follows the approach originally taken in Laxon et al. (2013), and then used in subsequent studies (e.g. Tilling et al., 2018), thereby allowing the reader to place our new results in the context of the previously published statistics.

RC 1.2. Linked to the above comment, you suggest at Page 14, Line 19 that the satellite products do not capture the thinnest or thickest ice because of the grid-averaging process. It surely makes sense then, for a fair comparison between satellite obs and the ULS, to average the ULS measurements over some window, rather than binning individual obs in your histograms. I understand this is difficult because it means averaging over time and comparing it to grid averages over space. However, could you identify a time-averaging window based on recorded ice drift speeds (either a seasonally- representative drift speed or obtained from satellite ice motion data)?

We believe the Reviewer may have misunderstood a detail in the original text, regarding the satellite data. The monthly data sets as provided are actually gridded at 25 km resolution (without any additional gridding by us). Regarding the assessment of the satellite data products through comparison with the ULS analysis, we follow the methodology of Laxon et al. (2013) and Kwok and Cunningham (2015). This is to enable the reader to place our results in the context of the existing literature. We have revised the text in Section 4.1 for clarity.

RC 1.3. Your correlations between all the CS2 products are understandably very high (Page 12, Line 12) and there's not too much new analysis here on the differences between the CS2-only datasets. I would be very interested to know how closely each product picks up small-scale spatial variability in ice thickness. For instance, you could calculate the anomaly maps of each year compared to the climatological average, then assess the pixel-wise inter-product correlations between positive and negative anomalies. Do some of the CS2 products identify more or less of the smaller-scale thickness features than others? Are they realistic? This analysis would improve e.g. the first paragraph of the discussion.

Our results show for the first time the correlation between a variety of CS2-only products, which differ due to waveform retracking algorithms. We have also significantly extended the evaluation of the CS2 products, through comparisons with independent data over much longer time period (2010-2017) than has previously been published. We are unsure exactly what the Reviewer means by "small-scale spatial variability in ice thickness", since the majority of the satellite data products are provided on grids of resolution 25 km. Sea ice thickness varies on much smaller scales than this, but is not possible to assess it in this analysis of gridded data. However the results in Table 2 provide details of the spatial variability between products at the regional scale. APP-x shows the least variability from region to region, with mean ice thickness in spring across the product record falling within a range of 0.19 m for the southern regions of the central Arctic basin (regions 3-6). JPL shows the largest regional-scale variations, with a range of 0.97 m in average spring thickness across regions 3-6. Table 3 also provides some interesting results on differences between CS2-only products, in terms of seasonal variations relative to the 2011-2015 baseline. There are also differences in the growth rates (Table 5) and wintertime trends (Fig. 5) across the CS2-only products, where the GSFC product has the lowest daily growth rate. Based on these results, we have included additional discussion in Sections 4.2 and 5 about the regional differences.

RC 1.4. There is not a lot of analysis or discussion of regional variations in the products: whether any obvious differences are regionally-dependent etc., except from Table 2.
Can you add some deeper discussion on regional differences in the products, and offer some interpretation as to why these differences may exist? This would be a really valuable addition to the work.

Based on the results shown in Table 2 we have added new discussion to the second paragraph of Section 4.2, per this suggestion.

Minor comments/edits:

Page 2. Lines 15-16. By what mechanism does Laxon 13 speculate that lower ice volume could be a factor in recent minima in ice extent?

The direct statement is found in paragraph 27 of Laxon et al. (2013), which is as follows: "Finally, we can speculate that the lower ice thickness and volume in February/March 2012, as compared with February/March 2011, may have been one factor behind the record minimum ice extent reached in September 2012". The statement points out that the anomalously low ice volume at the end of winter in April 2012 may have contributed to the sea ice minimum observed in September 2012. We have revised the text (Section1, second paragraph) to clarify that we are referring to the September 2012 minimum.

P2. L29-31. Can you briefly introduce here some of the main differences in approach?

Section 2.1 is dedicated to the description of the satellite data products. We therefore elected to introduce information concerning the retrackers there, and we do not feel that this information is critical for inclusion in the introduction section.

P3. L17-18. This is confusing: you're not including the ICESat data in your evaluation right, so why do you discuss evaluating ICESat here? Do you mean they have been used as a tool for evaluating satellite products in the past?

We have revised the introduction section based on comment RC2.1 and subsequently removed this sentence about ICESat.

P4. L6. What is the time period under investigation?

The period under investigation is October 2010 - April 2018. For some analysis, e.g. the time-series of ice thickness shown in Fig. 5, data are shown based on the availability for each data product. The baseline period for comparisons is the period with full product overlap (spring 2011-fall 2015). We added a brief clarification in Section 2.1, paragraph 1.

P4. L8. What do you mean by mean scattering horizon?

We have revised Section 2.1.1, paragraph 2, based on updated information provided in Tilling et al., 2018 and the comments of Reviewer 2.

P4. L30+. I understand that estimates for snow depth are pretty similar between the CS-only datasets, but you do not refer to them here and they are critical. Can you add some info on how snow depth and density are estimated for each product?

For each product description in Section 2.1 we have added text about snow depths and densities, wherever that information was available.

P6. L16-18. The description of the waveform-fitting method used for the GSFC product needs to be improved. The physical model of Kurtz 14 does not calculate expected returns empirically. Their model is semi-analytical, producing a simulation of the expected echo shape based on statistical parameterizations of the sea ice backscattering properties and height distribution of surface roughness.

We have changed the explanation to better describe the method used to derive the GSFC thickness product, in Section 2.1.4, paragraph 2.

P7. L26-27. What is the physical relationship between the surface EB and ice thickness?

The OTIM model, used in the APP-x product, derives sea ice thickness as a function of heat fluxes, surface albedo and radiation, which all contribute to the surface energy budget. Furthermore, most of the flux and radiation parameters in the equations are functions of surface skin and air temperatures, surface pressure, surface air relative humidity, ice temperature, wind speed, cloud amount and snow depth, which are input parameters in the model. We refer the Reviewer to Wang et al. (2010) where the equations and a thorough explanation are provided. We revised the description of the APP-x data product, in Section 2.1.6, paragraph 2.

P7. L30-31. Why does sea ice roughness affect the surface EB?

We reference Wang et al. (2010), where this was stated in the conclusions. We added a reference to Wang et al. (2010) to the line in question, Section 2.1.6, paragraph 3.

P8. L 5. How is snow depth estimated in the model?

Based on personal communication with the data provider, the initial snow depth estimates are based on those from Warren et al. (1999), but these estimates are then adjusted using field observations and the final snow depth values have been chosen experimentally. We have added information about the treatment of snow in the APP-x sea ice thickness data product to Section 2.1.6.

P9. L2. What measurement of uncertainty is that? Bias? RMSE?

This measurement comes from the error propagation typically used in sea ice thickness remote sensing and follows Giles et al. (2007). The thickness uncertainty is calculated using an assumption of the probable errors associated with the variables used in the thickness equation. We added an explanatory subordinate clause in Section 2.2.2 and added the reference to Giles et al. (2007).

Giles, K. A., Laxon, S. W., Wingham, D. J., Wallis, D. W., Krabill, W. B., Leuschen, C. J., McAdoo, D., Manizade, S. S., Raney, R. K.: Combined airborne laser and radar altimeter measurements over the Fram Strait in May 2002. Remote Sens. Environ., 111(2–3), 182–194, doi:10.1016/j.rse.2007.02.037, 2007.

P9. L20. It is unconventional to subtract the focus product from the reference – typically it's the other way around. This makes the plots e.g. Fig 3 harder to interpret, because you have products with higher thickness than the CPOM data given as negative numbers.

We have changed the order of product differences in Equations 1 and 2, leading to changes in Figure 3 and the corresponding interpretations.

P10. L3. Why do you pick a 200 km radius? Is this an arbitrary threshold or based on estimated ice drift displacements over the season? Do you use the same threshold in winter and summer? Have you tried varying the threshold +/- 100 km to see if you get the same results?

We used a radius of 200 km, since it is consistent with the original evaluation presented in both Laxon et al. (2013) and Kwok and Cunningham (2015). This allows the reader to place our results in the context of the conclusions of prior studies.

P10. L9. So, one correlation coeff for the whole time period, using each monthly average as a single sample?

We have added clarification on paragraph 2, Section 3.2.

P10. L17. Above you refer to the issue of oversampling at the start or end of a seasonal period and then treating products as the same when comparing. The IceBridge spring data are collected in discrete campaigns, so the period of observations will often not be spread evenly over the March-April period. This seems like the same issue – so how might this be introducing uncertainty into your satellite-OIB product comparisons?

We have revised our methodology to calculate the correlations between the satellite data products and the IceBridge data. We now place the IceBridge and satellite-derived thickness data on a 0.4° latitude by 4° longitude grid to conduct the comparisons. The approach helps to mitigate the uneven spatial and temporal sampling of the sea ice along the IceBridge flight-lines compared to the monthly means obtained from the satellite data products. Also, as we have adopted the same approach originally used in Laxon et al. (2013) and subsequent studies (e.g. Tilling et al., 2018), this allows the reader to place our extended results in the context of the existing literature.

P11. L4. Can you give the mean and +/- (between the techniques) of the fall 2011/2012 ice thickness, to provide the reader with a bit more context? Here and in other similar places within the paper?

Added values for the mean thickness and its variation on the suggested line as well as in other places within Section 4.1.

P11. L18-20. Is there a way you can quantify this? e.g. by calculating the spatial covariance of the gridded data between years, or by simply calculating the average pixel standard deviation in ice thickness across the period? i.e. how sensitive are each of the products to spatial features of ice thickness like the thick ice tail in the Beaufort Gyre, at what scales?

This was meant as a brief discussion of results in Figure 2, based on visual inspection. Here we did not emphasize exact quantitative differences, as there are plentiful analyses included in the later figures and tables. We agree though that studying the sensitivity of the products, or even one product, to spatial features, would be a great study on its own!

P12. L3-4. For which season?

Good catch, these are for the spring season. We added the season in the second paragraph of Section 4.2.

P13. L 14. And indeed of the thicker sea ice too. . . It's very strange that the technique based on a thermodynamic model and surface EB suggests highest ice growth rates at the end of winter, when you'd expect such a method to be better at capturing the expected reduction in thermodynamic growth.

Unfortunately we have not discovered a reason behind this behaviour in the available references describing this data product. It seems like the air temperature, and other factors that likely support ice growth, outweigh some of the other physical phenomena in the EB model.

P13. L15-34. It is very interesting that the CS2 only products are more similar in spring than in fall, when intuitively you'd expect them to diverge more when the ice is thicker. Can you provide any insight on why this might be? In what ways may the processing differences induce variations in fall that are not apparent in spring?

After applying revisions to the method used to calculate seasonal averages, and obtaining a new version of the CS2SMOS data, we find that the CS2-only products tend to agree quite well in both the spring and the fall. The exception is the GSFC data product which diverges from the other CS2 products in the fall. Figure 3a shows that while ice thickness in the fall is on average thicker in the GSFC product, it is most apparent in regions 5 and 6 (East Siberian and Laptev Seas) and over the thickest multiyear ice in the central Arctic. Recall that the major difference between the GSFC product and the other CS2-only products is that the GSFC retracker uses a waveform fitting method (Table 1).

P14. L 34. What do you mean by 'varying dependency to draft'?

We were attempting to describe the varying correlation values for the three different buoys (Fig. 8). The line in question in Section 4.4 has been deleted from the revised manuscript.

P15. L2-3. Since the thickness is relatively low in this region (<1.5 m) this would imply a higher contribution of SMOS data to the optimal interpolation, right? So does this give you any information about the reliability of SMOS data?

The major difference between the CS2SMOS and AWI products is the inclusion of SMOS data, and this suggests that there are occasions when SMOS estimates higher thickness, which could be due to the relatively thin ice in this region, as you state. However, we do not make statements about the reliability of SMOS data and refer the Reviewer to Ricker et al., 2017 regarding the CS2SMOS interpolation scheme.

P16. L1-15. It is also worth noting that GSFC freeboard for both fall and spring is generally lower than the AWI freeboard, whereas the GSFC ice thickness was determined as the highest in Figs 2-3 etc. So the positive bias (whether an error or not) in GSFC ice thickness compared to CPOM and AWI is not coming from the freeboard measurements, and must be from the conversion to ice thickness (snow depth, snow or ice density differences, or perhaps filtering/processing chain differences).

We agree that, on average, GSFC freeboard observations are 0.14 meters thinner than corresponding observations from AWI in spring, and 0.08 meters thinner in fall (Fig. 10), despite GSFC indicating higher sea ice thickness for both seasons compared to the AWI product, as shown in Figures 2 and 3. We have included this result in section 4.5, but can only speculate about the cause.

P17. L7. 'freeboards of less than approximately 0.05 m'.

Added the word 'approximately' to the line in question, in Section 5, paragraph 3.

P17. L 23. You suggest the assumption is that SMOS data is too heavily weighted in the combined product, but where do you get this assumption from? Are you making this assumption, or have you got it from previous papers? (if so which ones?). Potentially the uneven thickness increase detected by SMOS is more realistic and the CS2 products cannot detect new thin ice..?

We based this statement on existing studies where it is stated that SMOS assumes 100% ice concentration in its thickness retrieval algorithm, which may cause underestimation of ice thickness in conditions with less than 100% ice concentration, as mentioned in Tian-Kunze et al. 2014 and Ricker et al. 2017. It is certainly possible that SMOS estimates are more realistic for some conditions, most notably thin ice in areas of high ice concentration. We have revised the text of Section 5 (discussion) to reflect the reprocessed CS2SMOS data set used in the revised manuscript.

P18. L 5-8. Re. variations in ice concentration thresholds, you could have defined your own conservative threshold (e.g. 75

The focus of this study is in the assessment of publicly available data products as they are provided to the community. Furthermore, a variety of ice concentration products are used in the product processing chains. Choosing one of those, and applying it routinely across all products, would likely favour some products and hinder others. Table 1 provides further details of the ice concentration data and thresholds applied to each product.

P19. Snow section. This is quite out of place here, as you do not pay much attention to variations in snow treatment between products, and the focus of the paper is on inter-

comparison rather than the systematic issue of estimating snow (common between all products). I suggest removing this passage and if desired, reference the challenge in a single sentence or so.

We have added text in Section 2 about the treatment of snow in each data product, specifying the differences. Although all products use snow climatology estimates derived from Warren et. al. (1999), they employ different ways to implement the climatology. Based on this, the paragraph about snow in Section 6 has been modified and as suggested by Reviewer 2, it is now shorter and most of the text is moved to Section 5, the discussion.

Table 1. This is great! Really useful compilation.

Thank you.

Fig 5. Can you add the linear per winter ice growth rate for each product and year to the plot? i.e. X cm/month.

We expanded our results to include growth rates in a new table, Table 5.

Fig 7. Can you emphasize more clearly on the plot that the bold distribution is ice draft, as a quick glance at the figure gives one the instant conclusion that CS2SMOS data must be best.

Added bold text to the legends in Figure 7 so as to highlight "ULS Draft".

Fig 8. Can you add the x = 0.9y line to the plot as a reference?

The 0.9 factor is an estimate of the ratio between thickness and draft based on Rothrock et al. (2008), and therefore we do not want to emphasise it as the true ratio. Whereas for Figure 7, where the distribution shapes are compared, plotting also ULS draft is crucial.

We have modified the text in the third paragraph of Section 4.4 to emphasize the fact that 0.9 is an estimate of the draft to thickness ratio.

<u>Response to Review #2 (Thomas Armitage)</u>

Review of "Assessment of Contemporary Satellite Sea Ice Thickness Products of Arctic Sea ice" by Sallila et al.

The authors present an intercomparison of several different sea ice thickness data products, principally derived from CryoSat-2, and compare these datasets against BGEP upward-looking sonar draft data and Operation IceBridge airborne thickness data. Unsurprisingly, they find good agreement between the three CS2 products which use empirical retrackers (CPOM/AWI/JPL), while the GSFC data appears to overestimate the mean thickness early in the season and shows quite different spatial variability, and the CS2SMOS data potentially underestimates the thickness throughout the season. The analysis is straightforward, presented clearly enough and, while there are no new science results, I believe this paper could be useful to the community. I'm happy to recommend publication once my points below have been addressed.

We thank the Reviewer for this recommendation for publication.
Here we review a range of publicly available, satellite-derived sea ice thickness data products, and show, for the first time, a side-by-side comparison of the thickness estimates. We disagree that there are "no new science results": we discuss the spatial and temporal variations in sea ice thickness and growth rates across the Arctic Ocean between 2010 and 2018, thereby providing scientific analyses of 8 years of CryoSat-2 observations, which is a longer time-series of satellite-derived thickness observations than has been previously published. In addition, the evaluation of the satellite products using in situ data (airborne and ULS) spans a time period that exceeds any previously-published result (e.g. Laxon et al., 2013; Kwok and Cunningham, 2015, Wang et al., 2016, or Tilling et al., 2018).

I have a few suggestions that I think will improve the paper:

RC 2.1. Introduction: I found the introduction cumbersome. I would combine the first two paragraphs into one, and reduce the length by ~half as some of the material you discuss here is not really relevant for the current manuscript. Most of the fifth paragraph (Page 3, lines 12-27) is unnecessary in the intro, particularly as you essentially repeat it all in section 2, so remove – the first sentence can go at the end of the previous paragraph, and combine this with the final intro paragraph.

We have modified the introduction (Section 1) according to this suggestion.

RC 2.2. Why don't you use the EM bird data? I believe it is a fairly extensive dataset, is commonly used by groups to assess CS2 thickness data, and is the only technique that directly measures thickness rather than a proxy (draft/freeboard).

We agree that the AEM data are very useful for assessment. However, one of our requirements for both the satellite sea ice thickness data products, as well as the validation data sets, is that they are publicly available to the user community. Unfortunately, we have not found a publicly available version, and the source information for the AEM data used in the previous studies you mention was not provided. It would have been possible to acquire

1

the relevant data from one of the data providers examined here (AWI), as they kindly offered it, but such data do not currently fit the requirements of open access and may have raised concerns about the independence of the study, since it comes from one of the data providers. While additional evaluation against EM bird data might potentially provide further insights, we conclude that its absence does not impact the results shown in our study, nor is the presented evaluation of the satellite data compromised in any way.

RC 2.3. You do not discuss the impact of the different treatments of snow on your product intercomparison. There is a paragraph on the general difficulty of snow in sea ice altimetry in the conclusions (Page 19, lines 14-34) – I don't think this belongs here, I would move it to the discussion – but I would also like to see some discussion on how the different snow treatments impact your results. Importantly, you say that all the CS2 data use the 'modified' Warren climatology (Table 1), however I believe that the CPOM processing uses the basin-mean value of the Warren snow depth over the central Arctic region for MYI and half this value for FYI (at least this is my understanding from Tilling et al. (2018)), whereas I believe AWI/JPL use the spatially varying 'modified' Warren snow depth. I don't know if you realized this but it is not clear in the manuscript? [It might be worth checking with Tilling et al.] This could explain some of the differences in Figure 2 (AWI/JPL overestimate thick ice, underestimate thin ice relative to CPOM) as well as why the AWI and JPL maps look quite similar (they treat snow the same, but different to CPOM), and some of the differences in the in-situ comparison.

Thank you for bringing the Tilling et al. (2018) publication to our attention. While it was not available at the time of writing, it is now published, and we are happy to add this reference to the manuscript. Based on personal communication with the data provider we confirmed that the snow in the CPOM product is developed according to your description above.
We added details about snow treatment in the product descriptions in Section 2.1 and corrected Table 1 accordingly. We now describe how snow is treated per product in Section 2. We have also shortened and modified the paragraph about snow (page 19 lines 14-34 of the original paper) and moved it to the discussion section.

RC 2.4. The APP-x data is very poor, and I don't really see what value it brings to the paper, other than to say that it is very poor (you could be more unequivocal about this). It shows almost no interannual or spatial variability, and you find that the mean thickness grows by 1m between Feb and Apr (page 13 line 10) which is physically unbelievable. After showing the APP-x data in Figures 2 and 3, maybe Figure 4, I would consider dropping it from the rest of the manuscript and say in your conclusions that it is unrealistic and shouldn't be used over a CS2 product. I was very surprised that you suggest (page 19, line 8) users should use the APP-x data simply because it is available daily in NRT and Arctic-wide! Judging from your analysis, given that APP-x does not capture interannual or spatial variability, users would be better off simply using the monthly CS2 climatology because at least it captures the spatial variability and does not grow unrealistically thick ice at the end of winter.

Our goal is not to make decisions or assumptions about how potential end-users will use these data products. Rather our goal is to provide an independent assessment of the currently available satellite sea ice thickness data products so as to allow end users to make reliable and informed decisions, based on their "use case". The APP-x product is an

operational data product, meaning that it is derived twice daily, provided in a routine data format, and is available year-round in a continuous data stream with consistent latency from a national data centre. Moreover, the APP-x ice thickness data is part of a climate data record maintained by the National Oceanic and Atmospheric Administration (NOAA). Because this is the only operational ice thickness data product currently available it is of wide interest to the community who rely on such observational inputs delivered in an operational setting. We have provided a balanced summary of both the advantages and disadvantages of the APP-x product in the second paragraph of the conclusions. However, we also acknowledge that none of the data products currently available meet the general needs of the end user, and, in the final paragraph of the manuscript text, we provide a set of recommendations for improvements.

Minor comments:

It's not clear what you mean by "contemporary" in the title – occurring at the same time, or occurring in the present? Either way I don't think it's necessary; consider revising: "An assessment of satellite-derived Arctic sea ice thickness data"

By contemporary we mean data that belong to the present. The word "contemporary" is a widely used adjective in scientific publishing and we use it here to indicate "current" data sets. This is to alert the reader to the fact that we do not consider historical satellite thickness datasets that are otherwise good or interesting, such as from ICESat or Envisat.

Page 1, line 19-20: Should read "Among the data compared, the blended..."

Corrected.

Page 2, line 26: Re-write: "The most widely-used thickness datasets are derived from the radar altimeter..."

Sentence revised.

Page 2, line 29: "..retrack*ing*..."

Corrected.

Page 2, line 32: "*...basin-scale* gradients..."

Added "basin-scale".

Page 3, line 4: "Given the variety of sea ice thickness data..."

Corrected.

Page 4, line 8: You should reference Tilling et al. (2018) as well as Laxon, as the data you are using was produced by Tilling et al and is somewhat different from the original Laxon data.

The publication by Tilling et al. (2018) was not available at the time of writing. However, as it is now published, we are happy to add this reference to the manuscript in Section 2.1.1 and, where applicable, we have updated the Tilling et al., 2016 references to Tilling et al., 2018.

Page 5, line 9, Table 1: The CPOM processing uses separate retrackers for leads (Gaussian+Exponential model fit) and floes (threshold), and they apply a correction to account for this.

We added details about the separate approaches taken for ice floe and lead retracking in the CPOM product description in Section 2.1.1 and we modified the description in Table 1.

Page 5, lines 24-29: I don't think it is true to say that the AWI processing "does not differ significantly" from the CPOM processing. They are 'similar', but a number of important differences I can think of from the top of my head: 1) different retracking approaches (I believe AWI apply a 40% or 50% threshold retracker to all waveforms), 2) different sea level interpolation, 3) different waveform discrimination criteria (e.g., right and left sided peakiness), the MSS which you mention, different treatment of snow (see my comments above).

We have modified the lines in question, in paragraph 2 in Section 2.1.2. We added more detail about the AWI processing chain in Section 2.1.2, and we modified Table 1 for the CPOM product following the description in Tilling et al. (2018), page 1211, Section 4.2.2.

Page 6, line 9: I think Kwok and Cunningham (2015) is missing from the bibliography? Please check all references in the text actually appear in the bibliography.

Thank you for pointing this out. We added Kwok and Cunningham (2015) to the references and we checked that all of the other references were complete.

Section 3.1: Calculating a pixel-by-pixel correlation after oversampling the data to 5km doesn't make sense, as it may artificially increase the correlation stats. I would suggest sampling all the data onto an identical 25 km grid for the analysis. Also, why use a 50km search radius for the interpolation when the data are posted on 25 km grids? Surely this will act to smooth the data, also improving the correlation?

We thank the Reviewer for their comment which is consistent with RC1.1 and RC3.2. The gridding approach has been modified. We have clarified this in the methods, Section 3.1. We have also revised all text where the results of the gridded data are discussed. Please see further details in our response to RC1.1.

Page 9, line 13: You should provide some justification for using the CPOM data as your baseline, it seems rather arbitrary.

We choose the CPOM data product as the baseline against which to compare the other products partly because it was the first CryoSat-2 sea ice thickness data made publicly available and is thus the most widely used product. Also, according to the number of citations, the method is the best known with roughly four times the citations for the

Page 11, lines 5-9: The phrasing here jarred a bit, as it suggests that the "cooler summer = thicker ice in 2013/14" result is a "finding" of yours which was "also noted" by Tilling et al. (2015). Obviously, you haven't shown any linkages between temperature and thickness, and I know you didn't mean to imply this, so please rephrase this section to say that "Tilling et al found that... The thickness changes in Figure 2 are consistent with their result."

We modified the text in the first paragraph of Section 4.1 as suggested.

Page 11, line 25: The AWI CS2 data appears to show extensive thin ice (<0.5m) in fall but this is the noise floor of CS2, so not necessarily believable – is there something in the AWI algorithm that causes this? I would like some more discussion/explanation of this.

In response to Reviewer 3, we modified the range and increments of the colour scale used in Figure 2, to more clearly show the thickness gradients. We have also modified the text in the first paragraph of Section 4.2 to more appropriately describe the data shown in Figure 2.

Page 14, lines 1-10, Table 3: You should calculate the anomalies relative to the same baseline period to make the numbers comparable i.e., 2011-2015 considering this is the common baseline.

Thank you for this suggestion. We have modified the analysis so that now the baseline period is 2011-2015, and all anomalies are calculated relative to the mean over that period. We updated the results in Table 3 and the associated interpretations in the text, which occur mainly in Section 4.3, paragraph 4. Also, the term "climatological mean" has been changed to "baseline mean".

Page 14, line 16: "CS2-only products" – judging from figure 7, all of the satellite products miss the thickest/thinnest ice including CS2SMOS?

We agree, and we have modified the text in Section 4.4, paragraph 2.

Page 14, lines 20-21: by what measure does the AWI product most closely align with the ULS draft? Likewise, by what measure does the GSFC data least agree? I would say the CPOM/AWI/JPL show the same agreement from figure 7?

Our statements are based on the agreement between modal draft and modal thickness (assuming a ratio of 0.9), the width of the distributions (full width half maximum/standard deviation), the shape of the thickness distributions, and how well the distributions capture both the distribution of thin ice <1.5 m, and the decay in the distribution for thicker ice >1.5 m. The results shown in Figure 7 have been updated through the addition of ULS draft and satellite thickness data for the 2016-2017 season and the use of the revised CS2SMOS data set. Consequently, we have revised our interpretation of the draft to thickness comparisons and results. We have modified the text in Section 4.4, paragraph 2 to more adequately describe the comparisons.

Section 4.5, paragraph 1, Figure 10: I would think the negative freeboards in the GSFC data must have something to do with the different retracking, as this is the major difference between the data products? I would speculate that the GSFC floe retracking might be sensitive to off-nadir scattering, as they use a functional fit to the entire waveform and power in the trailing edge could skew the fit to later delay times.

Thank you for this very interesting insight. We agree that the abundance of negative freeboards in the GSFC product is anomalous compared with the other data products and may be a result of the empirical retracker. We also agree that it would be insightful for the community to be able to understand the sensitivity of the CryoSat-2 freeboard estimates to the specific retrackers. This analysis is unfortunately outside the scope of our current study which is targeted towards end users of the existing satellite thickness data products, rather than an assessment of retracker methodology and sensitivities

Response to Review #3 (Robert Ricker)

Review of Assessment of Contemporary Satellite Sea Ice Thickness Products for Arctic Sea Ice by Sallila et al.

Summary

The paper "Assessment of Contemporary Satellite Sea Ice Thickness Products for Arctic Sea Ice" evaluates and compares 6 different sea ice thickness data products that are publicly available, including data from CryoSat-2 (CS2), SMOS and the Advanced Very-High-Resolution Radiometer (AVHRR). For the evaluation, they use independent ice draft and thickness measurements obtained from the Beaufort Gyre Exploration Project (BGEP) upward looking sonars (ULS) and Operation IceBridge. The authors find that products utilizing CS2-only measurements are reliable for sea ice thickness between ~0.5 m and 4 m, while the merged CS2-SMOS product was the most reliable for thin ice. In contrast to the other products, the AVHRR dataset does not seem to represent a reliable ice thickness distribution at the end of the winter season.

I think this study is potentially very useful for the user community of those sea ice thickness data sets, since so far, such a review of current sea-ice thickness data sets is not existing to my knowledge. It is well written, has a clear structure and generally is easy to follow. It presents useful information bundled in one paper. Table 1, for example, is great as an overview. However, I also find that the paper lacks some crucial information regarding the products and the way the data are analyzed. I also have the feeling that there are some inconsistencies in the data analysis. My 3 major concerns are:

RC 3.1. I find that the description of the sea-ice thickness data sets in section 2.1 is incomplete. The AWI and the CPOM processing is quite different in some aspects, e.g. retracking of ice and lead waveforms. I think this should be described in more detail. See therefore also my comments below.

Thank you for pointing out that the AWI and the CPOM processing is quite different. Text in Sections 2.1.1 and 2.1.2 has been substantially revised to present additional details of the CPOM and AWI processing chain. We have also revised Table 1 to clarify the details of the CPOM retrackers.

RC 3.2. I am not really sure about the way the original data are re-gridded for the comparison. The paper lacks motivation regarding the chosen values for the re-gridding resolution of 5 km and the search radius of 50 km. I am also not sure if this is really the best way to do it, see detailed comments below. In any case, I suggest to revise section 3.1 in order to motivate your chosen method.

Thank you for your comment on gridding which is consistent with a comment from Reviewer 1 (see RC1.1). We have changed the gridding approach used to calculate correlation between data products and we have clarified this in the methods, Section 3.1. We have also revised all text where the results of the gridded data are discussed. Please see further details in our response to RC1.1.

RC 3.3. It also seems strange that CS2SMOS underestimates ice thickness in the MYI zone compared to the AWI product. CS2SMOS uses the AWI product for merging with SMOS, but in MYI areas, SMOS should not have a significant impact. See also my detailed comments below.

See our response to SC1. We discovered that there was an issue related to data used in the first version of our manuscript, where some corrupt data files were used accidentally. The results for CS2SMOS have changed now that we are using the correct data product files, as shown in Figure RC3.2 below. In agreement with your comment, the thickness of the MYI region in the CS2SMOS data product now resembles that of the AWI CryoSat-2 product.

Detailed Comments

P1 L18-19: "... are reliable for sea ice [thickness] between ~0.5 m and 4 m" - it should be mentioned that you are talking about thickness.

Added "thickness" on page 1, Abstract.

P2 L26: "One of the most widely used thickness data sets derives from the radar altimeter flown on CryoSat-2 ..." - this sentence sounds a bit odd to me.

Altered this text.

P4 L11-12: Please mention the version of the CS2SMOS weekly ice thickness product.

To avoid being repetitive and keeping the introductory paragraph of 2.1 compact, we decided to include the details of the product versions in the product descriptions that appear in the subsections. Specifically, please see Section 2.1.5, where we added the version number (v1.3) for the CS2SMOS data product.

P5 L8-9: I don't think that the CPOM processor uses the TFMRA algorithm (AWI). They also use a different retracking method for lead returns: a gaussian function is fitted to the waveform, as far as I know. See therefore, Tilling et al. (2017) - "Estimating Arctic sea ice thickness and volume using CryoSat-2 radar altimeter data".

Thank you for pointing this out. This has been corrected in Section 2.1.1. and also in Table 1.
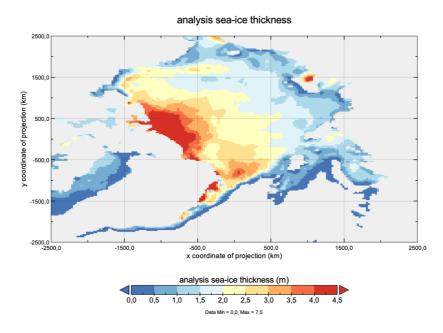
analysis sea-ice thickness

Figure RC3.1: CS2SMOS Analysis Thickness from 20150323-20150329, version 1.3.

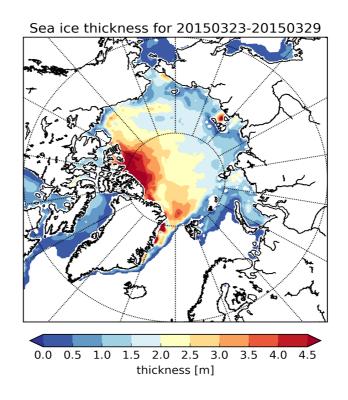
Sea ice thickness for 20150323-20150329

Figure RC3.2: Response to comment RC 3.3 and Figure RC3.1: Revised mapping of sea ice thickness from the CS2SMOS data product, version 1.3, for the week 20150323-20150329. The revised map is consistent with Figure RC3.1 and illustrates that data in the files we are now using in the revised manuscript are consistent with those from the data provider.

P5 L18: Please provide the version number of the AWI product. Is it 1.2?

The version of the AWI data product used in our manuscript is 2.0. This detail was added to Section 2.1.2.

P5 L18-28: Here it should be mentioned that the TFMRA is used. Moreover, I do not agree with the statement "The algorithm employed by AWI does not differ significantly from that used in the derivation of the CPOM product". While I agree that there are some similarities in the processing and also the derived ice thickness distributions look similar, some significant differences exist. For example, AWI uses the same retracking for lead and sea ice echoes, while CPOM does not (see above). Moreover, the surface type classification is different. CPOM also applies a retracker bias correction (Tilling et al. (2017).

Thank you for this clarification and please also see our response to RC3.1. We have substantially revised the data product descriptions in Section 2 and we now include details of the TFMRA retracking algorithm in the second paragraph of Section 2.1.2.

P7 L16: What is the reason to include the APP-x data here? They do not seem to represent the entire Arctic sea ice thickness distribution and ice growth during the season (see Figure 2 a,b). Then, one could argue to also include the SMOS sea ice thickness product.

Regarding the APP-x data product, please see our response to RC2.4.

Regarding the SMOS-only sea ice thickness data product: While we agree that the SMOS data product differs to the ones presented in the manuscript, it is only reliable for studies of thin ice <1 m thick (Ricker et al., The Cryosphere, 11, 1607–1623, 2017). Thus it is not representative of overall ice thickness in regions 1-6 of the Arctic, which is the focus area of our study. Also, for the sake of readability we limit the comparison to six data products. We believe that by including the CS2SMOS data set in our analysis, which differs from the CS2-only products and highlights the most relevant aspects of the SMOS observations, it provides the most useful comparison and insight for the reader.

P9 L9-11: Why do you choose a 5 km grid and a 50 km search radius? What happens at the ice edges or at the coasts. For example, areas, where the original grid contains NaN's in case of open water. Does this mean that you obtain ice thickness estimates on your new grid, where there were NaN's before, if valid ice thickness grid cells are found within the 50 km search radius? Wouldn't that erroneously enlarge the ice area? If there are more than one grid cell found within the search radius, are you only considering the closest one? Or is there some weighting applied?

We agree with the Reviewer that the gridding approach used in the original version of the manuscript (a 5 km grid with a 50 km search radius) was not appropriate for comparing the data sets against each other. We examined the results of the original gridding approach and found issues similar to those you describe, in particular in the Central Arctic region (at the northern limits of region 1), and also around the ice edge. We have therefore changed our approach and now follow the method originally introduced in Laxon et al. (2013), and used

again in Tilling et al. (2018), to conduct the product vs. product comparisons and to calculate the correlations (Figures 4 and 9). In the revised manuscript we use a 0.4° latitude by 4° longitude grid. We have revised the relevant areas of the manuscript text (in particular, Section 3) to clarify this change. Please also see our response to RC1.1.

P10 L3: Why do you choose such a large radius (200 km)? Especially in the Beaufort Sea, where those moorings are deployed, you may have very mixed ice regimes with some chunks of MYI surrounded by FYI. Wouldn't it be better to only choose the nearest one or at least a smaller radius?

We used a radius of 200 km, since it is consistent with the original evaluation presented in Laxon et al. (2013) and subsequently in Kwok and Cunningham (2015). This allows the reader to place our results in the context of the conclusions of prior studies.

P10 L16: Why are the IceBridge data interpolated on a 50 km grid? Why not 5 km as the re-gridded satellite products?

We have revised our methodology to calculate the correlations between the satellite data products and the IceBridge data. In order to place our results in the context of the existing literature, we have adopted the same approach as that of Laxon et al. (2013), who used a 0.4° latitude by 4° longitude grid to compare IceBridge and CryoSat-2 thickness estimates. This approach accounts for the uneven spatial and temporal sampling of the sea ice along the IceBridge flight-lines compared to the monthly means obtained from the satellite data products.

P11 L15-16: I am a bit confused about the CS2SMOS ice thickness maps. Here, it seems that they consequently underestimate ice thickness. Especially in MYI areas, the thickness should be very similar to the AWI CS-2 Product, since this is used for the data merging, and SMOS data should have almost no impact in the MYI zones. When I plot the CS2SMOS thickness for the week 2015-03-23 - 2015-03-29 (see Figure R1 (now Figure RC3.1), data from Meereisportal, version 1.3), using a similar color scale with the switch from light blue to yellow at 2 m, the MYI tail in the Beaufort Sea appears yellow, indicating thickness above 2 m. In your map it seems to be below 2 m. Of course, your map shows the average from March - April, but I would assume that this is not much different.

Your confusion is well justified! Based on your remarks, and comment SC1, we re-examined the CS2SMOS data and found that the CS2SMOS dataset we had used in the original manuscript was incomplete, and some corrupted data files were accidentally used. For this revision, we have corrected the issue. We downloaded all of the CS2SMOS version 1.3 data from Meereisportal again, and we can confirm that the thickness data are now consistent with that shown in Figure RC3.1 (see Figure RC3.2 above, and Figure 2 in the manuscript).

Figure 2: I really would recommend to use a different color scale with a linear color gradient, starting at 0.0 m. Moreover I would also suggest to use a finer resolution, e.g. 0.25 m. 0.5 m is too coarse from my point of view.

We revised the colour bars used in Figures 2 and 3, to start at 0 m, and we changed the increment from 0.5 m to 0.25 m according to this suggestion.

P12 L12-14: As mentioned above, it seems a bit strange that CS2SMOS is significantly thinner than AWI CS-2 in the MYI regions, since it uses the AWI CS-2 data, while the SMOS ice thickness estimates are not valid in MYI areas.

We have corrected this. See response to SC1 and RC3.3.

P17 L2-3: See above.

We have corrected this. See response to SC1.

Figures 5: The CS2SMOS data seem very noisy and show some strange behavior, e.g. strong decreases in mean ice thickness in some months. This should be checked. It does not seem to occur in Ricker et al (2017): "Satellite-observed drop of Arctic sea ice growth in winter 2015–2016". Although they show sea-ice volume, I would expect similar behavior.

We have corrected this. See response to SC1.

Figure 5 and 6: Over which area are these averages calculated? Regions 1-6 as indicated in Figure 1?

Yes, the averages are calculated for the regions 1-6 as defined in Figure 1. While we had indicted this in the figure caption for Figure 6, we had omitted the information in the caption for Figure 5. We have thus modified the text of the figure caption for Figure 5.

Figure 6: "a)" and "b)" are missing in the figure.

We have added the pertinent letters to Figure 6.

Assessment of Contemporary Satellite Sea Ice Thickness Products for Arctic Sea Ice

Heidi Sallila[1], Sinéad Louise Farrell [2], Joshua McCurry[3] ~~Joshua McCurry[2,3], Sinéad Louise Farrell[2,3]~~ and Eero Rinne[1]

[1]Finnish Meteorological Institute, Helsinki, Finland
[2]Cooperative Institute for Climate and Satellites, University of Maryland, College Park, MD, USA
[3]Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, USA
~~[2]Cooperative Institute for Climate and Satellites, University of Maryland, College Park, MD, USA~~
~~[3]NOAA Laboratory for Satellite Altimetry, College Park, MD, USA~~

Correspondence to: Heidi Sallila (heidi.sallila@fmi.fi)

Abstract. Advances in remote sensing of sea ice over the past two decades have resulted in a wide variety of satellite-derived sea ice thickness data products becoming publicly available. Selecting the most appropriate product is challenging given end user objectives range from incorporating satellite-derived thickness information in operational activities, including sea ice forecasting, routing of maritime traffic, and search and rescue, to climate change analysis, longer-term modeling, prediction, and future planning. Depending on the use case, selecting the most suitable satellite data product can depend on the region of interest, data latency, and whether the data are provided routinely, for example via a climate or maritime service provider. Here we examine a suite of current sea ice thickness data products, collating key details of primary interest to end users. We assess eight years of sea ice thickness observations derived from sensors onboard the CryoSat-2 (CS2), Advanced Very-High-Resolution Radiometer (AVHRR) and Soil Moisture and Ocean Salinity (SMOS) satellites. We evaluate the satellite-only observations with independent ice draft and thickness measurements obtained from the Beaufort Gyre Exploration Project (BGEP) upward looking sonars (ULS) and Operation IceBridge (OIB), respectively. We find a number of key differences among data products, but find that products utilizing CS2-only measurements are reliable for sea ice thickness, particularly between ~0.5 m and 4 m. Among ~~those compare~~data compared, a blended CS2-SMOS product was the most reliable for thin ice. Ice thickness distributions at the end of winter appeared realistic when compared with independent ice draft measurements, with the exception of those derived from AVHRR. There is disagreement among the products in terms of the magnitude of the mean thickness trends, especially in spring 2017. ~~We noticed disagreement among the products in mean thickness trends especially in the winters of 2012–2013 and 2016–2017.~~ Regional comparisons reveal ~~large~~ noticeable differences in ice thickness between products, particularly in the ~~Kara Sea, an area~~marginal seas in areas of considerable ship traffic.

1

1 Introduction

With the observed decline in Arctic sea ice extent (Parkinson and Cavalieri, 2008; Markus et al., 2009; Perovich et al., 2018) and interests in the exploitation of regional natural resources, human activities in the Arctic have increased, alongside concerns for the state of the ice cover. Numerous objectives, ranging from sea ice forecasting and climate monitoring to navigation, require observations of Arctic sea ice, and it is recognised as an Essential Climate Variable (ECV) in the Global Climate Observing System (GCOS, Belward and Dowell, 2016). Given its remote location, seasonally-available sunlight, and inhospitable climate, remote sensing provides the only means to obtain routine, basin-scale, and sustained observations of the Arctic Ocean. ~~Objectives for exploiting satellite derived observations of the polar oceans are broad, and include improving sea ice forecasts, monitoring trends and variability in the polar climate system, advancing long term climate predictions and future planning, understanding changes in marine habitats for wildlife and fisheries, as well as supporting operational activities, including navigation, hazard monitoring, disaster response, search and rescue, and many commercial activities. Because of these objectives, interest in routine measurements of sea ice thickness has increased, and is concurrent with the availability of new satellite data products that are the direct result of advances in remote sensing over the past few decades.~~

~~Sea ice is a key component of the Earth's climate system, acting as an insulator between the ocean and atmosphere, and its importance is recognised by the Global Climate Observing System (GCOS) through its status as an Essential Climate Variable (ECV) (Belward and Dowell, 2016). Sea ice thickness is an important sub-variable of the sea ice ECV.~~ Although sea ice extent ~~has~~ is traditionally ~~been~~ the most widely discussed~~more exploited~~ variable, sea ice thickness measurements are just as~~, if not more,~~ important~~than extent~~, and ~~are~~ needed together with ice concentration to calculate sea ice volume, the best indicator of ~~the~~ change~~s~~ in the Arctic ice cover (e.g., Laxon et al., 2013; Song, 2016). Even if ice extent ~~would~~ remain~~s~~ stable in ~~the same for~~ consecutive years, if the thickness decreases, the ice cover will be less resilient and more unlikely to regain thickness, eventually leading to decreased extent and volume. ~~Furthermore~~For example, Laxon et al. (2013) speculated that lower ice thickness, and volume, may ~~be~~ have been a contributing factor ~~behind~~ to the ~~recent~~ September 2012 record minimum~~s of~~ sea ice extent. ~~One way to understand the complex interactions of sea ice and its sub-variables in the climate system is through models, and in turn models initialized with sea ice observations can be used to forecast future ice cover conditions. Recently satellite estimates of sea ice thickness have been used for the first time in model initialisations, revealing great improvements in model predictions (Yang et al., 2014; Allard et al., 2018; Xie et al., 2018), and these studies emphasize the importance of careful selection of the initialisation product. There have been other model studies with observational products from non-satellite sources such as airborne campaigns (e.g., Lindsay et al., 2012). However this type of data has restrictions in spatial and temporal coverage, and needs interpolation to emulate Arctic wide coverage.~~

As a result of advances over two decades in remote sensing of Arctic sea ice, a wide variety of satellite sea ice thickness data products have become available to the scientific community. Radar altimeter measurements from CryoSat-2 (Wingham et al., 2006) and laser altimeter measurements from ICESat and ICESat-2 (Markus et al., 2017) are major sources for estimating sea ice thickness, and due to high-inclination orbits, provide nearly full coverage of the Arctic Ocean (Laxon et al., 2013; Markus et al., 2017). One of the most widely used thickness data sets derives from the radar altimeter flown on CryoSat 2, the only satellite currently operating that provides nearly full coverage of the Arctic Ocean (Laxon et al., 2013). Sea ice tThickness is derived from carefully processed radar altimeter measurements of ice surface elevation (Laxon et al., 2013; Kurtz et al. 2014), which requires knowledge of snow loading, and the assumption of hydrostatic equilibrium (Giles et al. 2007). Because of differing approaches to for retracking radar altimeter waveforms, a variety of processing algorithms exist, resulting in an array of altimeter-derived thickness data products (e.g., Laxon et al., 2013; Kurtz et al., 2014; Ricker et al. 2014; Price et al. 2015; Tilling et al., 2015, 2016, 2018). These products broadly agree in the spatial distribution and basin-scale gradients of ice thickness across the Arctic Ocean, but differ in their absolute magnitude. In addition to the CryoSat 2 only ice thickness products, there are other productsThickness may also be derived from passive microwave radiometer measurements, as well as from visible and synthetic aperture radar imagery, that and these observations offer additional information such as coverage in the marginal ice zone or detection of sea ice conditions during the Arctic summer (e.g., Kaleschke et al., 2012; Key and Wang, 2015; Ricker et al., 2017a). Recent studies have demonstrated the utility of initialising numerical models with satellite-derived estimates of sea ice thickness to improve model predictions (Yang et al., 2014; Allard et al., 2018; Blockley and Peterson, 2018; Stroeve et al., 2018; Xie et al., 2018).

Given the variety of sea ice thickness data products that are now available, and broad observational objectives, ranging from advancing long-term climate predictions and future planning, to supporting operational activities, including navigation, commerce, hazard monitoring, search and rescue, and disaster response, identifying the most suitable product depends on the end user requirements. Without knowledge, means, or time to compare the different products, the end user may not select the product most appropriate for their needs. WhichDetermining which satellite data product is the most suitable appropriate depends upon a variety of factors, including the end user's region of interest, and data product characteristics including spatial coverage, temporal and spatial resolution, accuracy and quality, as well as data availability and latency.

In this paper we review a set of publicly-available satellite-derived sea ice thickness products, and compare them side by side for the first time, outlining their key attributes which are of interest to potential end users and for a range of applications. We evaluate sea ice thickness estimates derived from observations collected by CryoSat 2 (CS2), the Advanced Very High Resolution Radiometer (AVHRR), and the Soil Moisture and Ocean Salinity satellite (SMOS). Although previous studies have evaluated individual satellite products (e.g. Laxon et al., 2013; Ricker et al., 2017b; Wang et al., 2016) and assessed the impact of retracker differences (e.g. Xia and Xie, 2018; Yi et al., 2019), none have focused on product intercomparison and have therefore lacked the details sought by many end users. To address this need, We we compare data product attributes

and assess differences across products ~~both~~ from both a regional and seasonal perspective, across the central Arctic Ocean and peripheral seas (Figure 1). We also evaluate the satellite data products, through comparisons with independent thickness measurements obtained in situ. ~~There are several airborne and in situ observations suitable for validating the satellite measurements. Flight campaigns, such as Operation IceBridge (OIB) and CryoVEx, complement, and are used to evaluate, NASA's Ice Cloud and Elevation Satellite (ICESat) and CryoSat 2 observations, respectively. Comprehensive datasets covering a wide variety of Arctic variables are available from field measurements obtained during research cruises like the Norwegian young sea ice expedition (N-ICE2015), or the Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAIC), planned to commence in October 2019. Also, a good set of observational products are available from North Polar drifting stations, ice mass balance buoys, upward looking sonars (ULS) that form part of the Beaufort Gyre Exploration Project (BGEP) and airborne EM measurements gathered through the Canadian Arctic Sea Ice Mass Balance Observatory campaign (see for example further information in Ricker et al., 2017b and Allard et al., 2018). For our evaluations we make use of OIB and ULS observations, which provide thickness and ice draft, respectively, since these variables are most compatible with the satellite data, and their records are more extensive and sustained over the period of interest, compared to many of the other publicly available in situ data sets.~~

The paper is structured as follows: ~~section~~ Section 2 introduces the sea ice thickness products and the independent data sets used for validation. Section 3 describes the methods used for product comparison. In ~~section~~ Section 4 we present the results of the comparisons across satellite products, and a product evaluation against independent estimates of ice thickness obtained from in situ and airborne sensors. We provide a summary discussion in ~~section~~ Section 5. In ~~section~~ Section 6 we conclude with a look ahead and provide recommendations for future satellite-derived thickness products aimed at addressing operational needs.

2 Data

2.1 Satellite Data Products for Arctic Sea Ice Thickness

Here we assess six contemporaneous sea ice thickness data products, derived from satellite measurements collected by CryoSat-2 (CS2), the Advanced Very-High-Resolution Radiometer (AVHRR), and the Soil Moisture and Ocean Salinity satellite (SMOS). Since our focus is an assessment of product utility for a range of operational activities we required that data products were open access, had basin-wide coverage of the Arctic Ocean (Figure 1), and were available for the majority of the CS2 observation period (April 2010 to present)~~time period under investigation~~. Four ~~of these~~ products comprise CS2-only thickness estimates, and include the Centre for Polar Observation and Modeling (CPOM) seasonally-averaged thickness product (Laxon et al., 2013, Tilling et al., 2018), the Alfred Wegener Institute (AWI) ~~version 2~~ monthly thickness product

4

(Ricker et al., 2014), the NASA Jet Propulsion Laboratory (JPL) monthly thickness product (Kwok and Cunningham, 2015) and the NASA Goddard Space Flight Center (GSFC) 30-day thickness product (Kurtz and Harbeck, 2017). For comparison with CS2-only thickness estimates, we also consider blended CS2 - Soil Moisture Ocean Salinity (CS2SMOS) weekly ice thickness data (Ricker et al., 2017a), and the NOAA Climate Data Record (CDR) Extended Advanced Very High Resolution

5 Radiometer (AVHRR) Polar Pathfinder (APP-x) daily ice thickness product. Further details about each data product are provided below. Common characteristics including measurement technique, temporal and spatial coverage, latency, frequency, resolution, and algorithm-specific details, are outlined in Table 1. The products selected for assessment provide a representative sample of available sources. We ~~do however~~ acknowledge that the list of satellite-derived ice thickness products is not exhaustive and other sources of similar observations exist. For example two additional sources of CS2-only

10 sea ice data are now publicly available, one from the ESA Climate Change Initiative (CCI, Hendricks et al., 2018) and a second from the Laboratoire d'Études en Géophysique et Océanographie Spatiales (LEGOS) Center for Topographic studies of the Ocean and Hydrosphere (CTOH, Guerreiro et al., 2017). Details of these two products are included in Table 1 for completion but not included in the analysis. Sea ice age characteri~~s~~zation, and ice thickness, from the Visible Infrared Imager Radiometer Suite (VIIRS) instrument on the Suomi National Polar-orbiting Partnership (S-NPP) satellite, may also

15 be derived in a similar manner to the APP-x data product (Key et al., 2013). ~~These products were not included in this study since~~, however these data were not publicly available at the time of writing, ~~or required prior registration and approval with the data product provider~~, and are hence not included in the analysis. Since our focus is the availability of current sea ice thickness measurements for end users, our study period ~~ranges~~ spans the last eight years from fall 2010~~–~~ to spring 2018, facilitating an assessment of recent sea ice conditions in the Arctic Ocean. Thus we do not consider satellite thickness

20 records prior to the launch of CS2, such as from Envisat (e.g. Giles et al., 2008) or ICESat (e.g. Kwok et al., 2009).

2.1.1 CPOM

CPOM was first to produce publicly-available estimates of sea ice thickness from CS2 and they provide both near real time (NRT) thickness products for 28, 14, and 2-day observation periods (Tilling et al., 2016), which are updated on a daily basis

25 with a typical latency of 3 days, as well as a monthly and seasonally-averaged thickness data product~~s~~ (~~Laxon et al., 2013~~Tilling et al., 2018) ~~that is available for two seasonal periods only (October–November and March–April~~, Table 1). Archived data coverage begins in ~~March 2011~~November 2010, and ~~new seasonal~~monthly averages are available on an ad-hoc basis. Thickness data are available for the months of October through April on a 5 km grid for the full northern hemisphere (regions 1–14, Fig. 1), as well as on a 1 km grid for specific subpolar regions.

30

~~The basis for estimates of ice freeboard, and ultimately thickness, in the CPOM data product is level 1b CS2 waveforms, available from the European Space Agency (ESA) that~~CryoSat-2 radar waveforms delivered through the ESA Level 1b product are used to identify lead and ice floe elevations (~~Laxon et al., 2013~~Tilling et al., 2018). ~~The mean scattering horizon~~

of each waveform is identified through a threshold-first-maximum-retracking algorithm (TFMRA) with the threshold set at 70% of the power of the first maximum (Laxon et al., 2013). Radar waveforms associated with leads and ice floes are distinguished using fixed criteria for stack standard deviation and pulse peakiness (Laxon et al., 2013; Tilling et al., 2018). Leads and ice floes are distinguished using fixed criteria for stack kurtosis, stack standard deviation, and pulse peakiness. Ice elevation is defined as 70 % of the peak amplitude on the leading edge of the first peak in the radar waveform (Tilling et al., 2018). Lead elevations are determined separately (Tilling et al., 2018) through application of the retracking method developed in Giles et al. (2007) where a Gaussian plus exponential model is fit to each waveform. The CPOM algorithm utilizes the UCL13 mean sea surface (MSS) for the calculation of sea surface height anomalies, and to reduce the impact of geoid slope on freeboard estimates. This step is especially important in areas of low lead fraction (Skourup et al., 2017). Auxiliary information including the location of the ice edge and a product distinguishing first-year ice (FYI) from multi-year ice (MYI) is needed in the CPOM thickness algorithm. The sea ice edge is defined as the 75 % ice concentration contour in the For sea ice concentration the NSIDC Near Real Time DMSP SSMI/S Daily Polar Gridded Sea Ice Concentrations Concentration data set (available from the National Snow and Ice Data Center, NSIDC), while are used and the sea ice type is derived from the EUMETSAT's Ocean and Sea Ice Satellite Application Facility (OSI-SAF) product (Tilling et al., 2016 2018). The depth and density of snow on sea ice are based on the snow climatology of Warren et al. (1999) and applied by ice type. On MYI, the monthly mean climatological value is applied, and this value is halved for FYI (Kurtz and Farrell, 2011), meaning constant snow depth and density values for MYI and FYI are applied each month (Tilling et al., 2018). At the time of writing, only seasonal averages for two seasonal periods (October-November and March-April) were available, and this data version was used as the baseline data set in our analysis. Undefined thickness estimates are indicated by a value of 0.0000 in the CPOM product and these were removed before further processing.

2.1.2 AWI

AWI also provides monthly CS2 data products for October through April. Archived data begins in November 2010 and new monthly data are made available on a variable basis, but typically with one month latency (Table 1). In addition to thickness, AWI offers a number of additional geophysical and instrument parameters in their data product, including sea ice freeboard and concentration. It Of the products analysed here, it is the only product that provides uncertainty estimates for thickness and freeboard values (Hendricks et al., 2016). All parameters are provided on a 25 km grid for all regions of the Arctic (i.e., regions 1–14).

The AWI product uses the ESA L1b baseline C CryoSat-2 radar waveforms, applying a 50 % Threshold-First-Maximum Retracker (TFMRA, Ricker et al., 2014) to derive elevation for all surface types. The algorithm employed by AWI does not differ significantly from that used in the derivation of the CPOM product, although In contrast to the CPOM product, AWI currently uses the DTU15 MSS for the calculation of sea surface height anomalies (Hendricks et al., 2016). Skourup et al. (2017) demonstrated that processing with either the DTU15 or UCL13 MSS results in consistent freeboard estimates, with

only small deviations in areas of low lead concentration. For both sea ice concentration and type, AWI uses OSI-SAF products (Hendricks et al., 2016), and the ice edge is defined at the 70 % ice concentration contour. The AWI product applies snow depth based on the climatology of Warren et al. (1999), following the method described in Laxon et al. (2013), hereafter referred to as the "modified Warren climatology" (MWC). In the MWC, snow depth on FYI is 50 % of the climatology, and snow density is a monthly constant per ice type. The AWI product allows mixed ice types at the boundaries between MYI and FYI, based on confidence levels given in the ice-type product (Ricker et al., 2014). For this analysis AWI product version 2.0 was used and ice concentration, freeboard and thickness were assessed.

2.1.3 JPL

The JPL product provides monthly CS2 thickness data from January 2011 to December 2015, on a 25 km grid (Table 1). Coverage is limited geographically to the central Arctic (regions 1–6). Information on the ice concentration threshold defining the area within which sea ice thickness data are calculated is not provided. The JPL retracker for deriving surface elevation is based on the waveform centroid, rather than a leading-edge approach used in the AWI TFMRA algorithm. The JPL freeboard algorithm also uses the EGM2008 geoid, rather than a MSS model, which can result in anomalous sea ice thickness estimates in areas of steep ocean topography, such as near the Lomonosov and Gakkel Ridges (Skourup et al., 2017). Information about which sea ice concentration product is applied in the JPL algorithm is not provided in the available literature, nor the ice concentration threshold used to define the ice edge. Kwok and Cunningham (2015) state that the sea ice type is derived from analysed fields of Advanced Scatterometer (ASCAT) data. The data are currently only accessible to the public after registration on a product website. Snow depth and density in the JPL product is based on the MWC and applied similarly to Laxon et al. (2013), except that where Laxon et al. (2013) used a 50 % reduction in the climatology over FYI, Kwok and Cunningham (2015) examined both a 50 % and 70 % reduction, it was left unclear which version was selected for the final online product. Information about which sea ice concentration product is applied in the JPL algorithm is not provided in the available literature, but Kwok and Cunningham (2015) state that the sea ice type is derived from analysed fields of Advanced Scatterometer (ASCAT) data. The data are currently only accessible to the public after registration on a product website. Undefined thickness values corresponding to 9999.0 and -1.0 were removed before processing.

2.1.4 GSFC

GSFC provides 30-day ice thickness averages derived from CS2. Coverage begins in October 2010 and new data continue to be made available on a time-varying basis, but typically with a six week latency (Table 1). In addition to ice thickness, the GSFC product includes estimates of freeboard and surface roughness (Kurtz and Harbeck, 2017). All parameters are available on a 25 km Polar Stereographic SSM/I Gridgrid, for regions 1–8.

Unlike other CS2 freeboard algorithms, the GSFC product is derived using a waveform-fitting method ~~based on empirical calculation of expected returns from various surface types~~. The surface type, elevation, and other properties of the received radar waveform are derived through statistical comparison to analytically pre-computed waveforms, ~~Received waveforms are matched to a pre-calculated waveform by means of~~ using a least squares error minimization. Kurtz et al. (2014) note that this approach should nominally result in freeboard values lower than those derived from TFMRA-based methods. ~~Additionally, t~~The GSFC algorithm utili~~s~~zes the DTU10 MSS for freeboard calculation. Skourup et al. (2017) have demonstrated that more recent MSS models (e.g., UCL13, DTU15), ~~that~~ which incorporate sea surface height data from CS2, enhance the definition of gravity features, resulting in a more accurate freeboard derivation. Indeed Skourup et al. (2017) found that the DTU10 MSS in particular was not sufficient for freeboard processing due to decimeter-level discontinuities at 81.5°N and 86°N as well as at the ice edge, which resulted in erroneous freeboard measurements at these locations. The GSFC product uses a 70 % ice concentration threshold derived from the NSIDC Near-Real-Time DMSP SSMI/S Daily Polar Gridded Sea Ice Concentrations available at NSIDC to define the sea ice edge, and the OSI-SAF product is used for sea ice type (Kurtz and Harbeck, 2017). An additional unique attribute of the GSFC algorithm is the use of a single ice density value (915 kg/m3) for all ice types in the thickness derivation step (Kurtz et al., 2014). In the case of the other CS2-only thickness products a dual ice density approximation is made, with the assumption of a lower density for MYI (882~~.0~~ kg/m3) than for FYI (91~~6.~~7 kg/m3). Despite using a single density for both ice types, the GSFC algorithm employs the MWC for~~a~~ snow depth ~~correction to the Warren et al. (1999) climatology for seasonal ice, as described in Laxon et al., (2013) and determined from OIB snow depth measurements provided in Kurtz and Farrell (2011). This approach is consistent with all other CS2 thickness products (Table 1). For sea ice concentration GSFC uses the NSIDC Near-Real-Time DMSP SSMI/S Daily Polar Gridded Sea Ice Concentrations and the OSI-SAF product for sea ice type (Kurtz and Harbeck, 2017)~~. Here we used the GSFC product version 1.0 ice freeboard and thickness data. Undefined thickness estimates are indicated by a value of -9999 in the GSFC product and these were removed before further processing.

2.1.5 CS2SMOS

The CS2SMOS sea ice thickness product, developed by AWI and the University of Hamburg, is a blended product of thickness estimates from CS2 and SMOS. It provides weekly data for the Arctic northward of 50 °N on an EASE2 grid, with 25 km grid resolution, across regions 1–14. It is available for a period starting in ~~January 2011~~November 2010, ending April 2017.

The SMOS mission provides L-band observations of brightness temperature, which may be used to derive ice thickness in areas where thin sea ice exists (Kaleschke et al., 2012). CS2 exploits radar altimetry to measure the difference in height between the snow/ice surface and sea surface, which is then used to derive sea ice thickness through the assumption of hydrostatic equilibrium. Since CS2 was designed to measure ice thicker than 0.5 m, it may be advantageous to blend CS2 estimates with complementary estimates from SMOS. Due to the satellites having different spatial and temporal coverage,

optimal interpolation is used to merge the two data sets (Ricker et al., 2017a). The algorithm includes weighting the data based on the known uncertainties of the products and modelled spatial covariances (Ricker et al., 2017a; 2017b). For sea ice concentration and type the OSI-SAF Arctic daily products are used. Snow depth and density follow the MWC. The CS2 product used is the AWI CS2 product with processor version 1.2 (Ricker et al., 2014; Hendricks et al. 2016) and the SMOS thicknesses are from the University of Hamburg processor version 3.1 (Tian-Kunze et al., 2014; Kaleschke et al., 2016). For sea ice concentration and type OSI-SAF Arctic daily products are used. For our analysis thickness data from CS2SMOS product version 1.3 were used.

2.1.6 APP-x

The NOAA extended AVHRR Polar Pathfinder (APP-x) Thematic Climate Data Record (CDR)product provides sea ice thickness estimates, along with 18 other geophysical variables, in a climate data record (Key and Wang, 2015). Thickness estimates are available for both the Arctic Ocean (regions 1-14) and Southern Oceans, spanning 1982 to present. Data are provided twice daily, with a typical latency of approximately 4 days (Table 1). In contrast to the other satellite-derived thickness products, yYear-round thickness estimates are available, including throughout the summer, and are provided on a 25 km grid. We note a gap in the thickness record at the time of writing for the period 8 March to 1 May 2017 and hence no assessment of APP-x data exists forin spring 2017 in the following analyseswas not possible.

Sea ice thickness estimates are derived from AVHRR satellite radiometer measurements using a one-dimensional thermodynamic model (OTIM), based on a surface energy budget equation (Key and Wang, 2015). The OTIM model derives sea ice thickness as a function of surface heat fluxes, surface albedo and radiation, which all contribute to surface energy budget (Wang et al., 2010). Furthermore, most of the flux and radiation parameters in the equations are functions of surface skin and air temperatures, surface air pressure, surface air relative humidity, ice temperature, wind speed, cloud amount and snow depth, which are input parameters in the model (Wang et al., 2010). Ice thickness is calculated as a function of solar radiation and surface heat fluxes, surface albedo, and transmittance, where the surface fluxes are calculated as a function of surface skin and air temperature, surface air pressure, relative humidity, ice temperature, wind speed, cloud cover and snow depth. Thus OTIM relies on the following input for the ice thickness calculation: percentage cloud cover, surface skin temperature, surface broadband albedo and surface shortwave radiation fluxes, of which the latter two are obtained for daytime retrievals only. The sea ice concentration source for the APP-x product is Nimbus-7 SMMR and DMSP SSM/I data processed with the NASA Team Algorithm (Key and Wang, 2015) and the ice edge is defined at the 15 % ice concentration contour. The sea ice type is converted from the reflectances measured directly by AVHRR. Additionally the following input is needed for ice thickness: percentage cloud cover, surface skin temperature, surface broadband albedo and surface shortwave radiation fluxes, of which the latter two are obtained for daytime retrievals only. For the APP-x product the snow depth estimates are based on the snow depth climatology of Warren et al. (1999) but combined with field observations through experimentation and applied using monthly look-up tables (Wang, pers. comm., 2018). For our analysis thickness

data from APP-x CDR version 2.0 were used, and undefined ice thickness values identified as "9.96920996839e+36" were removed before processing.

Since sea ice dynamics are not included in the OTIM model, the thickness errors in the APP-x product are larger where the ice surface is not smooth, i.e. in regions with pressure ridges, hummocks and melt ponds (Wang et al., 2010). Generally OTIM tends to overestimate ice thickness, in particular for thin ice while underestimating thick ice, as the energy budget approach is less sensitive for thick ice (Wang et al., 2010). ~~Then again, APP-x tends to underestimate thick ice, as the energy budget approach is less sensitive for thick ice.~~ Moreover since the satellite sensor retrieves 2 m air temperature, ice surface temperature is derived from the 2 m measurement. Wang et al. (2010) state that the thickness estimates are more accurate for nighttime retrievals, when 2 m air temperature and ice surface temperature are closer, resulting in a smaller model error. For this reason, we ~~chose~~ utilise the nighttime estimates ~~to be used~~ in our study. Wang et al. (2010) note that errors due to uncertainties in snow depth and cloud fraction are the primary sources of error in the OTIM thickness estimates. We also note that OTIM is applicable to other optical satellite data including observations from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) instrument and EUMETSAT's Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument (Wang et al., 2010). ~~The sea ice concentration source for the APP-x product is Nimbus 7 SMMR and DMSP SSM/I data processed with the NASA Team Algorithm (Key and Wang, 2015). The sea ice type is converted from the reflectances measured directly by AVHRR.~~

2.2 Evaluation Data Sets

We evaluate the satellite-derived ice thickness estimates using independent upward looking sonar (ULS) observations of ice draft from Beaufort Gyre Exploration Project (BGEP) moorings, and airborne observations of ice thickness from Operation IceBridge (OIB). These data sets represent the most extensive and sustained record over the evaluation period, compared to many of the other publicly-available in situ thickness data sets.

2.2.1 Beaufort Gyre Exploration Project

Since August 2003, ~~the Beaufort Gyre Exploration Project~~ (BGEP) has operated a series of moorings in the Beaufort Sea, which have included an ~~upward looking sonar~~ (ULS) instrument. From 2003 to 2014 the ULS instrument produced a range estimate every two seconds, increasing in frequency to once per second starting with the 2014–2015 deployment. By subtracting the ULS range estimate from instrument depth, draft is measured to an accuracy of +/- 0.05 m per individual measurement (Krishfield et al., 2006).

Here we utilize ULS draft measurements from three mooring locations (A, B, and D, Fig. 1) in the Beaufort Sea over the six-month period spanning November–April, for all years from 2010-2017~~2011–2016~~. Since the ULS measures ice draft as floes drift across the mooring location, the data represent a high-resolution, time-varying measurement of many individual leads

and ice floes, thus providing a more complete picture of the regional ice thickness distribution. The draft to thickness ratio is approximately 0.9 (e.g., Rothrock et al., 2008), but to accurately compute thickness from draft, knowledge of ice type, ice density and snow loading are required. Here we do not convert draft to thickness, since that would introduce additional uncertainties. Rather we use the characteristics of the ice draft distribution to evaluate the satellite-derived thickness distribution.

2.2.2 Operation IceBridge

Operation IceBridgeOIB was launched in 2009, and is a sustained airborne mission designed to continue the collection of sea ice and land ice elevation measurements in the temporal gap between the end of the ICESat mission in 2009 and the launch of ICESat-2 in 2018 (Koenig et al., 2010). The mission includes an airborne altimeter, the Airborne Topographic Mapper (ATM), which provides high-resolution measurements of sea ice plus snow freeboard, and a snow radar instrument for derivation of snow depth (Newman et al., 2014). Together these allow for estimation of sea ice thickness (e.g., Farrell et al., 2012; Kurtz et al., 2013). Thickness uncertainty, calculated by propagation of estimated errors in the contributing variables, associated with IceBridge estimates is approximately 0.66 m (Richter-Menge and Farrell, 2013). Here we make use of the IDCSI4 (spring 2011) and NSIDC-0708 (spring 2012–2017) IceBridge thickness products available at the National Snow and Ice Data Center Distributed Active Archive Center websiteNSIDC (Kurtz et al., 2015; Kurtz, 2016). Due to the geographical layout of airborne flight surveys, the majority of IceBridge measurements sample multi-year sea ice in the Canada Basin (e.g., Richter-Menge and Farrell, 2013).

3 Methodology

3.1 Satellite Product Intercomparison

In order to compare the satellite data products, all ice thickness datasets are regridded to a 5 km grid, using a nearest neighbour algorithm with a 50 km search radius. Seasonally-averaged ice thicknesses are is computed for each product over two periods: fall (October and November), and spring (March and April). Seasonal averages are calculated by taking the arithmetic mean value of all available thickness estimates across the Arctic region (Fig. 1) within the two month period. Results for fall and spring are shown for each product over the period 2011 - 2017 in Figures 2a and 2b, respectively. Seasonal averages are only calculated using grid cells that contain valid thickness estimates for both months. This prevents the average from being biased by measurements taken at the beginning or end of an averaging period (i.e., measurements only available in one of the two months). For each product we also calculate a climatological mean ice thickness in the central Arctic (Fig. 1), averaged over the product record (as described in Table 1). Climatological means are calculated separately for the fall and spring seasons. For each season we then compute deviation from the climatological mean for each product.

To evaluate spatial variations in thickness within products, mean regional ice thickness is computed for the central Arctic Ocean and each peripheral sea (Table 2) using the polygons shown in Figure 1, for both the fall and spring seasons. Thickness values are reported for any region containing valid data points in the relevant months for at least one year of the product record. Although all regions with data are reported, regions outside the main product coverage areas defined in Table 1 may contain only a few data points and hence the reported regional ice thickness for some of the peripheral seas may represent only a small portion of the region as a whole (see Fig. 2).

To assess temporal variations across products, we calculate a baseline mean ice thickness for the central Arctic for the period common to all products (2011-2015), for both the fall and spring seasons. Subject to the time span over which each product is available (Table 1) we compute the anomaly with respect to the baseline mean for each season and report these results in Table 3. We also assess trends in winter ice growth for each product by calculating monthly mean ice thickness in the central Arctic (regions 1-6) during the period of study spanning 2010 to 2018 (Figure 5).

For point-to-point comparisons of the satellite data products, and to compute correlations across products, all ice thickness datasets are placed on a common 0.4° latitude by 4° longitude grid. We define the CPOM seasonally-averaged thickness product as the reference dataset against which the other products are compared. This approach is justified since the CPOM product was the first publicly-available CryoSat-2 thickness dataset (Laxon et al., 2013), and has been widely used in studies by end users (e.g. Allard et al., 2018; Blockley and Peterson, 2018; Stroeve et al., 2018). We therefore compute product differences as follows:

$$\Delta SIT = SIT_{cP} - SIT_{pC} \;, \tag{1}$$

where SIT_c is the CPOM sea ice thickness, SIT_p is the thickness of the product in question, and $\mathbf{\Delta}SIT$ is the difference between the two. Correlation statistics are calculated utilising grid cells in which both data sets contain thickness estimates. These grid cell pairs are used to compute the Pearson correlation coefficient (r), the product difference across common cells according to Eq. 1, and the standard deviation of this difference (Figure 4).

To evaluate spatial variations in thickness across products regional ice thickness averages are computed for the central Arctic Ocean and each peripheral sea (Table 2) using the polygons shown in Fig. 1, for both the fall and spring seasons. Due to the limited ice extent in fall, the availability of observations in some peripheral seas is not sufficient to calculate a robust regional average in October–November (or in the spring season in the case of the JPL and GSFC products due to their limited spatial coverage).

3.2 Satellite Product Evaluation

Since the BGEP moorings are tethered to the seafloor, ULS measurements are representative of ~~all~~ discrete ice floes drifting over the mooring location. To facilitate comparison with the satellite data, we select all product data points within 200 km of the mooring, following the approach described in Laxon et al. (2013), creating a comparison region centred on the mooring location. In order to avoid influence from areas outside the comparison region, we use the original satellite thickness data as provided, rather than the gridded data described in Section 3.1 above.

Both the ULS draft data and satellite thickness data within the comparison region are averaged over one month intervals from November to April for the period of overlap between the product record and ULS data spanning fall 2010 to spring 2017. The correlation coefficient for ULS draft measurements and product thickness observations is calculated at each mooring based on these paired monthly data points (Figure 8) ~~between each monthly-averaged remote sensing product and the average ULS measurements for same period is calculated for each mooring~~. Any individual ULS draft measurement thinner than 0.1 m is not included in the averaging, as these measurements may represent leads rather than ice floes (Krishfield et al., 2006; Krishfield et al., 2014). In the case of the CPOM product, ~~only seasonally averaged thickness data are available, so~~ correlations with the ULS ice draft are based on seasonal averages, processed in a similar manner to the~~rather than~~ monthly averages as described above.

The sea ice thickness data from Operation IceBridge consists of flightlines across the Canada Basin and the central Arctic Ocean (e.g. Richter-Menge and Farrell, 2013). Following Tilling et al. (2018)~~Laxon et al. (2013)~~, data acquired during IceBridge spring campaigns in 2011-2017 were placed on a 0.4° latitude by 4°longitude grid~~interpolated onto a 50 km grid~~ and compared against the satellite thickness data for March–April, for all six products ~~interpolated~~ at the same~~common~~ grid cell locations~~, using a nearest neighbor algorithm with a 100 km search radius~~. This allows the calculation of thickness differences between each product and IceBridge, as follows:

$$\Delta SIT_{eval} = SIT_{\sout{oibp}} - SIT_{\sout{poib}} \ , \tag{2}$$

where SIT_{oib} is the Operation IceBridge sea ice thickness, SIT_p is the thickness of the product in question, and $\boldsymbol{\Delta} SIT_{eval}$ is the difference between the two. Using the gridded data, we also calculate correlation coefficients, and standard deviation between each product and the IceBridge thickness observations (Figure 9).

4 Results

Here we present, for the first time, a side-by-side comparison of a suite of available CS2-only products, alongside a blended CS2-SMOS (CS2SMOS) product and one altimetry-independent sea ice thickness product, APP-x. The results are presented

13

in three parts. First we provide a review of Arctic sea ice thickness variability during the last eight years. Next we compare regional and temporal differences between the satellite products across the Arctic regions. Finally we evaluate the satellite-derived thicknesses through comparisons with independent measurements.

4.1 State of the Arctic Sea Ice Thickness

5   Seasonal ice thickness for fall and spring is shown in Fig. ure 2 for the period 2011–2017. Following the observed low summer sea ice minimum extents in 2011 and 2012 (Parkinson and Comiso, 2013), we find that the lowest ice thickness was recorded in fall 2011 and 2012 (Fig. 2a). Mean ice thickness was 1.12 m (1.19 for the CS2-only products) in fall 2011, with variations of +0.3 to -0.23 m (+0.24 to -0.11 for CS2-only) among the products. For fall 2012 the mean thickness was 1.18 m (1.19 for CS2-only), varying from +0.34 to -0.36 m (+0.24 and -0.11 for CS2-only) among products. The loss of multi-year

10   ice in the summer of 2012 due to the record sea ice minimum resulted in an overall thinner ice cover during the following winter/spring (i.e. spring 2013) when the mean ice thickness was 2.05 m (2.0 m for CS2-only) among products, which is visible in (Fig. 2b). Consistent with the results first noted in Tilling et al. (2015), we see that following Following a cool summer in 2013, survival of ice through the melt season resulted in a rebound in thickness in fall 2013 with mean thickness of 1.46 m (1.58 for CS2-only), and a thicker winter mean sea ice thickness of 2.34 m in spring 2014 (all products), which has

15   persisted in the central Arctic for subsequent seasons (Fig. 2). This finding is consistent with earlier results shown in Tilling et al. (2015), who also noted a slight recovery in 2013–2014 following low winter-time ice thickness in 2011 and 2012. The thickest ice at the end of winter was observed in 2014 (Fig. 2b) in a region stretching from northern Greenland near Kap Morris Jesup, to Banks Island in the Canadian Arctic Archipelago. This region of thick ice in the central Arctic has persisted throughout the following seasons. As of spring 2018 (not shown) the area of ice more than 3 meters thick adjacent to the

20   northern coasts of Greenland and the Canadian Arctic Archipelago is still greater than that of Spring spring 2012 and 2013, although the spatial extent of this thick ice area has diminished since 2014. In multiple springs (2012, 2014, 2015, 2016) an outflow of thick ice extends from the southern Canada Basin into the southern Beaufort Sea, due to the dynamic action of the Beaufort Gyre circulation (Fig. 2b). While this band of thick sea ice is captured in all of the CS2 products, it is not as easy to identify less distinct in the CS2SMOS product, and does not appear in the APP-x product, apart from spring 2012 (Fig. 2b).

25   This overall picture of the state of Arctic sea ice thickness over the last several years is consistent across the CS2-only products and the CS2SMOS blended product. Although the APP-x product captures the spatial gradient in thick to thin ice, from the northern coasts of Greenland and Canada to the Siberian coastline, respectively, the product does not resolve many of the recent major changes in sea ice thickness conditions in either the spring or fall.

4.2 Regional Differences

30   The satellite-derived thickness products differ in their regional coverage and the availability of thickness estimates across the northern hemisphere. APP-x has the most widespread coverage, although CPOM, AWI and CS2SMOS all provide thickness estimates in the sub-polar seas. The JPL product only provides estimates for an area approximately contiguous with regions

14

1–6, while the GSFC product provides estimates for an area approximated by regions 1–8 (Table 1, and Fig. 2). Three of the six products (AWI, CS2SMOS, and APP-x) consistently resolve thin ice (<= 0.5 m) at the periphery of the ice pack (regions 3–7) during the fall, but only CS2SMOS and APP-x resolve thin ice in these regions in spring (Fig. 2b). In addition APP-x does not resolve the thicker ice of the central Arctic Ocean, as evident in the other products, especially in the fall.

5

Maps of differences in ice thickness across products, as defined in Eq. 1., are shown in Fig. 3, and distributions of the differences between products are provided in Fig. S1. Average regional seasonal ice thickness for each region of the Arctic (as defined in Fig. 1) over the period 2011–2017available product record is provided in Table 2. In general, differences in ice thickness are larger in the fall than spring, although for spring even across the CS2-only products, differences range from
10 0.25 m to 2.12 m in regions 6-14 (Table 2). As expected, given strong consistencies in algorithmic approach, theThe closest agreement across products is found between the CPOM and ., AWI and JPL products (Fig. 3, 4, Fig. S1). The thickness estimates of these products have a correlation of 0.91-0.92 in the spring and 0.88 in the fall, with a mean difference of 0.03 - 0.08 m 0.88 for both seasons and a mean difference of 0.02–0.03 m (Fig. 4). However, there are noticeable regional differences in mean thickness during the observation period (Table 2), for example in spring in the Greenland Sea the AWI
15 product is 0.32 m thicker than the CPOM product, while in(0.42 m difference) and the Canadian Archipelago the CPOM product is 0.37 m thicker(0.53 m difference). From an Arctic-wide perspective, the mean thickness estimates in the JPL product also differ little from the CPOM product (0–0.02 m, Fig. 4) and are highly correlated with the CPOM product, 0.86 in spring and 0.83 in fall (Fig. 4). There are however also differences in the spatial gradients of ice thickness (Fig. 3), particularly in the central Arctic Ocean, where. For example the JPL product estimates thicker ice close to Greenland, and
20 thinner ice near the North Pole, and along the Siberian shelf zone than the CPOM product (see also Table 2). Of the CS2-only data products, Figure Figs. 3 and S1 demonstrates that the GSFC thickness product is the most dissimilar to the CPOM data, with thickness in both the fall and spring periods being higher on average, though with year-to-year spatial variation (Fig. 3). Mean ice thickness differences range from 0.02 to 0.250.08 to 0.27 m (Fig. 4), though despite these differences, the GSFC product is still highly correlated with the CPOM product (R = 0.85 in both spring and fall)., with correlations of 0.80
25 and 0.83 in the spring and fall, respectively (Fig. 4).

From a regional perspective, the GSFC product agrees closely with the CPOM product in spring (especially in regions 1-7), but larger thickness differences are recorded in fall across all regions (Table 2). The CS2SMOS product suggests thinner ice thicknesses than the CS2-only products, with mean differences of 0.08 to 0.36 0.27 to 0.60 m (Fig. 4) and differences
30 increasing towards the ice edge. For example in the Barents Sea, average spring thickness for the period 2011–2017 was only 0.380.41 m, over a meterexactly 1 m lower than the CPOM product suggests estimate (Table 2). Only in the MYI zone in spring does the CS2SMOS product provide estimates of thicker iceice thicker than the CPOM product in some years (Fig. 3b), though never for a whole region (Table 2). Despite its lower thickness estimates, the CS2SMOS product correlates well with CPOM, 0.86 and 0.88 in the fall and spring0.84 and 0.86 in the spring and fall, respectively (Fig. 4). We also note that

15

in fall 2017 the APP-x product shows thickness data covering an area south of the typical ice edge in regions 8–10 (Fig. 2a). Anomalous data in this region may be due to errors in the sea ice concentration field (not shown). Ice concentration is passed to the APP-x product from the lower-level APP product. In spring For ice in the central Arctic, in the APP-x product is consistently thicker in spring over FYI, and thinner across both the MYI zone and the thick outflow along the northern coast of Greenland and into the Canada Basin (Fig. 3b, Table 2). In the fall, the APP-x product contains mainly thinner ice for all regions (Table 2), except in 2011 in the Canada Basin, where the APP-x product suggests thicker ice than in the CPOM product, and in 2017 around the ice edge in regions 3–6 (Fig. 3a). Also, in fall, the APP-x product estimates thicker ice for regions 3-7, and 11, compared to the blended CS2SMOS product (Table 2). With correlations of 0.49 and 0.53 0.38 and 0.53 in the spring and fall, respectively, the APP-x thickness does data do not correlate as well with the CPOM product as the other data sources (Fig. 4). We also note that in fall 2017 the APP-x product shows thickness data covering an area south of the typical ice edge in regions 8-10 (Fig. 2a). Anomalous data in this region may be due to errors in the sea ice concentration field (not shown). Ice concentration is passed to the APP-x product from the lower-level APP product.

4.3 Differences in Winter-time Growth Rates

We now consider the winter-time growth rates across the central Arctic. The evolution of monthly mean ice thickness during winter is shown in Figure 5 for the entire study period spanning fall 2010 to spring 2018, and growth rates are provided in Table 5. The results are dependent on the product availability (Table 1), and in the case of the CPOM product, only seasonal means are available assessed. Monthly mean ice thickness (Fig. 5) in the central Arctic can differ by up to 1.2 m across products. As we might expect, the CPOM, AWI and CPOM JPL products are the most similar similar in terms of both the monthly mean trends in ice thickness and growth rates (Fig. 5, Table 5). While the maximum difference between the CPOM and AWI products is 0.1 m, the JPL product differs most noticeably from the AWI product by ~ 0.14 m in October - December 2013 (Fig. 5). Ice in the GSFC product is consistently thicker than the other CS2-only products, with ice thickness estimates beginning each season by up to 0.4 m thicker in fall, before converging towards the other CS2-only products by the end of the winter (Fig. 5). of the CS2-only products, having maximum difference of 0.1 m. Ice thickness in the JPL product differs very little from these two, most noticeably in January and February 2013 (0.2 m higher than AWI). Ice thickness in the GSFC product is almost constantly higher than the other CS2-only products, with fall sea ice thickness estimates beginning each season up to 0.4 m thicker, before converging towards the other CS2-only products by the end of the winter growth season. In December 2012 November 2010, February 2013 and January 2016, November 2016, and April 2018 the GSFC product indicates a small (maximum 0.1 m) decrease in thickness, whereas CPOM and AWI have a constant upward trend within during the ice growth period (Fig. 5). The daily ice growth rates are consistently lowest for the GSFC product (Table 5). Interestingly, for some early seasons, e.g. fall 2011, CS2SMOS estimates slightly higher mean sea ice thickness than the AWI product (Fig. 5). CS2SMOS has daily growth rates higher than those of GSFC, but lower than other CS2-only products, differing from the AWI product by a maximum of 0.001 m/day (Table 5). It is interesting to note that the CS2SMOS product also suggests a temporary decrease in ice thickness in December 2012 and February 2013. CS2SMOS is

16

consistently lower than CS2-only products for all months except October 2015 when it converges with the JPL product. Like GSFC, CS2SMOS does not exhibit a constant upward trend during the ice growth season, instead showing significant month to month variability. Indeed monthly mean sea ice thickness typically decreases from the previous month once or twice per growth season. APP-x has the highest growth rate (i.e. the smallest minimum and largest maximum thickness) exceeding the CS2-only products, and the ice cover can gain 2 m within the an ice growth season and up to 0.0108 m/day. The strongest increase in the APP-x mean thickness takes place at the end of winter between February and April, when the ice grows by up to 1 m (Fig. 5). This differs significantly from the winter-time evolution of the ice as shown in the CS2-only products, which suggest stability very little growth in ice thickness at the end of winter (March–April), with very little growth, likely due to the insulating properties of the overlying snow cover. There is agreement across all products containing CryoSat-2 data that the largest daily growth rates occurred in winter 2011-2012 (Table 5). Although the GSFC product has the lowest daily growth rate, it has the highest interannual variability in growth rate (0.0026 m/day), whereas the JPL product has the smallest (0.0005 m/day).

The year-to-year seasonal trends in central Arctic ice thickness (Fig. 6) are very coherent among the CS2-only products, with an increase in mean ice thickness in the fall between 2011 and 2014, followed by a slight decrease and leveling off (Fig. 6a).from 2011 to 2014, followed by a slight decrease and then a levelling off in mean thickness. In fall 2016 and 2017 (Fig. 6a) there is a minor divergence between the CPOM and AWI products, but both suggest a slight upward trend in fall ice thickness. The GSFC product in fall (Fig. 6a) shows a similar year-to-year trend in fall to the other CS2-only products, but is 0.3-0.50.2–0.4 m thicker on average (Fig. 6a). CS2SMOS follows the CS2-only products with maximum difference of -0.06 m compared with the AWI product and almost exactly the same mean ice thickness as the JPL product in fall 2011 and 2012 (Fig. 6a)appears to be about 0.3 m thinner than the CS2-only products, but follows a similar trend (Fig. 6a). The fall averages in the APP-x product are lower than for the other products by 0.2-0.6 m.and the(Fig. 6a), except in 2012, when the CS2SMOS product also suggests an average of ~1 m. The mean thickness trend in seasonal mean thickness of the APP-x product does not follow the onethose of the CS2 products, in particular for the 2012 and 2016 seasons, where the APP-x product hast an opposite trend (Fig. 6a).that have opposite directions for the trend line from the CS2 products.

In spring the central Arctic mean ice thickness differs little across the CS2-only products (Fig. 6b), with the strongest similarities again between the CPOM and, AWI and JPL products. The CS2-only products show a drop in thickness in 2012 and continuing to 2013, followed by a slight recovery in 2014, preceding another drop in mean thickness in spring 2015, which has persisted since then. The exception to the similar direction of CS2 products is in 2017 when the GSFC product suggests an increase in thickness, and is over 0.2 m higher than the other CS2-only products. By spring 2018 ice thickness in the GSFC product has decreased and is once again in line with the AWI and CPOM estimates, though slightly lower than the CPOM estimate (Fig. 6b). CS2SMOS product is very similar to the other CS2 products, the trend line being consistently approximately 0.2 m lower than the AWI one,. During this period the CS2SMOS mean is lower than the other products,

17

differing from the CS2-only products by up to 0.7 m. The mean thickness trend, however, follows that of the CS2-only products, except that the CS2SMOS product suggests that the recovery in thickness occurred in 2013 rather than in 2014 (Fig. 6b). APP-x estimates are higher than for the other products in spring with an almost constant mean thickness of ~2.5–2.6 m, and very little year-to-year variability (Fig. 6b). In 2014–2016 this is in line with the CS2-only mean (differing by only 0.2 m).

We have calculated the annual deviations from the mean thickness across the central Arctic for the baseline period 2011-2015, and present the results in Table 3. the overall central Arctic mean thickness for the period 2011–2017 (referred to as the climatological mean) and present the results in Table 3. The departure from the spring climatologicalbaseline means thickness in spring for theall CS2-only data products was 0.230.17 m in spring 2013 (0.12 m in 2012), and 0.200.17 m thicker in 2014, which is much in line with the thickness increase shown in Figure 6b. The CS2-only products are similar in the direction of annual departures from the climatological baseline mean except for GSFC in spring 2017, which shows a positive departure (0.05 m), whereas the CPOM and AWI product thicknesses are lower than their climatological baseline means (0.1 m and 0.08 m, respectively). A very similar pattern of departures can be seen for fall CS2-only thickness data, where there is a mean departure of 0.21 m in fall 2011, 0.13 m in 2012 and thickening of 0.18 m in fall 2013. The CS2SMOS deviations differ very little from the CS2-only product annual deviations, and are actually more in line with them in fall 2015 and 2016 compared to the GSFC product. CS2SMOS and APP-x shows no noticeable deviation from their its fall climatological baseline means, as previously observed in Figure 6b, although there is clearly more deviation in the fall compared to the spring. The changes in APP-x annual deviations follow those of the CS2 products in falls 2012-2015, as well as 2017, except for . For fall, APP-x shows again little variation, except for 2016, when it deviates by -0.21 m (see also Figure 6a). The record low ice thickness records in fall 2011 and 2012 (Fig. 2a) are also evident in the results shown in Table 3.

4.4 Comparison against Independent Observations

Next we consider the satellite-derived sea ice thickness products in the context of independent measurements from ULS and IceBridge to evaluate the utility of the satellite products for providing information on the full thickness distribution. As mentioned previously, the draft to thickness ratio is approximately 0.9 (Rothrock et al., 2008). Therefore we do not expect the modal thickness and draft to be equivalent, but we do expect the distributions to have the same characteristic shapes.

Histograms of the draft/thickness distributions are shown in Figure 7, and suggest that none of the CS2-only satellite products capture neither the thickest nor thinnest ice, although the CPOM, AWI and CS2SMOS products do have some observations of ice thickness below 0.5 m, with CS2SMOS performing best in this regard. This general result may be due to the fact that the monthly/seasonal satellite-derived products have been provided at 25 km resolution (Table 1), thus

potentially averaging out the thinnest and thickest satellite observations per grid cell. Modal ice draft is 1.3 m while modal ice thickness ranges 1.7 m to 2.3 m (Fig. 7). Comparing the characteristics of the draft/thickness distributions, including modal values and distribution width, the CS2SMOS, CPOM and AWI data sets most closely align with the ULS data. The JPL product distribution reveals slightly thicker ice than the CPOM and AWI products. The GSFC product has a bimodal thickness distribution for the study period, with modes at 1.7 and 2.3 m, and the APP-x product also has a modal ice thickness of 2.3 m~,~. This may in part be due to the fact that the satellite derived products have been averaged to 25 km, thus masking out the thinnest and thickest measurements per grid cell. Among the CS2 only products, the thickness distribution of the AWI product most closely aligns with the ULS draft distribution, while the GSFC thickness distribution exhibits the least agreement with the draft distribution. CS2SMOS thickness directly overlaps with ULS draft, implying that the CS2SMOS product underestimates ice thickness. APP-x has an exceptionally narrow thickness distribution that captures far less variation than CS2 thickness products and is inconsistent with the characteristic shape of the ULS ice draft distribution.

The correlations between the monthly-averaged, satellite-derived ice thickness and ULS ice draft are shown in Figure 8 for the months of November through April for each of the three BGEP moorings. The results are consistent with the approximation of a ~ 0.9 ice draft to thickness ratio (Rothrock et al., 2008). The correlations were calculated Figure 8 shows the correlation between satellite and ULS monthly averages, which were further combined across years 2010-2017 to aid visualization (e.g. November 2011 2010 – 2016 2017, December 2011 2010 – 2016 2017, etc.) with the following exceptions: JPL product data were assessed over the period 2011-2015, spring 2017 data were not available for the APP-x comparison, and seasonal averages were evaluated for the CPOM product. Correlation results for the CPOM, AWI and JPL products are strongest, followed by APP-x, GSFC and CS2SMOS (Fig. 8), but all satellite thickness products display very strong correlations with the ULS draft data, in line with previous results from Laxon et al., (2013), Kwok and Cunningham (2015), and Tilling et al., (2018). The exception to this is for ice drafts > 1.3 m, when there is an observable divergence in the results for the CS2SMOS and APP-x products. The APP-x thickness estimates are ~ 0.4 m higher than the CS2-only products, while the CS2SMOS data are 0.1-0.3 m thinner, suggesting a thickness overestimation in APP-x and an underestimation in CS2SOS, with respect to the thickest ice. in Fig. 8 to aid visualization. Since the comparisons in Figure 8 are again illustrating thickness against draft, data should optimally fall above the y=x line, when the ~0.9 draft to thickness ratio is considered. With this assumption, and considering the mean values as outlined in Table 4, AWI performs best, confirmed by high correlation values and with mean thickness values 0.2 m greater than draft. Although the best correlation values are recorded by the JPL product, the thickness values appear to be biased slightly high, when keeping the ~0.9 ratio as reference. Logically, the products with varying dependency to draft have lower correlation values, CS2SMOS being the lowest, although with relatively good correlation values. CS2SMOS thickness is roughly 10% higher than draft for ice around 1m, but is lower than draft when draft is greater than 1.3 m, which suggests underestimation of thickness in CS2SMOS estimates across this range. GSFC has comparatively low correlation values, and is roughly 0.5 m thicker than draft. APP-x thickness is increasingly greater than the ULS drafts with 0.5 m difference to draft of 1 m and 0.9 m to draft of 1.5 m, suggesting that it

~~is overestimating ice thickness at this draft range. APP-x hardly shows any deviation from its climatological mean at each mooring location (Table 4), but for other products years of ice gain and loss are visible.~~ The yearly means for ULS ice draft and satellite-derived ice thickness are provided in Table 4. The results suggest the mean ice thickness was largest in 2014 and 2015, and lowest in 2011 and 2012 (Table 4). In years with lower ice thickness, there is some disagreement between the ULS observations and the satellite products. 2013 appears to have the thinnest ULS ice draft results for the observation period, whereas the CPOM product suggests that this was one of the thickest years, while other CS2 products appear to have low ice thickness in 2013 only for buoy D.

Next, the satellite-derived thickness estimates for spring are compared with seven years of independent OIB thickness data (Fig. 9). The JPL product has the highest correlation (r=0.~~73~~76) with the OIB data and we~~. We~~ note that this is ~~slightly~~ higher than the correlation value of 0.~~56~~ 53 noted in Kwok and Cunningham (2015), who only considered data for March and April 2011–2012. Likewise, our correlation value of 0.~~67~~ 70 between CPOM and OIB thickness is ~~in line with~~slightly better than the results ~~from~~ shown in Laxon et al. (2013) and Tilling et al. (2018), who found ~~a~~ value~~s~~ of 0.61 and 0.67 when assessing CS2 against the 2011~~and~~-2012 and 2011-2014 OIB campaign~~s~~ data, respectively. Similar to the results observed with the ULS comparisons, the satellite products seem to be missing the thickest ice seen by the OIB measurements, but overall the agreement between the CS2 products and OIB is good. In terms of absolute differences, JPL ~~and GSFC~~ thickness estimates are ~~both~~ slightly higher on average than those of OIB by ~~0.07~~0.10 m, whereas CPOM and AWI estimates are lower by 0.11 m, GSFC by 0.05 m and CS2SMOS by 0.22 m (Fig. 9)~~0.14 m and 0.13 m, respectively~~. Although CS2SMOS differs twice as much from the OIB estimates compared to AWI and CPOM, for ice thickness between 0 m and 1.5 m the CS2SMOS estimates appear to agree best with OIB. Thus the greater difference would be explained by the thicker ice, where CS2SMOS estimates are lower than OIB. These findings for CS2SMOS are in line with previous validation studies (Ricker et al., 2017b) who evaluated CS2SMOS using observations from an airborne electromagnetic (AEM) induction thickness sounding device. APP-x has the smallest correlation of 0.54 and a peculiar vertical concentration of data~~line~~ in the scatterplot (Fig. 9), where a majority of APP-x sea ice thickness estimates fall into a thickness category between 2.25 and 2.5 m~~2.15 m and 2.25 m. Unlike other products, CS2SMOS produces thickness estimates that are significantly thinner than those of OIB, with an average difference of 0.66 m. Our findings, that for thinner ice, CS2SMOS differs less from the validation product, whereas for thicker ice it tends to underestimate thickness, are in line with previous validation studies (Ricker et al., 2017b) who evaluated CS2SMOS using observations from an airborne electromagnetic (AEM) induction thickness sounding device.~~

4.5 Ice Freeboard

Recall (from Section 2) that only two products provide freeboard estimates: AWI and GSFC. Figure 10 compares sea ice freeboard across these two products for spring and fall. Our assessment reveals a prevalence of negative freeboard estimates in the GSFC product that do not appear in the AWI product (Fig. 10). For the period October 2010 to April 2018, an average of 29.5 % of the freeboard measurements provided in the GSFC data product are negative, in contrast to 0.9 % of the

freeboard measurements in the AWI product. Negative freeboard estimates in the GSFC product often correspond with significantly higher freeboards in the AWI product for the same grid cell locations. An example of the GSFC freeboard product for April 2014 is shown in Figure 10c and highlights the spatial prevalence of anomalous, negative freeboard estimates, especially in the Kara and Barents Seas, where negative values persist throughout the winter. However, we note that negative freeboard estimates also occur in the Beaufort, Chukchi and Greenland Seas. This suggests that some negative estimates may be related to regional masking in the processing algorithm, or they could be due to the use of the DTU10 MSS in the ice thickness derivation (e.g. see Fig. 3 in Skourup et al., 2017).

Despite the high proportion of negative freeboard values in the GSFC product, it does not contain any negative thickness values. While GSFC freeboard observations are on average 0.08 m and 0.14 m thinner than corresponding observations from AWI in fall and spring, respectively (Fig. 10), the GSFC product has higher sea ice thickness for both seasons, and almost all regions, compared to the AWI product, as shown in Figures 2 and 3, and Table 2. Closer examination of individual data points in the GSFC product indicates that ice thickness in grid cells ~~with~~ containing negative freeboard values is not significantly lower than adjacent data points, suggesting that a filter may be applied to remove negative freeboard values before calculating ice thickness, and/or that ice thickness values are derived from interpolations across many grid cells.

We furthermore note that the AWI and CPOM data products are the only two data products that include negative sea ice thickness estimates. Approximately 0.8 % and 0.2 % of the thickness estimates are negative in the CPOM and AWI data products, respectively. The locations of negative ice thickness estimates for the month of April 2014 are shown in Figure S2, and are representative of the general pattern observed in other years of the study period. Negative data points are found along the ice edge (as defined by the ice concentration threshold of 70 % in the OSI-SAF product, plotted from the variable included in the AWI product, Table 1), suggesting that these thickness values are anomalous and are a result of edge effects in the sea surface height interpolation scheme.

5 Discussion

We expect thickness estimates across the CS2-only products to be similar, since their ~~only~~ primary difference~~s~~ are~~should be~~ due to the algorithmic approach and some of the auxiliary data inputs. These expectations are borne out as can be seen in the basin-scale maps shown in Fig~~s~~. 2 and 3. The CPOM and AWI products differ very little, while the JPL thickness estimates are also generally in close agreement. The GSFC product is among the products with thickest sea ice overall, particularly at the beginning of the growth season (Figs. 4, 5, 8), despite containing a very high percentage of negative freeboard values (Fig. 10). On the other hand, due to the inclusion of SMOS data, the CS2SMOS product is weighted for thinner ice, such that we expect overall thickness in this data product to be lower than in the CS2-only datasets(Fig. 2, Fig. 3), with a more realistic representation of areas with thin ice in the peripheral seas (regions 3-10).~~–~~ Since the APP-x product relies on a

thermodynamical model to derive thickness, Wang et al. (2010, 2016) state that the product is expected to perform best over level ice. We find that in fall, APP-x indeed has similar ice thickness to the CS2SMOS product, except over the thickest MYI in the central Arctic (Fig. 2a, Fig. 3a). In spring, however, APP-x appears to overestimate ice thickness across the entire Arctic Ocean (Fig. 2b, Fig. 3b). Despite similarities, there are also major regional differences, as seen in Figs. 2 and 3, and Table 2. Even among CS2-only products, ice thickness in the peripheral seas (regions 3-10) on average varies by 0.33 m in the fall and by 0.6 m in spring.

We have shown that the CS2-only satellite data products include reliable estimates for sea ice between ~0.5–4 m thick, depending on the product. In general, all satellite products capture a realistic winter-time ice thickness distribution, when compared to independent ice draft measurements, as demonstrated in Figure 7, with the exception of APP-x, which overestimates the sea ice thickness and underestimates the thickness variation~~does a very poor job at resolving thickness variations~~ in the Beaufort Sea. Ice in the Beaufort Sea is characterised by mixed amounts of deformed sea ice, which may partially explain the poor results from the APP-x product in this area. The CS2SMOS product has the best representation of the thin ice thickness, which would be expected as SMOS-based estimates are used in the majority of the thin ice regions (Ricker et al., 2017). ~~Although the width of the CS2SMOS thickness distribution suggests that the product captures a range of ice thickness in the region, it underestimates the absolute ice thickness (Fig. 7). This may be due to too strong a weighting on the SMOS observations relative to the CS2 data, and the impact is most clearly seen in the multi-year ice zone (Fig. 2), where the CS2SMOS data product is consistently lower than the CS2-only data sets.~~

However our results also show that no product adequately captures the thinnest sea ice in the thickness distribution at the end of winter (Fig. 7). The CS2-only products ~~also~~ do not resolve the thickness of sea ice less than 0.5 m thick, equivalent to freeboards of less than approximately 0.05 m (Figs. 4, 7, 9, 10), and the CS2SMOS product is the most reliable product in this regard (Fig. 7, Fig. 9). Ricker et al. (2014) found that uncertainties in sea ice thickness estimates are large for CS2 in areas where ice is less than 1 m thick. In contrast, the sensitivity is lost for SMOS when the ice is thicker than 1 m. While we have demonstrated that CS2-only products provide good results in the central Arctic ice pack, they lack robust estimates in some regions, particularly around the ice margins (regions 10-12), and in areas of new ice formation, where thin ice is expected. CS2SMOS, and APP-x to some extent, perform better in the peripheral seas. ~~The opposite argument applies to the CS2SMOS product.~~

The biggest difference in the temporal trend among the thickness products occurs in spring 2012 and 2013, when ice thickness was at its lowest during the observation period (Table 3, Fig. 5). The winter growth rate among the CS2-only products is very consistent, with the exception of the GSFC product (Fig. 5, Table 5). The GSFC product also shows a decrease in mean thickness between March and April in both 2016 and 2018, while all other products show an increase in thickness during the same periods. The CS2SMOS growth rate for March to April is lower than for the other products,

resulting in a mean ice thickness that is ~0.2 m lower than the CS2-only products. The decrease of the CS2SMOS growth rates could be due to the areas of less than 100 % ice concentration contributing to the growth in others, falling within the SMOS weighted area of the product. As stated by Ricker et al. (2017) and Tian-Kunze et al. (2014), SMOS assumes 100 % ice concentration in the thickness retrieval algorithm, which could cause underestimation of ice thickness in areas with lower concentration.. While the CS2 only products show a drop in average Arctic ice thickness following the record September 2012 sea ice minimum, the CS2SMOS product continues to suggest overall thickening between 2012 and 2014 (Fig. 2, Fig. 5). In addition both GSFC and CS2SMOS products suggest a decrease in mean ice thickness for the central Arctic in both December 2012 and February 2013 (Fig. 5, Fig. 6). During spring 2013 significant episodes of ice divergence were reported, with observations of pronounced lead signals in the multi-year ice zone (Willmes and Heinemann, 2015). This may be one reason why the CS2SMOS product departs from the trends shown in the CS2 only products, but further study is needed to fully understand the impact of ice dynamics on the CS2SMOS product and differences between it and the CS2 only products. The CS2SMOS product also shows uneven growth and decline in the mean ice thickness each winter that could be due to changes in the extent of thin ice, if the assumption holds true that SMOS is too heavily weighting in the blended product. Another period with deviation from the overall trends was the ice season between fall 2016 and spring 2017. In October-November 2016 a decrease in the mean ice thickness can be seen in the APP-x product, and to very small extent also in the CS2SMOS and CPOM products, whereas in March-April 2017 the GSFC products shows an increase in mean ice thickness. The APP-x product did not resolve any notable year-to-year variability in mean thickness. During the ice growth season the APP-x product shows the largest magnitude of ice growth, over 1.5 m between fall and spring (Fig. 6), with very little inter-annual variability (Fig. 2, Fig. 5). The largest growth in the APP-x product occurs between February and AprilJanuary and March, at the end of winter (Fig. 5), when in situ measurements typically show inhibited ice growth due to the insulating effects of the overlying snow cover. This suggests that thermodynamic assumptions in the OTIM algorithm for the end of winter may need further refinement.

Finally, we also note that all of the satellite-derived sea ice thickness products depend on additional auxiliary data sets (Table 1) in the derivation of ice thickness. As we have outlined in Table 1, there is great variation in the source of the auxiliary products and how they are used, particularly the mean sea surface model, ice type delineation and ice concentration. Detailed comparison of most use different sources for the mean sea surface model, ice type delineation, ice concentration, and the treatment of ice density and snow depth on ice. Comparing the auxiliary products is outside the scope of this study, but these could give rise to differences across products, in addition to the algorithmic differences. There is also a large range in the ice concentration threshold used to indicate the ice edge across products, varying from 15 % (in the APP-x and CS2SMOS products) to 75 % (in the CPOM product). With regards to the APP-x product we believe that an erroneous ice concentration threshold could be one possible explanation for the peculiar extent of the ice thickness estimates in fall 2017 (Fig. 2a). Additional differences across products may arise due to the treatment of ice density and snow depth on ice. Even though all of the satellite-derived thickness products assessed here make use of the MWC, they vary in their implementation method, as

23

described in Section 2.1 and none of the products resolve year-to-year variations in snow depth. Although there are multiple approaches proposed to obtain seasonal snow depth estimates on sea ice, they have yet to be routinely incorporated into a publicly-available, satellite-derived thickness data product. Potential solutions include utilizing model simulations (e.g., Blanchard-Wrigglesworth et al., 2015), atmospheric reanalysis data (e.g., Blanchard-Wrigglesworth et al., 2018), extrapolating in situ observations (e.g. Shalina and Sandven, 2018), or widely expanding the spatial and temporal coverage of current airborne measurement techniques (e.g. Kurtz and Farrell, 2011; King et al., 2015). Satellite passive microwave radiometer observations have also been used to derive snow on first-year sea ice (Brucker and Markus, 2013), as well as snow on thick ice (Maaß et al., 2013; Rostosky et al., 2018). One additional promising remote-sensing method is to combine two satellite altimeter observations retrieved at different wavelengths, enabling snow retrieval due to differences in penetration (Shepherd et al., 2018). For example this could be achieved through a combination of dual-band radar freeboard observations (e.g. Armitage and Ridout, 2015; Guerreiro et al., 2016; Lawrence et al., 2018) or by comparing freeboard measurements from laser and radar altimeters, to obtain an estimate of year-to-year changes in snow depth (e.g., Kwok and Markus, 2017).

6 Conclusions and Future Outlook

Satellite techniques have revolutionized our ability to measure the thickness of ice in the Arctic Ocean, providing critical information for scientists conducting studies of environmental change in the region, as well as a new source of data for forecasters, modelers, operators, and decision makers. Here we assessed a suite of existing satellite-derived, publicly-available, Arctic sea ice thickness data products, conducting a comprehensive examination of regional and seasonal differences over an eight-year period. As expected, the CS2-only products were similar, particularly at the end of winter (Fig. 6b). In March/April of each year2011-2016, APP-x reached a mean Arctic-wide thickness of ~2.6-2.7 m, which was thicker than any other satellite product (Fig. 5), and showed little to no inter-annual variability (Fig. 6). On the other hand, likely due to because of its inclusion of thin sea ice thickness, derived from passive microwave radiometer data, the CS2SMOS data product is on average 1-0.2 m thinner than the other satelliteCS2-only estimates at the end of the winter growth season. In fall, there was a larger spread in mean thickness across the products, and the GSFC thickness product diverged from the three other CS2-only products by approximately 0.2-~0.3-0.4 m (Fig. 6a). Evaluation of the satellite data products through comparisons with OIB and ULS measurements revealed that all products were well correlated with the independent ice draft/thickness estimates, with correlations of 0.54 and higher (Figs. 8 and 9). Five of the six products resolved an accurate winter-time sea ice thickness distribution for the Beaufort Sea when compared with ULS observations of ice draft, though with the AWI and CS2SMOS data sets produced producing the most robust results (Fig. 7). The APP-x data product did not resolve ice thickness variability in this region, and was biased thick compared to both the ULS ice draft observations and the alternative satellite thickness products. However in fall, APP-x sea ice thickness estimates in the peripheral seas (regions 3-10) resembled those of CS2SMOS (Fig. 2a). Our study revealed some other remarkable

differences across the products utilizing CS2 data: there were occasional reductions in mean ice thickness during the winter growth season in the GSFC product, and it diverged from the other CS2 products in winter 2016-2017 by approximately 0.3 m (Fig. 5), despite the prevalence of negative freeboard estimates in this product (Fig. 10). ~~Both the GSFC and CS2SMOS data products suggest reductions in Arctic mean ice thickness during the ice growth season. There was an unexpected increase in mean ice thickness in spring 2017 in the GSFC product that differed from the other CS2 products. There was also an upward mean ice thickness trend in spring 2013 in the CS2SMOS data product, when the CS2 only products showed a continued decrease in mean thickness.~~ Such anomalies require further study to evaluate their actual causes, and this can be accomplished for example through more detailed, along-orbit comparisons between CS2-derived data products and ~~with~~ coincident observations collected during ~~OIB and CryoVex~~aircraft underflights (e.g., Connor et al., 2011).

In terms of end user applications, the suitability of a particular product depends on the region of interest, as well as data latency and availability (Table 1). Moreover the purpose for which the data are used is critical in selecting the most suitable satellite product. For example climate assessments favour accuracy, while those engaged in operational or forecasting activities require low-latency, high-frequency observations. The frequency of the satellite data products evaluated here varies from twice a day (APP-x) to monthly (AWI), and latency varies from three / four days (APP-x and CPOM NRT) to products that are updated seasonally, or on an ad-hoc basis (GSFC and the CPOM seasonally-averaged thickness data product, Table 1). If access to NRT measurements, with year-round availability is required, APP-x would be the first choice, since it is the only product that provides daily coverage across the Arctic Ocean in both summer and winter, and provides a reasonable measure of mean ice thickness, especially in some of the peripheral seas (regions 3-10) in the fall. However, it does not resolve the cross-basin ice thickness gradient nor the location of the thickest ice, and overestimates FYI thickness at the end of winter. Should basin-scale gradients in ice thickness (i.e., the thickness distribution) be important to the end user, then the ~~APP-x product is not recommended, and~~ CPOM NRT product is preferable, although it is only available for the winter-growth season (Tilling et al., ~~2016~~2018). In terms of climatological studies, or model initialization and hindcast studies, the CS2-only data products are appropriate options, but for navigation in the Arctic, none of these products are suitable as a single source of information, and the utility of the observations would only be realised when combined with additional ice charting analyses.

A remaining challenge for the satellite-derived thickness products is the treatment of snow depth on sea ice. All of the satellite-derived thickness products assessed here make use of the modified Warren et al. (1999) snow climatology, as outlined in Laxon et al., 2013, but there is variation in the implementation method, as described in Section 2.1. In addition, ~~We do not have a reliable data source that provides observations of the interannual variability of snow (Blanchard-Wrigglesworth et al., 2018). Most of the satellite-derived thickness products assessed here make use the modified Warren et al. (1999) snow climatology (as outlined in Laxon et al., 2013), which does not~~none of the implementations resolve year-to-year variations in snow depth. ~~The need for improved knowledge of snow depth holds especially true for the APP-x, wherein~~

25

In light of the current challenges presented by the treatment of snow in the existing data products, This has led a selection of end users, particularly those conducting data assimilation experiments, are interested into use satellite-derived sea ice freeboard measurements, rather than ice thickness, since freeboard represents the remote sensing observation (rather than derived ice thickness which is a derivation). Currently there are only two products that provide the freeboard parameter, AWI and GSFC. Our analysis suggests that the AWI data set is preferable for ice freeboard due to a more realistic representation of ice freeboard measurements across the Arctic (Fig. 10). We found a high prevalence of erroneous, negative freeboard estimates throughout the Arctic in the GSFC product, that were especially concentrated in the peripheral seas, particularly in regions 7 and 8, as well as in the Beaufort Sea. The source of these anomalies is most likely associated with aspects of the GSFC algorithm and interpolation of the mean sea surface between lead tie-points.

In conclusion we suggest that low-latency, monthly composites, derived from CS2 data, or similar, but updated daily with the latest-available measurements, would benefit many sea ice thickness applications and provide an ideal solution to address many end-user needs. Further, it may be possible to obtain a more robust thickness distribution, through the inclusion of passive microwave observations of thin-ice thickness in the marginal ice zone. Although the CS2SMOS results are promising, most of our evaluations with independent data were focused on the MYI zone. Our results show however that aApplying an appropriate ratio, that adequately combines microwave and altimeter observations, is challenging and requires further evaluation with independent observations collected over FYI and in the peripheral seas. Higher resolution (<= 5 km) along-orbit and gridded data products would advance the utility of the observations at the regional scale. We also recommend that future products include both ice thickness and freeboard parameters, as well as an estimate of thickness

uncertainty and/or data quality flags, so that the satellite observations may be used in data assimilation experiments aimed at improving ice forecasting.

Author contribution

SLF designed the study, HS and JM carried out the data analysis under the guidance of SLF, and JM was responsible for
5   producing the figures. HS led the manuscript preparation, and SLF, ER and JM contributed to the writing and internal review.

Competing interests

The authors declare that they have no conflicts of interest.

Acknowledgements

References

Allard, R. A., Farrell, S. L., Hebert, D. H., Johnston, W. F., Li, L., Kurtz, N. T., Phelps, M. W., Posey, P. G., Tilling, R., Ridout, A., and Wallcraft, A. L.: Utilizing CryoSat-2 sea ice thickness to initialize a coupled ice-ocean modeling system, Advances in Space Research, doi:10.1016/j.asr.2017.12.030, 2018.
20

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

Belward, A. and Dowell, M. (eds.): The Global Observing System for Climate (GCOS): Implementation Needs, GCOS 2016
25 Implementation Plan, Global Ocean Observing System Report (GCOS 200, GOOS 214), World Meteorological Organisation, 341, https://library.wmo.int/opac/doc_num.php?explnum_id=3417, 2016.

Blanchard-Wrigglesworth, E., Farrell, S., Newman, T., and Bitz, C.: Snow cover on Arctic sea ice in observations and an Earth system model, Geophysical Research Letters, 42, 10342–10348. https://doi.org/10.1002/2015GL066049, 2015.

Blanchard-Wrigglesworth, E., Webster, M. A., Farrell, S. L. and Bitz, C. M.: Reconstruction of Snow on Arctic Sea Ice, Journal of Geophysical Research: Oceans, 123(5), 3588-3602, doi:https://doi.org/10.1002/2017JC013364, 2018.

Blockley, E. W. and Peterson, K. A.: Improving Met Office seasonal predictions of Arctic sea ice using assimilation of CryoSat-2 thickness, The Cryosphere, 12, 3419-3438, https://doi.org/10.5194/tc-12-3419-2018, 2018.

Brucker, L., and Markus, T.: Arctic-scale assessment of satellite passive microwave-derived snow depth on sea ice using Operation IceBridge airborne data, J. Geophys. Res. - Oceans, 118, 2892–2905, doi:10.1002/jgrc.20228, 2013.

Connor L. N., Laxon S. W., McAdoo D. C., Ridout A., Cullen R., Farrell S., and Francis R.: Arctic sea ice freeboard from CryoSat-2: Validation using data from the first IceBridge underflight, AGU Fall Meeting Abstracts, 2011.

Farrell, S. L., Kurtz, N., Connor, L. N., Elder, B. C., Leuschen, C., Markus, T., McAdoo, D. C., Panzer, B., Richter-Menge, J., and Sonntag, J.G.: A first assessment of IceBridge snow and ice thickness data over Arctic sea ice, IEEE Transactions on Geoscience and Remote Sensing, 50(6), 2098-2111, 2012.

Giles, K. A., Laxon, S. W., Wingham, D. J., Wallis, D. W., Krabill, W. B., Leuschen, C. J., McAdoo, D., Manizade, S. S., and Raney, R. K.: Combined airborne laser and radar altimeter measurements over the Fram Strait in May 2002, Remote Sens. Environ., 111, 182–194, 2007.

Giles, K. A., Laxon, S. W., and Ridout, A. L.: Circumpolar thinning of Arctic sea ice following the 2007 record ice extent minimum, Geophys. Res. Lett., 35(22), doi:10.1029/2008GL035710, 2008.

Grosfeld, K., Treffeisen, R., Asseng, J., Bartsch, A., Bräuer, B., Fritzsch, B., Gerdes, R., Hendricks, S., Hiller, W., Heygster, G., Krumpen, T., Lemke, P., Melsheimer, C., Nicolaus, M., Ricker, R. and Weigelt, M.: Online sea-ice knowledge and data platform <www.meereisportal.de>, Polarforschung, Bremerhaven, Alfred Wegener Institute for Polar and Marine Research & German Society of Polar Research, 85 (2), 143-155, doi:10.2312/polfor.2016.011, 2016.

Guerreiro, K., Fleury, S., Zakharova, E., Rémy, F., and Kouraev, A.: Potential for estimation of snow depth on Arctic sea ice from CryoSat-2 and SARAL/AltiKa missions, Remote Sensing of Environment, 186, 339-349, doi:10.1016/j.rse.2016.07.013, 2016.

Guerreiro, K., Fleury, S., Zakharova, E., Kouraev, A., Rémy, F., and Maisongrande, P.: Comparison of CryoSat-2 and ENVISAT radar freeboard over Arctic sea ice: toward an improved Envisat freeboard retrieval, The Cryosphere, 11, 2059–2073, https://doi.org/10.5194/tc-11-2059-2017, 2017.

5

Hendricks, S., Ricker, R., and Helm, V.: User Guide - AWI CryoSat-2 Sea Ice Thickness Data Product (v1.2), 10013/epic.48201, 2016.

Hendricks, S.: Sea Ice Climate Change Initiative Phase 2, D3.4 Product User Guide (PUG) for Sea Ice Thickness Dataset, SICCI-SIT-PUG-P2-17-02, v.1.0, 16 pp, URL

10 http://data.ceda.ac.uk/neodc/esacci/sea_ice/docs/SICCI_P2_SIT_PUG_D3.3_Issue_1.0.pdf, 2017.

Hendricks, S.; Paul, S.; Rinne, E.: ESA Sea Ice Climate Change Initiative (Sea_Ice_cci): Northern hemisphere sea ice thickness from the CryoSat-2 satellite on a monthly grid (L3C), v2.0. Centre for Environmental Data Analysis,

15 doi:10.5285/ff79d140824f42dd92b204b4f1e9e7c2. http://dx.doi.org/10.5285/ff79d140824f42dd92b204b4f1e9e7c2, 2018.

Kaleschke, L., Tian-Kunze, X., Maaß, N., Makynen, M., & Drusch, M.: Sea ice thickness retrieval from SMOS brightness temperatures during the Arctic freeze-up period. Geophysical Research Letters, 39: L05501. doi:10.1029/2012GL050916, 2012.

20

Kaleschke, L., Tian-Kunze, X., Maaß, N., Beitsch, A., Wernecke, A., Miernecki, M., Müller, G., Fock, B. H., Gierisch, A. M. U., Schlünzen, K. H., Pohlmann, T., Dobrynin, M., Hendricks, S., Asseng, J., Gerdes, R., Jochmann, P., Reimer, N., Holfort, J., Melsheimer, C., Heygster, G., Spreen, G., Gerland, S., King, J., Skou, N., Søbjærg, S. S., Haas, C., Richter, F., and Casal, T.: SMOS sea ice product: Operational application and validation in the Barents Sea marginal ice zone, Remote

25 Sensing of Environment, 180, 264 – 273, doi:10.1016/j.rse.2016.03.009, 2016.

Key, J. R., R. Mahoney, Y. Liu, P. Romanov, M. Tschudi, I. Appel, J. Maslanik, D. Baldwin, X. Wang and P. Meade: Snow and ice products from Suomi NPP VIIRS, J. Geophys. Res. Atmos., 118, 12,816–12,830, doi:10.1002/2013JD020459, 2013.

30 Key, J. and Wang, X.: Extended AVHRR Polar Pathfinder (APP-x) Climate Algorithm Theoretical Basis Document, NOAA Climate Data Record Program, CDRP-ATBD-0573, 01B-24b, Rev. 1, 82, 2015.

King, J., Howell, S., Derksen, C., Rutter, N., Toose, P., Beckers, J. F., Haas, C., Kurtz, N. and Richter‑Menge, J.: Evaluation of Operation IceBridge quick-look snow depth estimates on sea ice, Geophysical Research Letters, 42(21), 9302-9310, 2015.

5 Koenig, L., Martin, S., Studinger, M., and Sonntag, J.: Polar airborne observations fill gap in satellite data, Eos, Transactions American Geophysical Union, 91(38), 333-334, 2010.

Krishfield, R. A. and Proshutinsky, A.: BGOS ULS Data Processing Procedure. Woods Hole Oceanographic Institute report, available at: http://www.whoi.edu/fileserver.do?id=85684&pt=2&p=100409, last access 26 April 2018, 2006.

10

Krishfield, R.A., Proshutinsky, A., Tateyama, K., Williams, W.J., Carmack, E.C., McLaughlin, F.A. and Timmermans, M.L.: Deterioration of perennial sea ice in the Beaufort Gyre from 2003 to 2012 and its impact on the oceanic freshwater cycle, Journal of Geophysical Research: Oceans, 119(2), 1271-1305, 2014.

15 Kurtz, N. T., and S. L. Farrell: Large-scale surveys of snow depth on Arctic sea ice from operation IceBridge, Geophys. Res. Lett., 38, L20505, doi:10.1029/2011GL049216, 2011.

Kurtz, N. T., Farrell, S. L. , Studinger, M., Galin, N., Harbeck, J. P., Lindsay, R., Onana, V. D., Panzer, B., and Sonntag, J. G.: Sea ice thickness, freeboard, and snow depth products from Operation IceBridge airborne data. The Cryosphere, 7, 1035-
20 1056,
doi:10.5194/tc-7-1035-2013, 2013.

Kurtz, N. T., Galin, N., and Studinger, M.: An improved CryoSat-2 sea ice freeboard retrieval algorithm through the use of waveform fitting, The Cryosphere, 8, 1217-1237. doi:10.5194/tc-8-1217-2014, 2014.

25

Kurtz, N., Studinger, M., Harbeck, J., Onana, V., and Yi, D.: IceBridge L4 Sea Ice Freeboard, Snow Depth, and Thickness, Version 1 [March–April 2011], NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado USA, [accessed 2018], doi:10.5067/G519SHCKWQV6, 2015.

30 Kurtz, N.: IceBridge Sea Ice Freeboard, Snow Depth, and Thickness Quick Look, Version 1 [March–April, 2012–2017], NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado USA, [accessed 2018], doi:10.5067/GRIXZ91DE0L9, 2016.

Kurtz, N. and Harbeck, J.: CryoSat-2 Level-4 Sea Ice Elevation, Freeboard, and Thickness, Version 1 [October–April, 2010–2018], NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado USA, [accessed 2018], doi:10.5067/96JO0KIFDAS8, 2017.

Kwok, R., Cunningham, G. F., Wensnahan, M., Rigor, I., Zwally, H. J., and Yi, D.: Thinning and volume loss of the Arctic Ocean sea ice cover: 2003–2008, J. Geophys. Res. Oceans, 114, doi:10.1029/2009JC005312, 2009.

Kwok, R. and Cunningham, G.F.: Variability of Arctic sea ice thickness and volume from CryoSat-2, Phil. Trans. R. Soc. A., 373, 20140157, doi:10.1098/rsta.2014.0157pmid:26032317, 2015.

Kwok, R. and Markus, T.: Potential Basin-Scale Estimates of Arctic Snow Depth with Sea Ice Freeboards from CryoSat-2: An Exploratory Analysis, Adv. Space Res., doi:10.1016/j.asr.2017.09.007, 2017.

Lawrence, I. R., Tsamados, M. C., Stroeve, J. C., Armitage, T. W. K., and Ridout, A. L.: Estimating snow depth over Arctic sea ice from calibrated dual-frequency radar freeboards, The Cryosphere, 12, 3551-3564, https://doi.org/10.5194/tc-12-3551-2018, 2018.

Laxon S. W., Giles, K.A., Ridout, A. L., Wingham, D.J., Willatt, R., Cullen, R., Kwok, R., Schweiger, A., Zhang, J., Haas, C., Hendricks, S., Krishfield, R., Kurtz, N., Farrell, S. and Davidson, M.: CryoSat-2 estimates of Arctic sea ice thickness and volume, Geophys. Res. Lett., 40, 732–737, doi:10.1002/grl.50193, 2013.

~~Lindsay, R., Haas, C., Hendricks, S., Hunkeler, P., Kurtz, N., Paden, J., Panzer, B., Sonntag, J., Yungel, J., Zhang, J.: Seasonal forecasts of Arctic sea ice initialized with observations of ice thickness, Geophys. Res. Lett., 39, L21502, doi:10.1029/2012GL053576, 2012.~~

Markus, T., Stroeve, J. C., and Miller, J.: Recent changes in Arctic sea ice melt, onset, freezeup, and melt season length, J. Geophys. Res., 114, C12024, doi:10.1029/2009JC005436, 2009.

Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Farrell, S., Fricker, H., Gardner, A., Harding, D., Jasinski, M., Kwok, R., Magruder, L., Lubin, D., Luthcke, S., Morison, J., Nelson, R., Neuenschwander, A., Palm, S., Popescu, S., Shum, CK, Schutz, B. E., Smith, B., Yang, Y. and Zwally, J.: The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation, Rem. Sens. Environ., 190, 260-273, doi:10.1016/j.rse.2016.12.029, 2017.

31

Maaß, N., Kaleschke, L., Tian-Kunze, X., and Drusch, M.: Snow thickness retrieval over thick Arctic sea ice using SMOS satellite data, The Cryosphere, 7, 1971-1989, https://doi.org/10.5194/tc-7-1971-2013, 2013.

Newman, T., Farrell, S. L., Richter-Menge, J., Connor, L. N., Kurtz, N. T., Elder, B. C. and McAdoo, D.: Assessment of radar-derived snow depth over Arctic sea ice, Journal of Geophysical Research: Oceans, 119(12), 8578-8602, 2014.

Parkinson, C.L., and Cavalieri, D.J.: Arctic sea ice variability and trends, 1979-2006, J. Geophys. Res., 113, C07003. doi:10.1029/2007JC004558, 2008.

Parkinson, C.L., and Comiso, J.C.: On the 2012 record low Arctic sea ice cover: Combined impact of preconditioning and an August storm, Geophys. Res. Lett., doi:10.1002/grl.50349, 2013.

Paul, S., Hendricks, S. and Rinne, E.: Sea Ice Climate Change Initiative Phase 2, D2.1 Sea Ice Thickness Algorithm Theoretical Basis Document (ATBD), SICCI-P2-ATBD(SIT), v.1.0, 50 pp. https://icdc.cen.uni-hamburg.de/fileadmin/user_upload/ESA_Sea-Ice-ECV_Phase2/SICCI_2_ATBD_D2.1__SIT__Issue_1.0.pdf, 2017.

Perovich, D. K., Meier, W., Tschudi, M., Farrell, S., Hendricks, S., Gerland, S., Haas, C., Krumpen, T., Polashenski, C., Ricker, R. and Webster, M.: Sea ice cover [in "State of the Climate in 2018"], Bull. Amer. Meteor. Soc., 99(8), S147–152, doi:10.1175/2018BAMSStateoftheClimate.1., 2018.

Price, D., Beckers, J., Ricker, R., Kurtz, N., Rack, W., Haas, C., Helm, V., Hendricks, S. Leonard, G., and Langhorne, P. J.: Evaluation of CryoSat-2 derived sea-ice freeboard over fast ice in McMurdo Sound, Antarctica, J. Glaciol. 61 (226), doi:10.3189/2015JoG14J157, 2015.

Richter-Menge, J. A., and Farrell, S. L.: Arctic sea ice conditions in spring 2009–2013 prior to melt, Geophys. Res. Lett., 40, 5888–5893, doi:10.1002/2013GL058011, 2013.

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 (4), 1607-1622, doi:10.5194/tc-8-1607-2014, 2014.

Ricker, R., Hendricks, S., Kaleschke, L., and Tian-Kunze, X.: CS2SMOS User Guide v3.0, Alfred Wegener Institute, 10013/epic.51136, 2017a.

Ricker, R.; Hendricks, S.; Kaleschke, L.; Tian-Kunze, X.; King, J. and Haas, C.: A weekly Arctic sea-ice thickness data record from merged CryoSat-2 and SMOS satellite data, The Cryosphere, 11, 1607-1623, doi:10.5194/tc-11-1607-2017, 2017b.

5 Rostosky, P., Spreen, G., Farrell, S. L., Frost, T., Heygster, G., & Melsheimer, C.: Snow depth retrieval on Arctic sea ice from passive microwave radiometers - Improvements and extensions to multiyear ice using lower frequencies, J. Geophys. Res., 123, 7120–7138, https://doi.org/10.1029/2018JC014028, 2018.

Rothrock, D. A., Percival, D. B. and Wensnahan, M.: The decline in arctic sea-ice thickness: Separating the spatial, annual,
10 and interannual variability in a quarter century of submarine data, Journal of Geophysical Research: Oceans, 113(C5), 2008.

Shalina, E. V. and Sandven, S.: Snow depth on Arctic sea ice from historical in situ data, The Cryosphere, 12, 1867-1886, doi:10.5194/tc-12-1867-2018, 2018.

15 Shepherd, A., Fricker, H. A., and Farrell S. L.: Trends and connections across the Antarctic cryosphere, Nature, 558(7709), 223-232, 2018.

Skourup, H., Farrell, S. L., Hendricks, S., Ricker, R., Armitage, T. W. K., Ridout, A., Andersen, O. B., Haas, C., and Baker, S.: An assessment of state-of-the-art mean sea surface and geoid models of the Arctic Ocean: Implications for sea ice
20 freeboard retrieval, Journal of Geophysical Research: Oceans, 122, 8593–8613, doi:10.1002/2017JC013176, 2017.

Song, M.R.: Change of Arctic sea ice volume and its relationship with sea ice extent in CMIP5 simulations, Atmospheric and Oceanic Science Letters, 9, 22–30, doi:10.1080/16742834.2015.1126153, 2016.

25 Stroeve, J. C., Schröder, D., Tsamados, M. and Feltham, D. L.: Warm winter, thin ice? The Cryosphere, doi: https://doi.org/10.5194/tc-2017-287, 2018.

Tian-Kunze, X., Kaleschke, L., Maaß, N., Mäkynen, M., Serra, N., Drusch, M., and Krumpen, T.: SMOS-derived thin sea ice thickness: algorithm baseline, product specifications and initial verification, The Cryosphere, 8, 997–1018,
30 doi:10.5194/tc-8-997-2014, URL http://www.the-cryosphere.net/8/997/2014/, 2014.

Tilling, R.L., Ridout, A., Shepherd, A., and Wingham, D.J.: Increased Arctic sea ice volume after anomalously low melting in 2013, Nature Geoscience, 8(8), 643, doi:10.1038/NGEO2489, 2015.

33

Tilling, R.L., Ridout, A., and Shepherd, A.: Near-real-time Arctic sea ice thickness and volume from CryoSat-2, The Cryosphere, 10(5), 2003-2012, 2016.

Tilling, R.L., Ridout, A., and Shepherd, A.: Estimating Arctic sea ice thickness and volume using CryoSat-2 radar altimeter data, Advances in Space Research, Volume 62, Issue 6, Pages 1203-1225, doi:10.1016/j.asr.2017.10.051, 2018.

Wang, X., Key, J. R., and Liu, Y.: A thermodynamic model for estimating sea and lake ice thickness with optical satellite data, J. Geophys. Res., 115, C12035, doi:10.1029/2009JC005857, 2010.

Wang, X., Key, J., Kwok, R., and Zhang, J.: Comparison of Arctic Sea Ice Thickness from Satellites, Aircraft, and PIOMAS Data, Remote Sens., 8, 713, 2016.

Warren, S. G., Rigor, I. G., Untersteiner, N., Radionov, V. F., Bryazgin, N. N., Aleksandrov, Y. I., and Colony, R.: Snow depth on Arctic sea ice, J. Clim., 12, 1814–1829, 1999.

Willmes, S. and Heinemann, G.: Sea Ice Wintertime Lead Frequencies and Regional Characteristics in the Arctic, 2003-2015, Remote Sens., 8(1), 4; doi:10.3390/rs8010004, 2015.

Wingham, D. J., Francis, C. R., Baker, S., Bouzinac, C., Brockley, D., Cullen, R., de Chateau-Thierry, P., Laxon, S. W., Mallow, U., Mavrocordatos, C., Phalippou, L., Ratier, G., Rey, L., Rostan, F., Viau, P., and Wallis, D. W.: CryoSat: A mission to determine the fluctuations in Earth's land and marine ice fields, Adv. Space Res., 37, 841–871, doi:10.1016/j.asr.2005.07.027, 2006.

Xia, W. and Xie, H.: Assessing three waveform retrackers on sea ice freeboard retrieval from Cryosat-2 using Operation IceBridge Airborne altimetry datasets. Remote Sens. Environ., 204, 456–471, 2018.
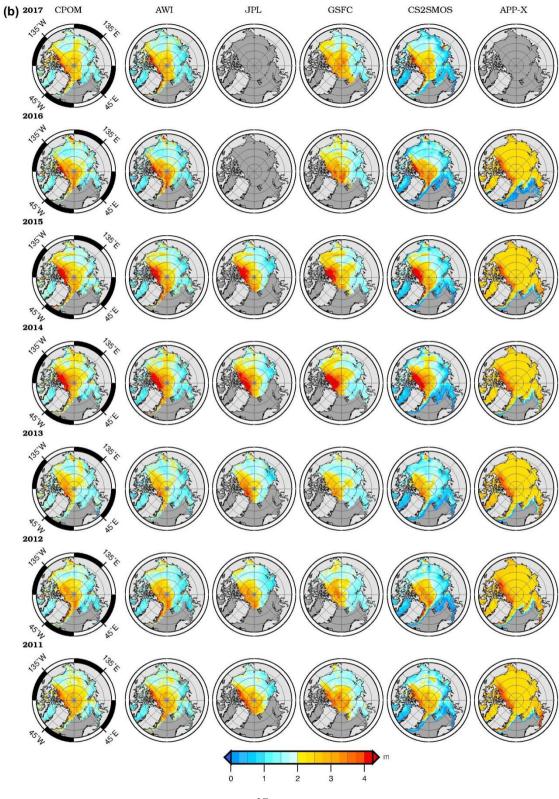
Xie, J., Counillon, F., and Bertino, L.: Impact of assimilating a merged sea ice thickness from CryoSat-2 and SMOS in the Arctic reanalysis, The Cryosphere, 12, 3671-3691, https://doi.org/10.5194/tc-12-3671-2018, 2018. The Cryosphere Discuss., doi:10.5194/tc-2018-101, in review, 2018.
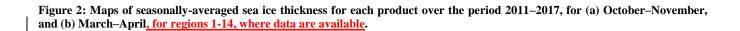
Yang, Q., Losa, S. N., Losch, M., Tian-Kunze, X., Nerger, L., Liu, J., Kaleschke, L., and Zhang, Z.: Assimilating SMOS sea ice thickness into a coupled ice-ocean model using a local SEIK filter, J. Geophys. Res. Oceans, 119, 6680-6692, doi:10.1002/2014JC009963, 2014.
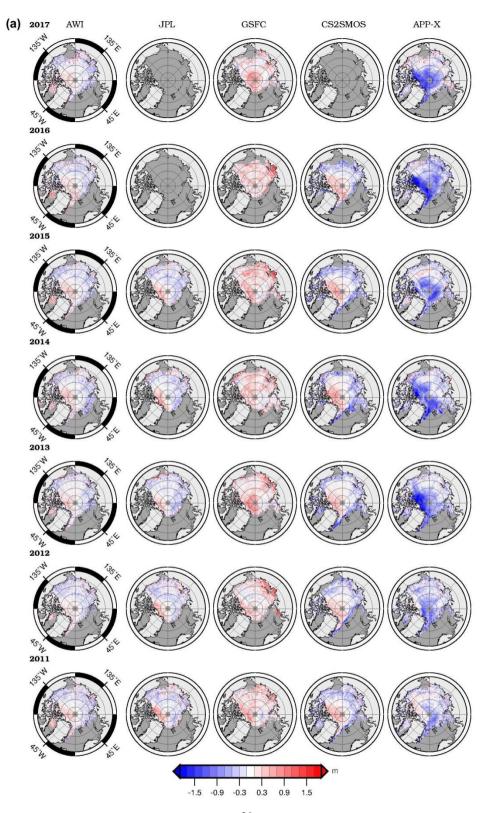
5

Figure 1: Regions utilised in data analysis: (1) Central Arctic Ocean, (2) Lincoln Sea, (3) Beaufort Sea, (4) Chukchi Sea, (5) East Siberian Sea, (6) Laptev Sea, (7) Kara Sea, (8) Barents Sea, (9) Norwegian Sea, (10) Greenland Sea, (11) Baffin Bay, (12) Davis Strait, (13) Hudson Strait, and (14) Canadian Arctic Archipelago. The dotted area represents the central Arctic region (1–6) within which all data products are available. The locations of Beaufort Gyre Exploration Project (BGEP) moorings A, B, and D are also indicated (white circles with mooring designation).

10

Figure 2: Maps of seasonally-averaged sea ice thickness for each product over the period 2011–2017, for (a) October–November, and (b) March–April, for regions 1-14, where data are available.

Figure 3: Maps of seasonally-averaged sea ice thickness differences for the period 2011–2017, where ~~the reference data set (CPOM) is subtracted from each data product~~each data product is subtracted from the reference data set (CPOM), for (a) October–November and (b) March–April, for regions 1-14, where data are available. Red (blue) regions indicate areas where the seasonally-averaged thickness is greater (less) than the reference data product.

5

Figure 4: <ins>Comparisons of s</ins>~~S~~ea ice thickness for each product <ins>and</ins> ~~compared to~~ the reference data set for <ins>(a) October–November and (b) March–April, for the period 2010-2018, regions 1-14 north of 65°N, subject to data availability (Table 1)</ins>~~March–April (top) and October–November (bottom)~~. Colour indicates measurement density, derived from the number of data points within each 0.05 m cell. Statistics for correlation (r), mean difference (std. dev.), and number of data points are provided. ~~The period over which correlations are calculated is based on product availability, as outlined in Table 1.~~
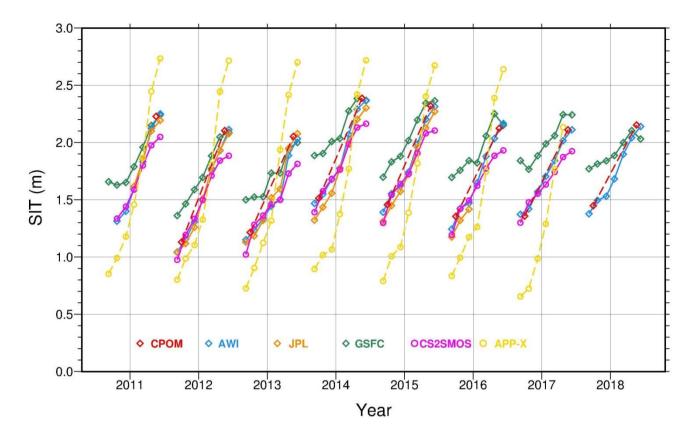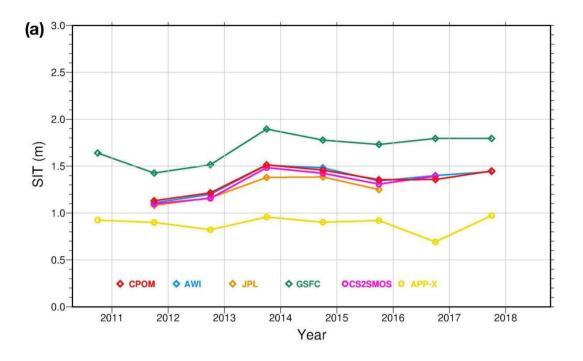
Figure 5: Sea ice thickness growth curves for October–April (monthly averages) ~~during~~ for the central Arctic (regions 1-6) for the period 2010–2018, ~~for each product per~~subject to data availability (~~solid lines~~Table 1), indicating interannual variability in the winter-time thickness evolution. For the CPOM data product, ~~only the~~ October–November and March–April seasonal averages are shown~~,~~ (red diamonds connected by a dashed red line).
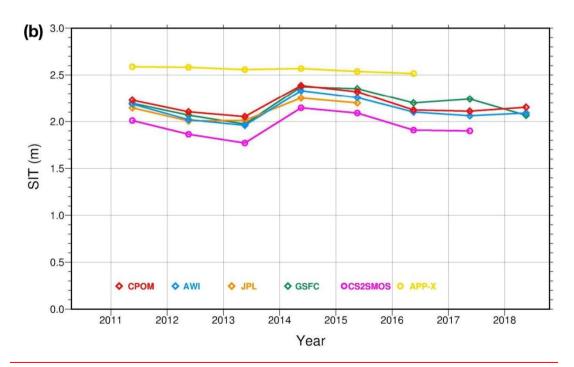
Figure 6: Time series of seasonally-averaged sea ice thickness ~~for each product,~~ for (a) October–November, and (b) March–April, over the central Arctic (regions 1–6) during the period 2010–2018, ~~based on product~~subject to data availability ~~as outlined in~~ (Table 1~~)~~.
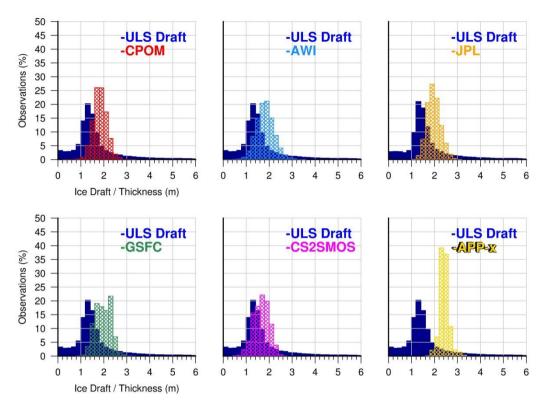
44

Figure 7: Spring (March–April) sea ice thickness distributions (cross-hatched) for each data product within 200 km radius of BGEP mooring locations, averaged for the period 2011 to 2017*2016, overlaid on the corresponding BGEP upward looking sonar (ULS) ice draft distribution (dark blue, solid). *The JPL data product averaging period is 2011 to 2015, and the APP-x data product average does not include thickness data for April 2017. Histogram bin width is 0.2 m.
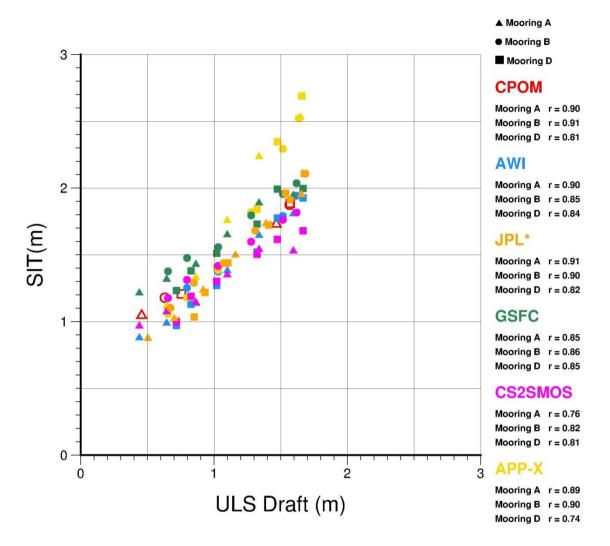
Figure 8: Correlation between monthly-averaged, satellite-derived ice thickness and ULS ice draft, for six months spanning November to April. Correlation coefficients are provided per ULS mooring (~~mooring~~ location~~s~~ indicated by symbols), for the period ~~2010–2017~~2010–2016, with the exception of the JPL product, wherein monthly averages and correlation coefficients are calculated for the period 2011–2015, and the APP-x product, where spring 2017 was excluded. In the case of the CPOM product the correlation coefficients are calculated based on seasonal (October–November, March–April) rather than monthly averages. ~~To aid visualisation~~Here, monthly averages were further combined across years (e.g. November ~~2011~~2010–~~2016~~2017, December ~~2011~~2010–~~2016~~2017, etc.) to provide six data points per mooring-product comparison.
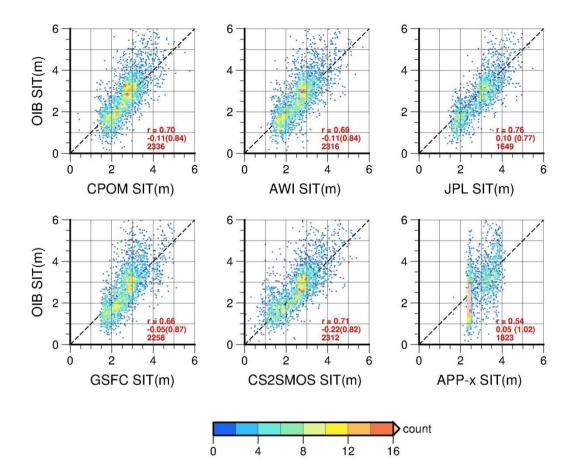
Figure 9: Comparison of satellite-derived ice thickness with Operation IceBridge thickness estimates at the end of the winter growth season (March–April). Comparisons were conducted by gridding satellite and aircraft data onto a common 0.4° latitude by 4° longitude50 km grid and using grid cells in which both data sets contained thickness estimates. Colour indicates measurement density (number of data points within each 0.01 m cell). Statistics for correlation (r), mean difference (std. dev.), and number of data points are calculated for the period 2011–2017, with the exception of the JPL product, wherein statistics are calculated for the period 2011–2015, and the APP-x product, where spring 2017 was excluded.
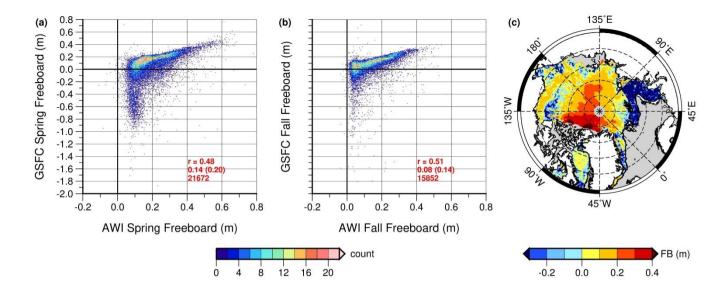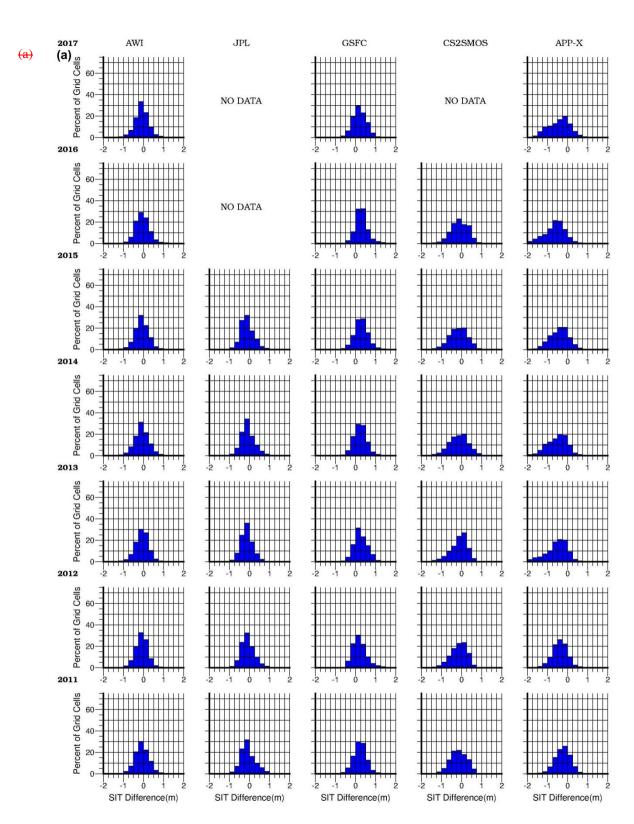
Figure 10: Comparison of sea ice freeboard in the GSFC and AWI data products for (a) March–April and (b) October–November, for the period 2011–2018, regions 1-14, north of 65°N. ~~Data were gridded at 25 km and data common to both grids were compared.~~ Colour indicates measurement density, derived from the number of data points within each 0.01 m cell. Statistics for correlation (r), mean difference (std. dev.), and number of data points are provided. (c) GSFC product freeboard (FB) for April 2014.
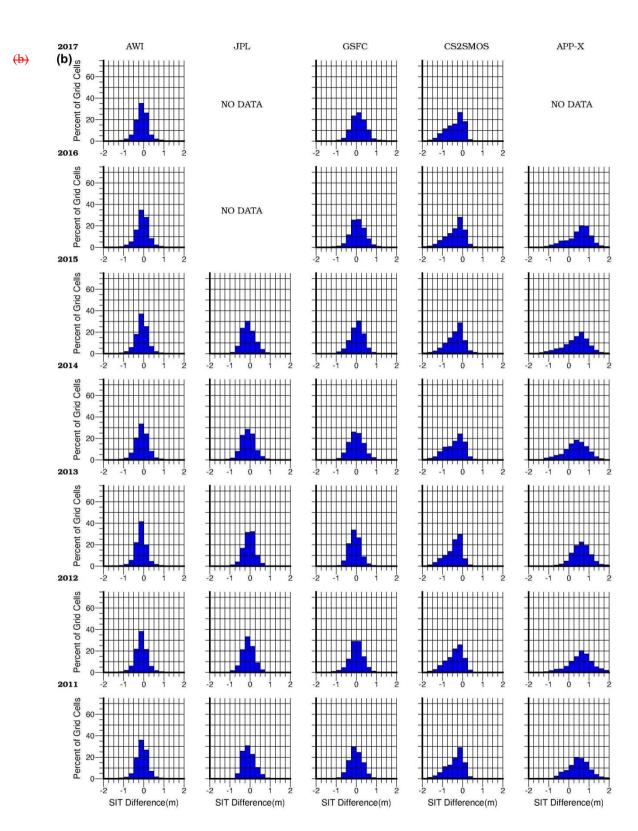
Figure S1: Histograms of seasonally-averaged sea ice thickness, where the reference data set (CPOM) is subtracted from each product~~Sea ice thickness deviation (in meters) for each product subtracted from reference data set (CPOM)~~, for (a) October–November, and (b) March–April, in the central Arctic (regions 1–6), for the period 2011–2017, subject to data availability (Table 1).
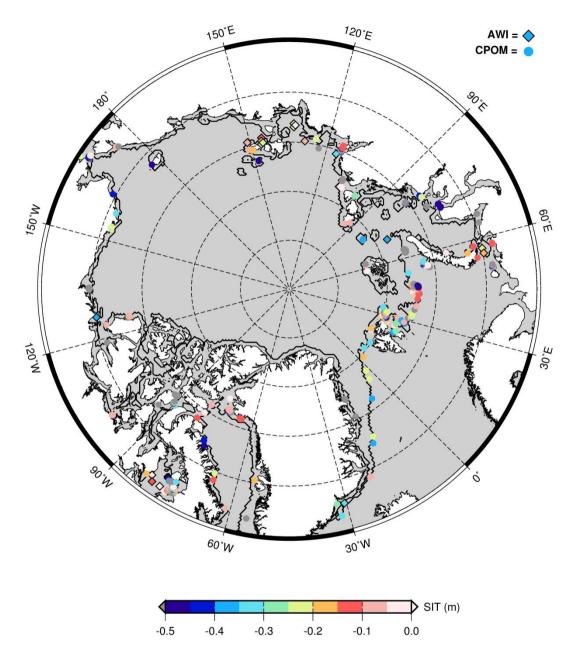
5

Figure S2: Locations of negative sea ice thickness values in the CPOM (circles) and AWI (diamonds with black outline) data products, for March–April 2014 (CPOM) and April 2014 (AWI). A contour representing the 70 % ice concentration level is also shown (thick black line).

Table 1. Characteristics of satellite-derived ice thickness products ("quality ~~flag~~ information" refers to the availability of a thickness uncertainty estimate or quality flag). Latency is the difference between the date of data acquisition and data delivery, estimated from data portal time stamps at the time of writing. Frequency refers to the temporal offset between consecutive datasets (e.g. daily, weekly, monthly). *The northern limit of coverage (81.5 °N) impacts the availability of the Envisat product, especially in the Central Arctic (region 1).

Product Name	CPOM	AWI	~~NASA /~~ JPL	~~NASA /~~ GSFC	CS2SMOS	APP-x	LEGOS TFMRA	ESA CCI
Temporal Range	~~March 2011-date (October-April only)~~November 2010-date	November 2010--date	January 2011--December 2015	October 2010--date	November 2010--April 2017	January 1982--date	November 2002-March 2012 * November 2010-April 2017	October 2002-March 2012 * November 2010-April 2017
Frequency	Daily (NRT)/monthly / seasonally	Monthly	Not updated post 2015	Daily	Weekly	Twice daily	Monthly	Daily
~~Geograpical~~ Geographical Coverage (by region number)	(1)–(14), ~~full Northern Hemisphere~~poleward of 40 °N	(1)–(14), poleward of 60 °N	(1)–(6)	(1)–(8), (10, partial), poleward of 55 °N	(1)–(14), poleward of 50 °N	(1)–(14), poleward of 50 °N	65 °N-81.5 °N*; poleward of 65 °N	(1), (3)-(14) south of 81.5 °N*; (1)-(14), poleward of 16.6 °N
Averaging Period	NRT data product: 2, 14, 28 days. ~~Seasonally-averaged product: 2 months~~Final products: monthly, seasonally-averaged	1 month	1 month	30 day	7 day	12 hours	1 month	1 month (moving average)
Latency	3 days (NRT) / seasonally	Variable	Not updated post 2015	~6 weeks	Variable	~4 days (as of April 2018)	~1 year	Variable
Grid Resolution	25 km, 5 km (full Arctic), 1 km (individual regions)	EASE2, 25 km	25 km	Polar Stereographic SSM/I, 25 km	EASE2, 25 km	EASE, 25 km ~~(EASE grid)~~	12.5 km	EASE2, 25 km
Freeboard	No	Yes	No	Yes	No	No	Yes	Yes
Quality ~~Flag~~Information	No	Yes	No	No	Yes	No	Yes	Yes
~~Technique~~Retracking technique	~~70% Threshold First Maxima Retracker Algorithm (TFMRA)~~Retracking based on ice type (lead, floe); Lead retracker	50% TFMRA	Waveform centroid retracker	Waveform fitting using waveform model	Blended SMOS and CS2 data	One-dimensional Thermodynamic Ice Model (OTIM)	60% TFMRA	50% TFMRA

	follows Giles et al. 2007; Floe retracker based on 70 % of peak amplitude on the waveform leading edge							
Mean Sea Surface	UCL2013	DTU15 MSS	EGM2008 geoid	DTU10 MSS	DTU15 MSS (for CS-2 data)	N/A	DTU15 MSS	DTU15 MSS
Snow Depth	Monthly constants for FYI and MYI based on the Modified Warren Climatologymodified climatology of Warren et al., 1999 (MWC)	Modified Warren Climatology MWC	Modified Warren Climatology MWC	Modified Warren Climatology MWC	Modified Warren climatologyMWC (CS2 data); linear relation with ice thickness (SMOS data)	0.2 mClimatology of Warren et al. (1999) combined with other available estimates	MWC	MWC
Ice Density	FYI: 916.7 kg/m^3; MYI: 882.0 kg/m^3	FYI: 916.7 kg/m^3; MYI: 882.0 kg/m^3	FYI: 917 kg/m^3; MYI: 882.0 kg/m^3	915 kg/m^3	FYI: 916.7 kg/m^3; MYI: 882.0 kg/m^3	N/A	FYI: 916.7 kg/m^3; MYI: 882.0 kg/m^3	FYI: 916.7 kg/m^3; MYI: 882.0 kg/m^3
Ice Concentration, Threshold	NSIDC Near-Real-Time DMSP SSMIS Daily Polar Gridded Sea Ice Concentrations, 75%	OSI-SAF, 70%	Not specified	NSIDC Near-Real-Time DMSP SSMIS Daily Polar Gridded Sea Ice Concentrations, 70%	OSI-SAF, 15%	NSIDC Nimbus-7 SMMR and DMSP SSM/I passive microwave data with NASA Team Algorithm applied, 15%	Sea ice age (NSIDC), 50%	OSI-SAF, 70%
Ice Type	OSI-SAF	OSI-SAF	ASCAT	OSI-SAF	OSI-SAF	Converted from reflectances	Sea ice age (NSIDC)	Integrated Climate Data Center (ICDC) ice type fraction
References	Laxon et al., 2013; Tilling et al., 2016; Tilling et al., 2018	Ricker et al., 2014; Hendricks et al., 2016	Kwok and Cunningham, 2015	Kurtz et al., 2014; Kurtz and Harbeck, 2017	Tian-Kunze et al., 2014; Kaleschke et al., 2015, Ricker et al., 20162017b	Wang et al., 2010; Key and Wang, 2015; Wang et al. 2016	Guerreiro et al., 2017	Hendricks, S., 2017; Paul et al., 2017; Hendricks et al., 2018
Public Data Source	http://www.cpom.ucl.ac.uk/csopr/seaice.html	Meereisportal, http://data.meereisportal.de/ (Grosfeld et al. 2016)	https://rkwok.jpl.nasa.gov/cryosat/download.html	https://nsidc.org/data/RDEFT4	Meereisportal, http://data.meereisportal.de/ (Grosfeld et al. 2016)	https://www.ncei.noaa.gov/data/avhrr-polar-pathfinder-extended/access/nhem/	http://ctoh.legos.obs-mip.fr	http://cci.esa.int/data

	October–November						March–April					
	CPOM	AWI	JPL	GSFC	CS2SMOS	APP-x	CPOM	AWI	JPL	GSFC	CS2SMOS	APP-x
Arctic Ocean [1]	~~1.64~~1.48	~~1.75~~1.48	~~1.6~~1.39	~~2.03~~1.77	~~1.44~~1.49	~~1.01~~0.96	~~2.46~~2.3	~~2.42~~2.25	~~2.45~~2.29	~~2.5~~2.33	~~1.93~~2.15	~~2.6~~2.58
Lincoln Sea [2]	~~2.04~~1.99	~~2.13~~2.11	~~2.58~~2.52	~~2.45~~2.42	~~1.99~~2.43	~~1.18~~1.19	~~3.8~~3.61	~~3.65~~3.51	~~3.68~~3.7	~~3.39~~3.34	~~2.94~~3.47	~~3.64~~3.66
Beaufort Sea [3]	~~1.01~~1.03	~~1.04~~0.79	~~1.06~~0.89	~~1.5~~1.2	~~0.66~~0.57	~~0.89~~0.77	~~2.06~~2.03	~~1.99~~1.97	~~2.1~~2.09	~~2.04~~2.00	~~1.31~~1.59	~~2.6~~2.59
Chukchi Sea [4]	-0.84	-0.61	-0.96	-1.13	-0.21	-0.47	~~1.95~~1.91	~~1.83~~1.79	~~1.82~~1.81	~~1.94~~1.92	~~1.18~~1.46	2.51
East Siberian Sea [5]	~~0.91~~0.92	~~0.76~~0.77	~~0.82~~0.81	~~1.27~~1.34	~~0.4~~0.52	~~0.67~~0.68	~~1.8~~1.83	~~1.7~~1.72	~~1.59~~1.56	~~1.74~~1.72	~~1.14~~1.39	2.44
Laptev Sea [6]	~~0.81~~0.82	~~0.52~~0.55	~~0.69~~0.67	~~0.99~~1.26	~~0.35~~0.43	0.61	~~1.5~~1.53	~~1.25~~1.27	~~1.15~~1.12	~~1.55~~1.56	~~0.78~~0.97	2.4
Kara Sea [7]	-0.85	-0.48	-0.87	-1.01	-0.37	-0.49	~~1.6~~1.59	~~1.39~~1.38	-1.63	~~1.56~~1.55	~~0.69~~0.79	~~2.49~~2.52
Barents Sea [8]	-1.31	-0.78	-1.26	-1.51	-0.37	-0.36	~~1.45~~1.41	~~1.45~~1.44	-1.99	~~1.67~~1.73	~~0.38~~0.41	~~2.3~~1.99
Greenland Sea [10]	~~1.68~~1.69	~~1.89~~1.78	-2.14	-2.4	~~1.14~~1.16	~~0.75~~0.56	~~2.62~~2.47	~~3.04~~2.79	-3.02	-2.97	~~1.7~~1.65	~~2.39~~2.18
Baffin Bay [11]	-0.91	-1.28	-	-2.47	-0.26	-0.63	~~1.53~~1.5	~~1.26~~1.27	-	-3.39	~~0.54~~0.64	~~2.48~~2.47
Davis Strait [12]	-1.01	-	-	-	-0.29	-0.24	~~1.59~~1.55	~~1.21~~1.15	-	-	~~0.48~~0.47	~~2.7~~2.54
Hudson Strait [13]	-1.12	-	-	-	-0.12	-0.22	~~1.49~~1.53	~~0.81~~0.82	-	-	~~0.44~~0.53	~~2.91~~2.93
Canadian Archipelago [14]	~~1.29~~1.25	~~1.42~~1.26	-1.19	-2.03	~~0.85~~0.81	~~0.92~~0.84	~~1.97~~1.95	~~1.44~~1.58	-2.09	-2.27	~~1.32~~1.4	~~2.66~~2.64

Table 3. ~~Deviation~~ Anomalies of seasonally-averaged annual sea ice thickness (in meters) ~~from~~ relative to the seasonally-averaged ~~climatological~~ baseline mean within the central Arctic (regions 1–6). The ~~climatological~~ baseline mean is calculated ~~per product availability during~~ for the period 2011– ~~2017~~ 2015 ~~(i.e. the JPL product climatological mean is calculated for the period 2011–2015, while the fall CS2SMOS and spring APP-x product climatological means are calculated for the period 2011–2016)~~. Anomalies relative to the 2011-2015 mean are included for later seasons according to the product availability.

October–November~~March–April~~	2011	2012	2013	2014	2015	2016	2017	2018
CPOM	~~−0.27~~ 0.2	−0.12	~~0.2~~ 0.18	~~0.07~~ 0.12	~~0.03~~ 0.02	0.02	~~0.04~~ 0.11	
AWI	~~−0.34~~ -0.22	~~−0.18~~ 0.13	~~0.22~~ 0.18	~~0.07~~ 0.15	~~0.06~~ 0.01	~~0.02~~ 0.06	~~0.14~~ 0.1	
JPL	~~−0.24~~ -0.17	−0.09	~~0.18~~ 0.13	~~0.11~~ 0.13	~~0.05~~ 0.0	-	-	
GSFC	~~−0.40~~ -0.25	~~−0.25~~ 0.16	~~0.36~~ 0.22	~~0.02~~ 0.1	~~0.03~~ 0.06	~~−0.01~~ 0.13	~~0.25~~ 0.13	
CS2SMOS	~~−0.30~~ 0.2	~~−0.16~~ 0.14	~~0.21~~ 0.19	~~0.18~~ 0.12	~~0.07~~ 0.01	~~0~~ 0.1	-	
APP-x	~~0.04~~ 0.0	~~−0.03~~ 0.08	~~0.08~~ 0.06	~~0.03~~ 0.0	~~0.04~~ 0.02	−0.21	~~0.04~~ 0.07	

~~October–November~~ March-April	2011	2012	2013	2014	2015	2016	2017	2018
CPOM	~~0.05~~ 0.01	~~−0.07~~ 0.11	~~−0.21~~ 0.17	~~0.2~~ 0.17	~~0.12~~ 0.1	~~0~~ 0.09	~~−0.09~~ -0.1	-0.06
AWI	~~0.06~~ 0.04	~~−0.10~~ 0.13	~~−0.25~~ 0.19	~~0.22~~ 0.18	~~0.12~~ 0.11	~~0.04~~ -0.04	~~−0.09~~ -0.08	-0.05
JPL	~~0.06~~ 0.02	~~−0.11~~ 0.12	~~−0.15~~ 0.11	~~0.16~~ 0.13	~~0.05~~ 0.08	-	-	
GSFC	~~0.07~~ 0.0	~~−0.18~~ 0.12	~~−0.30~~ 0.22	~~0.21~~ 0.18	~~0.11~~ 0.16	~~0.02~~ 0.01	~~0.07~~ 0.05	-0.13
CS2SMOS	~~−0.02~~ 0.03	~~−0.12~~ 0.11	~~0.02~~ -0.21	~~0.2~~ 0.17	~~0.04~~ 0.11	~~−0.05~~ -0.07	~~−0.07~~ -0.08	
APP-x	~~0.03~~ 0.02	~~0.02~~ 0.01	~~0.01~~ -0.01	~~0.02~~ 0.0	~~−0.04~~ -0.03	−0.05	-	

57

Table 4. Annual winter (January–April*) sea ice draft/thickness (meters, top value) and deviation from the ~~climatological average~~2011-2017 mean (meters, bottom value) for each BGEP ULS mooring (A, B, D) and ~~five~~ six satellite products (using measurements within 200 km of BGEP moorings). ~~Climatological mean is calculated per product for the period 2011–2016.~~ *March–April for CPOM statistics.

		ULS	CPOM	AWI	JPL	GSFC	CS2SMOS	APP-x
	2011–2016	1.23~~1.27~~	1.74	1.5~~1.51~~	1.61	1.74~~1.75~~	1.4~~1.17~~	1.99
Mooring A	**2011**	1.15 -0.12 -0.8	1.66~~1.70~~ -0.04 -0.07	1.4 -0.11 -0.1	1.45 -0.16	1.5 -0.25 -0.24	1.43~~1.12~~ -0.04 0.03	2.01 0.02
	2012	1.39 0.12 0.16	1.66~~1.75~~ 0.01 -0.08	1.43 -0.09 -0.07	1.42 -0.19	1.68 -0.07 -0.06	1.2~~0.96~~ -0.2	2.04 0.05
	2013	1.05 -0.22 -0.18	1.82~~2.06~~ 0.32 0.08	1.51 0.0 0.01	1.61 -0.01	1.48 -0.27 -0.26	1.27~~1.16~~ -0.01 -0.13	2.02 0.03
	2014	1.51 0.24 0.28	1.98~~2.12~~ 0.37 0.24	1.71 0.2 0.21	1.85 0.24	2.09 0.34 0.35	1.68~~1.41~~ 0.25 0.28	1.96 -0.03
	2015	1.32 0.05 0.1	1.9~~2.13~~ 0.39 0.16	1.73 0.21 0.23	1.74 0.12	2.09 0.34 0.36	1.69~~1.38~~ 0.21 0.29	1.99 0.0
	2016	1.2 -0.07 -0.03	1.44~~1.61~~ -0.13 -0.3	1.3 -0.22 -0.21	-	1.67 -0.08 -0.07	1.21~~0.96~~ -0.21 -0.19	1.91 -0.08
	2017	0.96 -0.26	1.72 -0.02	1.43 -0.07		1.66 -0.08	1.33 -0.07	
		ULS	CPOM	AWI	JPL	GSFC	CS2SMOS	APP-x
	2011–2016	1.36~~1.38~~	1.88~~1.9~~	1.68~~1.69~~	1.79~~1.78~~	1.84~~1.86~~	1.65~~1.36~~	2.01~~2.02~~
Mooring B	**2011**	1.42 0.04 0.06	1.9~~0 -~~ 0.01 0.02	1.61 -0.08 -0.07	1.61 -0.17	1.61 -0.25 -0.22	1.64~~1.27~~ -0.09 -0.01	2.03 0.02
	2012	1.4 0.02 0.04	1.89 -0.01 0.02	1.67 -0.01 0.0	1.72 -0.07	1.7 -0.16 -0.13	1.65~~1.3~~ -0.06 0.0	2.06 0.05
	2013	1.21~~1.2~~ -0.19 -0.16	2.04 0.14 0.17	1.77 0.08 0.09	1.83 0.04	1.75 -0.11 -0.09	1.68~~1.51~~ 0.15 0.04	2.03 0.02
	2014	1.47 0.09 0.11	2.02 0.12 0.14	1.77 0.09 0.1	1.85 0.06	2.09~~2.09~~ 0.23 0.25	1.82~~1.53~~ 0.17	1.98 -0.04
	2015	1.55 0.17 0.19	2.02 0.12 0.15	1.83 0.14 0.16	1.92 0.14	2.06 0.2 0.22	1.74~~1.41~~ 0.05 0.1	2.02 0.0
	2016	1.26 -0.13 -0.1	1.55 -0.36 -0.33	1.47 -0.22 -0.21	-	1.95 0.09 0.12	1.42~~1.14~~ -0.22 -0.23	1.96 -0.05

		ULS	CPOM	AWI	JPL	GSFC	CS2SMOS	APP-x
	2017	1.21 -0.15	1.71 -0.17	1.61 -0.06		1.69 -0.15	1.59 -0.06	
		ULS	**CPOM**	**AWI**	**JPL**	**GSFC**	**CS2SMOS**	**APP-x**
	2011–2016	1.37~~1.41~~	1.88~~1.9~~	1.62~~1.63~~	1.81	1.81~~1.82~~	1.52~~1.26~~	2.08
	2011	1.48 ~~0.07~~ 0.11	1.7~~1.70~~ ~~-0.20~~ -0.18	1.48 -0.15	1.54 -0.27	1.66 ~~-0.16~~ -0.15	1.44~~1.15~~ ~~-0.11~~ -0.09	2.15 0.07
	2012	1.39 -0.02	1.75 ~~-0.14~~ -0.13	1.51 ~~-0.12~~ -0.11	1.6 -0.21	1.66 ~~-0.16~~ -0.14	1.41~~1.12~~ ~~-0.14~~ -0.11	2.15 0.06
Mooring D	**2013**	1.06 ~~-0.35~~ -0.31	2.06 ~~0.16~~ 0.17	1.51 ~~-0.12~~ -0.11	1.59 -0.21	1.55 ~~-0.27~~ -0.26	1.28~~1.18~~ ~~-0.08~~ -0.25	2.16 0.08
	2014	1.64 ~~0.23~~ 0.27	2.12 ~~0.22~~ 0.23	1.89 0.26	2.13 0.33	2.15 ~~0.32~~ 0.34	1.86~~1.55~~ ~~0.29~~ 0.34	2.09 0.01
	2015	1.59 ~~0.18~~ 0.22	2.13 ~~0.24~~ 0.25	1.97 ~~0.34~~ 0.35	2.16 0.36	2.26 ~~0.43~~ 0.45	1.93~~1.57~~ ~~0.31~~ 0.4	2.0 -0.08
	2016	1.31 ~~-0.1~~ -0.06	1.61 ~~-0.28~~ -0.27	1.42 ~~-0.21~~ -0.2	-	1.67 ~~-0.15~~ -0.14	1.26~~0.99~~ -0.27	1.95 -0.14
	2017	1.13 -0.24	1.82 -0.07	1.57 -0.05		1.69 -0.11	1.49 -0.03	

5

Table 5. Winter (October-April) sea ice thickness growth rate (md^{-1}) for the period 2010-2018 in the central Arctic (regions 1-6), where rates are calculated based on product availability during each growth season.

Growth Rate (m/d)	CPOM	AWI	JPL	GSFC	CS2SMOS	APP-x
2010-2011				0.0032		0.0103
2011-2012	0.0064	0.0058	0.0057	0.0040	0.0050	0.0105
2012-2013	0.0055	0.0048	0.0052	0.0027	0.0044	0.0108
2013-2014	0.0058	0.0049	0.0054	0.0026	0.0042	0.0100
2014-2015	0.0057	0.0051	0.0053	0.0037	0.0044	0.0103
2015-2016	0.0051	0.0050		0.0025	0.0040	0.0099
2016-2017	0.0050	0.0041		0.0022	0.0034	
2017-2018	0.0047	0.0042		0.0014		

Assessment of Contemporary Satellite Sea Ice Thickness Products for Arctic Sea Ice

Heidi Sallila[1], Sinéad Louise Farrell [2], Joshua McCurry[3] ~~Joshua McCurry[2,3], Sinéad Louise Farrell[2,3]~~ and Eero Rinne[1]

[1]Finnish Meteorological Institute, Helsinki, Finland
[2]Cooperative Institute for Climate and Satellites, University of Maryland, College Park, MD, USA
[3]Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, USA
~~[2]Cooperative Institute for Climate and Satellites, University of Maryland, College Park, MD, USA~~
~~[3]NOAA Laboratory for Satellite Altimetry, College Park, MD, USA~~

Correspondence to: Heidi Sallila (heidi.sallila@fmi.fi)

Abstract. Advances in remote sensing of sea ice over the past two decades have resulted in a wide variety of satellite-derived sea ice thickness data products becoming publicly available. Selecting the most appropriate product is challenging given end user objectives range from incorporating satellite-derived thickness information in operational activities, including sea ice forecasting, routing of maritime traffic, and search and rescue, to climate change analysis, longer-term modeling, prediction, and future planning. Depending on the use case, selecting the most suitable satellite data product can depend on the region of interest, data latency, and whether the data are provided routinely, for example via a climate or maritime service provider. Here we examine a suite of current sea ice thickness data products, collating key details of primary interest to end users. We assess eight years of sea ice thickness observations derived from sensors onboard the CryoSat-2 (CS2), Advanced Very-High-Resolution Radiometer (AVHRR) and Soil Moisture and Ocean Salinity (SMOS) satellites. We evaluate the satellite-only observations with independent ice draft and thickness measurements obtained from the Beaufort Gyre Exploration Project (BGEP) upward looking sonars (ULS) and Operation IceBridge (OIB), respectively. We find a number of key differences among data products, but find that products utilizing CS2-only measurements are reliable for sea ice thickness, particularly between ~0.5 m and 4 m. Among ~~those compare~~data compared, a blended CS2-SMOS product was the most reliable for thin ice. Ice thickness distributions at the end of winter appeared realistic when compared with independent ice draft measurements, with the exception of those derived from AVHRR. There is disagreement among the products in terms of the magnitude of the mean thickness trends, especially in spring 2017. ~~We noticed disagreement among the products in mean thickness trends especially in the winters of 2012–2013 and 2016–2017.~~ Regional comparisons reveal ~~large~~ noticeable differences in ice thickness between products, particularly in the ~~Kara Sea, an area~~marginal seas in areas of considerable ship traffic.

1 Introduction

With the observed decline in Arctic sea ice extent (Parkinson and Cavalieri, 2008; Markus et al., 2009; Perovich et al., 2018) and interests in the exploitation of regional natural resources, human activities in the Arctic have increased, alongside concerns for the state of the ice cover. Numerous objectives, ranging from sea ice forecasting and climate monitoring to navigation, require observations of Arctic sea ice, and it is recognised as an Essential Climate Variable (ECV) in the Global Climate Observing System (GCOS, Belward and Dowell, 2016). Given its remote location, seasonally-available sunlight, and inhospitable climate, remote sensing provides the only means to obtain routine, basin-scale, and sustained observations of the Arctic Ocean. ~~Objectives for exploiting satellite derived observations of the polar oceans are broad, and include improving sea ice forecasts, monitoring trends and variability in the polar climate system, advancing long term climate predictions and future planning, understanding changes in marine habitats for wildlife and fisheries, as well as supporting operational activities, including navigation, hazard monitoring, disaster response, search and rescue, and many commercial activities. Because of these objectives, interest in routine measurements of sea ice thickness has increased, and is concurrent with the availability of new satellite data products that are the direct result of advances in remote sensing over the past few decades.~~

~~Sea ice is a key component of the Earth's climate system, acting as an insulator between the ocean and atmosphere, and its importance is recognised by the Global Climate Observing System (GCOS) through its status as an Essential Climate Variable (ECV) (Belward and Dowell, 2016). Sea ice thickness is an important sub-variable of the sea ice ECV.~~ Although sea ice extent ~~has~~ is traditionally ~~been~~ the most widely discussed~~more exploited~~ variable, sea ice thickness measurements are just as~~, if not more,~~ important~~than extent~~, and ~~are~~ needed together with ice concentration to calculate sea ice volume, the best indicator of ~~the~~ change~~s~~ in the Arctic ice cover (e.g., Laxon et al., 2013; Song, 2016). Even if ice extent ~~would~~ remains stable in ~~the same for~~ consecutive years, if the thickness decreases, the ice cover will be less resilient and more unlikely to regain thickness, eventually leading to decreased extent and volume. ~~Furthermore~~For example, Laxon et al. (2013) speculated that lower ice thickness, and volume, may ~~be~~ have been a contributing factor ~~behind~~ to the ~~recent~~ September 2012 record minimum~~s of~~ sea ice extent. ~~One way to understand the complex interactions of sea ice and its sub-variables in the climate system is through models, and in turn models initialized with sea ice observations can be used to forecast future ice cover conditions. Recently satellite estimates of sea ice thickness have been used for the first time in model initialisations, revealing great improvements in model predictions (Yang et al., 2014; Allard et al., 2018; Xie et al., 2018), and these studies emphasize the importance of careful selection of the initialisation product. There have been other model studies with observational products from non-satellite sources such as airborne campaigns (e.g., Lindsay et al., 2012). However this type of data has restrictions in spatial and temporal coverage, and needs interpolation to emulate Arctic wide coverage.~~

As a result of advances over two decades in remote sensing of Arctic sea ice, a wide variety of satellite sea ice thickness data products have become available to the scientific community. Radar altimeter measurements from CryoSat-2 (Wingham et al., 2006) and laser altimeter measurements from ICESat and ICESat-2 (Markus et al., 2017) are major sources for estimating sea ice thickness, and due to high-inclination orbits, provide nearly full coverage of the Arctic Ocean (Laxon et

5 al., 2013; Markus et al., 2017). One of the most widely used thickness data sets derives from the radar altimeter flown on CryoSat 2, the only satellite currently operating that provides nearly full coverage of the Arctic Ocean (Laxon et al., 2013). Sea ice tThickness is derived from carefully processed radar altimeter measurements of ice surface elevation (Laxon et al., 2013; Kurtz et al. 2014), which requires knowledge of snow loading, and the assumption of hydrostatic equilibrium (Giles et al. 2007). Because of differing approaches to for retracking radar altimeter waveforms, a variety of processing algorithms exist, resulting in an array of altimeter-derived thickness data products (e.g., Laxon et al., 2013; Kurtz et al., 2014; Ricker et

10 al. 2014; Price et al. 2015; Tilling et al., 2015, 2016, 2018). These products broadly agree in the spatial distribution and basin-scale gradients of ice thickness across the Arctic Ocean, but differ in their absolute magnitude. In addition to the CryoSat 2 only ice thickness products, there are other productsThickness may also be derived from passive microwave radiometer measurements, as well as from visible and synthetic aperture radar imagery, that and these observations offer

15 additional information such as coverage in the marginal ice zone or detection of sea ice conditions during the Arctic summer (e.g., Kaleschke et al., 2012; Key and Wang, 2015; Ricker et al., 2017a). Recent studies have demonstrated the utility of initialising numerical models with satellite-derived estimates of sea ice thickness to improve model predictions (Yang et al., 2014; Allard et al., 2018; Blockley and Peterson, 2018; Stroeve et al., 2018; Xie et al., 2018).

20 Given the variety of sea ice thickness data products that are now available, and broad observational objectives, ranging from advancing long-term climate predictions and future planning, to supporting operational activities, including navigation, commerce, hazard monitoring, search and rescue, and disaster response, identifying the most suitable product depends on the end user requirements. Without knowledge, means, or time to compare the different products, the end user may not select the product most appropriate for their needs. WhichDetermining which satellite data product is the most suitable appropriate

25 depends upon a variety of factors, including the end user's region of interest, and data product characteristics including spatial coverage, temporal and spatial resolution, accuracy and quality, as well as data availability and latency.

In this paper we review a set of publicly-available satellite-derived sea ice thickness products, and compare them side by side for the first time, outlining their key attributes which are of interest to potential end users and for a range of applications. We

30 evaluate sea ice thickness estimates derived from observations collected by CryoSat 2 (CS2), the Advanced Very High Resolution Radiometer (AVHRR), and the Soil Moisture and Ocean Salinity satellite (SMOS). Although previous studies have evaluated individual satellite products (e.g. Laxon et al., 2013; Ricker et al., 2017b; Wang et al., 2016) and assessed the impact of retracker differences (e.g. Xia and Xie, 2018; Yi et al., 2019), none have focused on product intercomparison and have therefore lacked the details sought by many end users. To address this need, We we compare data product attributes

3

and assess differences across products ~~both~~ from both a regional and seasonal perspective, across the central Arctic Ocean and peripheral seas (Figure 1). We also evaluate the satellite data products, through comparisons with independent thickness measurements obtained in situ. ~~There are several airborne and in situ observations suitable for validating the satellite measurements. Flight campaigns, such as Operation IceBridge (OIB) and CryoVEx, complement, and are used to evaluate, NASA's Ice Cloud and Elevation Satellite (ICESat) and CryoSat 2 observations, respectively. Comprehensive datasets covering a wide variety of Arctic variables are available from field measurements obtained during research cruises like the Norwegian young sea ice expedition (N-ICE2015), or the Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAIC), planned to commence in October 2019. Also, a good set of observational products are available from North Polar drifting stations, ice mass balance buoys, upward looking sonars (ULS) that form part of the Beaufort Gyre Exploration Project (BGEP) and airborne EM measurements gathered through the Canadian Arctic Sea Ice Mass Balance Observatory campaign (see for example further information in Ricker et al., 2017b and Allard et al., 2018). For our evaluations we make use of OIB and ULS observations, which provide thickness and ice draft, respectively, since these variables are most compatible with the satellite data, and their records are more extensive and sustained over the period of interest, compared to many of the other publicly available in situ data sets.~~

The paper is structured as follows: ~~section~~ Section 2 introduces the sea ice thickness products and the independent data sets used for validation. Section 3 describes the methods used for product comparison. In ~~section~~ Section 4 we present the results of the comparisons across satellite products, and a product evaluation against independent estimates of ice thickness obtained from in situ and airborne sensors. We provide a summary discussion in ~~section~~ Section 5. In ~~section~~ Section 6 we conclude with a look ahead and provide recommendations for future satellite-derived thickness products aimed at addressing operational needs.

2 Data

2.1 Satellite Data Products for Arctic Sea Ice Thickness

Here we assess six contemporaneous sea ice thickness data products, derived from satellite measurements collected by CryoSat-2 (CS2), the Advanced Very-High-Resolution Radiometer (AVHRR), and the Soil Moisture and Ocean Salinity satellite (SMOS). Since our focus is an assessment of product utility for a range of operational activities we required that data products were open access, had basin-wide coverage of the Arctic Ocean (Figure 1), and were available for the majority of the CS2 observation period (April 2010 to present)~~time period under investigation~~. Four ~~of these~~ products comprise CS2-only thickness estimates, and include the Centre for Polar Observation and Modeling (CPOM) seasonally-averaged thickness product (Laxon et al., 2013, Tilling et al., 2018), the Alfred Wegener Institute (AWI) ~~version 2~~ monthly thickness product

(Ricker et al., 2014), the NASA Jet Propulsion Laboratory (JPL) monthly thickness product (Kwok and Cunningham, 2015) and the NASA Goddard Space Flight Center (GSFC) 30-day thickness product (Kurtz and Harbeck, 2017). For comparison with CS2-only thickness estimates, we also consider blended CS2 - Soil Moisture Ocean Salinity (CS2SMOS) weekly ice thickness data (Ricker et al., 2017a), and the NOAA Climate Data Record (CDR) Extended Advanced Very High Resolution Radiometer (AVHRR) Polar Pathfinder (APP-x) daily ice thickness product. Further details about each data product are provided below. Common characteristics including measurement technique, temporal and spatial coverage, latency, frequency, resolution, and algorithm-specific details, are outlined in Table 1. The products selected for assessment provide a representative sample of available sources. We ~~do however~~ acknowledge that the list of satellite-derived ice thickness products is not exhaustive and other sources of similar observations exist. For example two additional sources of CS2-only sea ice data are now publicly available, one from the ESA Climate Change Initiative (CCI, Hendricks et al., 2018) and a second from the Laboratoire d'Études en Géophysique et Océanographie Spatiales (LEGOS) Center for Topographic studies of the Ocean and Hydrosphere (CTOH, Guerreiro et al., 2017). Details of these two products are included in Table 1 for completion but not included in the analysis. Sea ice age characteris~~s~~zation, and ice thickness, from the Visible Infrared Imager Radiometer Suite (VIIRS) instrument on the Suomi National Polar-orbiting Partnership (S-NPP) satellite, may also be derived in a similar manner to the APP-x data product (Key et al., 2013)~~. These products were not included in this study since~~, however these data were not publicly available at the time of writing~~, or required prior registration and approval with the data product provider~~, and are hence not included in the analysis. Since our focus is the availability of current sea ice thickness measurements for end users, our study period ~~ranges~~ spans the last eight years from fall 2010~~–~~ to spring 2018, facilitating an assessment of recent sea ice conditions in the Arctic Ocean. Thus we do not consider satellite thickness records prior to the launch of CS2, such as from Envisat (e.g. Giles et al., 2008) or ICESat (e.g. Kwok et al., 2009).

2.1.1 CPOM

CPOM was first to produce publicly-available estimates of sea ice thickness from CS2 and they provide both near real time (NRT) thickness products for 28, 14, and 2-day observation periods (Tilling et al., 2016), which are updated on a daily basis with a typical latency of 3 days, as well as a monthly and seasonally-averaged thickness data product~~s~~ (~~Laxon et al., 2013~~Tilling et al., 2018) ~~that is available for two seasonal periods only (October - November and March - April~~, Table 1). Archived data coverage begins in ~~March 2011~~November 2010, and ~~new seasonal~~monthly averages are available on an ad-hoc basis. Thickness data are available for the months of October through April on a 5 km grid for the full northern hemisphere (regions 1–14, Fig. 1), as well as on a 1 km grid for specific subpolar regions.

~~The basis for estimates of ice freeboard, and ultimately thickness, in the CPOM data product is level 1b CS2 waveforms, available from the European Space Agency (ESA) that~~CryoSat-2 radar waveforms delivered through the ESA Level 1b product are used to identify lead and ice floe elevations (~~Laxon et al., 2013~~Tilling et al., 2018). ~~The mean scattering horizon~~

of each waveform is identified through a threshold-first-maximum-retracking algorithm (TFMRA) with the threshold set at 70% of the power of the first maximum (Laxon et al., 2013). Radar waveforms associated with leads and ice floes are distinguished using fixed criteria for stack standard deviation and pulse peakiness (Laxon et al., 2013; Tilling et al., 2018). Leads and ice floes are distinguished using fixed criteria for stack kurtosis, stack standard deviation, and pulse peakiness. Ice elevation is defined as 70 % of the peak amplitude on the leading edge of the first peak in the radar waveform (Tilling et al., 2018). Lead elevations are determined separately (Tilling et al., 2018) through application of the retracking method developed in Giles et al. (2007) where a Gaussian plus exponential model is fit to each waveform. The CPOM algorithm utilizes the UCL13 mean sea surface (MSS) for the calculation of sea surface height anomalies, and to reduce the impact of geoid slope on freeboard estimates. This step is especially important in areas of low lead fraction (Skourup et al., 2017). Auxiliary information including the location of the ice edge and a product distinguishing first-year ice (FYI) from multi-year ice (MYI) is needed in the CPOM thickness algorithm. The sea ice edge is defined as the 75 % ice concentration contour in the For sea ice concentration the NSIDC Near Real Time DMSP SSMI/S Daily Polar Gridded Sea Ice Concentrations Concentration data set (available from the National Snow and Ice Data Center, NSIDC), while are used and the sea ice type is derived from the EUMETSAT's Ocean and Sea Ice Satellite Application Facility (OSI-SAF) product (Tilling et al., 2016 2018). The depth and density of snow on sea ice are based on the snow climatology of Warren et al. (1999) and applied by ice type. On MYI, the monthly mean climatological value is applied, and this value is halved for FYI (Kurtz and Farrell, 2011), meaning constant snow depth and density values for MYI and FYI are applied each month (Tilling et al., 2018). At the time of writing, only seasonal averages for two seasonal periods (October-November and March-April) were available, and this data version was used as the baseline data set in our analysis. Undefined thickness estimates are indicated by a value of 0.0000 in the CPOM product and these were removed before further processing.

2.1.2 AWI

AWI also provides monthly CS2 data products for October through April. Archived data begins in November 2010 and new monthly data are made available on a variable basis, but typically with one month latency (Table 1). In addition to thickness, AWI offers a number of additional geophysical and instrument parameters in their data product, including sea ice freeboard and concentration. It Of the products analysed here, it is the only product that provides uncertainty estimates for thickness and freeboard values (Hendricks et al., 2016). All parameters are provided on a 25 km grid for all regions of the Arctic (i.e., regions 1–14).

The AWI product uses the ESA L1b baseline C CryoSat-2 radar waveforms, applying a 50 % Threshold-First-Maximum Retracker (TFMRA, Ricker et al., 2014) to derive elevation for all surface types. The algorithm employed by AWI does not differ significantly from that used in the derivation of the CPOM product, although In contrast to the CPOM product, AWI currently uses the DTU15 MSS for the calculation of sea surface height anomalies (Hendricks et al., 2016). Skourup et al. (2017) demonstrated that processing with either the DTU15 or UCL13 MSS results in consistent freeboard estimates, with

only small deviations in areas of low lead concentration. For both sea ice concentration and type, AWI uses OSI-SAF products (Hendricks et al., 2016), and the ice edge is defined at the 70 % ice concentration contour. The AWI product applies snow depth based on the climatology of Warren et al. (1999), following the method described in Laxon et al. (2013), hereafter referred to as the "modified Warren climatology" (MWC). In the MWC, snow depth on FYI is 50 % of the climatology, and snow density is a monthly constant per ice type. The AWI product allows mixed ice types at the boundaries between MYI and FYI, based on confidence levels given in the ice-type product (Ricker et al., 2014). For this analysis AWI product version 2.0 was used and ice concentration, freeboard and thickness were assessed.

2.1.3 JPL

The JPL product provides monthly CS2 thickness data from January 2011 to December 2015, on a 25 km grid (Table 1). Coverage is limited geographically to the central Arctic (regions 1–6). Information on the ice concentration threshold defining the area within which sea ice thickness data are calculated is not provided. The JPL retracker for deriving surface elevation is based on the waveform centroid, rather than a leading-edge approach used in the AWI TFMRA algorithm. The JPL freeboard algorithm also uses the EGM2008 geoid, rather than a MSS model, which can result in anomalous sea ice thickness estimates in areas of steep ocean topography, such as near the Lomonosov and Gakkel Ridges (Skourup et al., 2017). Information about which sea ice concentration product is applied in the JPL algorithm is not provided in the available literature, nor the ice concentration threshold used to define the ice edge. Kwok and Cunningham (2015) state that the sea ice type is derived from analysed fields of Advanced Scatterometer (ASCAT) data. The data are currently only accessible to the public after registration on a product website. Snow depth and density in the JPL product is based on the MWC and applied similarly to Laxon et al. (2013), except that where Laxon et al. (2013) used a 50 % reduction in the climatology over FYI, Kwok and Cunningham (2015) examined both a 50 % and 70 % reduction, it was left unclear which version was selected for the final online product. Information about which sea ice concentration product is applied in the JPL algorithm is not provided in the available literature, but Kwok and Cunningham (2015) state that the sea ice type is derived from analysed fields of Advanced Scatterometer (ASCAT) data. The data are currently only accessible to the public after registration on a product website. Undefined thickness values corresponding to 9999.0 and -1.0 were removed before processing.

2.1.4 GSFC

GSFC provides 30-day ice thickness averages derived from CS2. Coverage begins in October 2010 and new data continue to be made available on a time-varying basis, but typically with a six week latency (Table 1). In addition to ice thickness, the GSFC product includes estimates of freeboard and surface roughness (Kurtz and Harbeck, 2017). All parameters are available on a 25 km Polar Stereographic SSM/I Gridgrid, for regions 1–8.

Unlike other CS2 freeboard algorithms, the GSFC product is derived using a waveform-fitting method ~~based on empirical calculation of expected returns from various surface types~~. The surface type, elevation, and other properties of the received radar waveform are derived through statistical comparison to analytically pre-computed waveforms, ~~Received waveforms are matched to a pre-calculated waveform by means of~~ using a least squares error minimization. Kurtz et al. (2014) note that this approach should nominally result in freeboard values lower than those derived from TFMRA-based methods. ~~Additionally, t~~The GSFC algorithm utili~~s~~zes the DTU10 MSS for freeboard calculation. Skourup et al. (2017) have demonstrated that more recent MSS models (e.g., UCL13, DTU15), ~~that~~ which incorporate sea surface height data from CS2, enhance the definition of gravity features, resulting in a more accurate freeboard derivation. Indeed Skourup et al. (2017) found that the DTU10 MSS in particular was not sufficient for freeboard processing due to decimeter-level discontinuities at 81.5°N and 86°N as well as at the ice edge, which resulted in erroneous freeboard measurements at these locations. The GSFC product uses a 70 % ice concentration threshold derived from the NSIDC Near-Real-Time DMSP SSMI/S Daily Polar Gridded Sea Ice Concentrations available at NSIDC to define the sea ice edge, and the OSI-SAF product is used for sea ice type (Kurtz and Harbeck, 2017). An additional unique attribute of the GSFC algorithm is the use of a single ice density value (915 kg/m3) for all ice types in the thickness derivation step (Kurtz et al., 2014). In the case of the other CS2-only thickness products a dual ice density approximation is made, with the assumption of a lower density for MYI (882~~.0~~ kg/m3) than for FYI (91~~6.~~7 kg/m3). Despite using a single density for both ice types, the GSFC algorithm employs the MWC for~~a~~ snow depth ~~correction to the Warren et al. (1999) climatology for seasonal ice, as described in Laxon et al., (2013) and determined from OIB snow depth measurements provided in Kurtz and Farrell (2011). This approach is consistent with all other CS2 thickness products (Table 1). For sea ice concentration GSFC uses the NSIDC Near-Real-Time DMSP SSMI/S Daily Polar Gridded Sea Ice Concentrations and the OSI-SAF product for sea ice type (Kurtz and Harbeck, 2017)~~. Here we used the GSFC product version 1.0 ice freeboard and thickness data. Undefined thickness estimates are indicated by a value of -9999 in the GSFC product and these were removed before further processing.

2.1.5 CS2SMOS

The CS2SMOS sea ice thickness product, developed by AWI and the University of Hamburg, is a blended product of thickness estimates from CS2 and SMOS. It provides weekly data for the Arctic northward of 50 °N on an EASE2 grid, with 25 km grid resolution, across regions 1–14. It is available for a period starting in ~~January 2011~~November 2010, ending April 2017.

The SMOS mission provides L-band observations of brightness temperature, which may be used to derive ice thickness in areas where thin sea ice exists (Kaleschke et al., 2012). CS2 exploits radar altimetry to measure the difference in height between the snow/ice surface and sea surface, which is then used to derive sea ice thickness through the assumption of hydrostatic equilibrium. Since CS2 was designed to measure ice thicker than 0.5 m, it may be advantageous to blend CS2 estimates with complementary estimates from SMOS. Due to the satellites having different spatial and temporal coverage,

8

optimal interpolation is used to merge the two data sets (Ricker et al., 2017a). The algorithm includes weighting the data based on the known uncertainties of the products and modelled spatial covariances (Ricker et al., 2017a; 2017b). For sea ice concentration and type the OSI-SAF Arctic daily products are used. Snow depth and density follow the MWC. The CS2 product used is the AWI CS2 product with processor version 1.2 (Ricker et al., 2014; Hendricks et al. 2016) and the SMOS thicknesses are from the University of Hamburg processor version 3.1 (Tian-Kunze et al., 2014; Kaleschke et al., 2016). For sea ice concentration and type OSI SAF Arctic daily products are used. For our analysis thickness data from CS2SMOS product version 1.3 were used.

2.1.6 APP-x

The NOAA extended AVHRR Polar Pathfinder (APP-x) Thematic Climate Data Record (CDR)product provides sea ice thickness estimates, along with 18 other geophysical variables, in a climate data record (Key and Wang, 2015). Thickness estimates are available for both the Arctic Ocean (regions 1-14) and Southern Oceans, spanning 1982 to present. Data are provided twice daily, with a typical latency of approximately 4 days (Table 1). In contrast to the other satellite-derived thickness products, yYear-round thickness estimates are available, including throughout the summer, and are provided on a 25 km grid. We note a gap in the thickness record at the time of writing for the period 8 March to 1 May 2017 and hence no assessment of APP-x data exists forin spring 2017 in the following analyseswas not possible.

Sea ice thickness estimates are derived from AVHRR satellite radiometer measurements using a one-dimensional thermodynamic model (OTIM), based on a surface energy budget equation (Key and Wang, 2015). The OTIM model derives sea ice thickness as a function of surface heat fluxes, surface albedo and radiation, which all contribute to surface energy budget (Wang et al., 2010). Furthermore, most of the flux and radiation parameters in the equations are functions of surface skin and air temperatures, surface air pressure, surface air relative humidity, ice temperature, wind speed, cloud amount and snow depth, which are input parameters in the model (Wang et al., 2010). Ice thickness is calculated as a function of solar radiation and surface heat fluxes, surface albedo, and transmittance, where the surface fluxes are calculated as a function of surface skin and air temperature, surface air pressure, relative humidity, ice temperature, wind speed, cloud cover and snow depth. Thus OTIM relies on the following input for the ice thickness calculation: percentage cloud cover, surface skin temperature, surface broadband albedo and surface shortwave radiation fluxes, of which the latter two are obtained for daytime retrievals only. The sea ice concentration source for the APP-x product is Nimbus-7 SMMR and DMSP SSM/I data processed with the NASA Team Algorithm (Key and Wang, 2015) and the ice edge is defined at the 15 % ice concentration contour. The sea ice type is converted from the reflectances measured directly by AVHRR. Additionally the following input is needed for ice thickness: percentage cloud cover, surface skin temperature, surface broadband albedo and surface shortwave radiation fluxes, of which the latter two are obtained for daytime retrievals only. For the APP-x product the snow depth estimates are based on the snow depth climatology of Warren et al. (1999) but combined with field observations through experimentation and applied using monthly look-up tables (Wang, pers. comm., 2018). For our analysis thickness

data from APP-x CDR version 2.0 were used, and undefined ice thickness values identified as "9.96920996839e+36" were removed before processing.

Since sea ice dynamics are not included in the OTIM model, the thickness errors in the APP-x product are larger where the ice surface is not smooth, i.e. in regions with pressure ridges, hummocks and melt ponds (Wang et al., 2010). Generally OTIM tends to overestimate ice thickness, in particular for thin ice while underestimating thick ice, as the energy budget approach is less sensitive for thick ice (Wang et al., 2010). ~~Then again, APP-x tends to underestimate thick ice, as the energy budget approach is less sensitive for thick ice.~~ Moreover since the satellite sensor retrieves 2 m air temperature, ice surface temperature is derived from the 2 m measurement. Wang et al. (2010) state that the thickness estimates are more accurate for nighttime retrievals, when 2 m air temperature and ice surface temperature are closer, resulting in a smaller model error. For this reason, we ~~chose~~ utilise the nighttime estimates ~~to be used~~ in our study. Wang et al. (2010) note that errors due to uncertainties in snow depth and cloud fraction are the primary sources of error in the OTIM thickness estimates. We also note that OTIM is applicable to other optical satellite data including observations from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) instrument and EUMETSAT's Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument (Wang et al., 2010). ~~The sea ice concentration source for the APP-x product is Nimbus 7 SMMR and DMSP SSM/I data processed with the NASA Team Algorithm (Key and Wang, 2015). The sea ice type is converted from the reflectances measured directly by AVHRR.~~

2.2 Evaluation Data Sets

We evaluate the satellite-derived ice thickness estimates using independent upward looking sonar (ULS) observations of ice draft from Beaufort Gyre Exploration Project (BGEP) moorings, and airborne observations of ice thickness from Operation IceBridge (OIB). These data sets represent the most extensive and sustained record over the evaluation period, compared to many of the other publicly-available in situ thickness data sets.

2.2.1 Beaufort Gyre Exploration Project

Since August 2003, ~~the Beaufort Gyre Exploration Project~~ (BGEP) has operated a series of moorings in the Beaufort Sea, which have included an ~~upward looking sonar~~ (ULS) instrument. From 2003 to 2014 the ULS instrument produced a range estimate every two seconds, increasing in frequency to once per second starting with the 2014–2015 deployment. By subtracting the ULS range estimate from instrument depth, draft is measured to an accuracy of +/- 0.05 m per individual measurement (Krishfield et al., 2006).

Here we utilize ULS draft measurements from three mooring locations (A, B, and D, Fig. 1) in the Beaufort Sea over the six-month period spanning November–April, for all years from 2010-2017~~2011–2016~~. Since the ULS measures ice draft as floes drift across the mooring location, the data represent a high-resolution, time-varying measurement of many individual leads

and ice floes, thus providing a more complete picture of the regional ice thickness distribution. The draft to thickness ratio is approximately 0.9 (e.g., Rothrock et al., 2008), but to accurately compute thickness from draft, knowledge of ice type, ice density and snow loading are required. Here we do not convert draft to thickness, since that would introduce additional uncertainties. Rather we use the characteristics of the ice draft distribution to evaluate the satellite-derived thickness distribution.

2.2.2 Operation IceBridge

~~Operation IceBridge~~OIB was launched in 2009, and is a sustained airborne mission designed to continue the collection of sea ice and land ice elevation measurements in the temporal gap between the end of the ICESat mission in 2009 and the launch of ICESat-2 in 2018 (Koenig et al., 2010). The mission includes an ~~airborne~~ altimeter, the Airborne Topographic Mapper (ATM), which provides high-resolution measurements of sea ice plus snow freeboard, and a snow radar instrument for derivation of snow depth (Newman et al., 2014). Together these allow for estimation of sea ice thickness (e.g., Farrell et al., 2012; Kurtz et al., 2013). Thickness uncertainty, calculated by propagation of estimated errors in the contributing variables, associated with IceBridge estimates is approximately 0.66 m (Richter-Menge and Farrell, 2013). Here we make use of the IDCSI4 (spring 2011) and NSIDC-0708 (spring 2012–2017) IceBridge thickness products available at the ~~National Snow and Ice Data Center - Distributed Active Archive Center website~~NSIDC (Kurtz et al., 2015; Kurtz, 2016). Due to the geographical layout of airborne flight surveys, the majority of IceBridge measurements sample multi-year sea ice in the Canada Basin (e.g., Richter-Menge and Farrell, 2013).

3 Methodology

3.1 Satellite Product Intercomparison

~~In order to compare the satellite data products, all ice thickness datasets are regridded to a 5 km grid, using a nearest neighbour algorithm with a 50 km search radius.~~ Seasonally-averaged ice thickness~~es~~ ~~are~~ is computed for each product over two periods: fall (October and November), and spring (March and April). Seasonal averages are calculated by taking the arithmetic mean value of all available thickness estimates across the Arctic region (Fig. 1) within the two month period. Results for fall and spring are shown for each product over the period 2011 - 2017 in Figures 2a and 2b, respectively. ~~Seasonal averages are only calculated using grid cells that contain valid thickness estimates for both months. This prevents the average from being biased by measurements taken at the beginning or end of an averaging period (i.e., measurements only available in one of the two months). For each product we also calculate a climatological mean ice thickness in the central Arctic (Fig. 1), averaged over the product record (as described in Table 1). Climatological means are calculated separately for the fall and spring seasons. For each season we then compute deviation from the climatological mean for each product.~~

11

To evaluate spatial variations in thickness within products, mean regional ice thickness is computed for the central Arctic Ocean and each peripheral sea (Table 2) using the polygons shown in Figure 1, for both the fall and spring seasons. Thickness values are reported for any region containing valid data points in the relevant months for at least one year of the product record. Although all regions with data are reported, regions outside the main product coverage areas defined in Table 1 may contain only a few data points and hence the reported regional ice thickness for some of the peripheral seas may represent only a small portion of the region as a whole (see Fig. 2).

To assess temporal variations across products, we calculate a baseline mean ice thickness for the central Arctic for the period common to all products (2011-2015), for both the fall and spring seasons. Subject to the time span over which each product is available (Table 1) we compute the anomaly with respect to the baseline mean for each season and report these results in Table 3. We also assess trends in winter ice growth for each product by calculating monthly mean ice thickness in the central Arctic (regions 1-6) during the period of study spanning 2010 to 2018 (Figure 5).

For point-to-point comparisons of the satellite data products, and to compute correlations across products, all ice thickness datasets are placed on a common 0.4° latitude by 4° longitude grid. We define the CPOM seasonally-averaged thickness product as the reference dataset against which the other products are compared. This approach is justified since the CPOM product was the first publicly-available CryoSat-2 thickness dataset (Laxon et al., 2013), and has been widely used in studies by end users (e.g. Allard et al., 2018; Blockley and Peterson, 2018; Stroeve et al., 2018). We therefore compute product differences as follows:

$$\Delta SIT = SIT_{cP} - SIT_{pC} \; , \tag{1}$$

where SIT_c is the CPOM sea ice thickness, SIT_p is the thickness of the product in question, and ΔSIT is the difference between the two. Correlation statistics are calculated utilising grid cells in which both data sets contain thickness estimates. These grid cell pairs are used to compute the Pearson correlation coefficient (r), the product difference across common cells according to Eq. 1, and the standard deviation of this difference (Figure 4).

To evaluate spatial variations in thickness across products regional ice thickness averages are computed for the central Arctic Ocean and each peripheral sea (Table 2) using the polygons shown in Fig. 1, for both the fall and spring seasons. Due to the limited ice extent in fall, the availability of observations in some peripheral seas is not sufficient to calculate a robust regional average in October-November (or in the spring season in the case of the JPL and GSFC products due to their limited spatial coverage).

3.2 Satellite Product Evaluation

Since the BGEP moorings are tethered to the seafloor, ULS measurements are representative of ~~all~~ discrete ice floes drifting over the mooring location. To facilitate comparison with the satellite data, we select all product data points within 200 km of the mooring, following the approach described in Laxon et al. (2013), creating a comparison region centred on the mooring location. In order to avoid influence from areas outside the comparison region, we use the original satellite thickness data as provided, rather than the gridded data described in Section 3.1 above.

Both the ULS draft data and satellite thickness data within the comparison region are averaged over one month intervals from November to April for the period of overlap between the product record and ULS data spanning fall 2010 to spring 2017. The correlation coefficient for ULS draft measurements and product thickness observations is calculated at each mooring based on these paired monthly data points (Figure 8)~~between each monthly-averaged remote sensing product and the average ULS measurements for same period is calculated for each mooring~~. Any individual ULS draft measurement thinner than 0.1 m is not included in the averaging, as these measurements may represent leads rather than ice floes (Krishfield et al., 2006; Krishfield et al., 2014). In the case of the CPOM product, ~~only seasonally averaged thickness data are available, so~~ correlations with the ULS ice draft are based on seasonal averages, processed in a similar manner to the~~rather than~~ monthly averages as described above.

The sea ice thickness data from Operation IceBridge consists of flightlines across the Canada Basin and the central Arctic Ocean (e.g. Richter-Menge and Farrell, 2013). Following Tilling et al. (2018)~~Laxon et al. (2013)~~, data acquired during IceBridge spring campaigns in 2011-2017 were placed on a 0.4° latitude by 4°longitude grid~~interpolated onto a 50 km grid~~ and compared against the satellite thickness data for March–April, for all six products ~~interpolated~~ at the same~~common~~ grid cell locations~~, using a nearest neighbor algorithm with a 100 km search radius~~. This allows the calculation of thickness differences between each product and IceBridge, as follows:

$$\Delta SIT_{eval} = SIT_{oibp} - SIT_{poib} \ , \tag{2}$$

where SIT_{oib} is the Operation IceBridge sea ice thickness, SIT_p is the thickness of the product in question, and ΔSIT_{eval} is the difference between the two. Using the gridded data, we also calculate correlation coefficients, and standard deviation between each product and the IceBridge thickness observations (Figure 9).

4 Results

Here we present, for the first time, a side-by-side comparison of a suite of available CS2-only products, alongside a blended CS2-SMOS (CS2SMOS) product and one altimetry-independent sea ice thickness product, APP-x. The results are presented

13

in three parts. First we provide a review of Arctic sea ice thickness variability during the last eight years. Next we compare regional and temporal differences between the satellite products across the Arctic regions. Finally we evaluate the satellite-derived thicknesses through comparisons with independent measurements.

4.1 State of the Arctic Sea Ice Thickness

Seasonal ice thickness for fall and spring is shown in Fig. ~~ure~~ 2 for the period 2011–2017. Following the observed low summer sea ice minimum extents in 2011 and 2012 (Parkinson and Comiso, 2013), we find that the lowest ice thickness was recorded in fall 2011 and 2012 (Fig. 2a). Mean ice thickness was 1.12 m (1.19 for the CS2-only products) in fall 2011, with variations of +0.3 to -0.23 m (+0.24 to -0.11 for CS2-only) among the products. For fall 2012 the mean thickness was 1.18 m (1.19 for CS2-only), varying from +0.34 to -0.36 m (+0.24 and -0.11 for CS2-only) among products. The loss of multi-year ice in the summer of 2012 due to the record sea ice minimum resulted in an overall thinner ice cover during the following winter/spring (i.e. spring 2013) when the mean ice thickness was 2.05 m (2.0 m for CS2-only) among products~~, which is visible in~~ (Fig. 2b). Consistent with the results first noted in Tilling et al. (2015), we see that following ~~Following~~ a cool summer in 2013, survival of ice through the melt season resulted in a rebound in thickness in fall 2013 with mean thickness of 1.46 m (1.58 for CS2-only), and a thicker winter mean sea ice thickness of 2.34 m in spring 2014 (all products), which has persisted in the central Arctic for subsequent seasons (Fig. 2). ~~This finding is consistent with earlier results shown in~~ Tilling et al. (2015)~~, who~~ also noted a slight recovery in 2013–2014 following low winter-time ice thickness in 2011 and 2012. The thickest ice at the end of winter was observed in 2014 (Fig. 2b) in a region stretching from northern Greenland near Kap Morris Jesup, to Banks Island in the Canadian Arctic Archipelago. This region of thick ice in the central Arctic has persisted throughout the following seasons. As of spring 2018 (not shown) the area of ice more than 3 meters thick adjacent to the northern coasts of Greenland and the Canadian Arctic Archipelago is still greater than that of ~~Spring~~ spring 2012 and 2013, although the spatial extent of this thick ice area has diminished since 2014. In multiple springs (2012, 2014, 2015, 2016) an outflow of thick ice extends from the southern Canada Basin into the southern Beaufort Sea, due to the dynamic action of the Beaufort Gyre circulation (Fig. 2b). While this band of thick sea ice is captured in all of the CS2 products, it is ~~not as easy to identify~~less distinct in the CS2SMOS product, and does not appear in the APP-x product, apart from spring 2012 (Fig. 2b). This overall picture of the state of Arctic sea ice thickness over the last several years is consistent across the CS2-only products and the CS2SMOS blended product. Although the APP-x product captures the spatial gradient in thick to thin ice, from the northern coasts of Greenland and Canada to the Siberian coastline, respectively, the product does not resolve many of the recent major changes in sea ice thickness conditions in either the spring or fall.

4.2 Regional Differences

The satellite-derived thickness products differ in their regional coverage and the availability of thickness estimates across the northern hemisphere. APP-x has the most widespread coverage, although CPOM, AWI and CS2SMOS all provide thickness estimates in the sub-polar seas. The JPL product only provides estimates for an area approximately contiguous with regions

14

1–6, while the GSFC product provides estimates for ~~an area approximated by~~ regions 1–8 (Table 1, ~~and~~ Fig. 2). Three of the six products (AWI, CS2SMOS, and APP-x) consistently resolve thin ice (<= 0.5 m) at the periphery of the ice pack (regions 3–7) during the fall, but only CS2SMOS and APP-x resolve thin ice in these regions in spring (Fig. 2b). In addition APP-x does not resolve the thicker ice of the central Arctic Ocean, as evident in the other products, especially in the fall.

5

Maps of differences in ice thickness across products, as defined in Eq. 1~~.~~, are shown in Fig. 3, and distributions of the differences between products are provided in Fig. S1. Average ~~regional~~ seasonal ice thickness for each region of the Arctic (as defined in Fig. 1) over the ~~period 2011–2017~~available product record is provided in Table 2. In general, differences in ice thickness are larger in the fall than spring, although for spring even across the CS2-only products, differences range from 0.25 m to 2.12 m in regions 6-14 (Table 2). ~~As expected, given strong consistencies in algorithmic approach, the~~The closest agreement across products is found between ~~the~~ CPOM~~and~~, AWI and JPL products (Fig. 3, 4, Fig. S1). The thickness estimates of these products have a correlation of 0.91-0.92 in the spring and 0.88 in the fall, with a mean difference of 0.03 - 0.08 m ~~0.88 for both seasons and a mean difference of 0.02-0.03 m~~ (Fig. 4). However, there are noticeable regional differences in mean thickness during the observation period (Table 2), for example in spring in the Greenland Sea the AWI product is 0.32 m thicker than the CPOM product, while in~~(0.42 m difference) and~~ the Canadian Archipelago the CPOM product is 0.37 m thicker~~(0.53 m difference)~~. ~~From an Arctic-wide perspective, the mean thickness estimates in the JPL product also differ little from the CPOM product (0–0.02 m, Fig. 4) and are highly correlated with the CPOM product, 0.86 in spring and 0.83 in fall (Fig. 4).~~ There are ~~however~~ also differences in the spatial gradients of ice thickness (Fig. 3), particularly in the central Arctic Ocean~~, where~~. For example the JPL product estimates thicker ice close to Greenland, and thinner ice near the North Pole, and along the Siberian shelf zone than the CPOM product (see also Table 2). Of the CS2-only data products, ~~Figure~~Figs. 3 and S1 demonstrate~~s~~ that the GSFC ~~thickness~~ product is the most dissimilar to the CPOM data, with thickness in both the fall and spring periods being higher on average, though with year-to-year spatial variation ~~(Fig. 3)~~. Mean ice thickness differences range from 0.02 to 0.25~~0.08 to 0.27~~ m (Fig. 4), though despite these differences, the GSFC product is still highly correlated with the CPOM product (R = 0.85 in both spring and fall).~~, with correlations of 0.80 and 0.83 in the spring and fall, respectively (Fig. 4).~~

From a regional perspective, the GSFC product agrees closely with the CPOM product in spring (especially in regions 1-7), but larger thickness differences are recorded in fall across all regions (Table 2). The CS2SMOS product suggests thinner ice thicknesses than the CS2-only products, with mean differences of 0.08 to 0.36 ~~0.27 to 0.60~~ m (Fig. 4) and differences increasing towards the ice edge. For example in the Barents Sea, average spring thickness for the period 2011–2017 was only 0.38~~0.41~~ m, ~~over a meter~~exactly 1 m lower than the CPOM product ~~suggests~~ estimate (Table 2). Only in the MYI zone in spring does the CS2SMOS product provide estimates of ~~thicker ice~~ice thicker than the CPOM product in some years (Fig. 3b)~~, though never for a whole region (Table 2)~~. Despite its lower thickness estimates, the CS2SMOS product correlates well with CPOM, 0.86 and 0.88 in the fall and spring~~0.84 and 0.86 in the spring and fall~~, respectively (Fig. 4). We also note that

in fall 2017 the APP-x product shows thickness data covering an area south of the typical ice edge in regions 8–10 (Fig. 2a). Anomalous data in this region may be due to errors in the sea ice concentration field (not shown). Ice concentration is passed to the APP-x product from the lower-level APP product. In spring ~~For~~ ice in the central Arctic~~,~~ in the APP-x product is consistently thicker ~~in spring~~ over FYI~~,~~ and thinner across both the MYI zone and the thick outflow along the northern coast of Greenland ~~and~~ into the Canada Basin (Fig. 3b, Table 2). In the fall, the APP-x product contains mainly thinner ice for all regions (Table 2), except in 2011 in the Canada Basin, where the APP-x product suggests thicker ice than in the CPOM product, and in 2017 around the ice edge in regions 3–6 (Fig. 3a). Also, in fall, the APP-x product estimates thicker ice for regions 3-7, and 11, compared to the blended CS2SMOS product (Table 2). With correlations of 0.49 and 0.53 ~~0.38 and 0.53~~ in the spring and fall, respectively, the APP-x thickness ~~does~~ data do not correlate as well with the CPOM product as the other data sources (Fig. 4). ~~We also note that in fall 2017 the APP-x product shows thickness data covering an area south of the typical ice edge in regions 8-10 (Fig. 2a). Anomalous data in this region may be due to errors in the sea ice concentration field (not shown). Ice concentration is passed to the APP-x product from the lower-level APP product.~~

4.3 Differences in Winter-time Growth Rates

We now consider the winter-time growth rates across the central Arctic. The evolution of monthly mean ice thickness during winter is shown in Figure 5 for the entire study period spanning fall 2010 to spring 2018, and growth rates are provided in Table 5. The results are dependent on the product availability (Table 1), and in the case of the CPOM product, only seasonal means are ~~available~~ assessed. Monthly mean ice thickness (Fig. 5) in the central Arctic can differ by up to 1.2 m across products. As we might expect, the CPOM, AWI and ~~CPOM~~ JPL products are the most ~~similar~~ similar in terms of both the monthly mean trends in ice thickness and growth rates (Fig. 5, Table 5). While the maximum difference between the CPOM and AWI products is 0.1 m, the JPL product differs most noticeably from the AWI product by ~ 0.14 m in October - December 2013 (Fig. 5). Ice in the GSFC product is consistently thicker than the other CS2-only products, with ice thickness estimates beginning each season by up to 0.4 m thicker in fall, before converging towards the other CS2-only products by the end of the winter (Fig. 5). ~~of the CS2-only products, having maximum difference of 0.1 m. Ice thickness in the JPL product differs very little from these two, most noticeably in January and February 2013 (0.2 m higher than AWI). Ice thickness in the GSFC product is almost constantly higher than the other CS2-only products, with fall sea ice thickness estimates beginning each season up to 0.4 m thicker, before converging towards the other CS2-only products by the end of the winter growth season.~~ In ~~December 2012~~ November 2010, ~~February 2013 and~~ January 2016, November 2016, and April 2018 the GSFC product indicates a small (maximum 0.1 m) decrease in thickness, whereas CPOM and AWI have a constant upward trend ~~within~~ during the ice growth period (Fig. 5). The daily ice growth rates are consistently lowest for the GSFC product (Table 5). Interestingly, for some early seasons, e.g. fall 2011, CS2SMOS estimates slightly higher mean sea ice thickness than the AWI product (Fig. 5). CS2SMOS has daily growth rates higher than those of GSFC, but lower than other CS2-only products, differing from the AWI product by a maximum of 0.001 m/day (Table 5). ~~It is interesting to note that the CS2SMOS product also suggests a temporary decrease in ice thickness in December 2012 and February 2013. CS2SMOS is~~

~~consistently lower than CS2-only products for all months except October 2015 when it converges with the JPL product. Like GSFC, CS2SMOS does not exhibit a constant upward trend during the ice growth season, instead showing significant month to month variability. Indeed monthly mean sea ice thickness typically decreases from the previous month once or twice per growth season.~~ APP-x has the highest growth rate (i.e. the smallest minimum and largest maximum thickness) exceeding the CS2-only products, and the ice cover can gain 2 m within ~~the~~ an ice growth season and up to 0.0108 m/day. The strongest increase in the APP-x mean thickness takes place at the end of winter between February and April, when the ice grows by up to 1 m (Fig. 5). This differs significantly from the winter-time evolution of the ice as shown in the CS2 ~~only~~ products, which suggest ~~stability~~ very little growth in ice thickness at the end of winter (March–April)~~, with very little growth~~, likely due to the insulating properties of the overlying snow cover. There is agreement across all products containing CryoSat-2 data that the largest daily growth rates occurred in winter 2011-2012 (Table 5). Although the GSFC product has the lowest daily growth rate, it has the highest interannual variability in growth rate (0.0026 m/day), whereas the JPL product has the smallest (0.0005 m/day).

The year-to-year seasonal trends in central Arctic ice thickness (Fig. 6) are very coherent among the CS2-only products, with an increase in mean ice thickness in the fall between 2011 and 2014, followed by a slight decrease and leveling off (Fig. 6a).~~from 2011 to 2014, followed by a slight decrease and then a levelling off in mean thickness. In fall 2016 and 2017 (Fig. 6a) there is a minor divergence between the CPOM and AWI products, but both suggest a slight upward trend in fall ice thickness.~~ The GSFC product ~~in fall (Fig. 6a)~~ shows a similar year-to-year trend in fall to the other CS2-only products, but is 0.3-0.5~~0.2–0.4~~ m thicker on average (Fig. 6a). CS2SMOS follows the CS2-only products with maximum difference of -0.06 m compared with the AWI product and almost exactly the same mean ice thickness as the JPL product in fall 2011 and 2012 (Fig. 6a)~~appears to be about 0.3 m thinner than the CS2-only products, but follows a similar trend (Fig. 6a)~~. The fall averages in the APP-x product are lower than for the other products by 0.2-0.6 m.and the~~(Fig. 6a), except in 2012, when the CS2SMOS product also suggests an average of ~1 m. The mean thickness~~ trend in seasonal mean thickness of the APP-x product does not follow ~~the one~~those of the CS2 products, in particular for the 2012 and 2016 seasons, where the APP-x product hast an opposite trend (Fig. 6a).~~that have opposite directions for the trend line from the CS2 products.~~

In spring the central Arctic mean ice thickness differs little across the CS2-only products (Fig. 6b), with the strongest similarities again between the CPOM~~and~~, AWI and JPL products. The CS2 ~~only~~ products show a drop in thickness in 2012 ~~and~~ continuing to 2013, followed by a slight recovery in 2014, preceding another drop in mean thickness in spring 2015, which has persisted since then. The exception to the similar direction of CS2 products is in 2017 when the GSFC product suggests an increase in thickness, and is ~~over~~ 0.2 m higher than the other CS2-only products. By spring 2018 ice thickness in the GSFC product has decreased and is once again in line with the AWI ~~and CPOM~~ estimates, though slightly lower than the CPOM estimate (Fig. 6b). CS2SMOS product is very similar to the other CS2 products, the trend line being consistently approximately 0.2 m lower than the AWI one~~.~~. ~~During this period the CS2SMOS mean is lower than the other products,~~

differing from the CS2-only products by up to 0.7 m. The mean thickness trend, however, follows that of the CS2-only products, except that the CS2SMOS product suggests that the recovery in thickness occurred in 2013 rather than in 2014 (Fig. 6b). APP-x estimates are higher than for the other products in spring with an almost constant mean thickness of ~2.5–2.6 m, and very little year-to-year variability (Fig. 6b). In 2014–2016 this is in line with the CS2-only mean (differing by only 0.2 m).

We have calculated the annual deviations from the mean thickness across the central Arctic for the baseline period 2011-2015, and present the results in Table 3. the overall central Arctic mean thickness for the period 2011–2017 (referred to as the climatological mean) and present the results in Table 3. The departure from the spring climatologicalbaseline means thickness in spring for theall CS2-only data products was 0.230.17 m in spring 2013 (0.12 m in 2012), and 0.200.17 m thicker in 2014, which is much in line with the thickness increase shown in Figure 6b. The CS2-only products are similar in the direction of annual departures from the climatological baseline mean except for GSFC in spring 2017, which shows a positive departure (0.05 m), whereas the CPOM and AWI product thicknesses are lower than their climatological baseline means (0.1 m and 0.08 m, respectively). A very similar pattern of departures can be seen for fall CS2-only thickness data, where there is a mean departure of 0.21 m in fall 2011, 0.13 m in 2012 and thickening of 0.18 m in fall 2013. The CS2SMOS deviations differ very little from the CS2-only product annual deviations, and are actually more in line with them in fall 2015 and 2016 compared to the GSFC product. CS2SMOS and APP-x shows no noticeable deviation from their its fall climatological baseline means, as previously observed in Figure 6b, although there is clearly more deviation in the fall compared to the spring. The changes in APP-x annual deviations follow those of the CS2 products in falls 2012-2015, as well as 2017, except for . For fall, APP-x shows again little variation, except for 2016, when it deviates by -0.21 m (see also Figure 6a). The record low ice thickness records in fall 2011 and 2012 (Fig. 2a) are also evident in the results shown in Table 3.

4.4 Comparison against Independent Observations

Next we consider the satellite-derived sea ice thickness products in the context of independent measurements from ULS and IceBridge to evaluate the utility of the satellite products for providing information on the full thickness distribution. As mentioned previously, the draft to thickness ratio is approximately 0.9 (Rothrock et al., 2008). Therefore we do not expect the modal thickness and draft to be equivalent, but we do expect the distributions to have the same characteristic shapes.

Histograms of the draft/thickness distributions are shown in Figure 7, and suggest that none of the CS2 only satellite products capture neither the thickest nor thinnest ice, although the CPOM, AWI and CS2SMOS products do have some observations of ice thickness below 0.5 m, with CS2SMOS performing best in this regard. This general result may be due to the fact that the monthly/seasonal satellite-derived products have been provided at 25 km resolution (Table 1), thus

potentially averaging out the thinnest and thickest satellite observations per grid cell. Modal ice draft is 1.3 m while modal ice thickness ranges 1.7 m to 2.3 m (Fig. 7). Comparing the characteristics of the draft/thickness distributions, including modal values and distribution width, the CS2SMOS, CPOM and AWI data sets most closely align with the ULS data. The JPL product distribution reveals slightly thicker ice than the CPOM and AWI products. The GSFC product has a bimodal thickness distribution for the study period, with modes at 1.7 and 2.3 m, and the APP-x product also has a modal ice thickness of 2.3 m~~.~~. ~~This may in part be due to the fact that the satellite derived products have been averaged to 25 km, thus masking out the thinnest and thickest measurements per grid cell. Among the CS2 only products, the thickness distribution of the AWI product most closely aligns with the ULS draft distribution, while the GSFC thickness distribution exhibits the least agreement with the draft distribution. CS2SMOS thickness directly overlaps with ULS draft, implying that the CS2SMOS product underestimates ice thickness. APP-x has an exceptionally narrow thickness distribution that captures far less variation than CS2 thickness products and is inconsistent with the characteristic shape of the ULS ice draft distribution.~~

The correlations between the monthly-averaged, satellite-derived ice thickness and ULS ice draft are shown in Figure 8 for the months of November through April for each of the three BGEP moorings. The results are consistent with the approximation of a ~ 0.9 ice draft to thickness ratio (Rothrock et al., 2008). The correlations were calculated ~~Figure 8 shows the correlation~~ between satellite and ULS monthly averages, which were ~~further~~ combined across years 2010-2017 to aid visualization (e.g. November ~~2011~~2010 ~~2016~~2017, December ~~2011~~2010 ~~2016~~2017, etc.) with the following exceptions: JPL product data were assessed over the period 2011-2015, spring 2017 data were not available for the APP-x comparison, and seasonal averages were evaluated for the CPOM product. Correlation results for the CPOM, AWI and JPL products are strongest, followed by APP-x, GSFC and CS2SMOS (Fig. 8), but all satellite thickness products display very strong correlations with the ULS draft data, in line with previous results from Laxon et al., (2013), Kwok and Cunningham (2015), and Tilling et al., (2018). The exception to this is for ice drafts > 1.3 m, when there is an observable divergence in the results for the CS2SMOS and APP-x products. The APP-x thickness estimates are ~ 0.4 m higher than the CS2-only products, while the CS2SMOS data are 0.1-0.3 m thinner, suggesting a thickness overestimation in APP-x and an underestimation in CS2SOS, with respect to the thickest ice. ~~in Fig. 8 to aid visualization. Since the comparisons in Figure 8 are again illustrating thickness against draft, data should optimally fall above the y=x line, when the ~0.9 draft to thickness ratio is considered. With this assumption, and considering the mean values as outlined in Table 4, AWI performs best, confirmed by high correlation values and with mean thickness values 0.2 m greater than draft. Although the best correlation values are recorded by the JPL product, the thickness values appear to be biased slightly high, when keeping the ~0.9 ratio as reference. Logically, the products with varying dependency to draft have lower correlation values, CS2SMOS being the lowest, although with relatively good correlation values. CS2SMOS thickness is roughly 10% higher than draft for ice around 1m, but is lower than draft when draft is greater than 1.3 m, which suggests underestimation of thickness in CS2SMOS estimates across this range. GSFC has comparatively low correlation values, and is roughly 0.5 m thicker than draft. APP-x thickness is increasingly greater than the ULS drafts with 0.5 m difference to draft of 1 m and 0.9 m to draft of 1.5 m, suggesting that it~~

is overestimating ice thickness at this draft range. APP-x hardly shows any deviation from its climatological mean at each mooring location (Table 4), but for other products years of ice gain and loss are visible. The yearly means for ULS ice draft and satellite-derived ice thickness are provided in Table 4. The results suggest the mean ice thickness was largest in 2014 and 2015, and lowest in 2011 and 2012 (Table 4). In years with lower ice thickness, there is some disagreement between the ULS observations and the satellite products. 2013 appears to have the thinnest ULS ice draft results for the observation period, whereas the CPOM product suggests that this was one of the thickest years, while other CS2 products appear to have low ice thickness in 2013 only for buoy D.

Next, the satellite-derived thickness estimates for spring are compared with seven years of independent OIB thickness data (Fig. 9). The JPL product has the highest correlation (r=0.7376) with the OIB data and we. We note that this is slightly higher than the correlation value of 0.56 53 noted in Kwok and Cunningham (2015), who only considered data for March and April 2011–2012. Likewise, our correlation value of 0.67 70 between CPOM and OIB thickness is in line withslightly better than the results from shown in Laxon et al. (2013) and Tilling et al. (2018), who found a-values of 0.61 and 0.67 when assessing CS2 against the 2011 and 2012 and 2011-2014 OIB campaigns data, respectively. Similar to the results observed with the ULS comparisons, the satellite products seem to be missing the thickest ice seen by the OIB measurements, but overall the agreement between the CS2 products and OIB is good. In terms of absolute differences, JPL and GSFC thickness estimates are both slightly higher on average than those of OIB by 0.070.10 m, whereas CPOM and AWI estimates are lower by 0.11 m, GSFC by 0.05 m and CS2SMOS by 0.22 m (Fig. 9)0.14 m and 0.13 m, respectively. Although CS2SMOS differs twice as much from the OIB estimates compared to AWI and CPOM, for ice thickness between 0 m and 1.5 m the CS2SMOS estimates appear to agree best with OIB. Thus the greater difference would be explained by the thicker ice, where CS2SMOS estimates are lower than OIB. These findings for CS2SMOS are in line with previous validation studies (Ricker et al., 2017b) who evaluated CS2SMOS using observations from an airborne electromagnetic (AEM) induction thickness sounding device. APP-x has the smallest correlation of 0.54 and a peculiar vertical concentration of dataline in the scatterplot (Fig. 9), where a majority of APP-x sea ice thickness estimates fall into a thickness category between 2.25 and 2.5 m2.15 m and 2.25 m. Unlike other products, CS2SMOS produces thickness estimates that are significantly thinner than those of OIB, with an average difference of 0.66 m. Our findings, that for thinner ice, CS2SMOS differs less from the validation product, whereas for thicker ice it tends to underestimate thickness, are in line with previous validation studies (Ricker et al., 2017b) who evaluated CS2SMOS using observations from an airborne electromagnetic (AEM) induction thickness sounding device.

4.5 Ice Freeboard

Recall (from Section 2) that only two products provide freeboard estimates: AWI and GSFC. Figure 10 compares sea ice freeboard across these two products for spring and fall. Our assessment reveals a prevalence of negative freeboard estimates in the GSFC product that do not appear in the AWI product (Fig. 10). For the period October 2010 to April 2018, an average of 29.5 % of the freeboard measurements provided in the GSFC data product are negative, in contrast to 0.9 % of the

freeboard measurements in the AWI product. Negative freeboard estimates in the GSFC product often correspond with significantly higher freeboards in the AWI product for the same grid cell locations. An example of the GSFC freeboard product for April 2014 is shown in Figure 10c and highlights the spatial prevalence of anomalous, negative freeboard estimates, especially in the Kara and Barents Seas, where negative values persist throughout the winter. However, we note that negative freeboard estimates also occur in the Beaufort, Chukchi and Greenland Seas. This suggests that some negative estimates may be related to regional masking in the processing algorithm, or they could be due to the use of the DTU10 MSS in the ice thickness derivation (e.g. see Fig. 3 in Skourup et al., 2017).

Despite the high proportion of negative freeboard values in the GSFC product, it does not contain any negative thickness values. While GSFC freeboard observations are on average 0.08 m and 0.14 m thinner than corresponding observations from AWI in fall and spring, respectively (Fig. 10), the GSFC product has higher sea ice thickness for both seasons, and almost all regions, compared to the AWI product, as shown in Figures 2 and 3, and Table 2. Closer examination of individual data points in the GSFC product indicates that ice thickness in grid cells with containing negative freeboard values is not significantly lower than adjacent data points, suggesting that a filter may be applied to remove negative freeboard values before calculating ice thickness, and/or that ice thickness values are derived from interpolations across many grid cells.

We furthermore note that the AWI and CPOM data products are the only two data products that include negative sea ice thickness estimates. Approximately 0.8 % and 0.2 % of the thickness estimates are negative in the CPOM and AWI data products, respectively. The locations of negative ice thickness estimates for the month of April 2014 are shown in Figure S2, and are representative of the general pattern observed in other years of the study period. Negative data points are found along the ice edge (as defined by the ice concentration threshold of 70 % in the OSI-SAF product, plotted from the variable included in the AWI product, Table 1), suggesting that these thickness values are anomalous and are a result of edge effects in the sea surface height interpolation scheme.

5 Discussion

We expect thickness estimates across the CS2-only products to be similar, since their only primary differences areshould be due to the algorithmic approach and some of the auxiliary data inputs. These expectations are borne out as can be seen in the basin-scale maps shown in Figs. 2 and 3. The CPOM and AWI products differ very little, while the JPL thickness estimates are also generally in close agreement. The GSFC product is among the products with thickest sea ice overall, particularly at the beginning of the growth season (Figs. 4, 5, 8), despite containing a very high percentage of negative freeboard values (Fig. 10). On the other hand, due to the inclusion of SMOS data, the CS2SMOS product is weighted for thinner ice, such that we expect overall thickness in this data product to be lower than in the CS2-only datasets(Fig. 2, Fig. 3), with a more realistic representation of areas with thin ice in the peripheral seas (regions 3-10). Since the APP-x product relies on a

thermodynamical model to derive thickness, Wang et al. (2010, 2016) state that the product is expected to perform best over level ice. We find that in fall, APP-x indeed has similar ice thickness to the CS2SMOS product, except over the thickest MYI in the central Arctic (Fig. 2a, Fig. 3a). In spring, however, APP-x appears to overestimate ice thickness across the entire Arctic Ocean (Fig. 2b, Fig. 3b). Despite similarities, there are also major regional differences, as seen in Figs. 2 and 3, and Table 2. Even among CS2-only products, ice thickness in the peripheral seas (regions 3-10) on average varies by 0.33 m in the fall and by 0.6 m in spring.

We have shown that the CS2-only satellite data products include reliable estimates for sea ice between ~0.5–4 m thick, depending on the product. In general, all satellite products capture a realistic winter-time ice thickness distribution, when compared to independent ice draft measurements, as demonstrated in Figure 7, with the exception of APP-x, which overestimates the sea ice thickness and underestimates the thickness variation~~does a very poor job at resolving thickness variations~~ in the Beaufort Sea. Ice in the Beaufort Sea is characterised by mixed amounts of deformed sea ice, which may partially explain the poor results from the APP-x product in this area. The CS2SMOS product has the best representation of the thin ice thickness, which would be expected as SMOS-based estimates are used in the majority of the thin ice regions (Ricker et al., 2017). ~~Although the width of the CS2SMOS thickness distribution suggests that the product captures a range of ice thickness in the region, it underestimates the absolute ice thickness (Fig. 7). This may be due to too strong a weighting on the SMOS observations relative to the CS2 data, and the impact is most clearly seen in the multi-year ice zone (Fig. 2), where the CS2SMOS data product is consistently lower than the CS2-only data sets.~~

However our results also show that no product adequately captures the thinnest sea ice in the thickness distribution at the end of winter (Fig. 7). The CS2-only products ~~also~~ do not resolve the thickness of sea ice less than 0.5 m thick, equivalent to freeboards of less than approximately 0.05 m (Figs. 4, 7, 9, 10), and the CS2SMOS product is the most reliable product in this regard (Fig. 7, Fig. 9). Ricker et al. (2014) found that uncertainties in sea ice thickness estimates are large for CS2 in areas where ice is less than 1 m thick. In contrast, the sensitivity is lost for SMOS when the ice is thicker than 1 m. While we have demonstrated that CS2-only products provide good results in the central Arctic ice pack, they lack robust estimates in some regions, particularly around the ice margins (regions 10-12), and in areas of new ice formation, where thin ice is expected. CS2SMOS, and APP-x to some extent, perform better in the peripheral seas. ~~The opposite argument applies to the CS2SMOS product.~~

The biggest difference in the temporal trend among the thickness products occurs in spring 2012 and 2013, when ice thickness was at its lowest during the observation period (Table 3, Fig. 5). The winter growth rate among the CS2-only products is very consistent, with the exception of the GSFC product (Fig. 5, Table 5). The GSFC product also shows a decrease in mean thickness between March and April in both 2016 and 2018, while all other products show an increase in thickness during the same periods. The CS2SMOS growth rate for March to April is lower than for the other products,

resulting in a mean ice thickness that is ~0.2 m lower than the CS2-only products. The decrease of the CS2SMOS growth rates could be due to the areas of less than 100 % ice concentration contributing to the growth in others, falling within the SMOS weighted area of the product. As stated by Ricker et al. (2017) and Tian-Kunze et al. (2014), SMOS assumes 100 % ice concentration in the thickness retrieval algorithm, which could cause underestimation of ice thickness in areas with lower concentration.. While the CS2 only products show a drop in average Arctic ice thickness following the record September 2012 sea ice minimum, the CS2SMOS product continues to suggest overall thickening between 2012 and 2014 (Fig. 2, Fig. 5). In addition both GSFC and CS2SMOS products suggest a decrease in mean ice thickness for the central Arctic in both December 2012 and February 2013 (Fig. 5, Fig. 6). During spring 2013 significant episodes of ice divergence were reported, with observations of pronounced lead signals in the multi-year ice zone (Willmes and Heinemann, 2015). This may be one reason why the CS2SMOS product departs from the trends shown in the CS2 only products, but further study is needed to fully understand the impact of ice dynamics on the CS2SMOS product and differences between it and the CS2 only products. The CS2SMOS product also shows uneven growth and decline in the mean ice thickness each winter that could be due to changes in the extent of thin ice, if the assumption holds true that SMOS is too heavily weighting in the blended product. Another period with deviation from the overall trends was the ice season between fall 2016 and spring 2017. In October-November 2016 a decrease in the mean ice thickness can be seen in the APP-x product, and to very small extent also in the CS2SMOS and CPOM products, whereas in March-April 2017 the GSFC products shows an increase in mean ice thickness. The APP-x product did not resolve any notable year-to-year variability in mean thickness. During the ice growth season the APP-x product shows the largest magnitude of ice growth, over 1.5 m between fall and spring (Fig. 6), with very little inter-annual variability (Fig. 2, Fig. 5). The largest growth in the APP-x product occurs between February and AprilJanuary and March, at the end of winter (Fig. 5), when in situ measurements typically show inhibited ice growth due to the insulating effects of the overlying snow cover. This suggests that thermodynamic assumptions in the OTIM algorithm for the end of winter may need further refinement.

Finally, we also note that all of the satellite-derived sea ice thickness products depend on additional auxiliary data sets (Table 1) in the derivation of ice thickness. As we have outlined in Table 1, there is great variation in the source of the auxiliary products and how they are used, particularly the mean sea surface model, ice type delineation and ice concentration. Detailed comparison of most use different sources for the mean sea surface model, ice type delineation, ice concentration, and the treatment of ice density and snow depth on ice. Comparing the auxiliary products is outside the scope of this study, but these could give rise to differences across products, in addition to the algorithmic differences. There is also a large range in the ice concentration threshold used to indicate the ice edge across products, varying from 15 % (in the APP-x and CS2SMOS products) to 75 % (in the CPOM product). With regards to the APP-x product we believe that an erroneous ice concentration threshold could be one possible explanation for the peculiar extent of the ice thickness estimates in fall 2017 (Fig. 2a). Additional differences across products may arise due to the treatment of ice density and snow depth on ice. Even though all of the satellite-derived thickness products assessed here make use of the MWC, they vary in their implementation method, as

23

described in Section 2.1 and none of the products resolve year-to-year variations in snow depth. Although there are multiple approaches proposed to obtain seasonal snow depth estimates on sea ice, they have yet to be routinely incorporated into a publicly-available, satellite-derived thickness data product. Potential solutions include utilizing model simulations (e.g., Blanchard-Wrigglesworth et al., 2015), atmospheric reanalysis data (e.g., Blanchard-Wrigglesworth et al., 2018),

5 extrapolating in situ observations (e.g. Shalina and Sandven, 2018), or widely expanding the spatial and temporal coverage of current airborne measurement techniques (e.g. Kurtz and Farrell, 2011; King et al., 2015). Satellite passive microwave radiometer observations have also been used to derive snow on first-year sea ice (Brucker and Markus, 2013), as well as snow on thick ice (Maaß et al., 2013; Rostosky et al., 2018). One additional promising remote-sensing method is to combine two satellite altimeter observations retrieved at different wavelengths, enabling snow retrieval due to differences in

10 penetration (Shepherd et al., 2018). For example this could be achieved through a combination of dual-band radar freeboard observations (e.g. Armitage and Ridout, 2015; Guerreiro et al., 2016; Lawrence et al., 2018) or by comparing freeboard measurements from laser and radar altimeters, to obtain an estimate of year-to-year changes in snow depth (e.g., Kwok and Markus, 2017).

6 Conclusions and Future Outlook

15 Satellite techniques have revolutionized our ability to measure the thickness of ice in the Arctic Ocean, providing critical information for scientists conducting studies of environmental change in the region, as well as a new source of data for forecasters, modelers, operators, and decision makers. Here we assessed a suite of existing satellite-derived, publicly-available, Arctic sea ice thickness data products, conducting a comprehensive examination of regional and seasonal differences over an eight-year period. As expected, the CS2-only products were similar, particularly at the end of winter

20 (Fig. 6b). In March/April of each year2011-2016, APP-x reached a mean Arctic-wide thickness of ~2.6-2.7 m, which was thicker than any other satellite product (Fig. 5), and showed little to no inter-annual variability (Fig. 6). On the other hand, likely due to because of its inclusion of thin sea ice thickness, derived from passive microwave radiometer data, the CS2SMOS data product is on average 1-0.2 m thinner than the other satelliteCS2-only estimates at the end of the winter growth season. In fall, there was a larger spread in mean thickness across the products, and the GSFC thickness product

25 diverged from the three other CS2-only products by approximately 0.2 ~0.3-0.4 m (Fig. 6a). Evaluation of the satellite data products through comparisons with OIB and ULS measurements revealed that all products were well correlated with the independent ice draft/thickness estimates, with correlations of 0.54 and higher (Figs. 8 and 9). Five of the six products resolved an accurate winter-time sea ice thickness distribution for the Beaufort Sea when compared with ULS observations of ice draft, though with the AWI and CS2SMOS data sets produced producing the most robust results (Fig. 7). The APP-x

30 data product did not resolve ice thickness variability in this region, and was biased thick compared to both the ULS ice draft observations and the alternative satellite thickness products. However in fall, APP-x sea ice thickness estimates in the peripheral seas (regions 3-10) resembled those of CS2SMOS (Fig. 2a). Our study revealed some other remarkable

24

differences across the products utilizing CS2 data: there were occasional reductions in mean ice thickness during the winter growth season in the GSFC product, and it diverged from the other CS2 products in winter 2016-2017 by approximately 0.3 m (Fig. 5), despite the prevalence of negative freeboard estimates in this product (Fig. 10). Both the GSFC and CS2SMOS data products suggest reductions in Arctic mean ice thickness during the ice growth season. There was an unexpected increase in mean ice thickness in spring 2017 in the GSFC product that differed from the other CS2 products. There was also an upward mean ice thickness trend in spring 2013 in the CS2SMOS data product, when the CS2-only products showed a continued decrease in mean thickness. Such anomalies require further study to evaluate their actual causes, and this can be accomplished for example through more detailed, along-orbit comparisons between CS2-derived data products and with coincident observations collected during OIB and CryoVex aircraft underflights (e.g., Connor et al., 2011).

In terms of end user applications, the suitability of a particular product depends on the region of interest, as well as data latency and availability (Table 1). Moreover the purpose for which the data are used is critical in selecting the most suitable satellite product. For example climate assessments favour accuracy, while those engaged in operational or forecasting activities require low-latency, high-frequency observations. The frequency of the satellite data products evaluated here varies from twice a day (APP-x) to monthly (AWI), and latency varies from three / four days (APP-x and CPOM NRT) to products that are updated seasonally, or on an ad-hoc basis (GSFC and the CPOM seasonally-averaged thickness data product, Table 1). If access to NRT measurements, with year-round availability is required, APP-x would be the first choice, since it is the only product that provides daily coverage across the Arctic Ocean in both summer and winter, and provides a reasonable measure of mean ice thickness, especially in some of the peripheral seas (regions 3-10) in the fall. However, it does not resolve the cross-basin ice thickness gradient nor the location of the thickest ice, and overestimates FYI thickness at the end of winter. Should basin-scale gradients in ice thickness (i.e., the thickness distribution) be important to the end user, then the APP-x product is not recommended, and CPOM NRT product is preferable, although it is only available for the winter-growth season (Tilling et al., 2016 2018). In terms of climatological studies, or model initialization and hindcast studies, the CS2-only data products are appropriate options, but for navigation in the Arctic, none of these products are suitable as a single source of information, and the utility of the observations would only be realised when combined with additional ice charting analyses.

A remaining challenge for the satellite-derived thickness products is the treatment of snow depth on sea ice. All of the satellite-derived thickness products assessed here make use of the modified Warren et al. (1999) snow climatology, as outlined in Laxon et al., 2013, but there is variation in the implementation method, as described in Section 2.1. In addition, We do not have a reliable data source that provides observations of the interannual variability of snow (Blanchard-Wrigglesworth et al., 2018). Most of the satellite-derived thickness products assessed here make use the modified Warren et al. (1999) snow climatology (as outlined in Laxon et al., 2013), which does not none of the implementations resolve year-to-year variations in snow depth. The need for improved knowledge of snow depth holds especially true for the APP-x, wherein

25

In light of the current challenges presented by the treatment of snow in the existing data products, This has led a selection of end users, particularly those conducting data assimilation experiments, are interested into use satellite-derived sea ice freeboard measurements, rather than ice thickness, since freeboard represents the remote sensing observation (rather than derived ice thickness which is a derivation). Currently there are only two products that provide the freeboard parameter, AWI and GSFC. Our analysis suggests that the AWI data set is preferable for ice freeboard due to a more realistic representation of ice freeboard measurements across the Arctic (Fig. 10). We found a high prevalence of erroneous, negative freeboard estimates throughout the Arctic in the GSFC product, that were especially concentrated in the peripheral seas, particularly in regions 7 and 8, as well as in the Beaufort Sea. The source of these anomalies is most likely associated with aspects of the GSFC algorithm and interpolation of the mean sea surface between lead tie-points.

In conclusion we suggest that low-latency, monthly composites, derived from CS2 data, or similar, but updated daily with the latest-available measurements, would benefit many sea ice thickness applications and provide an ideal solution to address many end-user needs. Further, it may be possible to obtain a more robust thickness distribution, through the inclusion of passive microwave observations of thin-ice thickness in the marginal ice zone. Although the CS2SMOS results are promising, most of our evaluations with independent data were focused on the MYI zone. Our results show however that a Applying an appropriate ratio, that adequately combines microwave and altimeter observations, is challenging and requires further evaluation with independent observations collected over FYI and in the peripheral seas. Higher resolution (<= 5 km) along-orbit and gridded data products would advance the utility of the observations at the regional scale. We also recommend that future products include both ice thickness and freeboard parameters, as well as an estimate of thickness

uncertainty and/or data quality flags, so that the satellite observations may be used in data assimilation experiments aimed at improving ice forecasting.

Author contribution

SLF designed the study, HS and JM carried out the data analysis under the guidance of SLF, and JM was responsible for producing the figures. HS led the manuscript preparation, and SLF, ER and JM contributed to the writing and internal review.

Competing interests

The authors declare that they have no conflicts of interest.

References

Allard, R. A., Farrell, S. L., Hebert, D. H., Johnston, W. F., Li, L., Kurtz, N. T., Phelps, M. W., Posey, P. G., Tilling, R., Ridout, A., and Wallcraft, A. L.: Utilizing CryoSat-2 sea ice thickness to initialize a coupled ice-ocean modeling system, Advances in Space Research, doi:10.1016/j.asr.2017.12.030, 2018.

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

Belward, A. and Dowell, M. (eds.): The Global Observing System for Climate (GCOS): Implementation Needs, GCOS 2016 Implementation Plan, Global Ocean Observing System Report (GCOS 200, GOOS 214), World Meteorological Organisation, 341, https://library.wmo.int/opac/doc_num.php?explnum_id=3417, 2016.

Blanchard-Wrigglesworth, E., Farrell, S., Newman, T., and Bitz, C.: Snow cover on Arctic sea ice in observations and an Earth system model, Geophysical Research Letters, 42, 10342–10348. https://doi.org/10.1002/2015GL066049, 2015.

Blanchard-Wrigglesworth, E., Webster, M. A., Farrell, S. L. and Bitz, C. M.: Reconstruction of Snow on Arctic Sea Ice,
5   Journal of Geophysical Research: Oceans, 123(5), 3588-3602, doi:https://doi.org/10.1002/2017JC013364, 2018.

Blockley, E. W. and Peterson, K. A.: Improving Met Office seasonal predictions of Arctic sea ice using assimilation of CryoSat-2 thickness, The Cryosphere, 12, 3419-3438, https://doi.org/10.5194/tc-12-3419-2018, 2018.

10   Brucker, L., and Markus, T.: Arctic-scale assessment of satellite passive microwave-derived snow depth on sea ice using Operation IceBridge airborne data, J. Geophys. Res. - Oceans, 118, 2892–2905, doi:10.1002/jgrc.20228, 2013.

Connor L. N., Laxon S. W., McAdoo D. C., Ridout A., Cullen R., Farrell S., and Francis R.: Arctic sea ice freeboard from CryoSat-2: Validation using data from the first IceBridge underflight, AGU Fall Meeting Abstracts, 2011.
15

Farrell, S. L., Kurtz, N., Connor, L. N., Elder, B. C., Leuschen, C., Markus, T., McAdoo, D. C., Panzer, B., Richter-Menge, J., and Sonntag, J.G.: A first assessment of IceBridge snow and ice thickness data over Arctic sea ice, IEEE Transactions on Geoscience and Remote Sensing, 50(6), 2098-2111, 2012.

20   Giles, K. A., Laxon, S. W., Wingham, D. J., Wallis, D. W., Krabill, W. B., Leuschen, C. J., McAdoo, D., Manizade, S. S., and Raney, R. K.: Combined airborne laser and radar altimeter measurements over the Fram Strait in May 2002, Remote Sens. Environ., 111, 182–194, 2007.

Giles, K. A., Laxon, S. W., and Ridout, A. L.: Circumpolar thinning of Arctic sea ice following the 2007 record ice extent
25   minimum, Geophys. Res. Lett., 35(22), doi:10.1029/2008GL035710, 2008.

Grosfeld, K., Treffeisen, R., Asseng, J., Bartsch, A., Bräuer, B., Fritzsch, B., Gerdes, R., Hendricks, S., Hiller, W., Heygster, G., Krumpen, T., Lemke, P., Melsheimer, C., Nicolaus, M., Ricker, R. and Weigelt, M.: Online sea-ice knowledge and data platform <www.meereisportal.de>, Polarforschung, Bremerhaven, Alfred Wegener Institute for Polar and Marine Research
30   & German Society of Polar Research, 85 (2), 143-155, doi:10.2312/polfor.2016.011, 2016.

Guerreiro, K., Fleury, S., Zakharova, E., Rémy, F., and Kouraev, A.: Potential for estimation of snow depth on Arctic sea ice from CryoSat-2 and SARAL/AltiKa missions, Remote Sensing of Environment, 186, 339-349, doi:10.1016/j.rse.2016.07.013, 2016.

Guerreiro, K., Fleury, S., Zakharova, E., Kouraev, A., Rémy, F., and Maisongrande, P.: Comparison of CryoSat-2 and ENVISAT radar freeboard over Arctic sea ice: toward an improved Envisat freeboard retrieval, The Cryosphere, 11, 2059–2073, https://doi.org/10.5194/tc-11-2059-2017, 2017.

5

Hendricks, S., Ricker, R., and Helm, V.: User Guide - AWI CryoSat-2 Sea Ice Thickness Data Product (v1.2), 10013/epic.48201, 2016.

Hendricks, S.: Sea Ice Climate Change Initiative Phase 2, D3.4 Product User Guide (PUG) for Sea Ice Thickness Dataset,
10   SICCI-SIT-PUG-P2-17-02, v.1.0, 16 pp, URL
http://data.ceda.ac.uk/neodc/esacci/sea_ice/docs/SICCI_P2_SIT_PUG_D3.3_Issue_1.0.pdf, 2017.

Hendricks, S.; Paul, S.; Rinne, E.: ESA Sea Ice Climate Change Initiative (Sea_Ice_cci): Northern hemisphere sea ice thickness from the CryoSat-2 satellite on a monthly grid (L3C), v2.0. Centre for Environmental Data Analysis,
15   doi:10.5285/ff79d140824f42dd92b204b4f1e9e7c2. http://dx.doi.org/10.5285/ff79d140824f42dd92b204b4f1e9e7c2, 2018.

Kaleschke, L., Tian-Kunze, X., Maaß, N., Makynen, M., & Drusch, M.: Sea ice thickness retrieval from SMOS brightness temperatures during the Arctic freeze-up period. Geophysical Research Letters, 39: L05501. doi:10.1029/2012GL050916, 2012.

20

Kaleschke, L., Tian-Kunze, X., Maaß, N., Beitsch, A., Wernecke, A., Miernecki, M., Müller, G., Fock, B. H., Gierisch, A. M. U., Schlünzen, K. H., Pohlmann, T., Dobrynin, M., Hendricks, S., Asseng, J., Gerdes, R., Jochmann, P., Reimer, N., Holfort, J., Melsheimer, C., Heygster, G., Spreen, G., Gerland, S., King, J., Skou, N., Søbjærg, S. S., Haas, C., Richter, F., and Casal, T.: SMOS sea ice product: Operational application and validation in the Barents Sea marginal ice zone, Remote
25   Sensing of Environment, 180, 264 – 273, doi:10.1016/j.rse.2016.03.009, 2016.

Key, J. R., R. Mahoney, Y. Liu, P. Romanov, M. Tschudi, I. Appel, J. Maslanik, D. Baldwin, X. Wang and P. Meade: Snow and ice products from Suomi NPP VIIRS, J. Geophys. Res. Atmos., 118, 12,816–12,830, doi:10.1002/2013JD020459, 2013.

30   Key, J. and Wang, X.: Extended AVHRR Polar Pathfinder (APP-x) Climate Algorithm Theoretical Basis Document, NOAA Climate Data Record Program, CDRP-ATBD-0573, 01B-24b, Rev. 1, 82, 2015.

King, J., Howell, S., Derksen, C., Rutter, N., Toose, P., Beckers, J. F., Haas, C., Kurtz, N. and Richter‑Menge, J.: Evaluation of Operation IceBridge quick-look snow depth estimates on sea ice, Geophysical Research Letters, 42(21), 9302-9310, 2015.

5   Koenig, L., Martin, S., Studinger, M., and Sonntag, J.: Polar airborne observations fill gap in satellite data, Eos, Transactions American Geophysical Union, 91(38), 333-334, 2010.

Krishfield, R. A. and Proshutinsky, A.: BGOS ULS Data Processing Procedure. Woods Hole Oceanographic Institute report, available at: http://www.whoi.edu/fileserver.do?id=85684&pt=2&p=100409, last access 26 April 2018, 2006.

10

Krishfield, R.A., Proshutinsky, A., Tateyama, K., Williams, W.J., Carmack, E.C., McLaughlin, F.A. and Timmermans, M.L.: Deterioration of perennial sea ice in the Beaufort Gyre from 2003 to 2012 and its impact on the oceanic freshwater cycle, Journal of Geophysical Research: Oceans, 119(2), 1271-1305, 2014.

15   Kurtz, N. T., and S. L. Farrell: Large-scale surveys of snow depth on Arctic sea ice from operation IceBridge, Geophys. Res. Lett., 38, L20505, doi:10.1029/2011GL049216, 2011.

Kurtz, N. T., Farrell, S. L. , Studinger, M., Galin, N., Harbeck, J. P., Lindsay, R., Onana, V. D., Panzer, B., and Sonntag, J. G.: Sea ice thickness, freeboard, and snow depth products from Operation IceBridge airborne data. The Cryosphere, 7, 1035-
20   1056,
doi:10.5194/tc-7-1035-2013, 2013.

Kurtz, N. T., Galin, N., and Studinger, M.: An improved CryoSat-2 sea ice freeboard retrieval algorithm through the use of waveform fitting, The Cryosphere, 8, 1217-1237. doi:10.5194/tc-8-1217-2014, 2014.

25

Kurtz, N., Studinger, M., Harbeck, J., Onana, V., and Yi, D.: IceBridge L4 Sea Ice Freeboard, Snow Depth, and Thickness, Version 1 [March–April 2011], NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado USA, [accessed 2018], doi:10.5067/G519SHCKWQV6, 2015.

30   Kurtz, N.: IceBridge Sea Ice Freeboard, Snow Depth, and Thickness Quick Look, Version 1 [March–April, 2012–2017], NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado USA, [accessed 2018], doi:10.5067/GRIXZ91DE0L9, 2016.

Kurtz, N. and Harbeck, J.: CryoSat-2 Level-4 Sea Ice Elevation, Freeboard, and Thickness, Version 1 [October–April, 2010–2018], NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado USA, [accessed 2018], doi:10.5067/96JO0KIFDAS8, 2017.

Kwok, R., Cunningham, G. F., Wensnahan, M., Rigor, I., Zwally, H. J., and Yi, D.: Thinning and volume loss of the Arctic Ocean sea ice cover: 2003–2008, J. Geophys. Res. Oceans, 114, doi:10.1029/2009JC005312, 2009.

Kwok, R. and Cunningham, G.F.: Variability of Arctic sea ice thickness and volume from CryoSat-2, Phil. Trans. R. Soc. A., 373, 20140157, doi:10.1098/rsta.2014.0157pmid:26032317, 2015.

Kwok, R. and Markus, T.: Potential Basin-Scale Estimates of Arctic Snow Depth with Sea Ice Freeboards from CryoSat-2: An Exploratory Analysis, Adv. Space Res., doi:10.1016/j.asr.2017.09.007, 2017.

Lawrence, I. R., Tsamados, M. C., Stroeve, J. C., Armitage, T. W. K., and Ridout, A. L.: Estimating snow depth over Arctic sea ice from calibrated dual-frequency radar freeboards, The Cryosphere, 12, 3551-3564, https://doi.org/10.5194/tc-12-3551-2018, 2018.

Laxon S. W., Giles, K.A., Ridout, A. L., Wingham, D.J., Willatt, R., Cullen, R., Kwok, R., Schweiger, A., Zhang, J., Haas, C., Hendricks, S., Krishfield, R., Kurtz, N., Farrell, S. and Davidson, M.: CryoSat-2 estimates of Arctic sea ice thickness and volume, Geophys. Res. Lett., 40, 732–737, doi:10.1002/grl.50193, 2013.

~~Lindsay, R., Haas, C., Hendricks, S., Hunkeler, P., Kurtz, N., Paden, J., Panzer, B., Sonntag, J., Yungel, J., Zhang, J.: Seasonal forecasts of Arctic sea ice initialized with observations of ice thickness, Geophys. Res. Lett., 39, L21502, doi:10.1029/2012GL053576, 2012.~~

Markus, T., Stroeve, J. C., and Miller, J.: Recent changes in Arctic sea ice melt, onset, freezeup, and melt season length, J. Geophys. Res., 114, C12024, doi:10.1029/2009JC005436, 2009.

Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Farrell, S., Fricker, H., Gardner, A., Harding, D., Jasinski, M., Kwok, R., Magruder, L., Lubin, D., Luthcke, S., Morison, J., Nelson, R., Neuenschwander, A., Palm, S., Popescu, S., Shum, CK, Schutz, B. E., Smith, B., Yang, Y. and Zwally, J.: The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation, Rem. Sens. Environ., 190, 260-273, doi:10.1016/j.rse.2016.12.029, 2017.

31

Maaß, N., Kaleschke, L., Tian-Kunze, X., and Drusch, M.: Snow thickness retrieval over thick Arctic sea ice using SMOS satellite data, The Cryosphere, 7, 1971-1989, https://doi.org/10.5194/tc-7-1971-2013, 2013.

Newman, T., Farrell, S. L., Richter-Menge, J., Connor, L. N., Kurtz, N. T., Elder, B. C. and McAdoo, D.: Assessment of radar-derived snow depth over Arctic sea ice, Journal of Geophysical Research: Oceans, 119(12), 8578-8602, 2014.

Parkinson, C.L., and Cavalieri, D.J.: Arctic sea ice variability and trends, 1979-2006, J. Geophys. Res., 113, C07003. doi:10.1029/2007JC004558, 2008.

Parkinson, C.L., and Comiso, J.C.: On the 2012 record low Arctic sea ice cover: Combined impact of preconditioning and an August storm, Geophys. Res. Lett., doi:10.1002/grl.50349, 2013.

Paul, S., Hendricks, S. and Rinne, E.: Sea Ice Climate Change Initiative Phase 2, D2.1 Sea Ice Thickness Algorithm Theoretical Basis Document (ATBD), SICCI-P2-ATBD(SIT), v.1.0, 50 pp. https://icdc.cen.uni-hamburg.de/fileadmin/user_upload/ESA_Sea-Ice-ECV_Phase2/SICCI_2_ATBD_D2.1__SIT__Issue_1.0.pdf, 2017.

Perovich, D. K., Meier, W., Tschudi, M., Farrell, S., Hendricks, S., Gerland, S., Haas, C., Krumpen, T., Polashenski, C., Ricker, R. and Webster, M.: Sea ice cover [in "State of the Climate in 2018"], Bull. Amer. Meteor. Soc., 99(8), S147–152, doi:10.1175/2018BAMSStateoftheClimate.1., 2018.

Price, D., Beckers, J., Ricker, R., Kurtz, N., Rack, W., Haas, C., Helm, V., Hendricks, S. Leonard, G., and Langhorne, P. J.: Evaluation of CryoSat-2 derived sea-ice freeboard over fast ice in McMurdo Sound, Antarctica, J. Glaciol. 61 (226), doi:10.3189/2015JoG14J157, 2015.

Richter-Menge, J. A., and Farrell, S. L.: Arctic sea ice conditions in spring 2009–2013 prior to melt, Geophys. Res. Lett., 40, 5888–5893, doi:10.1002/2013GL058011, 2013.

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 (4), 1607-1622, doi:10.5194/tc-8-1607-2014, 2014.

Ricker, R., Hendricks, S., Kaleschke, L., and Tian-Kunze, X.: CS2SMOS User Guide v3.0, Alfred Wegener Institute, 10013/epic.51136, 2017a.

Ricker, R.; Hendricks, S.; Kaleschke, L.; Tian-Kunze, X.; King, J. and Haas, C.: A weekly Arctic sea-ice thickness data record from merged CryoSat-2 and SMOS satellite data, The Cryosphere, 11, 1607-1623, doi:10.5194/tc-11-1607-2017, 2017b.

5   Rostosky, P., Spreen, G., Farrell, S. L., Frost, T., Heygster, G., & Melsheimer, C.: Snow depth retrieval on Arctic sea ice from passive microwave radiometers - Improvements and extensions to multiyear ice using lower frequencies, J. Geophys. Res., 123, 7120–7138, https://doi.org/10.1029/2018JC014028, 2018.

Rothrock, D. A., Percival, D. B. and Wensnahan, M.: The decline in arctic sea-ice thickness: Separating the spatial, annual, 10   and interannual variability in a quarter century of submarine data, Journal of Geophysical Research: Oceans, 113(C5), 2008.

Shalina, E. V. and Sandven, S.: Snow depth on Arctic sea ice from historical in situ data, The Cryosphere, 12, 1867-1886, doi:10.5194/tc-12-1867-2018, 2018.

15   Shepherd, A., Fricker, H. A., and Farrell S. L.: Trends and connections across the Antarctic cryosphere, Nature, 558(7709), 223-232, 2018.

Skourup, H., Farrell, S. L., Hendricks, S., Ricker, R., Armitage, T. W. K., Ridout, A., Andersen, O. B., Haas, C., and Baker, S.: An assessment of state-of-the-art mean sea surface and geoid models of the Arctic Ocean: Implications for sea ice 20   freeboard retrieval, Journal of Geophysical Research: Oceans, 122, 8593–8613, doi:10.1002/2017JC013176, 2017.

Song, M.R.: Change of Arctic sea ice volume and its relationship with sea ice extent in CMIP5 simulations, Atmospheric and Oceanic Science Letters, 9, 22–30, doi:10.1080/16742834.2015.1126153, 2016.

25   Stroeve, J. C., Schröder, D., Tsamados, M. and Feltham, D. L.: Warm winter, thin ice? The Cryosphere, doi: https://doi.org/10.5194/tc-2017-287, 2018.

Tian-Kunze, X., Kaleschke, L., Maaß, N., Mäkynen, M., Serra, N., Drusch, M., and Krumpen, T.: SMOS-derived thin sea ice thickness: algorithm baseline, product specifications and initial verification, The Cryosphere, 8, 997–1018, 30   doi:10.5194/tc-8-997-2014, URL http://www.the-cryosphere.net/8/997/2014/, 2014.

Tilling, R.L., Ridout, A., Shepherd, A., and Wingham, D.J.: Increased Arctic sea ice volume after anomalously low melting in 2013, Nature Geoscience, 8(8), 643, doi:10.1038/NGEO2489, 2015.

Tilling, R.L., Ridout, A., and Shepherd, A.: Near-real-time Arctic sea ice thickness and volume from CryoSat-2, The Cryosphere, 10(5), 2003-2012, 2016.

Tilling, R.L., Ridout, A., and Shepherd, A.: Estimating Arctic sea ice thickness and volume using CryoSat-2 radar altimeter data, Advances in Space Research, Volume 62, Issue 6, Pages 1203-1225, doi:10.1016/j.asr.2017.10.051, 2018.

Wang, X., Key, J. R., and Liu, Y.: A thermodynamic model for estimating sea and lake ice thickness with optical satellite data, J. Geophys. Res., 115, C12035, doi:10.1029/2009JC005857, 2010.

Wang, X., Key, J., Kwok, R., and Zhang, J.: Comparison of Arctic Sea Ice Thickness from Satellites, Aircraft, and PIOMAS Data, Remote Sens., 8, 713, 2016.

Warren, S. G., Rigor, I. G., Untersteiner, N., Radionov, V. F., Bryazgin, N. N., Aleksandrov, Y. I., and Colony, R.: Snow depth on Arctic sea ice, J. Clim., 12, 1814–1829, 1999.

Willmes, S. and Heinemann, G.:. Sea Ice Wintertime Lead Frequencies and Regional Characteristics in the Arctic, 2003-2015, Remote Sens., 8(1), 4; doi:10.3390/rs8010004, 2015.

Wingham, D. J., Francis, C. R., Baker, S., Bouzinac, C., Brockley, D., Cullen, R., de Chateau-Thierry, P., Laxon, S. W., Mallow, U., Mavrocordatos, C., Phalippou, L., Ratier, G., Rey, L., Rostan, F., Viau, P., and Wallis, D. W.: CryoSat: A mission to determine the fluctuations in Earth's land and marine ice fields, Adv. Space Res., 37, 841–871, doi:10.1016/j.asr.2005.07.027, 2006.

Xia, W. and Xie, H.: Assessing three waveform retrackers on sea ice freeboard retrieval from Cryosat-2 using Operation IceBridge Airborne altimetry datasets. Remote Sens. Environ., 204, 456–471, 2018.
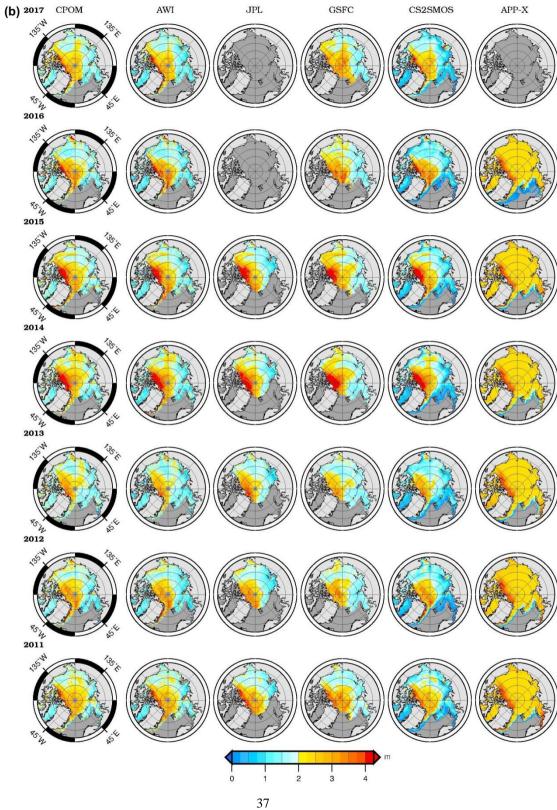
Xie, J., Counillon, F., and Bertino, L.: Impact of assimilating a merged sea ice thickness from CryoSat-2 and SMOS in the Arctic reanalysis, The Cryosphere, 12, 3671-3691, https://doi.org/10.5194/tc-12-3671-2018, 2018.The Cryosphere Discuss., doi:10.5194/tc-2018-101, in review, 2018.
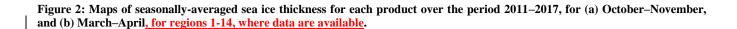
Yang, Q., Losa, S. N., Losch, M., Tian-Kunze, X., Nerger, L., Liu, J., Kaleschke, L., and Zhang, Z.: Assimilating SMOS sea ice thickness into a coupled ice-ocean model using a local SEIK filter, J. Geophys. Res. Oceans, 119, 6680-6692, doi:10.1002/2014JC009963, 2014.
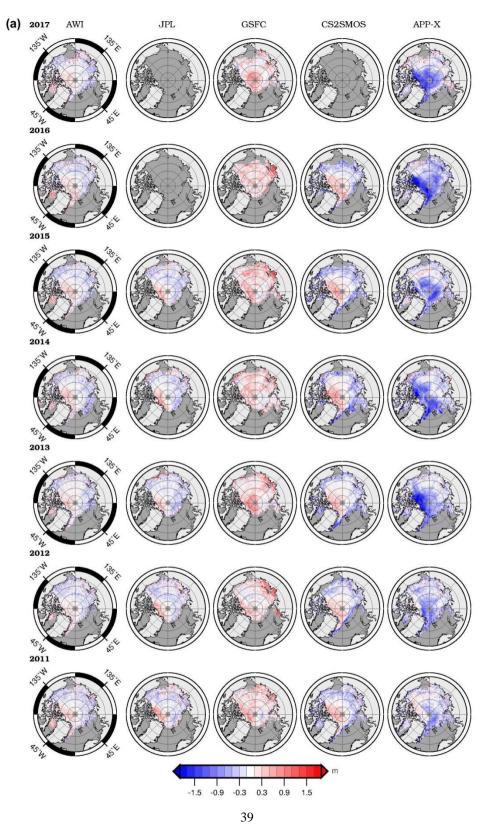
5

Figure 1: Regions utilised in data analysis: (1) Central Arctic Ocean, (2) Lincoln Sea, (3) Beaufort Sea, (4) Chukchi Sea, (5) East Siberian Sea, (6) Laptev Sea, (7) Kara Sea, (8) Barents Sea, (9) Norwegian Sea, (10) Greenland Sea, (11) Baffin Bay, (12) Davis Strait, (13) Hudson Strait, and (14) Canadian Arctic Archipelago. The dotted area represents the central Arctic region (1–6) within which all data products are available. The locations of Beaufort Gyre Exploration Project (BGEP) moorings A, B, and D
10 **are also indicated (white circles with mooring designation).**

(a)

Figure 2: Maps of seasonally-averaged sea ice thickness for each product over the period 2011–2017, for (a) October–November, and (b) March–April, for regions 1-14, where data are available.

(a)

Figure 3: Maps of seasonally-averaged sea ice thickness differences for the period 2011–2017, where ~~the reference data set (CPOM) is subtracted from each data product~~each data product is subtracted from the reference data set (CPOM), for (a) October–November and (b) March–April, for regions 1-14, where data are available. Red (blue) regions indicate areas where the seasonally-averaged thickness is greater (less) than the reference data product.
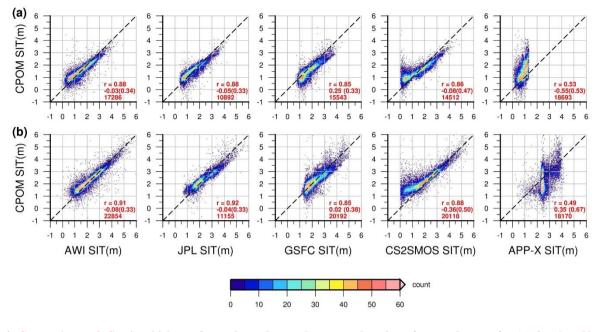
5

Figure 4: Comparisons of sSea ice thickness for each product and compared to the reference data set for (a) October–November and (b) March–April, for the period 2010-2018, regions 1-14 north of 65°N, subject to data availability (Table 1)March–April (top) and October–November (bottom). Colour indicates measurement density, derived from the number of data points within each 0.05 m cell. Statistics for correlation (r), mean difference (std. dev.), and number of data points are provided. The period over which correlations are calculated is based on product availability, as outlined in Table 1.
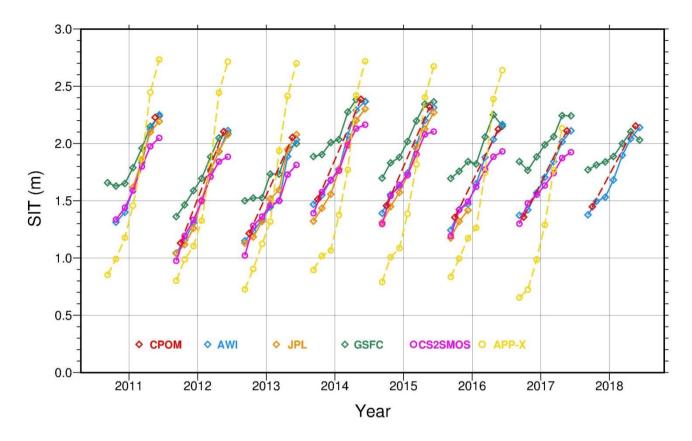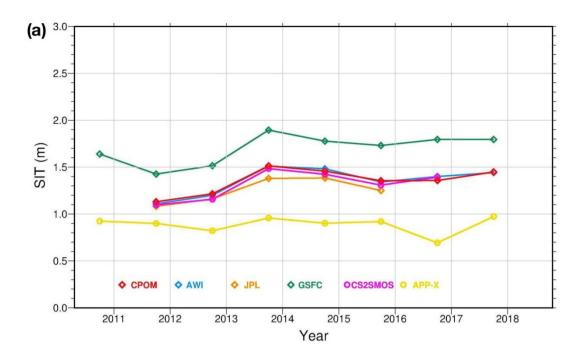
Figure 5: Sea ice thickness growth curves for October–April (monthly averages) ~~during~~ for the central Arctic (regions 1-6) for the period 2010–2018, ~~for each product per~~subject to data availability (~~solid lines~~Table 1), indicating interannual variability in the winter-time thickness evolution. For the CPOM data product, ~~only the~~October–November and March–April seasonal averages are shown, (red diamonds connected by a dashed red line).
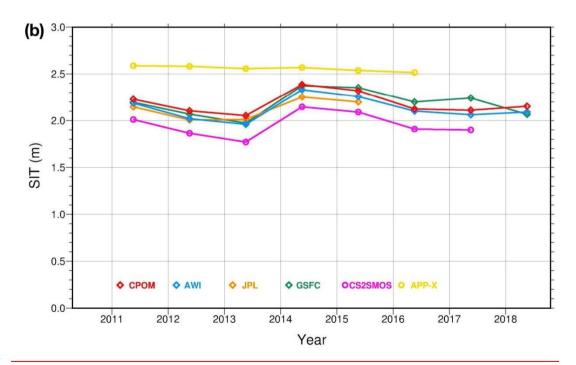
Figure 6: Time series of seasonally-averaged sea ice thickness ~~for each product,~~ for (a) October–November, and (b) March–April, over the central Arctic (regions 1–6) during the period 2010–2018, ~~based on product~~subject to data availability ~~as outlined in~~ (Table 1).
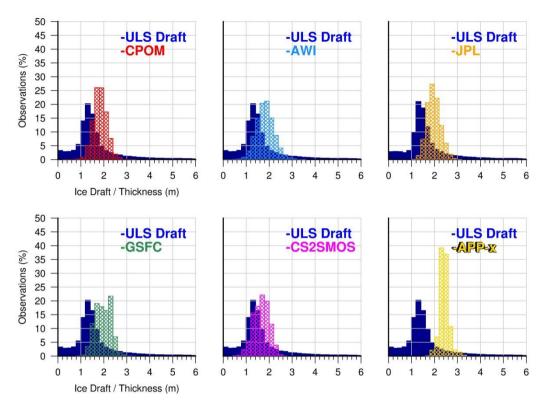
44

Figure 7: Spring (March–April) sea ice thickness distributions (cross-hatched) for each data product within 200 km radius of BGEP mooring locations, averaged for the period 2011 to 2017*2016, overlaid on the corresponding BGEP upward looking sonar (ULS) ice draft distribution (dark blue, solid). *The JPL data product averaging period is 2011 to 2015, and the APP-x data product average does not include thickness data for April 2017. Histogram bin width is 0.2 m.
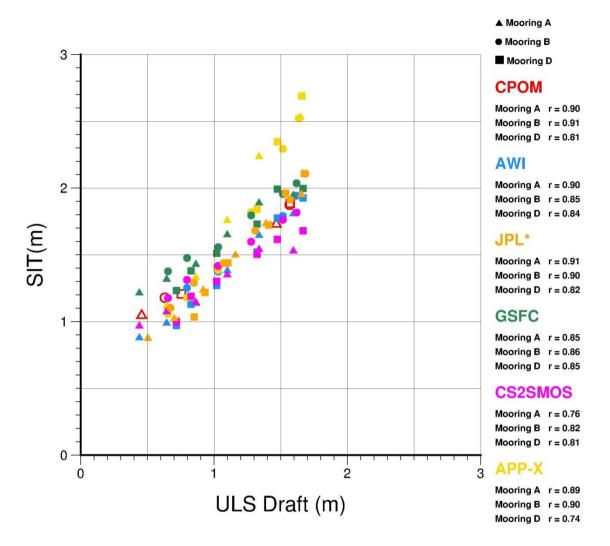
Figure 8: Correlation between monthly-averaged, satellite-derived ice thickness and ULS ice draft, for six months spanning November to April. Correlation coefficients are provided per ULS mooring (mooring locations indicated by symbols), for the period 2010–2017~~2010–2016~~, with the exception of the JPL product, wherein monthly averages and correlation coefficients are calculated for the period 2011–2015, and the APP-x product, where spring 2017 was excluded. In the case of the CPOM product the correlation coefficients are calculated based on seasonal (October–November, March–April) rather than monthly averages. To aid visualisation~~Here~~, monthly averages were further combined across years (e.g. November ~~2011~~2010–~~2016~~2017, December ~~2011~~2010–~~2016~~2017, etc.) to provide six data points per mooring-product comparison.
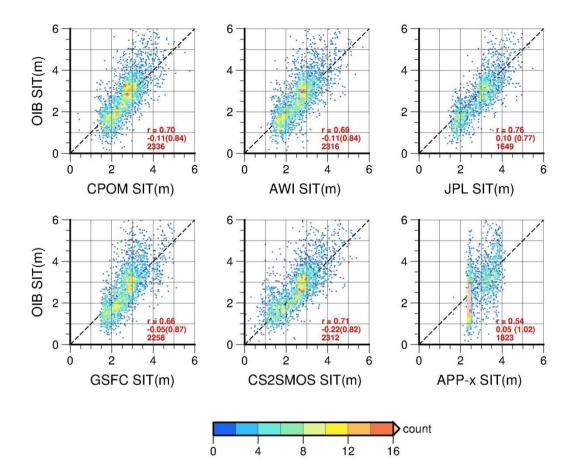
46

Figure 9: Comparison of satellite-derived ice thickness with Operation IceBridge thickness estimates at the end of the winter growth season (March–April). Comparisons were conducted by gridding satellite and aircraft data onto a common 0.4° latitude by 4° longitude~~50 km~~ grid and using grid cells in which both data sets contained thickness estimates. Colour indicates measurement density (number of data points within each 0.01 m cell). Statistics for correlation (r), mean difference (std. dev.), and number of data points are calculated for the period 2011–2017, with the exception of the JPL product, wherein statistics are calculated for the period 2011–2015, and the APP-x product, where spring 2017 was excluded.

Figure 10: Comparison of sea ice freeboard in the GSFC and AWI data products for (a) March–April and (b) October–November, for the period 2011–2018, regions 1-14, north of 65°N. ~~Data were gridded at 25 km and data common to both grids were compared.~~ **Colour indicates measurement density, derived from the number of data points within each 0.01 m cell. Statistics for correlation (r), mean difference (std. dev.), and number of data points are provided. (c) GSFC product freeboard (FB) for April 2014.**

Figure S1: Histograms of seasonally-averaged sea ice thickness, where the reference data set (CPOM) is subtracted from each product~~Sea ice thickness deviation (in meters) for each product subtracted from reference data set (CPOM)~~, for (a) October–November, and (b) March–April, in the central Arctic (regions 1–6), for the period 2011–2017, subject to data availability (Table 1).

5

Figure S2: Locations of negative sea ice thickness values in the CPOM (circles) and AWI (diamonds with black outline) data products, for March–April 2014 (CPOM) and April 2014 (AWI). A contour representing the 70 % ice concentration level is also shown (thick black line).

52

Table 1. Characteristics of satellite-derived ice thickness products ("quality ~~flag~~ information" refers to the availability of a thickness uncertainty estimate or quality flag). Latency is the difference between the date of data acquisition and data delivery, estimated from data portal time stamps at the time of writing. Frequency r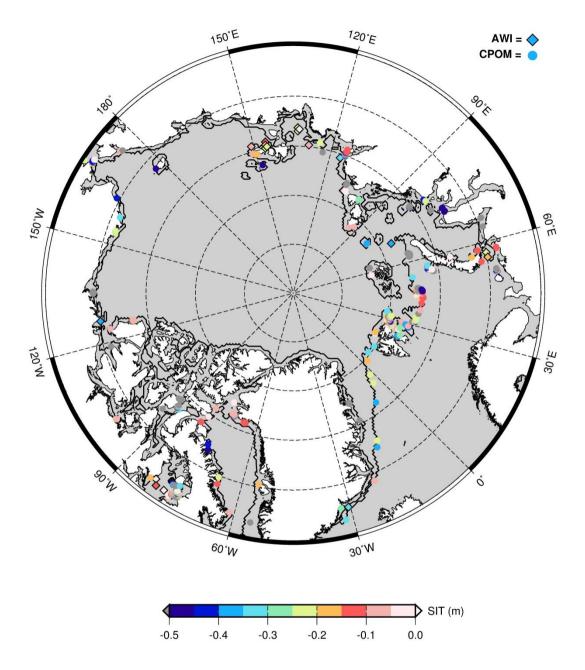efers to the temporal offset between consecutive datasets (e.g. daily, weekly, monthly). *The northern limit of coverage (81.5 °N) impacts the availability of the Envisat product, especially in the Central Arctic (region 1).

Product Name	CPOM	AWI	~~NASA /~~ JPL	~~NASA /~~ GSFC	CS2SMOS	APP-x	LEGOS TFMRA	ESA CCI
Temporal Range	~~March 2011-date (October-April only)~~ November 2010-date	November 2010--date	January 2011-- December 2015	October 2010--date	November 2010--April 2017	January 1982--date	November 2002-March 2012 * November 2010-April 2017	October 2002-March 2012 * November 2010-April 2017
Frequency	Daily (NRT)/monthly / seasonally	Monthly	Not updated post 2015	Daily	Weekly	Twice daily	Monthly	Daily
~~Geograpical~~ Geographical Coverage (by region number)	(1)–(14), ~~full Northern Hemisphere~~ poleward of 40 °N	(1)–(14), poleward of 60 °N	(1)–(6)	(1)–(8), (10, partial), poleward of 55 °N	(1)–(14), poleward of 50 °N	(1)–(14), poleward of 50 °N	65 °N-81.5 °N*; poleward of 65 °N	(1), (3)-(14) south of 81.5 °N*; (1)-(14), poleward of 16.6 °N
Averaging Period	NRT data product: 2, 14, 28 days. ~~Seasonally-averaged product: 2 months~~ Final products: monthly, seasonally-averaged	1 month	1 month	30 day	7 day	12 hours	1 month	1 month (moving average)
Latency	3 days (NRT) / seasonally	Variable	Not updated post 2015	~6 weeks	Variable	~4 days (as of April 2018)	~1 year	Variable
Grid Resolution	25 km, 5 km (full Arctic), 1 km (individual regions)	EASE2, 25 km	25 km	Polar Stereographic SSM/I, 25 km	EASE2, 25 km	EASE, 25 km ~~(EASE grid)~~	12.5 km	EASE2, 25 km
Freeboard	No	Yes	No	Yes	No	No	Yes	Yes
Quality ~~Flag~~ Information	No	Yes	No	No	Yes	No	Yes	Yes
~~Technique~~ Retracking technique	~~70% Threshold First Maxima Retracker Algorithm (TFMRA)~~ Retracking based on ice type (lead, floe); Lead retracker	50% TFMRA	Waveform centroid retracker	Waveform fitting using waveform model	Blended SMOS and CS2 data	One-dimensional Thermodynamic Ice Model (OTIM)	60% TFMRA	50% TFMRA

	Col1	Col2	Col3	Col4	Col5	Col6	Col7	Col8
	follows Giles et al. 2007; Floe retracker based on 70 % of peak amplitude on the waveform leading edge							
Mean Sea Surface	UCL2013	DTU15 MSS	EGM2008 geoid	DTU10 MSS	DTU15 MSS (for CS-2 data)	N/A	DTU15 MSS	DTU15 MSS
Snow Depth	Monthly constants for FYI and MYI based on the ~~Modified Warren Climatology~~ modified climatology of Warren et al., 1999 (MWC)	~~Modified Warren Climatology~~ MWC	~~Modified Warren Climatology~~ MWC	~~Modified Warren Climatology~~ MWC	~~Modified Warren climatology~~ MWC (CS2 data); linear relation with ice thickness (SMOS data)	~~0.2 m~~ Climatology of Warren et al. (1999) combined with other available estimates	MWC	MWC
Ice Density	FYI: 916.7 kg/m^3; MYI: 882.0 kg/m^3	FYI: 916.7 kg/m^3; MYI: 882.0 kg/m^3	FYI: 917 kg/m^3; MYI: 882.0 kg/m^3	915 kg/m^3	FYI: 916.7 kg/m^3; MYI: 882.0 kg/m^3	N/A	FYI: 916.7 kg/m^3; MYI: 882.0 kg/m^3	FYI: 916.7 kg/m^3; MYI: 882.0 kg/m^3
Ice Concentration, Threshold	NSIDC Near-Real-Time DMSP SSMIS Daily Polar Gridded Sea Ice Concentrations, 75%	OSI-SAF, 70%	Not specified	NSIDC Near-Real-Time DMSP SSMIS Daily Polar Gridded Sea Ice Concentrations, 70%	OSI-SAF, 15%	NSIDC Nimbus-7 SMMR and DMSP SSM/I passive microwave data with NASA Team Algorithm applied, 15%	Sea ice age (NSIDC), 50%	OSI-SAF, 70%
Ice Type	OSI-SAF	OSI-SAF	ASCAT	OSI-SAF	OSI-SAF	Converted from reflectances	Sea ice age (NSIDC)	Integrated Climate Data Center (ICDC) ice type fraction
References	Laxon et al., 2013; Tilling et al., 2016; Tilling et al., 2018	Ricker et al., 2014; Hendricks et al., 2016	Kwok and Cunningham, 2015	Kurtz et al., 2014; Kurtz and Harbeck, 2017	Tian-Kunze et al., 2014; Kaleschke et al., 2015, Ricker et al., ~~2016~~2017b	Wang et al., 2010; Key and Wang, 2015; Wang et al. 2016	Guerreiro et al., 2017	Hendricks, S., 2017; Paul et al., 2017; Hendricks et al., 2018
Public Data Source	http://www.cpom.ucl.ac.uk/csopr/seaice.html	Meereisportal, http://data.meereisportal.de/ (Grosfeld et al. 2016)	https://rkwok.jpl.nasa.gov/cryosat/download.html	https://nsidc.org/data/RDEFT4	Meereisportal, http://data.meereisportal.de/ (Grosfeld et al. 2016)	https://www.ncei.noaa.gov/data/avhrr-polar-pathfinder-extended/access/nhem/	http://ctoh.legos.obs-mip.fr	http://cci.esa.int/data

Table 2. Seasonally-averaged sea ice thickness for each Arctic region (in meters) for all six data products~~, where ice cover was greater than 10%. Regional averages are calculated over the period 2011-2017, but are based on product availability, as outlined in Table 1~~. Regional thickness is reported for all regions that contain valid thickness estimates. The data used in this table spans the period 2010-2018, based on product availability, as outlined in Table 1. Average thicknesses are calculated only based on years in which ice is present.

	October–November						March–April					
	CPOM	AWI	JPL	GSFC	CS2SMOS	APP-x	CPOM	AWI	JPL	GSFC	CS2SMOS	APP-x
Arctic Ocean [1]	~~1.64~~1.48	~~1.75~~1.48	~~1.6~~1.39	~~2.03~~1.77	~~1.44~~1.49	~~1.01~~0.96	~~2.46~~2.3	~~2.42~~2.25	~~2.45~~2.29	~~2.5~~2.33	~~1.93~~2.15	~~2.6~~2.58
Lincoln Sea [2]	~~2.04~~1.99	~~2.13~~2.11	~~2.58~~2.52	~~2.45~~2.42	~~1.99~~2.43	~~1.18~~1.19	~~3.8~~3.61	~~3.65~~3.51	~~3.68~~3.7	~~3.39~~3.34	~~2.94~~3.47	~~3.64~~3.66
Beaufort Sea [3]	~~1.01~~1.03	~~1.04~~0.79	~~1.06~~0.89	~~1.5~~1.2	~~0.66~~0.57	~~0.89~~0.77	~~2.06~~2.03	~~1.99~~1.97	~~2.1~~2.09	~~2.04~~2.00	~~1.31~~1.59	~~2.6~~2.59
Chukchi Sea [4]	~~0.84~~	~~0.61~~	~~0.96~~	~~1.13~~	~~0.21~~	~~0.47~~	~~1.95~~1.91	~~1.83~~1.79	~~1.82~~1.81	~~1.94~~1.92	~~1.18~~1.46	2.51
East Siberian Sea [5]	~~0.91~~0.92	~~0.76~~0.77	~~0.82~~0.81	~~1.27~~1.34	~~0.4~~0.52	~~0.67~~0.68	~~1.8~~1.83	~~1.7~~1.72	~~1.59~~1.56	~~1.74~~1.72	~~1.14~~1.39	2.44
Laptev Sea [6]	~~0.81~~0.82	~~0.52~~0.55	~~0.69~~0.67	~~0.99~~1.26	~~0.35~~0.43	0.61	~~1.5~~1.53	~~1.25~~1.27	~~1.15~~1.12	~~1.55~~1.56	~~0.78~~0.97	2.4
Kara Sea [7]	~~0.85~~	~~0.48~~	~~0.87~~	~~1.01~~	~~0.37~~	~~0.49~~	~~1.6~~1.59	~~1.39~~1.38	~~1.63~~	~~1.56~~1.55	~~0.69~~0.79	~~2.49~~2.52
Barents Sea [8]	~~1.31~~	~~0.78~~	~~1.26~~	~~1.51~~	~~0.37~~	~~0.36~~	~~1.45~~1.41	~~1.45~~1.44	~~1.99~~	~~1.67~~1.73	~~0.38~~0.41	~~2.3~~1.99
Greenland Sea [10]	~~1.68~~1.69	~~1.89~~1.78	~~2.14~~	~~2.4~~	~~1.14~~1.16	~~0.75~~0.56	~~2.62~~2.47	~~3.04~~2.79	~~3.02~~	~~2.97~~	~~1.7~~1.65	~~2.39~~2.18
Baffin Bay [11]	~~0.91~~	~~1.28~~	-	~~2.47~~	~~0.26~~	~~0.63~~	~~1.53~~1.5	~~1.26~~1.27	-	~~3.39~~	~~0.54~~0.64	~~2.48~~2.47
Davis Strait [12]	~~1.01~~	-	-	-	~~0.29~~	~~0.24~~	~~1.59~~1.55	~~1.21~~1.15	-	-	~~0.48~~0.47	~~2.7~~2.54
Hudson Strait [13]	~~1.12~~	-	-	-	~~0.12~~	~~0.22~~	~~1.49~~1.53	~~0.81~~0.82	-	-	~~0.44~~0.53	~~2.91~~2.93
Canadian Archipelago [14]	~~1.29~~1.25	~~1.42~~1.26	~~1.19~~	~~2.03~~	~~0.85~~0.81	~~0.92~~0.84	~~1.97~~1.95	~~1.44~~1.58	~~2.09~~	~~2.27~~	~~1.32~~1.4	~~2.66~~2.64

Table 3. ~~Deviation~~ Anomalies of seasonally-averaged annual sea ice thickness (in meters) ~~from~~ relative to the seasonally-averaged ~~climatological~~ baseline mean within the central Arctic (regions 1–6). The ~~climatological~~ baseline mean is calculated ~~per product availability during~~ for the period 2011–~~2017~~ 2015 ~~(i.e. the JPL product climatological mean is calculated for the period 2011-2015, while the fall CS2SMOS and spring APP-x product climatological means are calculated for the period 2011-2016)~~. Anomalies relative to the 2011-2015 mean are included for later seasons according to the product availability.

October–November~~March–April~~	2011	2012	2013	2014	2015	2016	2017	2018
CPOM	~~-0.27~~-0.2	−0.12	~~0.2~~0.18	~~0.07~~0.12	~~0.03~~0.02	0.02	~~0.04~~0.11	
AWI	~~-0.34~~-0.22	~~-0.18~~-0.13	~~0.22~~0.18	~~0.07~~0.15	~~0.06~~0.01	~~0.02~~0.06	~~0.14~~0.1	
JPL	~~-0.24~~-0.17	−0.09	~~0.18~~0.13	~~0.11~~0.13	~~0.05~~0.0	-	-	
GSFC	~~-0.40~~-0.25	~~-0.25~~-0.16	~~0.36~~0.22	~~0.02~~0.1	~~0.03~~0.06	~~-0.01~~0.13	~~0.25~~0.13	
CS2SMOS	~~-0.30~~-0.2	~~-0.16~~-0.14	~~0.21~~0.19	~~0.18~~0.12	~~0.07~~0.01	~~0~~0.1	-	
APP-x	~~0.04~~0.0	~~-0.03~~-0.08	~~0.08~~0.06	~~0.03~~0.0	~~0.04~~0.02	−0.21	~~0.04~~0.07	

~~October–November~~March-April	2011	2012	2013	2014	2015	2016	2017	2018
CPOM	~~0.05~~0.01	~~-0.07~~-0.11	~~-0.21~~-0.17	~~0.2~~0.17	~~0.12~~0.1	~~0~~-0.09	~~-0.09~~-0.1	-0.06
AWI	~~0.06~~0.04	~~-0.10~~-0.13	~~-0.25~~-0.19	~~0.22~~0.18	~~0.12~~0.11	~~0.04~~-0.04	~~-0.09~~-0.08	-0.05
JPL	~~0.06~~0.02	~~-0.11~~-0.12	~~-0.15~~-0.11	~~0.16~~0.13	~~0.05~~0.08	-	-	
GSFC	~~0.07~~0.0	~~-0.18~~-0.12	~~-0.30~~-0.22	~~0.21~~0.18	~~0.11~~0.16	~~0.02~~0.01	~~0.07~~0.05	-0.13
CS2SMOS	~~-0.02~~0.03	~~-0.12~~-0.11	~~0.02~~-0.21	~~0.2~~0.17	~~0.04~~0.11	~~-0.05~~-0.07	~~-0.07~~-0.08	
APP-x	~~0.03~~0.02	~~0.02~~0.01	~~0.01~~-0.01	~~0.02~~0.0	~~-0.04~~-0.03	−0.05	-	

Table 4. Annual winter (January–April*) sea ice draft/thickness (meters, top value) and deviation from the ~~climatological average~~2011-2017 mean (meters, bottom value) for each BGEP ULS mooring (A, B, D) and ~~five~~ six satellite products (using measurements within 200 km of BGEP moorings). ~~Climatological mean is calculated per product for the period 2011-2016.~~ *March–April for CPOM statistics.

		ULS	CPOM	AWI	JPL	GSFC	CS2SMOS	APP-x
	2011–2016	1.23~~1.27~~	1.74	1.5~~1.51~~	1.61	1.74~~1.75~~	1.4~~1.17~~	1.99
Mooring A	**2011**	1.15 -0.12 -0.8	1.66~~1.70~~ -0.04 -0.07	1.4 -0.11 -0.1	1.45 -0.16	1.5 -0.25 -0.24	1.43~~1.12~~ -0.04 0.03	2.01 0.02
	2012	1.39 0.12 0.16	1.66~~1.75~~ 0.01 -0.08	1.43 -0.09 -0.07	1.42 -0.19	1.68 -0.07 -0.06	1.20~~0.96~~ -0.2	2.04 0.05
	2013	1.05 -0.22 -0.18	1.82~~2.06~~ 0.32 0.08	1.51 0.0 0.01	1.61 -0.01	1.48 -0.27 -0.26	1.27~~1.16~~ -0.01 -0.13	2.02 0.03
	2014	1.51 0.24 0.28	1.98~~2.12~~ 0.37 0.24	1.71 0.2 0.21	1.85 0.24	2.09 0.34 0.35	1.68~~1.41~~ 0.25 0.28	1.96 -0.03
	2015	1.32 0.05 0.1	1.9~~2.13~~ 0.39 0.16	1.73 0.21 0.23	1.74 0.12	2.09 0.34 0.36	1.69~~1.38~~ 0.21 0.29	1.99 0.0
	2016	1.2 -0.07 -0.03	1.44~~1.61~~ -0.13 -0.3	1.3 -0.22 -0.21	-	1.67 -0.08 -0.07	1.21~~0.96~~ -0.21 -0.19	1.91 -0.08
	2017	0.96 -0.26	1.72 -0.02	1.43 -0.07		1.66 -0.08	1.33 -0.07	

		ULS	CPOM	AWI	JPL	GSFC	CS2SMOS	APP-x
	2011–2016	1.36~~1.38~~	1.88~~1.9~~	1.68~~1.69~~	1.79~~1.78~~	1.84~~1.86~~	1.65~~1.36~~	2.01~~2.02~~
Mooring B	**2011**	1.42 0.04 0.06	1.9~~0 -~~ 0.01 0.02	1.61 -0.08 -0.07	1.61 -0.17	1.61 -0.25 -0.22	1.64~~1.27~~ -0.09 -0.01	2.03 0.02
	2012	1.4 0.02 0.04	1.89 -0.01 0.02	1.67 -0.01 0.0	1.72 -0.07	1.7 -0.16 -0.13	1.65~~1.3~~ -0.06 0.0	2.06 0.05
	2013	1.21~~1.2~~ -0.19 -0.16	2.04 0.14 0.17	1.77 0.08 0.09	1.83 0.04	1.75 -0.11 -0.09	1.68~~1.51~~ 0.15 0.04	2.03 0.02
	2014	1.47 0.09 0.11	2.02 0.12 0.14	1.77 0.09 0.1	1.85 0.06	2.09~~2.09~~ 0.23 0.25	1.82~~1.53~~ 0.17	1.98 -0.04
	2015	1.55 0.17 0.19	2.02 0.12 0.15	1.83 0.14 0.16	1.92 0.14	2.06 0.2 0.22	1.74~~1.41~~ 0.05 0.1	2.02 0.0
	2016	1.26 -0.13 -0.1	1.55 -0.36 -0.33	1.47 -0.22 -0.21	-	1.95 0.09 0.12	1.42~~1.14~~ -0.22 -0.23	1.96 -0.05

		ULS	CPOM	AWI	JPL	GSFC	CS2SMOS OS	APP-x
	2017	1.21 -0.15	1.71 -0.17	1.61 -0.06		1.69 -0.15	1.59 -0.06	
Mooring D	**2011–2016**	1.37~~1.41~~	1.88~~1.9~~	1.62~~1.63~~	1.81	1.81~~1.82~~	1.52~~1.26~~	2.08
	2011	1.48 ~~0.07~~ 0.11	1.7~~1.70~~ ~~-0.20~~ -0.18	1.48 -0.15	1.54 -0.27	1.66 ~~-0.16~~ -0.15	1.44~~1.15~~ ~~-0.11~~ -0.09	2.15 0.07
	2012	1.39 -0.02	1.75 ~~-0.14~~ -0.13	1.51 ~~-0.12~~ -0.11	1.6 -0.21	1.66 ~~-0.16~~ -0.14	1.41~~1.12~~ ~~-0.14~~ -0.11	2.15 0.06
	2013	1.06 ~~-0.35~~ -0.31	2.06 ~~0.16~~ 0.17	1.51 ~~-0.12~~ -0.11	1.59 -0.21	1.55 ~~-0.27~~ -0.26	1.28~~1.18~~ ~~-0.08~~ -0.25	2.16 0.08
	2014	1.64 ~~0.23~~ 0.27	2.12 ~~0.22~~ 0.23	1.89 0.26	2.13 0.33	2.15 ~~0.32~~ 0.34	1.86~~1.55~~ ~~0.29~~ 0.34	2.09 0.01
	2015	1.59 ~~0.18~~ 0.22	2.13 ~~0.24~~ 0.25	1.97 ~~0.34~~ 0.35	2.16 0.36	2.26 ~~0.43~~ 0.45	1.93~~1.57~~ ~~0.31~~ 0.4	2.0 -0.08
	2016	1.31 ~~-0.1~~ -0.06	1.61 ~~-0.28~~ -0.27	1.42 ~~-0.21~~ -0.2	-	1.67 ~~-0.15~~ -0.14	1.26~~0.99~~ -0.27	1.95 -0.14
	2017	1.13 -0.24	1.82 -0.07	1.57 -0.05		1.69 -0.11	1.49 -0.03	

Table 5. Winter (October-April) sea ice thickness growth rate (md^{-1}) for the period 2010-2018 in the central Arctic (regions 1-6), where rates are calculated based on product availability during each growth season.

Growth Rate (m/d)	CPOM	AWI	JPL	GSFC	CS2SMOS	APP-x
2010-2011				0.0032		0.0103
2011-2012	0.0064	0.0058	0.0057	0.0040	0.0050	0.0105
2012-2013	0.0055	0.0048	0.0052	0.0027	0.0044	0.0108
2013-2014	0.0058	0.0049	0.0054	0.0026	0.0042	0.0100
2014-2015	0.0057	0.0051	0.0053	0.0037	0.0044	0.0103
2015-2016	0.0051	0.0050		0.0025	0.0040	0.0099
2016-2017	0.0050	0.0041		0.0022	0.0034	
2017-2018	0.0047	0.0042		0.0014		