Response to RC 1:

Dear Editors and Reviewers:

We would like to thank the reviewer for the detailed and helpful comments, suggestions, and careful checking. Comments are responded on point-by-point basis. The reviewer comments appear in black. The responses are in blue and the proposed changes to manuscript are in ***bold italics***.

Summary

This paper compares different sea ice type products currently available to the community. The products are based on passive microwave data, scatterometer data (C or Ku band), or a combination of both. The products have been developed empirically via training data. The type fields are inter-compared and evaluated against a widely-used sea ice age product and SAR retrievals. The products perform better in mid-winter than in early or late winter when melt/re-freeze may occur. Ku-band scatterometer generally is better at type discrimination. Combination of passive microwave and scatterometer data can yield better performance at times, but not in all situations.

General comment

This is a fairly comprehensive review of the primary sea ice type products available. There are notable differences in how the products are assembled, the input source data, and their performance in different conditions. Thus, this paper is a valuable contribution to the community be providing such an assessment. The paper is quite thorough and overall it does a good job in presenting the inter-comparison and evaluation of the products.

Reply:

Thank you for the thorough review. Your comments and suggestions are highly appreciated. For better comparison and evaluation of the sea ice type products, we revised the manuscript from the following two aspects:

1) The "Data" section was re-structured. This section includes four sub-sections: "2.1 Microwave remote sensing", "2.2 Sea ice type products", "2.3 Sea ice age product" and "2.4 Other data". In section 2.1, specifications of the sensors and the satellite data are introduced in a chronological order, with subsections of passive/active microwave remote sensing data. In section 2.2, theory of SIT classification is introduced at the beginning, followed with the overall description of the respective SIT products in terms of grid size, projection, availability period, a summary of the satellite data used and the algorithm with necessary details. In section 2.3, sea ice age product (with evaluations from previous studies) is introduced with the input data, algorithm and evaluation of the product. In section 2.4, the SAR images along with auxiliary data are described accordingly.

2) A new section of "Methods" was added, which includes "3.1 Estimation of MYI extent" and "3.2 Visual interpretation of SAR imagery". We modified the computation of MYI extent in the revision for consistent griding, projection among all the SIT products. In section 3.1, Information such as co-locating/re-griding the data and calculation of the MYI extent was introduced. In section 3.2, the theory and characteristics of sea ice classification in SAR images were introduced with references from previous studies and examples from our study. In addition, we interpreted all the entire SAR images, consulted with ice experts regarding the results, compared with the gridded ice classification results, and eventually gave the quantitative evaluation results (Kappa coefficient and overall accuracy for respective SAR image and SITY, SIA product).

Besides, case studies were presented in the chronological order with more discussions referring to the physical background and the algorithms of SIT products. Figures were modified for better presentation. A thorough edit of the language style and grammar were conducted. And all the references and citations have been double-checked and corrected accordingly.

Specific comments are below, but one overall comment is on the SAR data used for evaluation. In general, SAR is going to be the best "truth" for comparison. It is high resolution, so it can delineated even individual floes often. And it is all-sky, so retrievals of type are available anywhere the sensor collects imagery. However, the challenge with SAR is interpreting the imagery. The authors interpret the SAR imagery and classify various locations as a given ice type, but they don't give a particular rationale or provide references for their classification basis. Often, expert ice analysts interpret SAR fields for operational ice charts. They have deep experience in understanding the imagery and properly defining features. It appears the authors here classify the imagery themselves. This is okay, but I would like to see more substantial justification for their classification. Reply: Thank you for the advice. In the revision, we added a new section regarding the theory of SAR interpretation. The theory and characteristics of sea ice classification in SAR images was introduced in the section "3.2 visual interpretation of SAR imagery".

Another weakness with the SAR comparison is that it is just a few scenes in selected regions and selected periods. And even within the SAR scenes, a few specific locations are picked out as "pure types" for comparison. Ideally, a full SAR image would be classified and compared. I know automated SAR classification algorithms for sea ice are troublesome, so I can understand the approach taken, but it results is a fairly ad hoc and qualitative evaluation. Since this paper is otherwise quite comprehensive, I won't request more evaluation, but ideally (perhaps in a future paper), it would be good to get classified SAR images – perhaps from an expert ice analyst at an operational ice center – and conduct a more comprehensive and quantitative evaluation of the ice type

products.

Reply: Thank you for the advice. In the revision, we interpreted all the entire SAR images, consulted with ice experts regarding the results, compared with the gridded ice classification results, and eventually gave the quantitative evaluation results (Kappa coefficient and overall accuracy for respective SAR image and SITY, SIA product).

A final note is that there is a need for a thorough copy edit for English language style and grammar. The issues are mostly minor – in particular, there are numerous missing articles ("the", "a", "an") – but they are widespread throughout the manuscript. I don't bother to point them out individually as they are too numerous, but they need to be addressed before final publication.

Reply: Thank you for the advice. We went through the manuscript and conducted a thorough edit for the language style and grammar.

Specific comments (by line number):

11: The authors definite "sea ice type" as "SIT" here. This is fine and it is used consistently throughout the manuscript. However, as a sea ice scientist, "SIT" means "sea ice thickness" to me. And particularly with numerous thickness products coming out from altimeters, "SIT" is becoming quite common in the community to denote thickness. I can understand wanting to use an abbreviation and "SIT" makes sense for ice type, and the context is clear throughout the manuscript. So, I can't say it needs to be changed, but it might be something for the authors to consider. For me, every time I saw it, "thickness" popped into my mind first until I recalibrated. I can't think of another good abbreviation myself, but one could just use "type" or "Type" as a short-hand, instead of "SIT".

Reply: Thank you for the comment. We now use "SITY" to represent "sea ice type". All the abbreviations in the manuscript were modified accordingly.

28-30: I'm struck by the use of more than author listed and then "et al." in the citations – i.e., "Comiso, Parkinson, et al., 2008". Generally, if there are more than two authors, just the first author is listed followed by "et al." – i.e., it would be "Comiso et al., 2008". In looking at The Cryosphere guidance for citations, I don't see anything that indicates two authors should be listed, so I'm not sure of the rationale. This seems to be done throughout the manuscript. (If there are only two authors, you list both, e.g., if it were "Comiso and Parkingson, 2008".) Not a big deal and I assume the copy editing will decide the proper citation format. I just haven't seen this before and it struck me as odd.

Reply: Thank you for the advice. We carefully checked all the references and corrected the citations in the revision.

31-32: Be careful about terminology. "Thin" and "Young" ice are standard stage of

development classifications. I think here you mean "thinner and younger" for FYI, and then "thicker MYI". I'm also not sure what you mean by "firm" in relation to MYI?

Reply: Agree. The sentence was modified to (L37-L38):

"The Arctic sea ice has been increasingly dominated by thinner and younger first-year ice (FYI) instead of thicker and older multiyear ice (MYI), …"

57: "ergodic" is an obscure word – I was not familiar with it. Based on my understanding after looking it up, I'm not sure it is used properly here. Regardless, I think a simpler word is appropriated here or I wonder if it is needed at all – "combined use of both data" is clear to me.

Reply: Agree. The word "ergodic" was deleted. "combination use of both data" was replaced with "combined use of data".

62-63: "While ice chart…" is a confusing sentence – not sure what it is say. I would suggest revising.

Reply: Thanks. The sentence was modified to (L83-L85)

"While the ice chart is used as "ground truth" in some validation (Aaboe et al., 2021a), some areas of MYI in the ice charts correspond to areas with MYI concentration of approximately 50% or greater (Lindell and Long, 2016a), indicating the overestimation of MYI in ice charts."

72: Just one example of grammar/style issues: "…are detailed investigated." – It should be "are investigated in detail."

Reply: Done.

107-109: Is AMSR-E used in the product? The description indicates only AMSR2 is used. So, why describe AMSR-E characteristics? Why not just describe AMSR2 characteristics?

Reply: AMSR-E is used in Zhang-SIT from 2002 to 2011. In the revision, we added a new section "2.1 Microwave remote sensing data", where the satellite data are introduced in a chronological order meanwhile with same level of details. The microwave radiometer data were introduced as two series of dataset: one includes SSM/I and SSMIS, the other means AMSR-E and AMSR2.

109: Maybe another grammar/style issue: "working" is okay, but typically when describing sensors or satellites, "operating" or "collecting data" are more common. "working" seems a bit colloquial here.

Reply: Agree. "working" is replaced with "operating".

147: This goes for all products, but noting here because NSIDC products have specified

references that should be used. For SIA, it is:

Tschudi, M., W. N. Meier, J. S. Stewart, C. Fowler, and J. Maslanik. 2019. EASE-Grid Sea Ice Age, Version 4. [Indicate subset used]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: https://doi.org/10.5067/UTAV7490FEPB. [Date Accessed].

This should be cited in the manuscript text and listed in the references. I see that the dataset website is noted in the Acknowledgment section, but where a reference is provided, it should be included in the manuscript proper, including the dataset DOI. I know all datasets do not provide a formal citation and/or DOI – for example for OSI-SAF, their recommended citation is simply: "The type dataset shall be referred to as the Sea ice type product of the EUMETSAT Ocean and Sea Ice Satellite Application Facility (OSI SAF, osi-saf.eumetsat.int)." If that is all that is provided, that is fine, though I would also say that the product ID (OSI-403-d) and version (if provided) should be included. The other datasets used should be cited to the extent they properly can be.

Reply: Agree. References for all the datasets used in the manuscript were updated accordingly.

185-186: I think the potential for MYI increase could be explained better here. In practice, overall Arctic MYI cannot increase over the winter – it can only decrease via advection out of the Arctic. "Temporary" increases can happen within products due to divergence – e.g., a 100% MYI pixel diverging into two pixels with 50% ice each; if the threshold for detection is <50%, there will now be two pixels. And regionally, MYI can increase, both due to divergence or due to advection into the region from neighboring regions.

Reply: Thank you for the advice. The sentence was modified as below (L350-L351):

"However, it can temporarily or regionally increase due to ice divergence or advection from neighbouring regions (Kwok et al., 1999)."

191: This is discussed a bit more later, but this left me hanging: "why such a dramatic peak in the first half of winter?" Maybe provide a brief explanation and then say it will be discussed further later in the paper.

Reply: Thanks for your suggestion. This section has been modified thoroughly. The following sentence were added to give a brief explanation (L359-L361):

"For C3S-SITY and OSISAF-SITY, such pattern is caused by underestimation of MYI in October, while for KNMI-A it is mainly due to the overestimation of MYI in November in the peripheral seas of the Arctic and will be further discussed later in Section 5."

204: I would use "to" instead of a "-" because it looks like a minus sign. Or use an "em-dash" or "en-dash" with spaces on each side.

Reply: Agree. "-" was replaced with "to" or "en-dash" in the revision.

219: Figure 5 is mentioned quite cursorily here, but I notice the behavior of several products in BS during 2016-2017. That sticks out compared to other years and regions. Why was the performance so different?

Reply: Thanks for the comment. Explanations for the performance were added in the revision. This sentence was modified as below (Note that we modified the abbreviations of each region in the revision) (L412-L414)

"Overall, the MYI extent in the CAO and ESL regions shows a consistently negative trend, while that in the BCS region remains constant or is increasing. The former mainly results from the outflow of MYI to more southern areas."

224-225: Okay, the KNMI-SIT increase is mainly in the BS and ESS regions. But why? In general, this paragraph (223-229) feels like it needs to drill down a bit more and give more detail/explanation.

Reply: Thank you for the advice. We included more discussions in the previous section "4.1.1 Weekly MYI extent variation", where the reasons behind are similar. We therefore did not include more explanations here.

259: Kind of the same thing here. Okay, you have an overestimation of MYI, but that doesn't specifically explain the abnormal increase in MYI during 2016-2017. Why was the MYI overestimated in the one year versus others.

Reply: Thanks for the advice. The abnormal increase in MYI during 2016-2017 was explained in section "4.1.1 Weekly MYI extent variation". The sentences read as below (L353-L356):

"For instance, all the SITY products show increasing MYI extent in March/April 2017 except Zhang-SITY. This could be caused by the enhanced melting during this period (Raphael and Handcock, 2022; Ye et al., 2016a), which leads to the radiometric and scattering signatures of FYI similar as that of MYI therefore unsatisfactory performances of the SITY algorithms."

265: How are cases selected? Were they ad hoc? Random? Was it simply availability of imagery? Or was there some physical rationale to select the scenes? I understand in general wanting different regimes and different time periods, but why those specific images on those specific days at those specific regions? In other words, what "different conditions" were you selecting for here?

Reply: Thank you for the advice. The images were selected based on the availability,

time, region and overall SIT distribution. The sentence was modified to explain the rationale of selecting the scenes: (L469-L471)

"Five cases are addressed in this study to present SITY distributions under different conditions based on the availability of data and feasibility of visual interpretation."

268-271: Following from my general comment above, how were characteristics of the SAR images used for visual interpretation. What is the basis? There are no references here to justify the classifications.

Reply: Thank you for the comment. We added references to justify the rationale of the classification. More specifically, in the new section "3.2 Visual interpretation of SAR imagery", the theory and characteristics of sea ice classification in SAR images were introduced.

273-278: This paragraph illuminates the previous comment. The text is very "squishy". You say things like "appears to be MYI", "more likely to be MYI", "which could be interpreted as newly generated FYI". This is very qualitative and seemingly tentative. I think maybe you just need to say why something "appears to be MYI" – what is in the SAR image that leads to that conclusion and what is the basis for that?

Reply: Thank you for the advice. As mentioned in previous replies, the theory and characteristics of sea ice classification in SAR images will be introduced in the new section 3.2.

288-293: Same here. You have "could be identified as MYI", "are a typical feature of FYI".

Reply: Thanks. In the revision, we added a new section on the theory and characteristics of sea ice classification in SAR images, which become the basis of the visual interpretation from SAR.

299: I guess there is a thematic reason for the order – looking at early and late winter as "edge" cases, but it seems more logical to order these subsections chronologically: early winter, mid-winter, and then late winter.

Reply: Thank you. These cases were presented in the chronological order as suggested.

315-319: Again, very qualitative.

Reply: Thanks. Quantitative results such as Kappa coefficient and overall accuracy of the respective SIT products on each case were added.

391-446: I can see the logic of discussing the methods here – you are linking them to the performance assessed previously. However, to a large extent, this feels like it should go with the data product descriptions in Section 2.1. I guess moving this there would

make that section rather long. But I kind of feel like I get to here and I finally understand how the type products are created – after all of the evaluation and comparison. I'll leave it to the authors to decide if this fits better in 2.1 or should stay here. Or maybe, put some description in 2.1 and then the relation to the product evaluation here.

Reply: Thank you for the thoughtful comments. In the revision, we included more details of the algorithm used in the respective SIT products in the section "2.2 sea ice type products" and put the description of satellite data to section 2.1 "Microwave remote sensing data".

400: "[55]"? Is this a numbered reference?

Reply: Yes. We have corrected the citation in the revision.

434-436: Melt affects the performance in early and late winter. But melt also basically makes the algorithms ineffective in spring and summer. That is implied, but never really stated explicitly it seems.

Reply: Thank you for the advice. The sentence was modified to (L670-L671):

"In early and late winter larger uncertainties are likely to occur due to surface processes such as wet snow attenuation and changes in brine salinity (Barber and Thomas, 1998; Voss et al., 2003; Ye et al., 2016a; Ye et al., 2016b)."

Figure 5: Following up from above, I'm struck by the notable increase in many products in BS during 2016-2017. That is not noticeable in another region or year. There is some discussion, in relation to OSI-SAF in Figure 7, but the text doesn't really discuss this. I think this deserves some explanation in the text.

Reply: Thank you for the comment. We included discussions related to the notable increase during 2016-2017 in the section "4.1.1 Weekly MYI extent variation". The sentences read as below (L353-L356):

"For instance, all the SITY products show increasing MYI extent in March/April 2017 except Zhang-SITY. This could be caused by the enhanced melting during this period (Raphael and Handcock, 2022; Ye et al., 2016a), which leads to the radiometric and scattering signatures of FYI similar as that of MYI therefore unsatisfactory performances of the SITY algorithms."

Figure 5: What is the shaded green region that accompanies OSISAF? Maybe it is in the main text and I missed it, but regardless, it should also be included in the caption for the figure.

Reply: The upper and lower lines of shaded green region represent the MYI extent calculated under the assumptions of 1) regarding the "Amb" class as MYI, and 2) regarding the "Amb" class as FYI. In the new section "3.1 Estimation of MYI extent", we included such statements. (L300-L302)

"The Amb class in C3S-SITY and OSISAF-SITY could be regarded as either MYI or FYI thus the MYI extent is calculated under both circumstances. This results in two values for the respective SITY products, one for the pixels of MYI class and the other for the pixels of MYI and Amb classes."

Figure 6-7: This is a style/aesthetics thing, but the beige/brown for the OW seems odd. Such a color is more commonly used for land, and definitely not for ocean. I would suggest considering a different OW color – just swapping the land color (light gray) for the ocean color, would be more logical to me. Of course, you want to make sure that the colors contrast and are clearly delineated. But a good solution other than beige/brown for OW seems possible.

Reply: Thanks. The colors were modified as suggested in the revision.

Response to RC 2:

Dear Editors and Reviewers:

We would like to thank the reviewer for the detailed and helpful comments, suggestions, and careful checking. Comments are responded on point-by-point basis. The reviewer comments appear in black. The responses are in blue and the proposed changes to manuscript are in ***bold italics***.

General comment

This manuscript presents the inter-comparison of various SIT products from microwave remote sensing data. The performance of the SIT products was evaluated and the causes of differences in the products were analyzed. SIT has been used as an important information in research for global climate change and future prediction. Therefore, a comparative study on the performance of the operationally used SIT products is of high importance.

The manuscript is well written, and appropriate tables and figures are used to explain the results. It seems very meaningful to analyze the comparison results in time and space. However, in order for this manuscript to be published in the Cryosphere, more descriptions should be added about the data and methodology used in the study (requires a section on methodology). More discussion of the results is required.

Reply:

Thank you for the thorough review. Your comments and suggestions are highly appreciated. For better comparison and evaluation of the sea ice type products, we revised the manuscript from the following two aspects:

3) The "Data" section was re-structured. This section includes four sub-sections: "2.1 Microwave remote sensing", "2.2 Sea ice type products", "2.3 Sea ice age product" and "2.4 Other data". In section 2.1, specifications of the sensors and the satellite data are introduced in a chronological order, with subsections of passive/active microwave remote sensing data. In section 2.2, theory of SIT classification is introduced at the beginning, followed with the overall description of the respective SIT products in terms of grid size, projection, availability period, a summary of the satellite data used and the algorithm with necessary details. In section 2.3, sea ice age product (with evaluations from previous studies) is introduced with the input data, algorithm and evaluation of the product. In section 2.4, the SAR images along with auxiliary data are described accordingly.

4) A new section of "Methods" was added, which includes "3.1 Estimation of MYI extent" and "3.2 Visual interpretation of SAR imagery". We modified the computation of MYI extent in the revision for consistent griding, projection among

all the SIT products. In section 3.1, Information such as co-locating/re-griding the data and calculation of the MYI extent was introduced. In section 3.2, the theory and characteristics of sea ice classification in SAR images were introduced with references from previous studies and examples from our study. In addition, we interpreted all the entire SAR images, consulted with ice experts regarding the results, compared with the gridded ice classification results, and eventually gave the quantitative evaluation results (Kappa coefficient and overall accuracy for respective SAR image and SITY, SIA product).

Besides, case studies were presented in the chronological order with more discussions referring to the physical background and the algorithms of SIT products. Figures were modified for better presentation. A thorough edit of the language style and grammar were conducted. And all the references and citations have been double-checked and corrected accordingly.

Specific comments

Abstract: Specify the names of the SIT products (algorithms) analyzed in this study.

Reply: Thanks for the advice. We modified the sentence as follows.

"This study analyzed eight daily SITY products from five retrieval approaches covering the winters of 1999–2019, including purely radiometer-based (C3S-SITY), scatterometer-based (KNMI-SITY and IFREMER-SITY) and combined ones (OSISAF-SITY and Zhang-SITY)."

(Note that sea ice type is abbreviated as SITY in the revision)

Line 27: Please state clearly why sea ice is a sensitive indicator of climate change.

Reply: Thanks. The sentences were modified as follows in the revision (L28-L32).

"Sea ice is an important component of the earth system. Sea ice influences climate change through two primary processes: the ice-albedo feedback and the insulating effect. Sea ice reflects more solar radiation than the ocean due to its high albedo. In addition, sea ice hinders the heat exchange between the ocean and the atmosphere because of its low thermal conductivity. Through global warming, the loss of sea ice leads to increasing absorption of solar radiation and heat flux from the ocean to the atmosphere, which further enhances the loss of sea ice and global warming."

Line 29: It would be nice if it quantitatively indicated how much the thickness and volume of sea ice decreased.

Reply: Thanks. The sentences were modified to provide quantitative indication of thickness and volume decrease (L33-L36).

"Its extent has reduced by 40%–50% of its average in the 1980s (Perovich et al., 2020), whereas the average ice thickness has decreased by about 1.75 m in winter in the

central Arctic Ocean (Rothrock et al., 2008; Kwok and Cunningham, 2015), which eventually leads to a volume loss of roughly 66% since 1980 (Petty et al., 2020; Kwok, 2018)."

Line 35-37: Please specify how sea ice patterns affect Arctic and mid-high latitude regions and how they affect Arctic ecosystems.

Reply: Thanks. These sentences were modified as below (L40-L45).

"The change of sea ice type (SITY) distribution impacts the climate of the Arctic and mid-high latitude regions through changes in water vapor, cloud properties, as well as large-scale atmospheric circulations such as the Atlantic Meridional Overturning (Liu et al., 2012; Screen et al., 2013; Boisvert et al., 2016; Belter et al., 2021). In addition, it influences the Arctic ecosystems by changing the habitat conditions for various Arctic species and is crucial for human activities such as shipping, tourism and resource extraction (Emmerson and Lahn, 2012; Meier et al., 2014)."

Line 68: The authors did three scientific questions, but the second question (how we choose SIT product for different applications) is lacking in discussion.

Reply: Thank you for the advice. We deleted the second question in the revision.

Line 83: How are the microwave scattering and radiometric characteristics of MYI and FYI different?

Reply: Thank you for your advice. We revised the manuscript as suggested.

In the "Introduction" section, we included sentences about the physical background for microwave SITY classification as mentioned above. The sentences read as below (L55-L59):

"On one hand, brightness temperatures (Tbs) of MYI tend to be lower than that of FYI because of its low-loss, low-salinity properties (Vant et al., 1978; Weeks and Ackley, 1986). Such difference is generally larger at higher frequencies (i.e. smaller penetration depth), which reflects the distinguished physical properties of MYI and FYI at the sub-surface layer (Sinha and Shokr, 2015). On the other hand, due to the high volume scattering and low scattering loss, MYI has a relatively higher backscatter than FYI at the same frequency (Onstott, 1992)"

In the "Data" section (2.2 Sea ice type products), before introducing the individual sea ice types, more detailed description regarding the physical background of sea ice type classification from microwave observations was included (L126-L152). These paragraphs start with:

"FYI and MYI can be discriminated from microwave satellite observations based on their distinctive radiometric and scattering signatures …"

"Depending on the ambient conditions, sea ice at different stages of development undergoes different thermodynamic and dynamic processes, resulting in disparate microwave radiometric and scattering properties of different sea ice types …"

"In addition to Tbs at different channels, parameters of their combinations are also used in sea ice type discrimination …"

"Microwave scattering of sea ice is determined by the surface and volume scattering, which is influenced by factors …"

Line 77: Each product has a different grid size. It should be explained how it was dealt with in the comparative evaluation.

Reply: Thanks for the advice. We modified the computations in the revision to account for the factor of different grid size. SITY products in polar-stereographic projection were firstly re-gridded to the EASE2 grid before the computation of MYI extent. Details of such were presented in the new Section"3 Methods" (3.1 Estimation of MYI extent).

Line 78-143: Please describe in more detail how FYI and MYI are distinguished due to which characteristics in each SIT algorithm. For example, if a SIT product is produced based on PR and GR, an explanation is required for the differences between the values of PR and GR of ice types and why the differences occur.

Reply: Thank you for the advice. In the revision, the section of "Data" was modified as follows. The new section 2.1 was entitled with "Microwave remote sensing data", where the specification of the sensors and the satellite data used in all the SIT products were introduced. Definition (and equation) of PR and GR was introduced in this section. The new section 2.2 is "sea ice type products", where the grid resolution and details of the algorithm used in each SIT product were introduced.

Line 150: NSIDC-SIA was used as reference data. How accurate is NSIDC-SIA?

Reply: Thanks. In the revision, references regarding the evaluation of NSIDC-SIA were added and summarized in Section 2.3 "Sea ice age product" (L273-L281). The sentences read as below:

"The accuracy of NSIDC-SIA largely depends on the ice trajectories tracking technique and ice motion data. There are mainly two sources of error in NSIDC-SIA: the tracking errors related to the coarse resolution of microwave satellite data and those induced by ice motion data vacancy near the coast. The under-sampling of ice motion along with the scheme of oldest ice age assignment lead to an overall discontinuous sea ice age distribution and overestimation of old ice (Korosov et al., 2018). Besides, ice motion velocities from buoys are generally higher than those from satellite data (Schwegmann et al., 2011). Improper interpolation approach could lead to artificial divergence in ice motion when the buoy estimation differs significantly

from the satellite-based data. It could result in approximately 20% less MYI in the buoy-affected region according to a numerical experiment (Szanyi et al., 2016). Such impact is mainly found in the years 1983–2005 and has been largely mitigated by tuning the interpolation approach in the current version (Tschudi et al., 2020)"

Line 153: How and what information was retrieved from the SAR images for the SIT products evaluation should be described.

Reply: Thank you for the comment. In the revision, we added a new section regarding the theory of SAR interpretation. The theory and characteristics of sea ice classification in SAR images was introduced in the section "3.2 visual interpretation of SAR imagery". In addition, we interpreted all the entire SAR images, convert the sea ice classification results from ice types polygons to grided ice classification results, and eventually give the quantitative evaluation results.

Line 173: Is it the result of this study that different SIT distribution patterns were found in the regions selected by the authors?

Reply: It is a misleading sentence here. We will delete the sentence in the revision.

Line 185-186: Is 'divergent movements' the only cause of increase in the MYI extent?

Reply: "Advection from neighbouring regions" could also lead to increase in the MYI extent. The sentence will be modified as follows in the revision.

"However, it can temporarily or regionally increase due to ice divergence or advection from neighbouring regions (Kwok, Cunningham, et al. 1999)."

Line 263: The authors compared SIT daily products with the SAR images. It is necessary to discuss the comparison between the image captured at a specific time and the daily product.

Reply: Thank you for the advice. In the revision, the impact of such will be discussed in the comparison.

Line 263: The authors identified the distribution of MYI by visually analyzing the SAR image. It would be better if MYI could be determined by quantitatively analyzing backscattering or textures from the SAR images.

Reply: Thank you for the advice. In the revision, we will add a new section regarding the theory of SAR interpretation. The theory and characteristics of sea ice classification in SAR images will be introduced in the section "3.2 visual interpretation of SAR imagery". In addition, we will interpreted the entire SAR images, consulted with ice experts regarding the results, compared with the gridded ice classification results, and eventually gave the quantitative evaluation results (Kappa coefficient and overall accuracy for respective SAR image and SITY, SIA product).

Line 368-371: How are the input parameters affected by atmospheric factors and surface features? More discussion is needed.

Reply: Thank you for the advice. We further investigated the magnitude of cloud liquid water values (atmospheric factors) and its impact. It turns out that variation of the cloud liquid water path has little impact on GR. We therefore removed this sentence in the revised manuscript. Regarding the impact of surface features, more discussion was added. The sentences in the next few lines were (L590-L594) modified as below:

"However, this parameter can be impacted by surface features (e.g., snow properties) during the winter season (Rostosky et al., 2018; Ye et al., 2019; Comiso, 1983). In the beginning and ending stages of winter, the variability of GR_37v19v can be significant when air temperature exhibits warm-cold cycles, which trigger wet-dry cycles or melt-refreeze cycles of snow (Voss et al., 2003; Ye et al., 2016b; Ye et al., 2016a), or when wet/thick precipitation suddenly appears (Voss et al., 2003; Rostosky et al., 2018)"

Line 393: Explain clearly about the training dataset.

Reply: Thanks. To better explain the training dataset, we rephrased the sentences as below (L619-L622):

"Some algorithms use the thresholds derived from a training dataset that does not vary with time, region or satellite sensors, namely fixed thresholds, while others employ dynamic thresholds to account for the variability of training dataset. The former algorithms work relatively well under conditions similar to the training dataset, however it gives anomalous SITY distribution results in other conditions."

Technical comments

Line 157: SAR Wide B → SAR Wide Beam

Reply: Done

Line 199: What does (2000) mean?

Reply: It is a typo here. The reference and citation was corrected.

Line 402: Is [55] a reference number?

Reply: Yes. The citation was corrected.

Figure 4: Delete 'Jan' on the horizontal axis.

Reply: Done

Response to RC 3:

Dear Editors and Reviewers:
We would like to thank the reviewer for the detailed and helpful comments, suggestions, and careful checking. Comments are responded on point-by-point basis. The reviewer comments appear in black. The responses are in blue and the proposed changes to manuscript are in ***bold italics***.

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.
Reply: Thank you for your advice. The physical background of sea ice type classification from passive and active microwave remote sensing has be added in the revision. Such descriptions are presented in the following sections: 1) section "1 Introduction", where the basic theory of FYI/MYI discrimination from microwave observations is summarized; 2) section "2.2 sea ice type products", where the theory of

sea ice classification is introduced at the beginning of this section, before introducing all the SIT products; 3) section "3.2 visual interpretation of SAR image", which presents the physical background and sea ice scattering characteristics in SAR image.

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.

Reply: Thank you for the advice. To better illustrate the input satellite data and algorithms used in the products, we restructured the Section of Data. In the revised manuscript, Section "2.1 Microwave remote sensing data" describes the satellite data, whereas Section "2.2 Sea ice type products" provides details of the algorithms used in each ice type product.

As for the two products you mentioned, since this study focuses on the inter-comparison of sea ice type products that tells ice type classifications without information on the specific fraction. We therefore did not include NASA-Team and ECICE MYI concentration products. For clarifications, we modified the sentences in the section "Introduction" (P2L59-L78), which reads as below:

"There exist different algorithms which either provide a fractional MYI/FYI coverage or assignment of one or the other ice type (e.g. MYI and FYI) to a grid cell. The former, referred to as sea ice type concentration (SITC) algorithms, includes algorithms such as the NASA Team and ECICE algorithm …The latter, referred to as SITY algorithms, include …"

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.

Reply: Thank you for your advice. In the revision, we included more quantitative results and analysis in the manuscript as suggested. Regarding the two points you mentioned in the comments. 1) the NSIDC sea ice age data is used for overall comparison.

References regarding the evaluation of the sea ice age product were added in the revision. In addition, weekly MYI extent is calculated from sea ice type products for consistent comparison, where more quantitative comparison is included in the revised manuscript. 2) In the revision, we interpreted all the entire SAR images as suggested, compared with the gridded ice classification results, and eventually give the quantitative evaluation results (Kappa coefficient and overall accuracy for respective SAR image and SITY, SIA product).

For the manual interpretation, the theory and characteristics of sea ice classification in SAR images is introduced in the new section "3.2 visual interpretation of SAR imagery". In addition, the interpretation is conducted under the supervision and consult with sea ice experts who are experienced in SAR interpretation for more reliable results.

GC4: The discussion of the results is not well linked to the existing literature.
Reply: Thank you for the advice. In the revision, we included more recent literatures in the manuscript to be better linked to recent studies.

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.
Reply: Agree. We use "SITY" as the abbreviation for "sea ice type" in the revised manuscript.

- 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.
Reply: Thank you for the advice. We categorized the sea ice type products into three groups, which reads as below:
"..., including purely radiometer-based (C3S-SITY), scatteromter-based (KNMI-SITY and IFREMER-SITY) and combined ones (OSISAF-SITY and Zhang-SITY)."

- 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.

Reply: Agree. We added clarifications in Section "Introduction" in the following two parts:

1) P2L56-L57, *"… Among them, most algorithms focus on the discrimination of MYI and FYI. These algorithms identify SITY (i.e. the discrimination of MYI and FYI in this study) based on …"*

2) P3L90, *"This study aims to investigate differences among the SITY products and give comprehensive evaluations on the identification of MYI and FYI."*

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

- 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.
Reply: Agree. In the revision, we specified the temporal resolution of SITY and SIA product, e.g. P1L12 "… eight daily SITY product …", P1L15 "… a weekly sea ice age product (NSIDC-SIA) …", and only mentioned the "average" difference (P1L15).

- L14/15: Please also see my over-arching comment to the conclusions.
Reply: Thank you for the comment. We carefully checked all the numbers mentioned in the manuscript in the revision. For the exceptional large difference, we add statements in Section "4.1.2" (P13L382-L385) and refer to the discussions in Section "5 Discussion".

- 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).
Reply: Agree. Thank you for the comment. The sentence is rephrased as below:
"In the ASCAT period, KNMI-SITY tends to overestimate MYI (especially in early winter), whereas Zhang-SITY and IFREMER-SITY tend to underestimate MYI."

- 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?
Reply: Thank you for the comment. The retrieval of all products is not based on a classification approach. In the revised manuscript, we give more detailed description about the classification algorithms used in the respective SITY products. On the other hand, this study focuses on the comparison and evaluation of current SITY products. Further investigation of classification algorithms could be done in future however is out of the scope of this manuscript.

- 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.

Reply: Done.

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;

Reply: Agree. We add descriptive sentences/phrases at the beginning of this paragraph (P2L56-L57). The modified sentences are:

"…Among them, most algorithms focus on the discrimination of MYI and FYI. These algorithms identify SITY (i.e. the discrimination of MYI and FYI in this study) based on the distinct radiometric and scattering characteristics of different ice types."

ii) Write what the fundamental 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);

Reply: Thank you for the advice. We added sentences with brief description of the fundamental principles for microwave SITY classification in this section (More detailed description can be found in the section "2.2 Sea ice type products"). The new sentences in this section are as follows (L55-L59):

"On one hand, brightness temperatures (Tbs) of MYI tend to be lower than that of FYI because of its low-loss, low-salinity properties (Vant et al., 1978; Weeks and Ackley, 1986). Such difference is generally larger at higher frequencies (i.e. smaller penetration depth), which reflects the distinguished physical properties of MYI and FYI at the sub-surface layer (Sinha and Shokr, 2015). On the other hand, due to the high volume scattering and low scattering loss, MYI has a relatively higher backscatter than FYI at the same frequency (Onstott, 1992)"

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.

Reply: Agree. In the revised manuscript, we firstly summarize the approaches in terms of fractional and binary results, later mention the typical algorithms accordingly. The modified sentences are (L59-L66):

"There exist different algorithms which either provide a fractional MYI/FYI coverage or assignment of one or the other ice type (e.g. MYI and FYI) to a grid cell. The former referred to as sea ice type concentration (SITC) algorithms, includes

algorithms such as the NASA Team algorithm and ECICE algorithm, …. The latter referred to as SITY algorithms, includes …"

iv) Move information about evaluation results obtained by others so far into the next paragraph (see Lines 47/48: "By comparing ... Kwok 2004)."
Reply: Agree. We move this information to the next paragraph in the revised manuscript.

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).
Reply: Agree. In the revised manuscript, we summarize the approaches in terms of fractional and binary results, respectively. The modified sentences are (L59-L66):
"There exist different algorithms which either provide a fractional MYI/FYI coverage or assignment of one or the other ice type (e.g. MYI and FYI) to a grid cell. The former, referred to as sea ice type concentration (SITC) algorithms, includes algorithms such as the NASA Team and ECICE algorithm, …. The latter, referred to as SITY algorithms, include …"

- 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.
Reply: Thank you for the advice. In the revised manuscript, we specify that the sea ice type concentration algorithms include both kinds of approaches. We therefore keep all the references you mentioned in the comment.
      Regarding the MYI data sets created by Ron Kwok, it is not included in the inter-comparison since our manuscript focus on the inter-comparison and evaluation on SITY products, which refer to the binary products (whereas Kwok's dataset is a dataset of fractional MYI/FYI coverage). As mentioned in previous replies, we also added sentences in this paragraph on clarifying the focus of this manuscript.

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 distribution of, e.g. 50%, would have which implications is not yet given in a convincing way.

Reply: Thanks for the advice. Implications of large errors can be found in previous studies however have not been quantified to our knowledge. This is also one of the reasons that we need a more comprehensive evaluation of these products.

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)

Reply: Agree. "limited" is deleted in the revised manuscript.

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

Reply: Thanks for the comment. The sentence is rephrased to (L83-L85)

"While the ice chart is used as "ground truth" in some validation (Aaboe et al., 2016), some areas of MYI in the ice charts correspond to areas with MYI concentration of approximately 50% or greater (Lindell and Long, 2016a), indicating the overestimation of MYI in ice charts."

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?

Reply: We mean the products that are currently operationally produced and provided to the users. To avoid unnecessarily misunderstanding, we delete this word in the revision.

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.

Reply: Thank you for the advice. We have revised the manuscript as suggested.

In the "Introduction" section, we include sentences about the physical background for microwave SITY classification as mentioned above. The sentences included additionally are as follows (L55-L59):

"On one hand, brightness temperatures (Tbs) of MYI tend to be lower than that of FYI because of its low-loss, low-salinity properties (Vant et al., 1978; Weeks and Ackley, 1986). Such difference is generally larger at higher frequencies (i.e. smaller penetration depth), which reflects the distinguished physical properties of MYI and FYI at the sub-surface layer (Sinha and Shokr, 2015). On the other hand, due to the high volume scattering and low scattering loss, MYI has a relatively higher backscatter than FYI at the same frequency (Onstott, 1992)."

In the "Data" section (2.2 sea ice type products), before introducing the individual sea ice types, we included several paragraphs to describe the physical background for ice type classification from microwave observations (L126-L152). These paragraphs start with:

"FYI and MYI can be discriminated from microwave satellite observations based on their distinctive radiometric and scattering signatures …"

"Depending on the ambient conditions, sea ice at different stages of development undergoes different thermodynamic and dynamic processes, resulting in disparate microwave radiometric and scattering properties of different sea ice types …"

"In addition to Tbs at different channels, parameters of their combinations are also used in sea ice type discrimination …"

"Microwave scattering of sea ice is determined by the surface and volume scattering, which is influenced by factors …"

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.
Reply: Thank you for the advice. In the revised manuscript, the "Data" section includes four subsections: "2.1 Microwave remote sensing data", "2.2 Sea ice type products", "2.3 Sea ice age product" and "2.3 Other data". In Section 2.1, specification of the microwave sensors is introduced with the native resolution. Table A1 is included to show detailed info. In Section 2.2, the grid resolution of each sea ice type product is specified.

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 approach do discriminate open ocean and sea ice. In short: You need to rephrase these statements.

Reply: Thank you for the suggestion. We meant to say that the Tb measurements are corrected for the atmospheric influence by using auxiliary data and radiative transfer model. One of the auxiliary data is sea ice concentration, which is used to estimate the surface emissivity in the radiative transfer model. Note that the authors do not correct the Tb measurements for sea ice concentration. We rephrased these statement in the revision. In addition, description of each SITY product is organized in the following way: 1) summary of the product, including radiometer/scatterometer/combined-based, the resolution, availability, versions and so on. 2) three main procedures: pre-processing, classification approach, and post-processing. Description of C3S-SITY can be found in section 2.2.1 C3S-SITY.

Line 96: I suggest to remove the "further" and to also provide an equation of how the gradient ratio is computed.
Reply: "further" is deleted. In the revised manuscript, there are four subsections in the "Data" section. The equation of gradient ratio is included in the new section "2.1 microwave remote sensing".

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?
Reply: Thanks for the comment. The sentence is misleading. It was modified to the sentences as below (L184-L187):
"The daily updated probability density functions (PDFs) of the collected training data is dynamic in time and captures the seasonal and interannual variabilities. The pre-defined areas over which the data are collected are the climatological MYI and FYI regions, which are north of Greenland and Canada with longitude between 30°W and 120°W for MYI, and the Kara Sea, Baffin Bay, Laptev Sea and the Bay of Bothnia for FYI."

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 ...?
Reply: Thanks for the comment. These sentences were deleted here. In the revised manuscript, two versions of C3S-SITY are introduced in the first paragraph (L171-L174).

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;

Reply: Thank you for the advice. We made substantial modifications on the introduction of Data and Products. In the revised manuscript, there are 4 subsections in the new "Data" section. In the new section "2.1 microwave remote sensing", we introduce all the microwave remote sensing data used in the various SITY products. In the new section "2.2 Sea ice type products", we then describe each product in details. Introduction of each product will include the following information: 1) grid resolution, projection, the satellite data used in different periods; 2) the algorithm used in the product, where we could include more details.

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.);

Reply: Thanks for the advice. Substantial modifications have been made in the revision. More details regarding the product and satellite data have been included. As mentioned in the previous reply, details of the microwave remote sensing data were given in the new section 2.1, where it mentions that the incidence angle of ASCAT ranges from 25° to 65°. Backscatter from ASCAT is normalized to a reference incidence angle of 40° and later used as an input for the classification.

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

Reply: The native resolution of ASCAT is around 25km x 34 km, this information is included in the new Table A1.

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

Reply: Thanks for the comment. It is an inappropriate expression here. Description of OSISAF-SITY has been written and can be found in section 2.2.2. The relevant sentence is modified as below (L214-L215):

"Ice types and their probabilities are derived using classifiers based on the respective observational parameters (GR_19v37v and σ_0), where swath data of different sensors are used."

- 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.

Reply: Thank you for the suggestion. The description is misleading. In the new section 2.1, satellite data is introduced in a chronological order. Which data is used in the respective SITY product can be found in section 2.2.

- 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?

Reply: Thank you for the advice. In the new section 2.1 "Microwave remote sensing data", the satellite data is introduced in a chronological order meanwhile with same level of details. The microwave radiometer data is introduced as two series of dataset: one includes SSM/I and SSMIS, the other are AMSR-E and AMSR2. Spatial resolution of the sea ice type products is mentioned in the new section 2.2.

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

Reply: In the revised manuscript, details of ASCAT were presented in the new section 2.1 and Table A1. ASCAT provides only vertically polarized backscatter. The grid spacing of ASCAT data is 12.5 km.

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

Reply: The "given weights" are dependent on the distance between centers of grid and footprint. These sentences were modified as below (L215-L216):

"The probabilities are then gridded based on the distance between each footprint and the polar stereographic grid."

- 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 ...

Reply: Thank you for the advice. We include such expression in the revised manuscript. For instance, on L190-L192:

"Note that C3S-SITY defines an ambiguous ice type class (referred to as Amb) in addition to the pure MYI and FYI classes. The Amb class represents sea ice with a low classification probability. It may be both pure MYI, FYI or a mixture of FYI and MYI (Aaboe et al., 2021c)"

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.

Reply: Thank you for the advice. We added a new subsection "2.1 Microwave remote sensing" in the revision, where the sensors and specifications are introduced.

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.

Reply: Thank you for the advice. In the revision, we modified the sentences to underline more clearly the meaning of each acronym, and specified the two products we include in the evaluation is KNMI-Q and KNMI-A. The modified sentences are as below (L221-L225):

"KNMI-SITY is a series of purely scatterometer-based products with grid spacing of 12.5 km in a polar stereographic projection. The scatterometer data used includes ERS, QSCAT, OSCAT and ASCAT, which results in four respective SITY products, referred to as KNMI-E, KNMI-Q, KNMI-O and KNMI-A respectively, available during the periods of 1992–2001, 1999–2009, 2010–2013 and 2007–2016. In this study, KNMI-Q and KNMI-A are included in the comparison and evaluation considering the comparable input data as other products."

- 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.

Reply: Thank you for the advice. In the revision, specifications of the sensors are introduced in the new section 2.1 with same level of details, whereas the grid resolution and algorithm are introduced in the new section 2.2. For SITY product with several times of switching input data, we refer to the Table 1/Figure 1 when describing details such as periods.

- 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?

Reply: Yes, it applies to all the KNMI sea ice type products. The modified sentences are as follows (L229-L234):

"In the stage of classification. KNMI-SITY uses a refined Bayesian algorithm for ice/water discrimination, based on the probabilistic distances to the Geophysical Model Functions of ocean wind and sea ice. The sea ice pixels are then classified into FYI, second-year ice (SYI) and older MYI using VV polarized backscatter with two thresholds, which are determined from the data of March of each year (Belmonte Rivas et al., 2018)."

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

Reply: Thank you for the advice. We deleted this sentence to avoid misunderstanding.

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.

Reply: Thank you for the advice. These sentences were modified to (L236-L238)

"IFREMER-SITY is another series of purely scatterometer-based products, with grid spacing of 12.5 km in a polar stereographic projection. There are two SITY products in IFREMER-SITY, which use QSCAT and ASCAT data for the respective years of 1999–2009 and 2010–2015, referred to as IFREMER-Q and IFREMER-A, respectively."

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

Reply: Thank you for the comment. In the revised manuscript, we modified the sentence as below to avoid providing different details of the product (L239-L240).

"In the first stage, the backscatter coefficients at different incidence angles (e.g., ASCAT backscatter) are normalized to the value at a constant incidence angle of 40° to account for the influence of varying incidence angles."

- 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?

Reply: Thank you for the question. To answer the questions you raise here, we modified the sentences as below (L240-L242):

"In the core classification, a set of day-to-day-varying thresholds are then used for the discrimination between MYI and FYI. These thresholds are derived from the backscatter data of several winters and are found to be inter-annually consistent (Girard-Ardhuin, 2016)"

- While we learn here that the product is gridded to a polarstereographic grid, there is no information about the grid in the previous paragraphs.
Reply: Thanks. In the revision, The projection and grid spacing information of all the SIY products. Were introduced in the first paragraph of each SITY description.

Lines 132-137:
- "employs adaptive" --> "employs an adaptive"
Reply: Done

- "based on the thought of clustering" could possibly be re-phrased. What kind of clustering approach is used? K-means?
Reply: Thank you for pointing it out. It has been modified to "based on K-means clustering" in the revision.

- 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?
Reply: It uses horizontally polarized Tbs at 37 GHz and the backscatter from scatterometer. It is included in the beginning of section 2.2 about how the gradient ratio is calculated.

- 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?
Reply: As mentioned in the afore-mentioned replies, information regarding the microwave satellite data was introduced in the new section 2.1 in the revision.

- 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?
Reply: To answer the questions here, we added the following sentences in the revision (L249-L250).

"Scatterometer data from QSCAT and ASCAT is used successively in Zhang-SITY, with QSCAT data until November 23, 2009."

The ASCAT data is normalized to backscatters at constant incidence angle. In addition, it is found that the algorithm in Zhang-SITY is not highly influenced by the use of different scatterometer data. Such details have been included in the paragraph of describing the algorithm used in Zhang-SITY (L256-L258).

"It is an unsupervised classification approach thus does not require the selection of training dataset. In addition, the results from different sensors are generally consistent thus no further processing is conducted for the satellite data (Zhang et al., 2019)."

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.

Reply: Agree. We deleted this paragraph in the revised manuscript.

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).

Reply: Thanks. The suggested reference has been included. The sentences have been modified in the revision. Please find it on L266-L268.

"It is derived by tracking trajectories of virtual Lagrangian ice parcels of each grid cell. The ice motion data used in the tracking process is based on passive and active microwave observations as well as auxiliary data such as drifting buoys (Fowler et al., 2004; Maslanik et al., 2011; Tschudi et al., 2020)."

- 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.

Reply: We agree that the sea ice age product tends to overestimate the presence of old ice if we regard the pixels fully covered by the ice at "the age" (given by the product). Since SYI is regarded as MYI in this manuscript, the overestimation applies to the discrimination between FYI and SYI also. We removed the citation and rephrased the description here.

Line 157: "Images are" --> "All five SAR images are ..."
Reply: Done.

- Was any filtering (speckle?) applied?
Reply: No. And it was a typo for the pixel size of Sentinel-1, which should be "40 m" not "160 m".

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.

Reply: In the revision, we have included the acquisition date of the SAR images. The modified figure is as below:
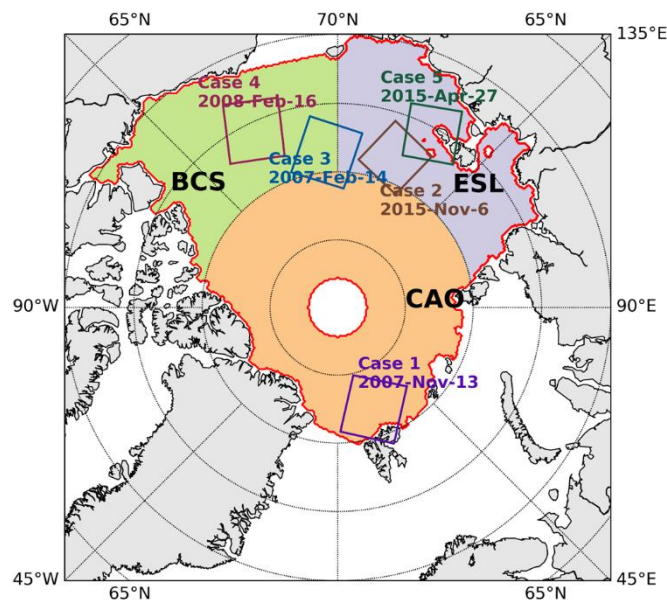
Lines 161-165:
- "For better interpretation of SAR images" --> This motivation needs to be explained better. It is not at all clear why, for the interpretation of the few SAR images used, these

two additional data sets are required. What is the problem with the SAR images that such data are needed?

Reply: The interpretation of the SAR images does not need these data. We will therefore delete these sentences in the revision. These sentences were modified as (L289-L292) *"Auxiliary data from atmospheric reanalysis is used in addition to the SAR images in the validation. The reanalysis data includes 2 m air temperature and 10 m wind from the ERA5 hourly dataset, produced using 4D-Var data assimilation and model forecasts in CY41R2 of the European Centre for Medium-Range Weather Forecasts integrated Forecast System (ECMWFs) (Hersbach et al., 2018)."*

- 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.

Reply: We admit that it is not appropriate to use such heterogeneous data in the manuscript. The data is also not required for the SAR interpretation. We therefore deleted these sentences.

- 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?

Reply: As mentioned in the previous replies, these sentences were deleted in the revision.

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.

Reply: Agree. Naming of the regions was modified to "the central Arctic Ocean (CAO), the East Siberian and Laptev Seas (ESL), along with the Beaufort and Chukchi Seas (BCS)" in the revised manuscript. Abbreviations were updated accordingly.

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."

Reply: Agree. In the revision, we added a new section regarding the estimation of MYI extent, Section 3.1 "Estimation of MYI extent". Details can be found on L294-302, the sentences start with:

"For the inter-comparison, the Arctic MYI extent is calculated from the respective SITY and SIA products …"

- 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.

Reply: Thank you for the comment. Here are the replies to the two questions: 1) No, we did not use a common land mask because of the reason you mentioned. 2) No, it was not accounted for in the original manuscript. We therefore corrected the computations in the revision. SITY products in polar-stereographic projection are firstly re-gridded to the EASE grid before the computation of MYI extent. Details of such have been presented in the new Section"3 Methods" (3.1 Estimation of MYI extent).

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.

Reply: Thank you for the advice. In the revised manuscript, we calculated the weekly averaged MYI extent from the SITY products with the NSIDC-SIA product. All the relevant numbers have been modified accordingly. Plots of the weekly averaged MYI extent was presented in Fig. 4. The modified figure is as below:
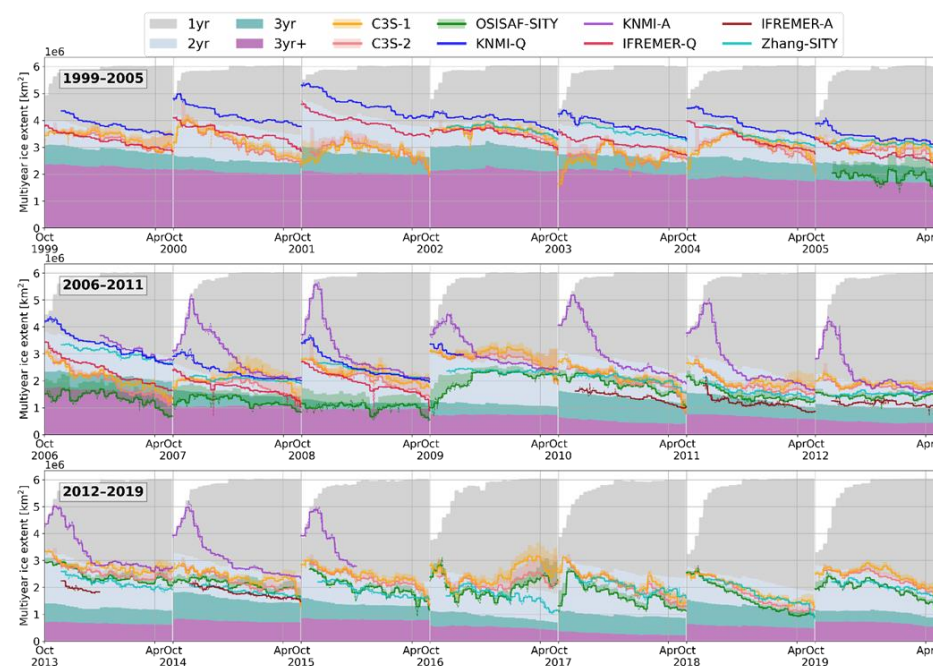
Line 184: "decreasing trend" --> What you possibly mean is a decrease or an increase in the MYI extent (over time) or a positive or negative trend of the MYI extent (over time). A decreasing (or increasing) trend, in contrast, is a trend that changes its value with time or, in other words, if this was 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.

Reply: Thank you for the advice. Such expressions have been corrected throughout the manuscript as suggested. This sentence has been modified as follows (L349-L350).

"…, the overall Arctic MYI extent cannot increase over the winter – it can only decrease through ice advection out of the Arctic."

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

Reply: Thank you for the advice. This sentence has been modified to (L350-L351)

"it can temporarily or regionally increase due to ice divergence or advection from neighbouring regions (Kwok et al., 1999)."

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."

Reply: Thanks. It has been modified as suggested.

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.

Reply: Thank you for the comment. In the revision, we check all the numbers presented in the manuscript, and confirm that the largest difference was correctly calculated. Explanations have been added accordingly (L387-L390).

"Between any two SITY products, the difference in weekly MYI extent is up to 4.5 ×10^6 km^2, which occurs between OSISAF-SITY and KNMI-A in late October 2008. Considering the size of the study region (about 6.5×10^6 km^2), such discrepancy is significant. This is caused by the relatively low MYI extent from OSISAF-SITY and the exceptional high value from KNMI-A in late October, the reason for which will be discussed in Section 5."

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.

Reply: Thank you for the comment. These are averages of the extent estimated from the "ambiguous" pixels. As explained in the first paragraph of the Section "Temporal reanalysis", "Amb class in C3S-SITY and OSISAF-SITY could be regarded as MYI and FYI thus the MYI extent is calculated under both circumstances". Large extent for the "ambiguous" class means large number of pixels with high uncertainties on ice type discrimination, which have atypical microwave signatures of MYI/FYI and usually lead to large discrepancies between the SITY products. In the revision, we included a separate paragraph in the section "4.1.1 Weekly MYI extent variation" (L375-L380), which reads as below:

"For the SITY products with the Amb class, the average extent of this class is 0.21 ×10^6 km^2, 0.26 ×10^6 km^2 and 0.26 ×10^6 km^2, respectively, for C3S-1, C3S-2 and OSISAF-SITY. As described in Section 2.2, these Amb pixels have atypical microwave signatures of MYI/FYI thus high uncertainties on ice type discrimination. Compared with the average Arctic MYI extent difference against NSIDC-SIA (0.42×10^6 km^2, 0.45×10^6 km^2, 0.79 ×10^6 km^2 for C3S-1, C3S-2 and OSISAF-SITY, respectively), the contribution of these pixels to the comparison is overall considerable. In addition, it could be large under situations that trigger the atypical microwave signatures, which will be further discussed in section 4.1.2."

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.

Reply: Thank you for the advice. In the section of "Introduction" and "Data", we have included descriptions about the physical background for microwave SITY classification and theory behind the classification algorithms.

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

Reply: We wanted to say that the inter-annual evolution of MYI extent from C3S- and OSISAF-SITY differs most with NSIDC-SIA. The sentence was modified to (L403-L405)

"For the inter-annual evolution of MYI extent, C3S-SITY and OSISAF-SITY differ most from other SITY products. The latter exhibit mild negative trend during 2000–2007 and rapid negative trend from 2007 to 2013, while the former show larger inter-annual variabilities."

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.

Reply: Thank you for the advice. In the revised manuscript, we added details as you mentioned and give possible explanations for the inter-comparison of SIA and SITY MYI extent in the first paragraph of this section (L414-L424).

"On one hand, MYI is extensively exported through the Fram Strait (Kuang et al., 2022) and, by small fractions, into the Barents Sea and through the Nares Strait following the Transpolar Drift Stream. In the ESL region, the MYI extent even decreases to zero in some winters (e.g. 2007–2009, 2012–2013), which is in line with the record low Arctic minimum sea ice extent in the previous Septembers. On the other hand, MYI is advected towards south along the Canadian Arctic Archipelago driven by the Beaufort Gyre. In the BCS region, large quantities of MYI are pushed out of this region following the anticyclonic current, the Beaufort Gyre, meanwhile replaced by the MYI from the CAO region. This eventually leads to the nearly constant or increasing MYI extent in the BCS region. In the ESL and BCS regions, it is found that the NSIDC-SIA MYI extent is usually considerably larger than the MYI extent from the SITY products, whereas it is less pronounced in the CAO region. This indicates that the mixture of MYI and FYI, which leads to the "overestimated" NSIDC-SIA MYI extent, occurs more frequently in the ESL and BCS regions than the CAO region, which could be explained by the more dynamic ice characteristics in the two regions."

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 mentioned in your description.

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

Reply: Thank you for the advice. In the revision, these sentences have been modified to provide more clear explanations (L414-L424):

"On one hand, MYI is extensively exported through the Fram Strait (Kuang et al., 2022) and, by small fractions, into the Barents Sea and through the Nares Strait following the Transpolar Drift Stream. In the ESL region, the MYI extent even decreases to zero in some winters (e.g. 2007–2009, 2012–2013), which is in line with the record low Arctic minimum sea ice extent in the previous Septembers. On the other hand, MYI is advected towards south along the Canadian Arctic Archipelago driven by the Beaufort Gyre. In the BCS region, large quantities of MYI are pushed out of this region following the anticyclonic current, the Beaufort Gyre, meanwhile replaced by the MYI from the CAO region. This eventually leads to the nearly constant or increasing MYI extent in the BCS region. In the ESL and BCS regions, it is found that the NSIDC-SIA MYI extent is usually considerably larger than the MYI extent from the SITY products, whereas it is less pronounced in the CAO region. This indicates that the mixture of MYI and FYI, which leads to the "overestimated" NSIDC-SIA MYI extent, occurs more frequently in the ESL and BCS regions than the CAO region, which could be explained by the more dynamic ice characteristics in the two regions."

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.

Reply: Thank you for the advice. We would like to refer to the evolution from month to month. The sentence has been modified to (L425-L426):

"most SITY products show similar intra-seasonal variation in the CAO region, while exhibiting disparate intra-seasonal evolutions in the BCS and ESL regions (especially in early and late winter)"

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.

Reply: Thank you for the advice. In fact, this issue can be found in different winter months, e.g. Figs. 7, 8 and A1. To illustrate it more clearly, we revised the sentence accordingly (L441-L443):

"In Fig. 7 a and b (along with Fig. A1 a–d, f–i in Appendix and Fig. 8 a–b, h–i), the discontinuous FYI delineation in the inner part of MYI pack is well demonstrated, which occurs in all winter months and could partly explain the MYI extent fluctuations in C3S-SITY".

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?

Reply: Thank you for the advice. The expression was not correct here. We were about to use "exceptional" to describe the MYI extent from KNMI-A not KNMI-Q. The sentence was therefore modified to (L448-L450):

"KNMI-Q has the second largest MYI coverage among the seven SITY products, with a slightly more finger-like structure of MYI extending through the Chukchi Sea into the ESL region."

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.

Reply: Thank you for the advice. It is true that MYI is underestimated in all the SITY products compared with NSIDC-SIA. Including Amb pixels as MYI could lead to larger MYI extent, however not necessarily better agreement on the spatial pattern. We therefore did not include additional descriptions here.

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).

Reply: Thank you for the advice. In the revised manuscript, we added a new section of "Methods". Information such as co-locating data and calculation of MYI extent was

placed in the "Methods" Section, more specifically in section "3.1 Estimation of MYI extent".

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

Reply: Thank you for the advice. In the revision, we made the boxes with same size in each SAR image. The physical background and principles of visual interpretation was presented in the new section "3.2 Visual interpretation of SAR imagery". Enlarged examples were given Fig. 3 as well as Figs. 9-13 to show more details of the SAR image thus more intuitive for the comparison. The boxes were selected from the locations where there are typical backscattering characteristics and could show the agreement/discrepancies among the various SITY products.

- 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.

Reply: Thanks. In the revision, we added a new section "3.2 Visual interpretation of SAR imagery" as suggested, where more details about the characteristics and interpretation principles were presented. In addition, we interpreted all the entire SAR images as suggested, compared with the gridded ice classification results, and eventually give the quantitative evaluation results (Kappa coefficient and overall accuracy for respective SAR image and SITY, SIA product).

- 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.

Reply: Thanks. In the revision, we added the polarization information in the respective figures. In addition, in the new section "3.2 Visual interpretation of SAR imagery", we included sentences as below (L332-L334):

"Images at HV polarization are prioritized for the visual interpretation if provided, since the cross-polarized backscattering signals have been shown to increase the separability between MYI and FYI (Gray et al., 1982; Onstott et al., 1979)"

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.

Reply: Thank you for the advice. In the revision, we added a section of "Method" in the revised manuscript, where the theory of interpreting SAR images is introduced. In addition, we made visual interpretation quantitatively based on the theory and discussed the quantitative comparison in the section of "4.3 Validation based on SAR".

As for the NSIDC-SIA product, it is true that the NSIDC-SIA MYI extent could be biased. However, NSIDC-SIA is used as a reference dataset for large-scale comparison. Discrepancies between the SITY products and NSIDC-SIA does not mean "bad". This requires further "validation" such as that based on SAR images. To better illustrate this logic, we modified the sentences at the beginning of the "Results" section (L338).

"This section starts with a temporal and spatial comparison of the SITY products, with NSIDC-SIA as a reference dataset. It then proceeds with validation against SAR images."

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

Reply: Done. Month and year of the scene is added to the first sentence of this section (L483), which reads as below:

"In Case 1, a typical scene of early winter (November 2017) in MIZ is shown in Fig. 9."

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.

Reply: Thank you for the comment. We totally understand the concern. However, if we do not comment on each product here, it would be difficult to summarize later. We therefore keep the comment as it is in the revised manuscript. In addition, we included a new section after presenting all the five cases, Section "4.3.4 Performances of sea ice type and age products".

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.

Reply: Thank you for the advice. Based on the interpretation results from SAR and the Kappa coefficient/OA, NSIDC-SIA generally overestimate MYI in this case although the general distribution is similar. The sentence was therefore modified to (L489-L490) *"NSIDC-SIA overestimates MYI generally thus yields a median Kappa coefficient and OA (0.56 and 0.73, respectively)."*

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.

Reply: Thank you for the advice. We have added a section of "3.2 Visual interpretation of SAR imagery" in the revised manuscript. In addition, quantitative interpretation results in terms of statistics such Kappa coefficient and overall accuracy were added, which can be found in the text and Table 4.

- 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.

Reply: Thank you for the advice. Marks of the azimuth and range direction were added to the SAR images, shown in Figs. 9-13.

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?

Reply: Thank you for the advice. MYI is slightly underestimated in NSIDC-SIA in this case. This is nearly negligible considering the temporal resolution difference. We therefore modify the sentence to (L500-L502)
"MYI is slightly underestimated in NSIDC-SIA, with the Kappa coefficient of 0.57 and OA of 0.80. Yet such difference is nearly negligible considering their different temporal resolutions and the mobility features of sea ice."

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.

Reply: Thank you for the advice. The FYI and MYI scattering characteristics in HV image are more distinct than that in HH image. We therefore presented the HV image instead of HH image in the manuscript. In fact, both HH and HV image were used for the visual interpretation. To illustrate the procedure better, we added a section of "3.2 Visual interpretation of SAR imagery" in the revised manuscript. The physical background and how the zoomed boxes are selected were explained accordingly in the new section. For the Sentinel-1 image, we included the HH image as well as the HV image in the corresponding cases.

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.

Reply: Thank you for the comment. We agree that this case actually show typical case of landfast ice (FYI) mixed with MYI. In the original manuscript, our interpretation about areas A and B is consistent with you (A as FYI and B as MYI). Regarding Area C, we agree that it should be landfast FYI instead of MYI. Interpretation results, the statistics and the texts were modified accordingly in the revision. The modified texts can be found in L542-L549. The modified figure (Fig.13) is shown as below.

"The MYI distribution pattern of KNMI-A resembles the SAR image except for a slight overestimation of MYI in the northern part of the image (area A) and nearly the island, which may be caused by ice deformation. The Kappa coefficient and OA is the largest for KNMI-A in this case. IFREMER-A and Zhang-SITY both completely ignore the MYI pack, and this error lasts for the whole winter (maps not shown). C3S-SITY (C3S-1 and C3S-2) and OSISAF-SITY manage to identify FYI in area A, and sporadically capture an elongated MYI feature northeast of the image (partly classified as Amb). However, they underestimate MYI in area B and overestimate MYI in the southern part (areas C and D), which leads to a near-zero level Kappa coefficient. NSIDC-SIA clearly captures the elongated MYI feature in this case though has slight underestimation of MYI in area B."
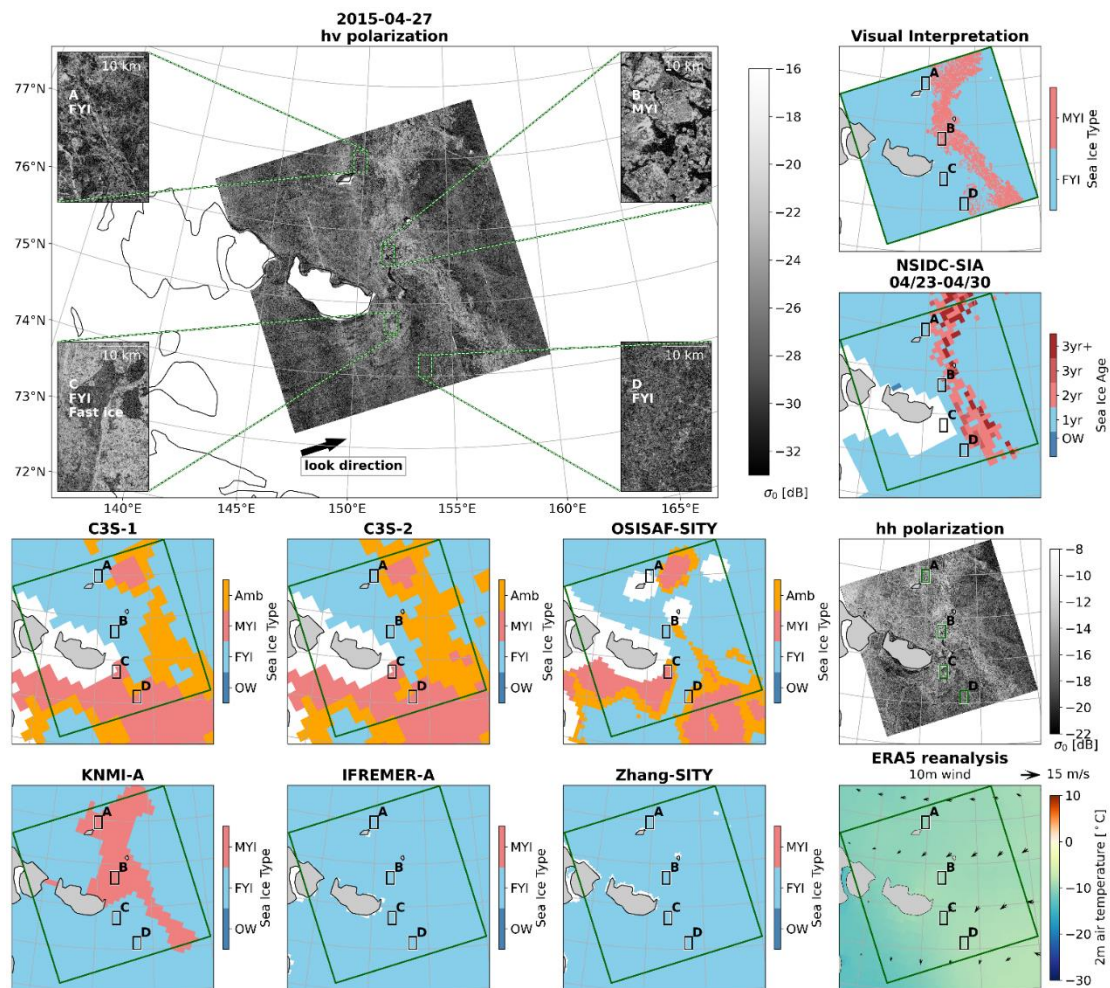
Line 312: Is there any reason why you put mid-winter after late-winter?

Reply: Thank you for the comment. In the revision, we presented the cases in the chronological order (early-, mid-, and 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.

Reply: Thank you for the advice. We agree with you on the comparison. To present the difference more clearly, we included the visual interpreted results in the same frame/subfigure as other SITY and SIA products. This sentence was modified to (L513-L516):

"… the MYI pack in this area, and this regional scale misclassification of MYI holds through the whole winter (maps not shown). Compared to the SAR image, the SITY

distribution in NSIDC-SIA has a distinct pattern, with overestimation of MYI in the northwest part of the image (area A) meanwhile underestimation in the northern part (east of area A) …"

Line 323: Please look at my comment to a similar statement made by you further above.
Reply: Thank you for the advice. We modified the sentence to (L515-L517)
"As mentioned previously, such discrepancies are mainly attributed to the mobility features of sea ice and the different temporal resolutions between NSDIC and the SAR image."

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.
Reply: Thank you for the advice. We agree that the NSIDC-SIA product is used for consistency check and inter-comparison. It is not proper to regard it as "cross-validation". In the revised manuscript, we modified such expressions meanwhile used SAR images in a quantitative way. We also added a section of "3.2 Visual interpretation of SAR imagery", where the theory for the visual interpretation is described.

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.
Reply: Thank you the advice. The sentence has been modified to (L588)
"The efficacy of input parameters depends on their capability to separate and physical properties of the sea ice types in question."

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.
Reply: Thank you for the advice. We further investigated the magnitude of cloud liquid water values and its impact. It turns out that variation of the cloud liquid water path has little impact on GR. We therefore removed this sentence in the revised manuscript.

Line 368: I am surprised 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.
Reply: Thank you for the advice. We confirm that one algorithm (i.e. OSISAF-SITY) uses this ratio the other way around. In the revised manuscript, we gave detailed description of each algorithm when introducing each SITY product (Sections 2.2.1-2.2.4).

Line 371: "ice layering" is one component of the snow properties and should not be mentioned as if it is a different thing.
Reply: Thank you for the advice. We deleted "ice layering" in the revised manuscript.

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 surprising 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.

Reply: Thank you for the advice. The sentence has been modified in the revised manuscript (see the end of this reply).

Regarding A), we agree that these are likely melt-refreeze cycles caused by increased solar radiation. It is also true that air temperatures during the whole cycle do not necessarily fluctuates freezing point. What we are trying to say is that, air temperatures that trigger these cycles are around freezing point. Evidence for this can be found in Voss et al., 2003 and Ye et al., 2016b. We therefore rephrased the sentence and included the references accordingly.

Regarding B), although solar radiation is not yet absent in the Arctic in October, we agree that the microwave signature changes are different from those in April. They could be caused by warm air intrusions or precipitation falling as wet snow, which lead to changes on snow wetness. We did not present evidence for the impact on the microwave signatures, since it can be found in previous studies such as Ye et al., 2016a. In the revised manuscript, we therefore rephrased the sentence and included the reference accordingly.

Regarding to C), we agree that thick snow load could lead to misidentification of MYI. This information was added in the revised manuscript.

The modified sentence is shown as below (L591-L593):

"In the beginning and ending stages of winter, the variability of GR_37v19v can be significant when air temperature exhibits warm-cold cycles, which trigger wet-dry cycles or melt-refreeze cycles of snow (Voss et al., 2003; Ye et al., 2016b; Ye et al.,

2016a), or when wet/thick precipitation suddenly appears (Voss et al., 2003; Rostosky et al., 2018)."

References:

Ye, Y., Heygster, G., and Shokr, M. (2016a). Improving multiyear ice concentration estimates with air temperatures. IEEE Transactions on Geoscience and Remote Sensing, 54(5), 2602 - 2614.

Ye, Y., Shokr, M., Heygster, G., and Spreen, G. (2016b). Improving multiyear sea ice concentration estimates with sea ice drift. Remote Sensing, 8(5), 397.

Voss, S., Heygster, G., and Ezraty, R. (2003). Improving sea ice type discrimination by the simultaneous use of SSM/I and scatterometer data. Polar Research, 22(1), 35-42.

Rostosky, P., Spreen, G., Farrell, S. L., Frost, T., Heygster, G., and Melsheimer, C. (2018). Snow depth retrieval on Arctic sea ice from passive microwave radiometers—Improvements and extensions to multiyear ice using lower frequencies. Journal of Geophysical Research: Oceans, 123(10), 7120-7138.

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.

Reply: Thank you for the advice. We would like to say that mitigation of the misclassification is caused by the different processing in C3S-2, which includes the temperature-based correction for improved ice type discrimination and the improved correction of Tb for the atmospheric influence. Details of the difference between C3S-1 and C3S-2 has been added in section 2.2.1. In the revised manuscript, the sentence was modified to (L595-L597):

"Such misclassification in C3S-1 is mitigated in C3S-2 due to the upgraded processing in C3S-2, which includes the temperature-based correction in the post-processing and the use of reanalysis data from ERA-5 instead of ERA-Interim in the atmospheric correction for Tb (see section 2.2)."

Line 377:

- "backscatter (sigma^o)" --> either "backscatter coefficient" or "sigma nought"

Reply: Done. "backscatter" was modified to "backscatter coefficient" in the revision.

- "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.

Reply: Thanks. The sentence was modified to (L598)

"which is commonly used in ice type discrimination due to the disparate scattering features of MYI and 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?

Reply: Thanks. The sentences have been modified as below. Explanations were added. (L600-L606)

"In comparison, the backscatter of MYI and FYI is more disparate at Ku-band than C-band (Rivas et al., 2018; Bi et al., 2020). Products using Ku-band backscatter generally perform better on identifying MYI, e.g. KNMI-Q, IFREMER-Q, and Zhang-SITY before 2009. This could be due to the fact that Ku-band scatterometer is more sensitive to the crystal structure of MYI since its wavelength (about 1.7 cm ~ 2.5 cm) is more consistent with the characteristic dimension of air bubbles in MYI (Ezraty and Cavanie, 1999). On the other hand, the dominant effect of surface scattering and the higher dependence on incidence angle makes C-band backscatter more suitable to distinguish the ice types with disparate surface roughness features, e.g. Cases 3 and 4 in Fig. 11, Fig. 12."

References:

Bi, H., Liang, Y., Wang, Y., Liang, X., Zhang, Z., Du, T., ... and Huang, H. (2020). Arctic multiyear sea ice variability observed from satellites: a review. Journal of Oceanology and Limnology, 38(4), 962-984.
Rivas, M., Otosaka, I., Stoffelen, A., & Verhoef, A. (2018). A scatterometer record of sea ice extents and backscatter: 1992–2016. The Cryosphere, 12(9), 2941-2953.
Ezraty, R. and Cavanié, A. (1999). Intercomparison of backscatter maps over Arctic sea ice from NSCAT and the ERS scatterometer, Journal of Geophysical Research: Oceans, 104, 11471–11483.

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

Reply: Thank you for the advice. The sentences were rephrased to (L612-L614)
"In the Beaufort and East Siberian Seas in late winter, employing Tb and backscatter measurements even leads to the worst SITY classification in OSISAF-SITY and Zhang-SITY (Case 5, Fig. 13). This indicates that simple data combination does not necessarily imply better classification results"

Line 394:
- Either: "employ a dynamic threshold" or "employ dynamic thresholds"
Reply: Done. It was modified to "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.

Reply: Thanks. The expression may not be clear in the original manuscript. We would like to say that the dynamic thresholds vary temporally or spatially, and depend on the specific satellite sensors. In the revised manuscript, we modified the sentence to (L619) *"while others employ dynamic thresholds to account for the variability of training dataset…"*

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 ...

Reply: Thank you for the advice. In the revised manuscript, it was modified to (L623-L624) *"On the other hand, the dynamic threshold approach considers the spatio-temporal variability of the microwave radiometric and scattering characteristics."*

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

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

Reply: "obscure" was replaced with "difficult"

- 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.

Reply: Thank you for the advice. We added descriptions of the physical background in the Introduction and Data Section as explained in previous replies. In the revised manuscript, the sentences were modified to (L638-L642)

"The large loss of old ice (e.g. older than four years) in the Arctic Ocean leads to a younger MYI regime in the Arctic, thus smaller microwave signature differences between MYI and FYI (Tschudi et al., 2020). On the other hand, because of the lower sensitivity of C-band scatterometer on MYI identification (as explained in section 5.1), the separation between FYI and MYI becomes more difficult, especially from ASCAT data (Belmonte Rivas et al., 2018; Zhang et al., 2019)."

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.

Reply: It is true that MYI can be found once in a while in these regions. However, as we understand, the mask mentioned here should have excluded the areas (part of the regions) where MYI could be present. Otherwise this mask would lead to obvious "wrong" results. Regarding the comment "Apart from that is the Chukchi Sea part of your region BS", we are not sure what the reviewer means. In the revised manuscript, the sentence was modified to (L649-L650)

"The first kind of correction scheme, a mask of the Arctic basin, has been used in C3S, OSISAF- and KNMI-SITY to remove the unphysical MYI signature in areas such as the Greenland, Kara, Barents and Chukchi Seas."

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.

Reply: Thank you for the advice. Having an additional plot to illustrate this issue would be too much for the example. In the revision, we modified the sentences to make it easy to understand (L651-L654).

"The thresholding filter in C3S-SITY and OSISAF-SITY exclude extreme values that are likely to cause misclassification, e.g., values beyond the simulated FYI PDF however within the wide simulated MYI PDF, which usually occurred in ice edge areas (Aaboe et al., 2021b; Aaboe et al., 2021c)"

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.

Reply: Thank you for the advice. The sentence was rephrased to (L654-L655)

"These two kinds of corrections exclude misclassification cases in regions outside the central Arctic thus have little impact on the overall SITY distributions."

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?
Reply: Thank you for the advice. In the revision, the sentence was modified to better explain the misclassification and the relation with the correction. In addition, references were added accordingly (L656-L657):
"The temperature-based correction in C3S-2 aims to reassign the erroneously classified FYI, which exhibits similar microwave signatures as FYI due to warm air intrusions (Ye et al., 2016a; Shokr and Agnew, 2013)."

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.
Reply: Thank you. We have added sentences here to explain the ice motion-based correction. The sentences were modified to (L658-L661)
"In Zhang-SITY, an ice motion confining procedure is introduced to eliminate overestimated MYI. The procedure builds upon the ice motion temporal records and confines the evolution of MYI according to the tolerance of ice motion. One drawback of this post-processing is that, the wrong reassignment of MYI to FYI could lead to continuous underestimation of MYI in consecutive days."

Lines 428/429: It is not sufficiently well described how the correction based on a median filter (spatial or temporal) works.
Reply: Thank you. The sentence was modified to (L661-L663)
"Another correction used in Zhang-SITY is the median filter correction, which considers spatial consistency and is employed to remove large unusual SITY spatial variations."

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.
Reply: Thank you. The sentence was modified to (L666-L667)

"Apart from the above three aspects (input parameters, classification methods and correction schemes), factors such as the covering period and spatial resolution make the five series SITY products different from each other."

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.

Reply: Thank you. In the revised manuscript, we deleted "Typically", added references and phrased the sentence. The sentence was modified to (L669-L672)

"In early and late winter larger uncertainties are likely to occur due to surface processes such as wet snow attenuation and changes in brine salinity (Barber and Thomas, 1998; Voss et al., 2003; Ye et al., 2016a; Ye et al., 2016b)."

Lines 440/441:
- It is sufficient to write "grid resolution", spatial can be omitted.
- "foot print" --> "footprint"
Reply: Done.

- 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.
Reply: Agree. In the revised manuscript, "true" is replaced with "nominal".

- 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.
Reply: Thank you. Yes, we should have mentioned how they (SITY products) process the ASCAT data. The ASCAT data used in the SITY products is already the normalized backscatter coefficient with same incidence angle of 40 degree. We therefore did not mention the influence of incidence angle in the discussion. In the revised manuscript, we added a section of "Microwave remote sensing data" (section 2.1) and "sea ice type products" (section 2.2), where how the characteristics and pre-processing of ASCAT data is introduced.

Line 452:
- Add "five" in front of "SAR images".
Reply: Done.

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

Reply: In the revised manuscript, we mention it when talking about the inter-comparison. The sentence is modified to (L682-L683)

"In this paper, eight SITY products based on five retrieval approaches were inter-compared through temporal and spatial analysis, with the NSIDC-SIA product as a comparative reference."

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.

Reply: Thank you for the advice. We agree that it is more appropriate to mention the average performance here in the conclusion. In the revision, these sentences were modified as below (L685-L694), with one paragraph summarizing their performances on the trends within winters, another paragraph about their average performance:

"The eight SITY products show overall negative trends of MYI extent as expected within most winters. Exceptions occur mainly in early and late winter months such as October/November and March/April. Compared to NSIDC-SIA, all the SITY products show smaller MYI extent and larger FYI extent except KNMI-SITY (KNMI-Q and KNMI-A). The bias of MYI extent between the SITY products and NSIDC-SIA varies from -1.32×10^6 km^2 (OSISAF-SITY, 2006–2009) to 0.49×10^6 km^2 (KNMI-A, 2009–2019). Among all the SITY products, Zhang-SITY in the QSCAT period and KNMI-Q agree best with NSIDC-SIA on the estimation of MYI and FYI extent, respectively.

Between any two SITY products, the difference in weekly MYI extent spans from 0.01×10^6 km^2 to 4.5 ×10^6 km^2. The largest discrepancy occurs between OSISAF-SITY and KNMI-A in late October 2008, while the smallest difference is found between KNMI-Q and IFREMER-Q in mid-winter months. It is in line with the spread of the SITY products, which is largest in early winter months such as November and smallest in mid-winter months like January."

- 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.

Reply: Thank you for the advice. We agree that it is important to summarize the performances of SITY products, therefore included two more paragraphs in this section. In addition, we believe that performances of the individual products could be useful information for the users.

- 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.
Reply: Thank you for the advice. We understand the concerns. However, since these are the main findings of this paper, it may be better to include such details in other sections, where more recent references were included.

Line 477: "extensive misclassification with higher uncertainties" --> So, the misclassification in itself is highly uncertain? Please re-write.
Reply: Thank you for the advice. "with higher uncertainties" was deleted in the revised manuscript.

Line 480: "Ku-Band ..." This statement is not new and has to be backed-up by existing literature.
Reply: Thank you for the comments. References "(Rivias et al., 2018; Ezraty and Cavanie, 1999)" were added in the revised manuscript.

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.
Reply: Thanks. The sentence was modified to (L733-L735)
"On the other hand, adaptive classification method that depends on the clustering pattern of the radiometric and backscattering signatures may be inefficient when the characteristic signatures of MYI and FYI have large overlaps;"

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.
Reply: As seen in the replies to previous comments, we gave more detailed description regarding the "ice motion correction". In the revised manuscript, the sentence was modified to (L736-L739)
"Excessive post-processing such as ice motion confining could lead to an over-correction problem, which becomes the basis for the subsequent corrections and eventually result in accumulative errors."

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

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

Reply: Thanks. "should be accounted with caution" was replaced with "should be considered with caution".

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.

Reply: Thank you for the advice. We agree that it is not that appropriate to call it "systematic evaluation". In the revised manuscript, the sentence was modified to

"This study inter-compares eight SITY products and provides …" On the other hand, performances of the SITY products were summarized in this paper, with discussions on the potential reasons and improvements (e.g., L711). We therefore think the expression "hints for improvement" is appropriate.

The sentence in L711 reads as below:

"IFREMER exhibits high frequency temporal variations in MYI extent, which could be caused by the day-to-day varying thresholds and improved by including appropriate post-processing;"

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

Reply: Thanks. The two frequencies are Ku- and C-band. In the revised manuscript, we added this information in the parenthesis "(WindRAD, Ku- and C-band)".

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

Reply: Thanks. It was implemented as suggested in the revised manuscript.

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"
Reply: Thanks. It was implemented as suggested in the revised manuscript.

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.
Reply: Thank you for the comment. Since the plots of the MYI extent from SITY products rarely go to the blocks in the lower part (older than 2/3 yr), and it is

Figure 4: Why are IFREMER-A data missing for January and April in 2014 & 2015?
Reply: Thank you for noticing the data gap. IFREMER-A data during this period was not available from the data provider, therefore was not included.

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.
Reply: Thanks. The figures were modified as suggested.

- 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.
Reply: Thanks. Figures and captions were modified.

Figure 8:
- What is the motivation to show boxes A to D with a different size?
Reply: Thanks. The boxes were made the same size in the revision.

- I have the same comment with respect to colors as for Figures 6 & 7.
Reply: Thanks. All the figures were modified accordingly.

- 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.
Reply: Thanks. We used similar colors in the revision.

- 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.
Reply: Thanks. Notes were modified accordingly.

- 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.
Reply: Thank you for the comment. In the comparison, the Amb class pixels were regarded as MYI and FYI, respectively. We believe that it is necessary to show the spatial distribution on the maps.

Figures 9 to 12:
- I have the same comments with respect to colors, legends, and ERA5 data naming as I had for Figure 8.
Reply: Thanks. Done as suggested.

- In addition, delineation of the boxes in the Sentinel-1 image in a different color than black would help to locate these better.
Reply: Thanks. Green boxes were used in the revision.

- 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.
Reply: In the revision, speed of the ice drift is marked with the length of arrows, whereas colors of the background indicate the 2-m air temperature.

Table 2: Cases where there is a "+" and a "-" indicate that both performances exist?
Reply: Yes. Both performances could exist in different locations for one SITY product in one case.

Table 3: I guess the GR listed in the context of OSISAT SIT is not correct?
Reply: We have checked that it is correct.

Typos / editoral comments:

Line 59; "indirect validation" --> perhaps better "inter-comparisons"?
Reply: "indirect validation" was modified to "inter-comparisons"

Line 89: "KNMI-" --> "KNMI-SIT"
Reply: Done.

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

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

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"
Reply: Done.

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

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.
Reply: Done.

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

Line 185: "The MYI" --> However, the MYI"
Reply: Done.

Line 185: "regional" --> "regionally"
Reply: Done.

Line 225: "is mainly resulted from" --> check grammar.
Reply: "is mainly resulted from" was modified to "is mainly attributed to".

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.

Reply: This part belongs to the top row of Figure 6. We added notes of Figures in the sentence for clarification. The sentence is modified to:

"On the other hand, IFREMER-Q (e.g. Fig. 7c) shows constantly less MYI than KNMI-Q (e.g. Fig. 7d) in the transition zone of MYI and FYI in BCS, in good agreement with their difference as shown in Fig. 6."

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

Reply: Done.

Line 288: Add the year.

Reply: Done.

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

Reply: Done.

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

Reply: "Slightly underestimation of MYI" was modified to "Slight underestimation of MYI".

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.

Reply: "thin MYI tongue" was modified to "narrow MYI tongue"

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

Reply: "can partly be resulted from" was modified to "can be partly attributed to ".

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."

Reply: Done.

Line 313:
- "transit zones" --> "transition zones" or "zones of mixed FYI - MYI coverage"

Reply: "transit zones" was modified to "transition zones".

- "steady discrepancies" --> re-phrase please.

Reply: "steady discrepancies" was modified to "constant discrepancies".

Line 315: Delete "validation and"
Reply: Done.

Lines 324-325: "The MYI feature ... round MYI floe" --> please check grammar.
Reply: the sentence was modified as follows (L518-L520):
"The bright MYI floe feature is clear in the northeast part of the SAR image, so as the dark FYI feature in the southwest part. Areas A and D exhibit high backscatter of round MYI floe, and areas B and C present typical characteristics of FYI with smooth texture and low backscatter."

Line 364: "serial" ???
Reply: "serial" is deleted.

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

Line 382: "confirmed" --> "shown"
Reply: Done.

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

Line 402: "speculate" --> "hypothesize" ?
Reply: "we speculate that this is because …" was modified to "The possible explanations could be that …"

Line 405: Either "to a sea ice type distribution" or "to sea ice type distributions"
Reply: "SIT distribution" was modified to "SIT distributions"

Line 406: Either: "An adavtive clustering algorithm is used" or "Adaptive clustering is used"
Reply: "Adaptive clustering algorithm is used" was modified to "An adaptive clustering algorithm is used"

Line 408: "thin ... seas" --> "narrow MYI tongues in the peripheral seas"
Line 427: "continuous underestimation" of what?
Reply: it was modified to "continuous underestimation of MYI"

Line 430:
- "over-correction problem" --> "over-correction.
Reply: Done.

- "thin MYI ... seas" --> We had that expression earlier. Please look up my comment there.
Reply: Agree. "thin MYI …" was modified to "narrow MYI …"

Line 436: "fully evaluated" --> "done"
Reply: Done.

Line 481: What is "small FYI in MYI pack"? Do you mean: "comparably small areas of FYI within a region dominated by MYI?"
Reply: Yes. "small FYI in MYI pack" was modified to "comparably small areas of FYI within a region dominated by MYI"

Line 487: "deep " --> "mid-"
Reply: Done.

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.
Reply: Thank you for the advice. We double-checked all the references in the manuscript and modified accordingly.