The Cryosphere Discussions

The Cryosphere Discuss., referee comment RC3
https://doi.org/10.5194/tc-2022-95-RC3, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on tc-2022-95

Anonymous Referee #3

Referee comment on "Inter-comparison and evaluation of Arctic sea ice type products" by Yufang Ye et al., The Cryosphere Discuss., https://doi.org/10.5194/tc-2022-95-RC3, 2022

Review of

Inter-comparison and evaluation of Arctic sea ice type products

by

Y. Ye, et al.

Summary:
A large, albeit shrinking portion of the Arctic Ocean sea ice cover is made of multiyear ice (MYI) that has survived at least one summer melt season. In order to more accurately assess the trend in Arctic Ocean MYI cover and the coverage of first-year ice, and to more reliably use these ice type fractions in other research areas, such as sea ice thickness retrieval, it is important to evaluate the existing sea ice type products. This study is an attempt into this direction. Nine different sea ice type products based on five different algorithms are compared with the NSIDC sea ice age data set and the MYI extent derived from it as well as with a set of five qualitatively interpreted satellite synthetic aperture radar (SAR) images. Time series of the MYI extent at daily and monthly temporal resolution are shown, inter-compared and discussed qualitatively in the light of the different algorithms, their potential limitations and post-processing steps. The performance of the different products is compared for specifically selected sub-regions of the SAR images.

I have a number of concerns with this manuscript which I summarize in my general comments and detail in my specific comments.

I also would like to note that the manuscript is difficult to read because of quite a number of strange formulations and problems with English grammar.

General comments:
GC1: As the authors state, this is one of the first (kind of) comprehensive evaluation of sea ice type products. This calls for provision of a solid physical background of the sea ice and its snow cover as relevant for its remote sensing using active and passive microwave instruments. This element is missing and jeopardizes the usefulness of the entire manuscript.
GC2: The description of the input satellite data and the algorithms used in the products as well as in the one major evaluation data set used is very heterogenuous and not complete for the understanding of the manuscript and its results. At least two products (NASA-Team MYI concentration and ECICE MYI concentration) are missing in addition.
GC3: The inter-comparison contains, if at all, little quantitative results. The results often appear to be quite hypothetical. As I see it, there are two main reasons for that. At first, the NSIDC sea ice age data set used as the main evaluation data set requires an evaluation that justifies its usage for the purpose of this manuscript. In addition, there is a methodological inconsistency behind comparing daily sea ice type products with weekly sea ice age data. Secondly, the SAR images used are only interpreted in a qualitative way. With that they can be used as a means for a consistency check of the general performance of the sea ice type products - but only within the error margin proposed by this manual interpretation. Both together clearly reduces the value of this manuscript, which has the character of a pure, qualitative inter-comparison study with little in-depth recommendations resulting from it for i) which product to pick and ii) how to improve which product in which way.
GC4: The discussion of the results is not well linked to the existing literature.

Specific comments (contain some typos / editoral comments):

Abstract:

- I recommend that you consider to find and use a different acronym for sea ice type because I find "SIT" very often used as an acronym for sea-ice thickness. A possible alternative could be SITY. Or, since "type" is not really that long compared to the words, e.g. thickness or concentration, you might also consider write the full expression all the time. But "SIT" is a bit unfortunate.

- I also recommend that you very briefly describe the various products named in the abstract. Perhaps they can be categorized into those products that rely solely on C-Band or Ku-Band data and/or products that use both active and passive microwave data? Please check the maximum allowed length of the abstract and perhaps delete details towards the end for more clarity of what types of products you did compare.

- I recommend to state upfront that by "sea ice type" you merely refer to multiyear ice and first-year ice. As you know, there is a number of other sea-ice types which you, however, not appear to take into account.

- L13/14: "towards sea ice ... images" --> "against a sea ice age product and compared with five Synthetic Aperture Radar images"

- While you write in Lines 14/15 about results found at daily and monthly temporal resolution it is not clear whether all products used come at daily temporal resolution. I also note that the sea ice age data set comes at weekly temporal resolution.

- L14/15: Please also see my over-arching comment to the conclusions.

- You might want to re-phrase "anomalous fluctuations" because it is not clear what you mean by that in the context of an underestimation (Line 17).

- Under (3) you write about details with respect to the classification (Line 23). Is the retrieval of all products investigated based on a classification approach?

- I have the feeling that the "Additionally, the change of separation pattern ... SIT method" (Lines 24/25) could be deleted for the sake of having more room for the above-mentioned suggestions.

Lines 41-57: I suggest to better structure this paragraph and in addition provide more background information. Specifically I recommend to

i) Tell the reader that by sea ice type discrimination you are referring to distinguishing between FYI and MYI;

ii) Write what the fundamendal differences in the physical properties of these ice types are that allow us to separate them by means of their microwave signature (be it for active or passive microwave sensors);

iii) Explain more clearly - but still briefly - what the different retrieval approaches are. It is

for instance not clear whether the main approach used is a classification. The NASA-Team algorithm (see below) does not use a classification, neither does ECICE.

iv) Move information about evaluation results obtained by others so far into the next paragraph (see Lines 47/48: "By comparing ... Kwok 2004)."

v) Mention that different methods exist which either provide a fractional MYI/FYI coverage or a binary classification (or assignment of one or the other ice class to a grid cell).

- In addition, I recommend to delete the Lomax et al. 1995 paper and instead include literature related to the NASA-Team algorithm and to the ECICE algorithm which both permit to compute FYI and MYI fractions and are both so far missing completely in your list. I am wondering why you are not considering these products as well in your inter-comparison. I am also wondering whether it would not make sense to get hands on the MYI data sets created by Ron Kwok and used in various publications of him and his group.

Lines 58/59: "Comparison ... methods" --> While an evaluation of products is per se an excellent idea and the improvement of the used retrieval methods a good motivation, I strongly suggest to provide 1-2 sentences that specify more clearly why it is important to (finally) provide a more comprehensive evaluation of these products. The first paragraph of your introduction only tells the reader that sea ice type is important. But requirements about the accuracy and a specific example where an error in the sea ice type discribution of, e.g. 50%, would have which implications is not yet given in a convincing way.

Line 61: Why "limited"? There are plenty of ship observations (see e.g.: https://www.cen.uni-hamburg.de/en/icdc/data/cryosphere/seaiceparameter-shipobs.html)

Lines 62-64: "... some MYI ... in ice charts." --> I don't understand this sentence; please consider to re-phrase it.

Line 67: "operational" --> please check what you mean here by operational. Do you mean existing? Or are you really referring to all sea-ice type products that are currently operationally (aka daily) produced and provided to the users?

Lines 83-86: "Microwave radiometer ... 2016)" --> As stated already in the context of the introduction, it would make a lot of sense to include a paragraph that clearly describes the relevant physical properties of the different sea ice types that are relevant for their discrimination in the different active and passive microwave signals. This is required to understand the algorithm details and to understand their limitations (also during the freezing and/or shoulder seasons) and would be important for the discussion section as well. Since this is the first paper of this kind it is certainly worth to dig into physics here.

Line 91: "on coarse resolution" --> Please write in the text more clearly the grid resolution of the data and, if relevant, also the native resolution of the data used as input.

Lines 93/94: I can understand that Tb measurements are corrected for the atmospheric influence because it disturbs the sea ice signal. I cannot understand why you need to correct the Tb measurements for sea ice concentration ... What I can imagine is that you use an additional sea ice concentration product to restrict the analysis of the sea ice type on the sea-ice covered area. If this is the case then please write it accordingly. However, admittedly this would contradict a bit the next sentence about the Bayesian appraoch do discriminate open ocean and sea ice. In short: You need to rephrase these statements.

Line 96: I suggest to remove the "further" and to also provide an equation of how the gradient ratio is computed.

Line 98: Where are these "fixed target areas" located? How are these selected? How large are these? Do these change annually? And: Why are these fixed?

Lines 100-102: "In 2021 ... scheme" --> My impression is that you are not including data of this new version into your comparison. Therefore I recommend to move this announcement towards the end of your paper, e.g. into the discussion where it could fit with your outlook / description of which improvements are (already) underway. But perhaps 2021 was a typo ...?

Lines 103-115: I recommend to comment on / give more details on:

i) the fact that the OSISAF-SIT is based on a very heterogeneous set of input parameters and on changes in the training data set (L111/112), which both could have an impact on the sea ice type product in terms of its consistency over time;

ii) what a "sigma_nought" is (Line 112) and in which way this variable is used (is it corrected towards a common incidence angle? for instance); what is the incidence angle range used? (Compare the next paragraph where you are comparably detailed as far as it concerns the Ku-Band scatterometers.);

iii) what the native resolution of the scatterometer data is;

iv) what a "swath projection" is (Line 112).

- I furthermore find the introduction of AMSR2 and AMSR-E the way done confusing. AMSR2 is available since July 2012 but it is included since 2016; whether AMSR-E data were at all used is not clear but AMSR-E is introduced.

- You describe the different sensors used with different degrees of detail; for instance you do not mention that SSM/I and SSM/IS are multi-channel radiometers with a number of frequencies while you do so for AMSR-E. You refer to "coarse" (previous paragraph) and "medium-resolution" (Line 106) as well as "higher spatial resolution" (Line 109) without a specific motivation. Why is it important to know the spatial resolution? How does the product (also the other products) actually deal with input data being available at different spatial resolutions?

- What is the spatial resolution achieved by ASCAT and what is the polarization used?

- What are "given weights" (Line 113)? How are these defined?

- L113-115 you could rephrase for improved clarity along the lines: Both C3S-SIT and OSISAF-SIT provide, in addition to the pure ice type classes FYI and MYI, an ambiguous ice type class that represents an unknown mixture of both ice types, referred to as "Amb". The products are provided with ...

Lines 120/121: "ASCAT is ..." certainly belongs either to the paragraph where you introduce ASCAT data for the first time. Or, alternatively, you could think about adding a sub-section wherein which you introduce all sensors and their specifications as far as relevant for this paper. Table 1 provides not enough information.

Lines 116-126:

- I recommend that within this paragraph you underline more clearly that KNMI-SIT is actually a synonym for three different sea ice type products of which you include two into your evaluation. I would then also avoid speaking of "the KNMI-SIT" but in general speak about KNMI sea-ice type products and then define KNMI-Q and KNMI-A as those you are referring to henceforth.

- While your refer to swath and grid in the previous paragraph you don't do this here. In which form are the data of the different scatterometers used within the sea-ice type retrieval? What is the grid resolution? What is the native resolution of the OSCAT and QuikSCAT data? Please refer to Table 1 / Figure 1 for clarification in terms of the time

periods the different satellite data and hence sea-ice type products are available. Reading the text it is not clear which time periods the different (?) products cover.

- L123/124: "In KNMI-SIT ..." --> Does this apply to all three products? Or is there a merged product? Is this classification done after FYI and MYI have been separated? What is the difference in the microwave signal that is exploited to separate SYI from older ice?

- L125/126: "In this study, backscatter ... SIT products." --> I don't understand this sentence; please re-phrase it.

Lines 127-131:

- Like for the previous paragraph it is not entirely clear whether IFREMER-SIT is again just a synonym for the two other products IFREMER-Q and IFREMER-A or whether these two are merged to form one product.

- I note that you give a few details about IFREMER-A but not about IFREMER-Q.

- I am not sure I understand what you mean by "series of time-varying thresholds" ... What is this "series"? Are you referring to a time series of backscatter data for several winters as written in Line 130? What do you mean by "seasonally consistent"? That the values agree with each other through the course of the freezing season?

- While we learn here that the product is gridded to a polarstereographic grid, there is no information about the grid in the previous paragraphs.

Lines 132-137:

- "employs adaptive" --> "employs an adaptive"

- "based on the thought of clustering" could possibly be re-phrased. What kind of clustering approach is used? K-means?

- For the other approaches listed above that utilize radiometer data you state that the gradient ratio of the 37 and 19 GHz channels with vertical polarization are used. Which channels are used here?

- Is it correct that the approach combines coarse resolution radiometer data (what is the resolution? How is the difference in spatial resolution between SSM/I / SSMIS and AMSR-E/2 taken into account?) with fine resolution scatterometer data? What kind of radiometer data are used? Daily gridded? Swath? Which grid?

- You write that QuikSCAT and ASCAT are used successively. Does this mean that you use QuikSCAT data until the very end of its nominal time with regular data provision in 2009 (?) and only afterwards ASCAT? How does the algorithm deal with the substantial difference in sensing geometry and coverage?

Lines 140/141: "climate consistent data record SIT products" --> Given the heteogeneity of the products described in terms of the spatial resolution of the input data and the various combinations of frequencies and potentially also polarizations used, I doubt that any of the above-mentioned products deserves yet an assignment into the group "climate consistent data record". Therefore, personally, I would skip this whole last paragraph (Lines 138-142); I don't think it is relevant for the paper.

Lines 147-150:

- The description of the sea-ice age product should be revised according to the information given in the more recent paper by Tschudi et al., Cryosphere, 14, 2020. In particular statements like "tracking of ice trajectories" should be avoided as should be wrong information about how the data set is derived like "passive and active microwave observations". This ice age data set is derived from the NSIDC sea ice motion data set (which in some way is described in the same paper).

- The paper by Korosov et al, 2018, is about the deficiencies and limitations of the NSIDC

sea ice age data set but should not be cited in the context of its description. I can kind of guess that you added this information "limited by the simple drift model and the oldest ice age assignment of grids" to illustrate that the NSIDC sea ice age data set may have its limitations but this would need to be explained in far more detail than in half a sentence. In fact, it is likely that the sea ice age product overestimates the presence of old ice and therefore is biased towards old ice. Whether this already applies to the discrimination between FYI and SYI I don't know; this I leave to you to think about.

Line 157: "Images are" --> "All five SAR images are ..."
- Was any filtering (speckle?) applied?

Lines 159/160: "the geolocations and acquiring dates of the SAR images" --> "the location of the five SAR images". There is no acquisition date given in Figure 2. Hence the acquisition dates are missing and the time difference with respect to the sea ice type products the SAR images are compared to is unknown. This needs to be included in the revised version of the manuscript.

Lines 161-165:

- "For better interpretation of SAR images" --> This motivation needs to be explained better. It is not at all clear why, for the interpretation of the few SAR images used, these two additional data sets are required. What is the problem with the SAR images that such data are needed?

- Why do you use the CERSAT/Ifremer product - which appears to be quite heterogeneous in terms of the input data when you could have used the NSIDC sea ice motion product coming at 25 km grid resolution on an EASE grid and with daily temporal resolution.

- What is the grid resolution of the ERA5 data and how did you co-locate these data with the sea ice type products and/or the SAR images?

Line 172: The naming of the regions is partly wrong and needs to be corrected. What you call ESS is actually the combined area of the East Siberian Sea and the Laptev Sea. What you call BS is not the Barents Sea but the combined area of the Beaufort Sea and the Chukchi Sea.

Line 180:

- Here, in this line your write "extent is calculated by general extent of pixels", in line 176 you write "MYI extent is estimated as the integral of all pixels specified ..." . Both formulations are not to the point and not specific enough. I recommend to re-phrase in both cases along the lines "We computed the MYI extent as the sum of the area of all grid cells classified as MYI."

- In this context I have two questions. 1) Did you use a common land mask? Or is this not required because the region of interest that is delineated by the red line in Fig. 2 is clear of any land influence? 2) Did you take into account that the grid cell area is only a constant in the EASE grid projection while it changes with latitude for the products in polar-stereographic projection? If you did not take this into account yet you must correct your computations.

Lines 183++: I have a conceptual difficulty with comparing daily sea ice type maps with weekly sea ice type maps derived from the NSIDC sea ice age. The comparison would be much more meaningful if you would average all daily products over every single week also used in the sea ice type product derived from the sea ice age. After all, this is your main data product for the inter-comparison.

Line 184: "decreasing trend" --> What you possibly mean is a decrease or an increase in the MYI extent (over time) or a positive or negative trend of the MYI extent (over time). A decreasing (or increasing) trend, in contrast, is a trend that changes its value with time or, in other words, if this was a linear trend then the slope of the trend line would

decrease (or increase) with time. Therefore, please correct your writing accordingly throughout the manuscript.

Line 186: "the divergent movements" --> Which movements? Movements of what? Are you referring to divergent sea ice motion?

Line 194: "to the NSIDC-SIA extents ... 2-3 years" --> You stated earlier that you compute the MYI extent from the NSIDC-SIA data set by summing over all grid cells exhibiting a sea ice age of 2 years or older. Hence you can simply write "to the MYI extent derived from the NSIDC SIA."

Lines 199/200: Given the fact that the entire Arctic Ocean (i.e. approximately the region of your study) has a size of about 7 x 10^6 km2, and the fact that rarely the entire Arctic Ocean is covered with MYI, this difference is far above being reasonable and requires more explanation. It is a 100% error.

Line 200: Are these extent estimates for class "ambiguous" reasonable? How do these values relate to the entire MYI extent? Please be more critical about and more specific within your interpretations.

Lines 204-210: "This is expected ... summer and winter." --> These lines call for the more careful delineation of the physics behind the various retrieval methods which I asked for earlier. Without that physical background these statements all remain hypothetical and are not sufficiently backed up by existing knowledge and hence not in line with good scientific practice.

Line 214: What do you mean by "most distinct variations"?

Subsection 3.2.1:

- One could have expected that you dedicate a bit more time to comment on the details such as the drop of the MYI extent to zero in some winters, when looking at the NSIDC SIA MYI extent of region ESS.

- What explains, to your opinion, the observation that especially for BS and ESS NSIDC SIA MYI extent is often considerably larger than the MYI extent offered by all other products? This is less pronounced for region CAO where we can also see numerous cases where the other MYI extent products exceed the NSIDC SIA MYI extent.

Lines 220/221:

- "The former ... Stream" --> This is not entirely correct and requires re-phrasing. You have chosen your region CAO such that the Transpolar Drift Stream goes right through it ... from the Pacific side towards the Atlantic side. Therefore, what explains the decrease in MYI extent is i) the export through Fram Strait and, by smaller fractions, into the Barents Sea and through Nares Strait, and is ii) the export driven by the Beaufort Gyre towards the South along the Canadian Arctic Archipelago.

- Note, in the sentence before "BS keeps constant or increasing" should also be re-phrased. You want to state something like the MYI extent in the BS/CS region remains constant or is increasing.

- Finally, what explains the decrease in MYI extent in region ESS? This is not clear yet. If it is exported towards the CAO, then this is going to be a northward flow ... a direction not yet mentiond in your description.

In short: Please be more accurate in the description of the results of your work.

Line 223: What are "varying evolution trends"? Either do trends vary between the different winters of years. Then this could be termed inter-annual variation of the trends describing the evolution of the MYI extent in the respective region. Or you want to comment that within a season, the evolution of the MYI extent from month to month differs between different winters or years. Then you need to specify that you are referring to the intra-seasonal variation of the MYI extent and need to drop the word "trend". Please be more clear in your writing.

Lines 238/239: "the discontinuous ... C3S-SIT" --> This is a too global statement because it reads as if daily MYI extent fluctuations are always explained by this discontinuous FYI delineation. You should not forget that this is a scene at the verge of freeze-up and therefore one cannot expect that all of the MYI has a "mature" microwave signature yet which would the algorithms let define it as such. In addition, I'd say this is an issue that is possibly limited to the late October / early November cases and is not of general validity. Please correct your writing accordingly.

Line 245: "with exceptional MYI distributed ... as ESS." --> I can agree on the 2nd largest MYI extent but there is only a quite small part where the finger-like structure of MYI extends through Chukchi Sea into the ESS region. The other finger-like structure at Severnaya Zemlya can be observed in basically all products and is hence not exceptional. Perhaps you want to state that this protrusion from the Chukchi Sea into ESS is not in agreement with the NSIDC SIA sea ice type?

Line 256: Looking back at this paragraph and the top two rows of Figure 7 you could also state that this is a good example where assigning the ambiguous ice type pixels to MYI actually improves the agreement in the spatial pattern with NSIDC SIA sea ice type.

Lines 268-271:

- This information should be placed in a section about methodologies of the inter-comparison, where you describe how you co-located data, and how you computed the MYI extent from the different data sets (and grids).

- There you also should reflect upon why and how you selected the boxes as you did and why these have a different size.

- It would be furthermore more than beneficial if you would elaborate on the way how you decided, based on the SAR images, which part of the ice is FYI and which MYI. Your sentence "Characteristics of brightness, texture, gemometric shape and context ..." is not sufficient for a journal such as "The Cryosphere"; it rather reads like written for a public science magazine, I am sorry. You have decibel values at hand and by digging into published literature you can get a much better, even quantitative handle on the interpretation of the SAR images.

- I note that you used HV-pol data from Sentinel-1 SAR. Why did you use cross-polarized images instead of co-polarized images? What is the advantage using those? Can I assume that the RADARSAT-1 images were HH-pol? You could note this additional information in the respective figures.

Lines 273-278:

- The description of what is seen in terms of ice types in the SAR image appears to be hypothetical and descriptive. There are tables and publications from which you can learn about the typical signatures (sigma nought) of MYI and FYI at C-Band HH-polarization. You should find and use these to put your assumptions on solid ground. Otherwise also these SAR images cannot serve as an evaluation or even validation data set but rather represent a vague inter-comparison source. And with that you can by no means adequately draw conclusions about the quality of the sea ice type products you are investigating here. You then also need to change the title of the manuscript, leaving out "evaluation". Also the usage of the NSIDC SIA MYI extent does not warrant so because it is known to be biased (this is visible in your manuscript as well) and is not a good source for evaluation in the way carried out by you.

- Please add month and year of the scene to the text.

Lines 279++: I am wondering whether it would make sense to not comment on / discuss every product here in the figures showing the comparison to the SAR images. Perhaps the most striking discrepancies would be enough to mention.

Lines 284/285: "which might be caused ... the product" --> It is not clear how these two issues can lead to an overestimation of the MYI extent derived from the sea ice age product when the delineation relevant to state whether a grid cell contains MYI or FYI is between FYI and SYI, and is therefore only influenced by the time from the last fall until the date this example is from ... hence basically 6 weeks in this case. Only a small amount of FYI is grown until then in that region and one can be confident that the majority of the grid cells is in fact predominantly MYI as seen in the sea ice age product.

Lines 289-293:

- Same comment as for Fig. 8 with respect to how to assign features and/or brightness distributions to ice types. This is a purely qualitative inter-comparison and not an evaluation.

- When I was working with SAR data during my PhD days I was always urged to denote the sensor flight and look direction by arrows. You could do so as well so that it is more clear where the low and where the high incidence angles are located.

Lines 297/298: "The MYI underestimation ... weekly temporal resolution." --> Why? What is the physical process required to have large discrepancies between a weekly ice type map and a daily ice type map? Is there evidence in the additional data used by you about this physical process?

Lines 301-305: Again the interpretation of the SAR image (and the boxes zoomed into) appears to be very hypothetical and is not well backed up by what could be taken from published literature (if the authors would have considered to use HH images instead of HV images). This would also have resulted in less processing artefacts in the image.

Lines 306-311: I would say this is a classical example where the NSIDC-SIA is one of the more useful data sets here. Looking carefully it is clear that regions C and D are both FYI. You can check the minimum extent end of summer 2014 please to check whether in that area close to Severnaya Zemlya sea ice survived the summer melt. I doubt so. Hence these two areas are located within the landfast sea ice (FYI) cover that develops there ususally - as is also well backed up by a sea ice drift speed of zero. While I could agree that region B is in fact MYI I doubt that region A is MYI. This is certainly an area where i) deformation and ii) deep snow plays a significant role in shaping the different microwave signals contributing to the (every) sea ice type classification.

Line 312: Is there any reason why you put mid-winter after late-winter?

Line 322: "Compared to the SAR image ... is overestimated ..." --> This is only part of the story. I would see this more differently and urge the authors to have another look to see that the agreement between NSIDC SIA and the supposedly FYI - MYI distribution in the SAR image is only acceptable in the bottom part of the SAR image whereas towards the top and top left there is both an underrepresentation of MYI and an overrepresentation of MYI, respectively.

Line 323: Please look at my comment to a similar statement made by you further above.

Lines 336-339:

- I don't see how your results underline or agree with the results of Korosov et al. and I also don't see how your results confirm that the NSIDC SIA data set is a cross-validation data set. It is at most a data set for consistency checks and inter-comparison. I will detail why below.

At first: none of the sea ice type products investigated has a finer temporal resolution than the NSIDC SIA product (and hence the MYI extent derived from it). Hence you cannot look at the sub-grid scale distribution of sea ice types (and age) in the NSIDC SIA maps and the information you claim to have at hand originates from the publication mentioned above and is not your own result.

Secondly, even though you have SAR images at hand you did not make the effort to first perform a high-level evaluation of the NSIDC SIA producte BEFORE you use it as a data set for inter-comparison. You could have carried out a dedicated pixel-wise comparison between the NSIDC SIA product and the SAR images used. But this would require i) more SAR images covering the same region over the weekly period represented by the NSIDC SIA product (i.e. ideally one at the beginning and one at the end of the 7-day period) for ... say ... 50 cases (which is a big project) and ii) using SAR images in a quantitative way, i.e. using the sigma nought values to delineate FYI from MYI, and in addition taking carefully the drift and deformation history of the respective regions into account to ensure that areas with a bright signature caused by deformation are not misinterpreted as MYI. Only with such a comparison, looking at the sub-grid scale distribution of the different ice types within single NSIDC SIA 12.5 km grid cells, you can shed more light about the "cross-validation" potential of these maps.

Line 367: "stability of the sea ice types" --> what do you mean with that? FYI will not disintegrate spontaneously and MYI will not become FYI over night. Please rewrite.

Lines 368-370: "This parameter ... or high frequency channels" --> I agree to this statement; however, I am wondering what the magnitude of cloud liquid water values typically observed during winter in the Arctic would be and what the impact would be specifically on the GR. I am pretty sure you can dig out this information in the available

literature and back up your statement adequately. There are sea ice concentration algorithms that specifically make use of the two channels that form this GR, e.g. the Comiso algorithm frequency mode; perhaps the paper by Andersen et al. from 2006 in Remote Sensing of Environment could enlighten you here. In short, unless the impact of atmospheric parameters such as cloud liquid water and water vapor on the GR at these two frequencies is really measurable I would remove this piece of information. If kept it needs to be backed up by adequate literature.

Line 368: I am surprized that one of the sea ice type algorithms uses this GR ratio the other way round, i.e. 19 V minus 37V. Please check. This (again) calls also for a better and more comprehensive description of the algorithm behind the products inter-compared in this study.

Line 371: "ice layering" is one component of the snow properties and should not be mentioned as if it is a different thing.

Lines 373-375: "when air temperatures fluctuates around freezing point and triggers snow metamorphism" --> Apart from the fact that this is another example of bad English grammar this statements needs to be formulated in a less global way.

A) What you call snow metamorphism with a likely impact on brightness temperatures particularly at 37 GHz are melt-refreeze cycles caused by elevated solar radiation during spring (April); during these cycles the air temperatures do not necessarily fluctuate around the freezing point.

B) In October solar radiation is absent, hence cannot be the trigger for snow metamorphism. Melt-refreeze cycles are also absent. What can happen in October is advection of warmer air masses and precipitation falling as wet snow or freezing rain - which admittedly can have an impact on the microwave signature of the sea ice cover. But without working with the theory (missing in your manuscript) you cannot explain it properly. Possibly wet snow masks MYI underneath, letting it look like FYI. But you don't present evidence for this in your data / results. While warm air and hence wet snow might be the reason for the underestimation of the MYI cover in the CAO using C3S-SIT it is not sufficiently clear why snow metamorphism should lead to an overestimaton in the BS and

ESS in late winter. What is the physical process that drives which change in the relevant microwave properties that cause the microwave observations to trick the algorithms, leading to an overestimation in MYI?

C) Another issue you did not yet bring up is the fact that parts of the ESS but also the parts of the CAO facing the Atlantic may experience particularly thick snow loads. Since the GR used here is not only sensitive to the sea ice type but it is also sensitive to the snow depth it is not surprizing that a sea ice type algorithm that uses the GR at 37 and 19 GHz tends to classify FYI as MYI as a result of a thick snow cover.

Lines 375/376: Please explain to the reader what the effect of the temperature correction scheme and the "upgraded tuning of atmospheric correction for Tb" [better --> the improved correction of the Tb for the atmospheric influence] is on the GR used so that the reader gets a credible piece of information here which you again ideally back up with appropriate literature.

Line 377:
- "backscatter (sigma^o)" --> either "backscatter coefficient" or "sigma nought"
- "which has good separability between MYI and FYI." A backscatter coefficient cannot have a good separability between MYI and FYI. A backscatter coefficient might be suitable to separate MYI from FYI.

Lines 378-381: "In comparison ... Fig. 12)" --> Also these lines should be re-written and re-phrased investing more space to describe the issues behind.

- In addition, you might want to provide an explanation why Ku-Band scatterometer measurements appear to be less sensitive to the surface roughness than C-Band scatterometer measurements. How about the sensitivity to the crystal structure of the MYI compared to the FYI? Is the contrast in the backscatter coefficient between MYI and FYI larger or smaller at Ku-Band compared to C-Band? Does this depend on the polarization? Does this depend in the incidence angle? What is the role of the different penetration depths into the snow and into the sea ice?

Lines 388/389: "In Beaufort and ... classification ..." --> please also see my comment for Line 373 further above.

Line 394:

- Either: "employ a dynamic threshold" or "employ dynamic thresholds"

- What do you mean by "variability of [the] training dataset"? Do you mean the spread of values around a chosen threshold brightness temperature or backscatter coefficient?

- What do you mean by "seasonality"? I recognize that sea ice type retrieval is limited to the freezing season, hence one season; you should be more specific here. It is also not clear to what the seasonality refers to ... to the MYI extent? to the physical properties of the sea ice and its snow cover? to the thresholds used?

- I don't understand what you mean by "shift in sensor type". Could you please elaborate on this in the text? I can guess that you perhaps mean the shift between using SSM/I or SSMIS data or between using ASCAT C-Band and QuikSCAT / OSCAT Ku-Band. But to me this is not a shift in sensor TYPE because it is either radiometers or scatterometers. Please be more specific here.

Lines 398/399: It is not clear what you mean by "takes sea ice variabilities into account". What "sea ice variabilities"? Are you referring to the spatiotemporal development of the physical properties of the sea ice and snow cover that influence it microwave backscattering characteristics and/or the microwave emission? Then please write it specifically. Currently, "sea ice variabilites" can mean anything from variations in sea ice thickness or concentration, different ice drift patterns, floe-size distributions, degree of deformation whatsoever ...

Lines 409/410: "can be partly ... more obscure"

- "obscure" --> "difficult" or "problematic"

- The statement as written is not conclusive because you are not providing the key message that the Arctic Ocean has lost a lot of its oldest ice AND that the difference in the radiometric and microwave backscattering properties is usually more pronounced between FYI and these older ice types than, e.g. second and third year sea ice. This feeds back again to the missing description of the physical background behind the sea ice type retrieval earlier in your manuscript.

Lines 418/419: In all of the regions mentioned here MYI ice can occur once in a while and hence a MYI ice signature in these regions certainly is not unphysical. Apart from that is the Chukchi Sea part of your region BS. This needs to be corrected in the text.

Lines 420-422: "Statistical thresholds ... the ice edge" --> Please provide a plot which illustrates how PDFs of the respective parameters used in the retrieval (i.e. backscatter coeffient or Tbs or GRs) of the MYI overlaps with the PDFs of ice types typically encountered along the ice edge so that the reader understands what you are referring to. Ideally, you have this figure along with the revised description of the sea ice type algorithms earlier in the manuscript so that here you simply need to refer to that figure.

Lines 422/423: "exclude ... distributions." --> I don't get what you want to state here. If the MYI extent in the above-mentioned peripheral seas and/or along the ice edge would be added to the MYI extent in your region of interest this would mean a considerable change in the overall SIT distribution. Therefore, please re-phrase your statement as it is currently not clear enough.

Lines 424/425: "reassign ... intrusions." --> Not sure what you want to state here. Do you mean "assign grid cells erroneously classified as FYI as the result of warm-air intrusion induced changes in the surface snow properties to the ice type MYI." ?

- How is this temperature based correction done?

- Aren't there other algorithms (published by one of the authers) that use this temperature based correction as well?

Line 425:

- What does an "ice motion confining procedure" do? I have no clue. Please explain it to the reader.

- "anomalous MYI overestimation" --> What is this? What is a "normal MYI overestimation" and what is the difference to an "anomalous" overestimation? Please re-phrase.

Lines 428/429: It is not sufficiently well described how the correction based on a median filter (spatial or temporal) works.

Line 432: "the five series SIT products ... are defined." --> I don't understand this sentence. Please re-write. It is possibly a problem of the grammar.

Line 434: "Typically ..." --> I don't see that your manuscript warrants yet to state the reason given for the larger spread in MYI extent during early and late winter as being typical.

Lines 440/441:

- It is sufficient to write "grid resolution", spatial can be omitted.

- "foot print" --> "footprint"

- What is the "true" spatial resolution of the ASCAT data? What is the "true" spatial resolution of the QuikSCAT data? You should please not forget that the finer grid resolution provided by the SIRF products (4.45 km) is the result of heavy smoothing and other signal reconstruction steps.

- Another issue that you did not take into account here are the different incidence angles of - especially - the ASCAT C-Band data compared to the microwave radiometer data and QuikSCAT / OSCAT.

Line 452:

- Add "five" in front of "SAR images".

- Any reason why you are not mentioning the NSIDC SIA product here?

Line 453 / the conclusions in general:

- I am not sure I would select a sea ice type product based on the maximum difference that a product might have compared to another independent data source. I would be interested in whether there are regions and time periods where there are systematic errors (and how large these are on average so that I might be able to correct them). In addition, I would be interested in the average performance of the product over a longer time period, i.e. whether there are artificial trends.

- I suggest to re-write your conclusions accordingly, focussing less on the individual products as you do in the list 1) to 4) (which should in any case contain 5 or even 9 entries according to what you write in Line 451), and instead concentrating on the larger picture provided by your qualitiative results. It might help in this context to again take a look at your time series plots and focus less on the inter-comparisons with the SAR images.

- I like the bullet point list further down on the next page. That one looks good but could be written even better by including specific details and referring to the existing literature.

Line 477: "extensive misclassification with higher uncertainties" --> So, the misclassification in itself is highly uncertain? Please re-write.

Line 480: "Ku-Band ..." This statement is not new and has to be backed-up by existing literature.

Lines 488-490: "On the other hand ... become obscure." needs to be re-written. The meaning is not clear and the grammar is not correct.

Line 492:

- Apart from the fact that we still don't know how "ice motion confining" works, it is not clear what "accumulative errors" are. Consider re-phrasing for improved understanding.

- "These post- ..." --> "Any post- ..."

Line 491: What is meant by "should be accounted with caution"? Please consider re-phrasing for improved understanding.

Lines 495/496:

- "This study ... of SIT retrieval approach" --> I don't agree. This study does not contain an "evaluation"; it is an inter-comparison study, mostly involving qualitative results. It provides hints of the quality of the sea ice type products investigated RELATIVE to the NSIDC SIA data set (which in itself is not well evaluated) and relative to only five SAR images which are not interpreted quantitatively.

- I further object to the notion "most popular". Please consider re-phrasing.

- I cannot see the "hints for further improvement". While you state where some of the sea ice type products have deficiencies, you neither come up with specific suggestions about how to improve (e.g. use a SIRF-like product as an input to the OSISAF sea-ice type product to improve the grid resolution) nor does the nature of your results being based on an inter-comparison to qualitative data support to draw conclusions into this direction. I warmly suggest to tone down the value and potential impact of your results.

Line 497: Please share with us which two frequencies WindRAD is going to use.

- "the potential of scatterometer on ice type discrimination" --> "the potential of scatterometer measurements for ice type discrimination"

Lines 499/500: "low frequency microwave measurements" --> "low frequency microwave radiometer measurements" because ASCAT already has been using C-Band for 15+ years which is also a "low frequency microwave measurement"

Figure 3: I suggest to reduce the number of colors used by getting rid of the NSIDC sea ice age and instead show the MYI extent derived from it as a black line - like you do in Fig. 4. If you want to show examples of how the different sea ice type products deal with different sea ice age then I suggest to show just the respective year - ideally a year where almost all sea ice type products provide MYI extent so that you can compare between the products. Alternatively, you could consider showing only MYI extent differences. If you want to keep the sea ice age information then I recommend to use shades of grey instead of colors for the sea ice age.

Figure 4: Why are IFREMER-A data missing for January and April in 2014 & 2015?

Figure 6 and 7:

- It is very counter-intuitive to show open water in brown, land in light grey and FYI in blue. Please use a more intuitive coloring such that, e.g. land is brown, open water is blue and FYI, Amb, and MYI are perhaps medium grey, light grey and white; the observation gap at the pole can then be colored black.

- I recommend to enlarge the figure as a whole.

- In the caption you could cross-ref to Table 1 or Figure 1 to make clear why there is a different number of maps for the two dates shown. In addition you need to refer one more time to the meaning of the red line and you need to comment on the different coloring of the observation hole.

Figure 8:

- What is the motivation to show boxes A to D with a different size?

- I have the same comment with respect to colors as for Figures 6 & 7.

- I suggest to show the NSIDC-SIA map in a different color code as well. What is important for you is to discriminate FYI from older ice which currently is difficult to delineate because the colors used for FYI and SYI are quite similar.

- I recommend to rename "sea surface wind" to "10m wind" because I guess this is what it is. Also make clear that "air temperature" possibly is the "2m air temperature". The additional information that these are daily averages would be appreciated as well.

- I would replace the legends for those sea ice type products that do not provide the ambiguous ice class with a legend which only shows the two ice classes present. It might make sense - in general - to then also include the class open water in the legend.

Figures 9 to 12:

- I have the same comments with respect to colors, legends, and ERA5 data naming as I had for Figure 8.

- In addition, delineation of the boxes in the Sentinel-1 image in a different color than black would help to locate these better.

- It might make sense to not use a continuous color table for the legend of the ice drift field, as the values are increments of 0.1 km/day.

Table 2: Cases where there is a "+" and a "-" indicate that both performances exist?

Table 3: I guess the GR listed in the context of OSISAT SIT is not correct?

Typos / editoral comments:

Line 59; "indirect validation" --> perhaps better "inter-comparisons"?

Line 89: "KNMI-" --> "KNMI-SIT"

Line 144: "study, sea ... were used ..." --> study, we used a sea ice age (SIA) product and five SAR images ..."

Line 154: "with SAR ... of HH" --> possibly better: "providing C-Band (5.3 GHz) SAR images at HH polarization."

Lines 155/156: "providing cross- ... ranging from" --> possibly better: "providing C-Band (5.4 GHz) SAR images at co- and cross-polarization (HV and HH) with incidence angles between"

Line 170: "polar hole of 87degN" --> "data acquisition gap north of 87degN centered at the pole"

Line 176: "within studied area" --> I picked this as one of the examples that underline the need for considerable English editing of the manuscript. The authors must check for usage of "the" and "a" which is often missing.

Line 181: You stated already in Line 170 that you excluded that area centred at the pole. Therefore you can delete this sentence.

Line 185: "The MYI" --> However, the MYI"

Line 185: "regional" --> "regionally"

Line 225: "is mainly resulted from" --> check grammar.

Lines 239-241: This part does not belong to the top row of Figure 6, right? It belongs to the data from 2007 and should be placed into the next paragraph.

Line 280: Typo: "boarder" --> "border"

Line 288: Add the year.

Line 289: "... in the western part were higher than in the eastern part."

Line 294: "Slightly underestimation of MYI" --> check grammar.

Line 296: I would say that "thin" could be misinterpreted as "thin MYI" in terms of its thickness. You might want to consider using "narrow" or "filament-like" or "finger-like" or similar.

Line 297: "can partly be resulted" --> check grammar.

Line 300: "A Sentinel-1 SAR image covering the southern part of the ESS near the coast acquired on April 27, 2015 is shown in Fig. 10."

Line 313:
- "transit zones" --> "transition zones" or "zones of mixed FYI - MYI coverage"
- "steady discrepancies" --> re-phrase please.

Line 315: Delete "validation and"

Lines 324-325: "The MYI feature ... round MYI floe" --> please check grammar.

Line 364: "serial" ???

Line 378: "when using backscatter" --> when using backscatter coefficient measurements of an active microwave instrument."

Line 382: "confirmed" --> "shown"

Lines 397/398: "vary ... ASCAT" --> "are different, especially at C-Band."

Line 402: "speculate" --> "hypothesize" ?

Line 405: Either "to a sea ice type distribution" or "to sea ice type distributions"

Line 406: Either: "An adavtive clustering algorithm is used" or "Adaptive clustering is used"

Line 408: "thin ... seas" --> "narrow MYI tongues in the peripheral seas"

Line 427: "continuous underestimation" of what?

Line 430:

- "over-correction problem" --> "over-correction.

- "thin MYI ... seas" --> We had that expression earlier. Please look up my comment there.

Line 436: "fully evaluated" --> "done"

Line 481: What is "small FYI in MYI pack"? Do you mean: "comparably small areas of FYI within a region dominated by MYI?"

Line 487: "deep " --> "mid-"

References: You need to check your reference list. For a considerable number of the entries the records are not complete; for instance is the year missing quite often. At least one of the references appears twice.