

Response to RC 3:

(The reviewer comments appear in black, the responses are in blue and the proposed changes to manuscript are in ***bold italics***.)

Comments are assigned to co-authors that we think could help to sort out.

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.

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 will revise the manuscript from the following two aspects:

- 1) The “Data” section will be re-structured. This section will include four sub-sections: “2.1 Microwave remote sensing”, “2.2 Sea ice type products” and “2.3 Other data”. In section 2.1, specifications of the sensors and the satellite data will be introduced in a chronological order, with subsections of passive/active microwave remote sensing data. In section 2.2, theory of SIT classification will be 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) and the SAR images will be described accordingly.
- 2) A new section of “Methods” will be added, which includes “3.1 Estimation of MYI extent” and “3.2 Interpretation of SAR imagery”. We will modify the computation of MYI extent in the revision for consistent gridding, projection among all the SIT products. In section 3.1, Information such as co-locating/re-gridding the data and calculation of the MYI extent will be introduced. In section 3.2, the theory and characteristics of sea ice classification in SAR images will be introduced with references from previous studies and examples from our study. In addition, we will interpret the entire SAR images, consult with ice experts regarding the results, convert the sea ice classification results from ice types polygons to grided ice classification results, and eventually give quantitative evaluation results.

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

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 will be added in the revision. Such descriptions will be presented in the following sections: 1) section “2.2 sea ice type products”, where the theory of sea ice classification will be introduced before all the SIT products; 2) section “3.2 visual interpretation of SAR imagery”, which is a subsection of “3 Methods”, and 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 heterogeneous 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.

As for the two products you mentioned, since this study focuses on 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. In order to clarify the focus, we will modify the sentences in the section “Introduction”, 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 will include more quantitative results and analysis in the manuscript. 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 will be added. In addition, weekly MYI extent will be calculated from sea ice type products for consistent comparison. 2) in the revision, we will interpret 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 (e.g. overall accuracy for each SAR image).

For the manual 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 conduct the interpretation 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. We will add more discussions that are linked with existing literature.

Specific comments (contain some typos / editorial 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 will categorize the sea ice type products into 1) and 2) those using solely active/passive microwave data and 3) those with combined microwave data.

- 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. On one hand, we will modify the first sentence in the last paragraph of “Introduction” Section. The sentence is modified to “... give comprehensive evaluations on the identification of MYI and FYI”. On the other hand, in the new section of “Method”, we will give clear statement that this paper focuses on the classification 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. We will specify the temporal resolution in the abstract. In the revision, we will mention "... nine daily SITY products", "... a weekly sea ice age product".

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

Reply: Thank you for the comment. We confirm that the numbers are correct. For the exceptional large difference, we will add discussions in the Results section.

- 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. The sentence is modified to

"...Zhang-SITY shows underestimation of MYI with relatively large fluctuations"

- 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 are not based on one classification approach. In the revised manuscript, we give more detailed description about the classification algorithms used in the SITY products. This study focuses on the comparison and evaluation of current SITY products. Further investigation of classification algorithms could be done in future studies.

- 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 (second and third sentences). The modified sentences are:

"...microwave satellite data. 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: Agree. We add sentences with brief description of the fundamental differences for microwave SITY classification. The new sentences are:

“On one hand, brightness temperatures (Tbs) of MYI tend to be lower than that of FYI because of high loss of radiation caused by scattering when going through the bubbly layer in the sub-surface of hummock area (Sinha and Shokr, 2015). Such difference depends on the wavelength of the radiation with respect to the typical dimensions of the bubbles (they should be comparable for the loss to be effective). On the other hand, due to the high volume scattering and low scattering loss, MYI have relatively higher backscatter than FYI at the same frequency (Carsey, 1992).”

References:

Sinha, N. K., & Shokr, M. (2015). Sea ice: physics and remote sensing. John Wiley & Sons.

Carsey, F. D. (Ed.). (1992). Microwave remote sensing of sea ice. American Geophysical Union.

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, respectively. We later mention the typical algorithms accordingly. The modified sentences are:

“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 ...”

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:

“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 will add sentences in this paragraph to clarify the focus of our 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. We will add sentences regarding the implications of large errors to highlight the need of evaluation.

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

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 will revise the manuscript as suggested.

In the “Introduction” section, we will include sentences about the physical background for microwave SITY classification as mentioned above. The sentences included additionally start with:

“On one hand, brightness temperatures (Tbs) of MYI tend to be lower than that of FYI because of high loss of radiation caused by scattering when going through the bubbly layer in the sub-surface of hummock area (Sinha and Shokr, 2015). Such difference depends on the wavelength of the radiation with respect to the typical dimensions of the bubbles (they should be comparable for the loss to be effective). On the other hand, due to the high volume scattering and low scattering loss, MYI have relatively higher backscatter than FYI at the same frequency (Carsey, 1992).”

References:

Sinha, N. K., & Shokr, M. (2015). Sea ice: physics and remote sensing. John Wiley & Sons.
Carsey, F. D. (Ed.). (1992). Microwave remote sensing of sea ice. American Geophysical Union.

In the “Data” section, before introducing the individual sea ice types, we will include a paragraph that describes the physical background. The paragraph starts with:

“Microwave radiometer and scatterometer are used to discriminate MYI and FYI due to their distinctive signatures. Microwave radiometer Microwave radiometers measure the upwelling radiation emitted by the Earth in terms of brightness temperature (Tb), which is linearly proportional to the physical temperature and microwave emissivity of the observed object...”

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 revision, the “Data” section will be divided into four subsections: “2.1 Microwave remote sensing data”, “2.2 sea ice type products” and “2.3 other data”. In the section of “2.1”, we will introduce the native resolution of the microwave remote sensing data. In the section of “2.2” we will clarify the grid resolution of the sea ice type products.

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 question. 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. To avoid misunderstanding, we modify the sentence as below:

“The Tb measurements are corrected for the atmospheric influence by using a Radiative Transfer Model function (Wentz, 1977) and auxiliary data such as sea ice concentration and atmosphere reanalysis data.”

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 will be 3 subsections in the "Data" section. The equation of gradient ratio will be 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 thus is modified to:
"The probability density functions (PDFs) of given ice classes used in the Bayesian classification are dynamically derived from training datasets, which are extracted from the pre-defined areas of MYI and FYI during the 15-day period centered on the specific day (Aaboe, Sørensen, et al. 2018). The pre-defined areas of MYI locate in the north of Greenland and Canadian Arctic Archipelago between 30°W and 120°W, whereas the FYI areas lie in the Kara Sea, Baffin Bay, Laptev Sea and the Bay of Bothnia.

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. Data of this new version, which is referred to as C3S-2, is indeed included in the comparison. One of the purpose of including this data is to see how much does it differ from the old version, C3S-1. We therefore would like to keep it in the manuscript.

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: As mentioned above, there will be 3 subsections in the new "Data" section. In the new section "2.1 microwave remote sensing", we will introduce all the microwave remote sensing data used in the various SITY products. In the new section "2.2 sea ice type products", we will 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. We will include more details regarding the product and satellite data. As mentioned in the previous reply, details of the microwave remote sensing data will be given in the new section 2.1. It will be stated that the incidence angle of ASCAT ranges from 25° to 65°. Backscatter from ASCAT is normalized to a reference incidence angle of 40 deg 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 10km x 25 km.

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

Reply: Thanks for the comment. The sentence is modified to:

"By using the Bayesian approach, sea ice class-probabilities are estimated for GR_{19v37v} and σ_0 , respectively."

- 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: Thanks. In the new section 2.1, satellite data will be introduced in a chronological order.

- 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 will be introduced in a chronological order meanwhile with same level of details. The microwave radiometer data will be introduced as two series of dataset: one includes SSM/I and SSMIS, the other means AMSR-E and AMSR2. Spatial resolution of the sea ice type products will be mentioned in the new section 2.2.

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

Reply: Details of ASCAT will be presented in the new section 2.1. ASCAT provides only vertically polarized backscatter. The spatial resolution achieved by ASCAT 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. Such details is not crucial for the product. This sentence is therefore deleted in the revision.

- 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: Done.

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 will add a new subsection "2.1 Microwave remote sensing" in the revision, where the sensors and specifications will be 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: Thanks. In the revision, we will modify the sentences to underline more clearly the meaning of each acronym, and specify the two products we include in the evaluation is KNMI-Q and KNMI-A. The modified sentences are as below:

"KNMI-SITY is a series of sea ice type products, which are purely based on scatterometer data, including ERS, QSCAT, OSCAT and ASCAT data. The corresponding four SITY products in KNMI-SITY are referred to as KNMI-E, KNMI-Q, KNMI-O, and KNMI-A, respectively. They are available during the periods of 1992-2001, 1999-2009, 2010-2013 and 2007-2016, respectively. In this study, KNMI-Q and KNMI-A are included in the comparison and evaluation."

- While you 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: Thanks. We will refer to the Table 1/Figure 1 when describing details such as periods. In the revision, specifications of the sensors will be introduced in the new section 2.1 with same level of details, whereas the grid resolution and algorithm will be introduced in the new section 2.2.

- 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 algorithm classifies surface into four categories (open water, first-year ice, second-year ice and older multiyear ice). We regard the second-year ice and the older multiyear ice as the MYI for comparable analysis. As explained in the previous sentence, "Fixed thresholds extracted from stable wintertime (March) data are adopted for sea ice type classification." The fixed thresholds mean all the thresholds used for ice type classification. Sea ice and open water is classified before the three ice types are further separated, while the three ice types are separated using the thresholds at the same time.

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

Reply: Since KNMI-O, which uses OSCAT data, is not included in the comparison of our study. We delete 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: Thanks. These sentences are modified to

“IFREMER-SITY is another series of scatterometer-only products. It includes two SITY products, which use QSCAT and ASCAT data for the 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: We modify the sentence as below to avoid providing different details for one of the product.

“To account for the varying incidence angle of ASCAT data, the backscatter coefficients are normalized to the backscatter at a constant incidence angle of 40°.”

- 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: To answer the questions you raise here, we modify the sentences as below:

“In IFREMER SITYs, day-to-day-varying thresholds are used for the separation between MYI and FYI. These thresholds are derived by analysing the backscatter data for 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: In the revision, we will add projection and grid spacing information of all the SIY products.

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: it is 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. We will add such information in the revision.

- 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 will be introduced in the new section 2.1.

- 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 will add the following sentences in the revision. *“Scatterometer data is obtained from QSCAT and ASCAT successively since QSCAT stopped functioning on 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 will be added in the paragraph of describing the algorithm used in Zhang-SITY.

“”

Lines 140/141: "climate consistent data record SIT products" --> Given the heterogeneity 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 delete 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, we will modify the description of the sea ice age product accordingly.

- 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 will remove the citation and rephrase the description here.

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

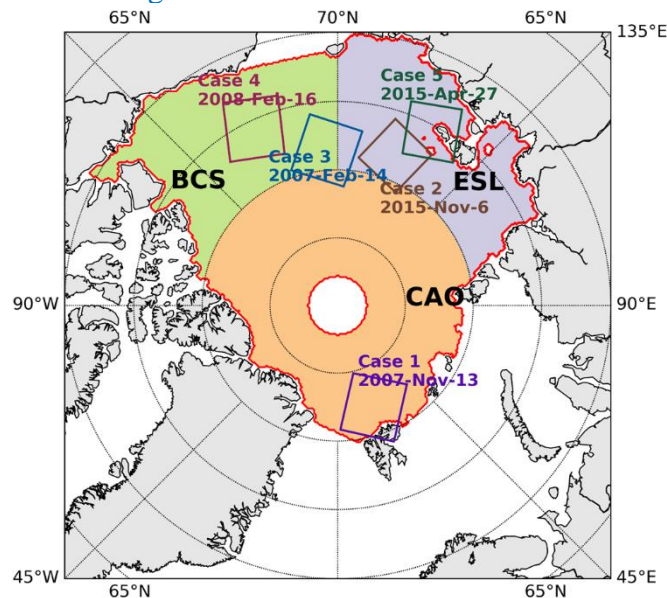
Reply: Done

- 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 will include the acquisition date of the SAR images. The modified figure will be 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.

- 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 will delete 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, we will delete these sentences 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 is 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 will be updated accordingly.

Line 180:

- Here, in this line you 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. We delete the sentence of Line 180 and modified the sentence in Line 176 to

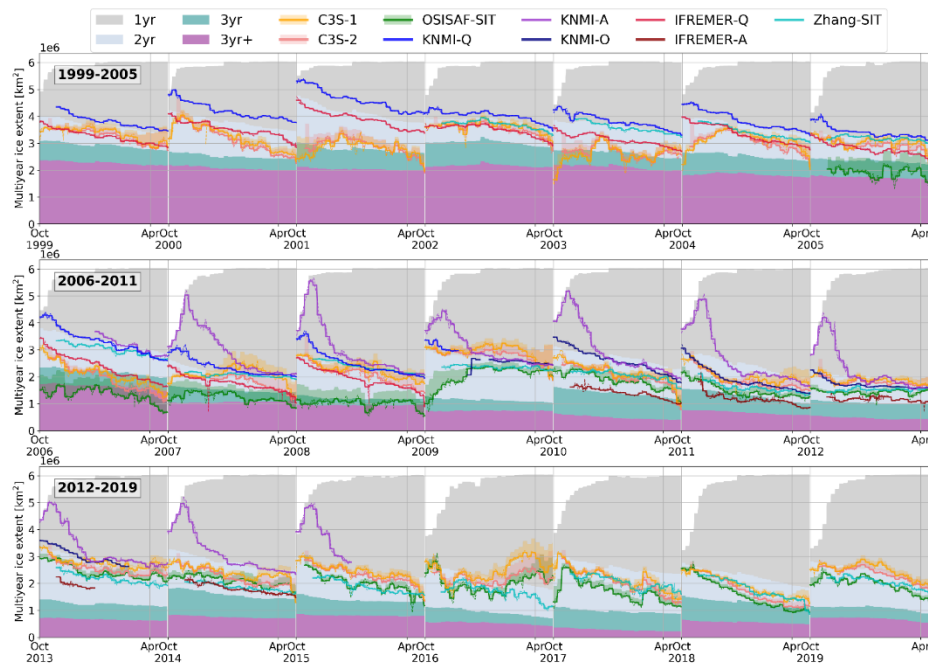
“For SITY and SIA products, we computed the MYI extent as the sum of the area of all grid cells classified as multiyear sea ice.”

- 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 is not accounted for in the original manuscript. We will therefore correct the computations in the revision. SITY products in polar-stereographic projection are firstly regridded to the EASE grid before the computation of MYI extent. Details of such will be 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 will compare the weekly averaged MYI extent from the SITY products with the NSIDC-SIA product. All the relevant numbers will be modified accordingly. Plots of the weekly averaged MYI extent will be presented in Fig. 3. The modified figure will be 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. This sentence will be modified as follows.

"since FYI can only turn to MYI when surviving a melting season, overall Arctic MYI extent cannot increase over the winter – it can only decrease through ice advection out of the Arctic"

Other problematic expressions will be modified accordingly, and all the "decreasing/increasing trend" will be modified to "negative/positive trend".

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 is modified to

"it can temporarily or regionally increase due to ice divergence or advection from neighbouring regions"

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: Done.

Lines 199/200: Given the fact that the entire Arctic Ocean (i.e. approximately the region of your study) has a size of about $7 \times 10^6 \text{ km}^2$, 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. We confirm that the largest difference is correctly calculated. Explanations will be added accordingly.

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: These are averages of the extent estimate from the "ambiguous" pixels. As explained in the first paragraph of 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. We will add a few sentences here to explain the meaning of these values and how they relate to the entire MYI extent. These sentences are

"As described in section 2.2, the Amb pixels have atypical microwave signatures of MYI/FYI thus high uncertainties on ice type discrimination. Compared with the average MYI extent difference between C3S-SITY and OSISAF-SITY, the contribution of these pixels to the comparison is overall small, however 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 will add 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 want to say that the inter-annual evolution of MYI extent from C3S- and OSISAF-SITY differs most with NSIDC-SIA. The sentence will be modified to *"Overall speaking, the inter-annual pattern of MYI extent from C3S-SITY and OSISAF-SITY differs most with NSIDC-SIA, with large discrepancies in the winter months of 2001-2003, 2006-2008 and 2016-2018."*

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 will add 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. The new sentences are as follows.

"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. In the ESL and BCS regions, it is found that the NSIDC-SIA MYI extent is usually considerably larger than the MYI extent from SITY"

products, whereas it is less pronounced in the CAO regions. This indicates that the mixture of MYI and FYI, which leads to “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.

Reply: Thank you for the advice. The sentence will be rephrased to

“The former mainly results from the outflow of MYI to more southern areas. On one hand, MYI is extensively exported through the Fram Strait and, by small fractions, into the Barents Sea and through the Nares Strait following the Transpolar Drift Stream. On the other hand, MYI is advected towards south along the Canadian Arctic Archipelago driven by the Beaufort Gyre.”

- 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.

Reply: Thank you for the advice. The sentence is modified to

“Overall, the MYI extent in the CAO and ESL regions shows consistently negative trend, while that in the BCS 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. This sentence is modified to

“In the BCS region, large quantities of MYI is 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 nearly constant or increasing MYI extent in the BCS region.”

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 will be modified to:

“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. 6, 7 and A1. To illustrate it more clearly, we will revise the sentence accordingly:

"In 错误! 未找到引用源。 a and b (along with 错误! 未找到引用源。 a-d, f-i in Appendix and Fig. 7 a-b, i-j), the discontinuous FYI delineation in the inner part of MYI pack is well demonstrated, which occurs in different winter months and could partly explain the daily 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 wrong here. We were about to use "exceptional" to describe the MYI extent from KNMI-A not KNMI-Q. The sentence will be therefore modified to:

"KNMI-Q has the second largest MYI coverage among the seven SITY products, with 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. As stated in the top two rows of Figure 7, "in the BCS region, MYI is overestimated compared to NSIDC-SIA", if we assign the ambiguous ice type pixels to MYI, this will actually worsen the agreement with NSIDC-SIA. This piece of information does not seem to be useful for SITY product improvement, we therefore did not include it in the manuscript.

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 will add a section of "Methods". Information such as co-locating data and calculation of MYI extent will be 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: It is difficult to see the details of SAR image without enlargement. The boxes are selected from the locations where there are typical backscattering characteristics and could show the agreement/discrepancies among the various SITY products. In the revision, we will adjust the boxes to ensure the similar sizes.

- 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, geometric 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: In the revision, we will add a new section "3.2 Visual interpretation of SAR imagery", where we will provide more details about the characteristics. In addition, we will interpret 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.

- 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: We will add the polarization information in the respective figures. In addition, in the new section "3.2 Visual interpretation of SAR imagery", we will add sentences as below:

"Images at HV polarization channel are prioritized if available. This is because the backscatter at cross-polarization channel is usually dominated by volume scattering thus is more disparate between FYI and MYI."

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. We add a section of "Method" in the revised manuscript, where the theory of interpreting SAR images is introduced. In addition, we make visual interpretation quantitatively based on the theory and discussed the quantitative comparison in the section of "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 will modify the sentences at the beginning of the “Results” section and the last paragraph of the “Introduction” section.

- 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. It reads:

“A typical scene of early winter (November, 2017) in MIZ is shown in 错误!未找到引用源。”

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: 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.

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. SITY distribution in NSIDC-SIA is generally consistent with the SAR image. The sentence is therefore modified to

“NSIDC-SIA shows generally consistent SITY distribution with the SAR image.”

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 will add a section of “SAR interpretation” in the revised manuscript. In addition, we will provide quantitative interpretation results in terms of statistics such overall accuracy.

- 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 will be added to the SAR images.

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. This is nearly negligible considering the temporal resolution difference. We therefore modify the sentence to

“MYI is slightly underestimated in NSIDC-SIA. 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 scattering characteristics are more separable in HV image than HH image. We therefore presented the HV image instead of HH image in the manuscript. In fact, both HH and HV image are used for the visual interpretation. To illustrate the procedure better, we add a section of “SAR interpretation” in the revised manuscript. The physical background and how the zoomed boxes are selected are explained accordingly in the new section. For the Sentinel-1 image, we will include additionally the HH 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 usually - 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). You might have doubted that area C is MYI. We also agree with on it. Area C should be landfast FYI instead of MYI. Therefore, we will modify the texts in this paragraph accordingly.

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

Reply: We put mid-winter after late-winter period since we wanted to show the cases with large discrepancies first and then later those with smaller discrepancies. The purpose is not well described in the original manuscript meanwhile not that necessary. In the revised manuscript, we will therefore present 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 overlay the visual interpreted results (contours) on each subfigure. This sentence is modified to:

“... the MYI pack in NSIDC-SIA is overestimated in the northeast part of the image (area A) meanwhile underestimated 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 will modify the sentence to

“As explained previously, such discrepancies are mainly attributed to the mobility features of sea ice and the different temporal resolutions between NSIDC 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 product 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 modify such expressions meanwhile use SAR images in a quantitative way. We add a “Methods” section, 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 is modified to

“The efficacy of input parameters is dependent on the 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 investigate 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 will therefore remove 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 (OSISAF-SITY) uses this ratio the other way around. In the revised manuscript, we give detailed description of each algorithm when introducing each SITY product.

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 delete "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 overestimation 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 will be 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 rephrase the sentence and include 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 snow wetness changes. 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 rephrase the sentence and include the reference accordingly.

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

The modified sentence is shown as below:

“In the beginning and ending stage of winter, the variability of GR_{37v19v} can be significant when air temperature exhibits warm-cold cycles which triggers wet-dry cycles or melt-refreeze cycles of snow (Ye et al., 2016a; Ye et al., 2016b; Voss et al., 2003), 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 will be added in section 2.2. In the revised manuscript, the sentence will be modified to:

“Such misclassification in C3S-1 is mitigated in C3S-2 due to the upgraded 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 (see section 2.2)”

Line 377:

- "backscatter (σ^0)" --> either "backscatter coefficient" or "sigma nought"

Reply: Done. "backscatter" is modified to "backscatter coefficient"

- "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 will be modified to

“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 will be modified as below. Explanations will be added.

"In comparison, the backscatter of MYI and FYI is more disparate at Ku-band than C-band (Bi et al., 2020; Rivas et al., 2018). Products using Ku-band backscatter generally performs better on identifying MYI, e.g. KNMI-, IFREMER-, and Zhang-SITY. 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 Cavanié, 1999). On the other hand, the greater importance of surface scattering and the higher dependence on incidence angle makes C-band backscatter more suitable to distinguish ice types with disparate surface roughness features, e.g. Case 5 in 错误! 未找到引用源。."

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., Otsuka, 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 will be rephrased to

"In Beaufort and East Siberian Seas in late winter, employing Tb and backscatter measurements even leads to worse SITY classification in OSISAF-SITY and Zhang-SITY (Case 3, 错误! 未找到引用源。). It 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 is 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: 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 modify the sentence to *“while others employ dynamic thresholds, which may vary with time, region and the satellite sensors”*

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 its microwave backscattering characteristics and/or the microwave emission? Then please write it specifically. Currently, "sea ice variabilities" 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 is modified to *“For the latter, the approach considers the spatio-temporal development of the physical properties of sea ice, which influences the microwave radiometric and scattering characteristics,”*

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

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

Reply: “obscure” is 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 will add descriptions of the physical background in the Introduction and Data Section. In the revised manuscript, the sentences will be modified to

“This could be attributed to the large loss of old ice (e.g. older than 4 years) in the Arctic Ocean (Tschudi et al., 2020), which leads to a younger MYI regime in the Arctic thus less pronounced microwave signature difference between MYI and FYI. It eventually makes the separation between FYI and MYI more difficult, especially from ASCAT data (Belmonte Rivas, Otosaka, et al. 2018, Zhang, Yu, 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 occur 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 will be modified to

“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 coefficient 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. Since the misclassification is not necessarily 'along the ice edge', we will delete it in the revision. On the other hand, including an additional plot to illustrate this issue would be too much for the example, with consideration of the large number of figures in the current manuscript. As mentioned in previous replies, we will provide more details in the section "2.2 sea ice type products" to make it easier to understand the sentence here.

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 will be rephrased to
"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." ?

Reply: Thank you for the advice. The sentence is modified to
"The temperature-based correction in C3S-2 aims to assign the erroneously classified grid cells of MYI due to the warm air intrusion induced changes of physical properties to FYI."

- How is this temperature based correction done?

Reply: In the revised manuscript, we briefly describe the temperature-based correction in the section of "2.2 Sea ice type products"

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

Reply: Other algorithms do not use the temperature-based correction.

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 will add one more sentence here to explain the ice motion-based correction. The sentences will be modified to

"In Zhang-SITY, an ice motion confining procedure is introduced to eliminate overestimated MYI. The procedure builds upon ice motion records and confines the evolution of MYI according to the tolerance of ice motion."

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 will be modified to

“Another correction used in Zhang-SITY is the median filter correction, which considers the spatial consistency and is employed to remove large unusual 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 will be modified to

“Apart from the above three aspects (input parameters, classification methods and correction schemes), factors such as the covering period and spatial resolution makes 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 delete “Typically”, add references and phrase the sentence. The sentence will be modified to

“In early and late winter larger uncertainties are likely to occur due to surface melting over sea ice and atmospheric influence (Voss et al., 2003; Ye et al., 2016a; 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 add 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

"In this paper, nine 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 understand that some readers may be interested in the periods and regions where there are systematic errors. However, as the NSIDC sea ice age data is only a 'reference' at large scale not the 'ground' truth, it is very unlikely that we could draw conclusions such as there are systematic errors in specific regions and periods given the evaluation with limited SAR images. Meanwhile, we will dig more into the average performance of the product over a longer period.

- 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 qualitative 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 will dig more into the larger picture of the results over a long period.

- 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 will include more details and refer to the existing literature in the revision.

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

Reply: "with higher uncertainties" is 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 "(Rivas et al., 2018; Ezraty and Cavanie, 1999)" are 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: The sentence is modified to

“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 give more detailed description regarding the “ice motion correction”. In the revised manuscript, the sentence is modified to

“Excessive post-processing such as ice motion confining could lead to over-correction problem, which becomes the basis for the subsequent corrections and eventually result in accumulative errors”

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

Reply: Implemented in the revised manuscript.

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

Reply: “should be accounted with caution” is 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.

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

“This study inter-compares nine SITY products and provides ...”

- 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: To tone down the value, the sentence is modified to

“This study ... and indicates the potentials for further improvement.”

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 add 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 is 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 is 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: Regarding your concerns of Figure 3, 1) since there are already many plots (from SITY products) in Figure 3, it is easier to identify the NSIDC sea ice age, thus to compare with the respective SITY. 2) this Figure is used to show an overall comparison we therefore include all the datasets of the entire period here. 3) we would like to show the sea ice age information. In the revision, we will use different shades of grey for the NSIDC sea ice age in Figure 3. In addition, we will include the weekly MYI extent from all the SITY products instead.

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

Reply: The IFREMER-A data is not available for the period at the time for the analysis. We will update the data in the revision.

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 will be 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. We will cross-ref Table 1 in the caption and add the meaning of red contour.

Figure 8:

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

Reply: Thanks. Size of the box does not need to be different. We will try to select similar size of boxes in the revision. (not yet modified in the example figure below)

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

Reply: It will be modified as suggested.

- 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: We will use a different color code for easier discrimination between FYI and older ice.](#)

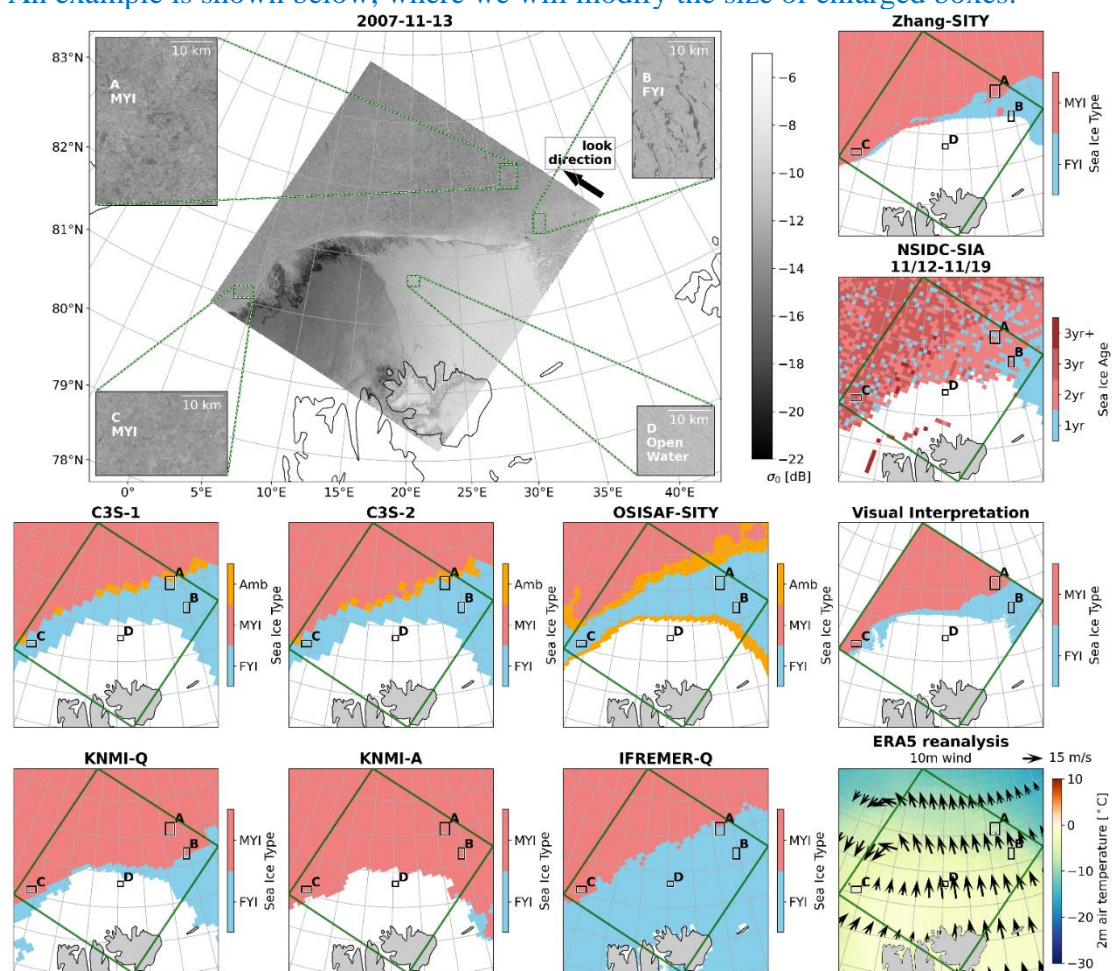
- I recommend to rename "sea surface wind" to "10m wind" because I guess this is what it is. Also make clear that "air temperature" possibly is the "2m air temperature". The additional information that these are daily averages would be appreciated as well.

[Reply: It will be modified as suggested.](#)

- 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: For SITY products that include the class of 'Amb', if we do not show the 'Amb' class, there will be pixels with missing information in the Figure. On the other hand, 'Amb' tells the pixels with large uncertainties of ice type classification. We therefore would like to keep the 'Amb' class in the figure.](#)

[An example is shown below, where we will modify the size of enlarged boxes.](#)



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: It will be modified accordingly.](#)

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

Reply: It will be modified as suggested.

- 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: The vectors represent ice drift, whereas the background colored field represents 2-m air temperature of the same day.

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

Reply: Yes. In the revision, we will add the overall accuracy of each SITY product for the respective SAR images. An example of the updated Table 2 is shown as below:

SITY products	Case 1		Case 2		Case 3		Case 4		Case 5	
	General pattern	Overall Accuracy	General pattern	Overall Accuracy	General pattern	Overall Accuracy	General pattern	Overall Accuracy	General pattern	Overall Accuracy
C3S-1*	-	0.76–0.82	○	0.83–0.86	--	0.47–0.47	+	0.71–0.73	+-	0.47–0.54
C3S-2	-	0.80–0.87	○	0.83–0.86	--	0.47–0.47	+-	0.72–0.77	+	0.40–0.49
OSISAF-SITY	--	0.52–0.64	-	0.76–0.79	--	0.47–0.47	--	0.78–0.82	+-	0.63–0.69
KNMI-Q	+	0.79	/	/	+	0.86	+	0.78	/	/
KNMI-A	++	0.70	++	0.66	○	0.89	○	0.78	○	0.87
IFREMER-Q	-	0.86	/	/	--	0.47	-	0.83	/	/
IFREMER-A	/	/	/	/	/	/	/	/	--	0.84
Zhang-SITY	○	0.91	--	0.60	-	0.84	○	0.82	--	0.84

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

Reply: It is correct, which is the ‘negative’ form of the other.

Typos / editorial comments:

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

Reply: “indirect validation” is 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" is 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 add notes of Figures in the sentence for clarification. The sentence is modified to:

"On the other hand, IFREMER-Q (e.g. Fig. 6c) shows constantly less MYI than KNMI-Q (e.g. Fig. 6d) in the transition zone of MYI and FYI in BS, in good agreement with their difference as shown in 错误! 未找到引用源。."

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" is 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" is modified to "narrow MYI tongue"

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

Reply: "can partly be resulted from" is 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" is modified to "transition zones".

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

Reply: "steady discrepancies" is 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 is modified as follows:

"The bright MYI 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"

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 ..." is 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" is modified to "SIT distributions"

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

Reply: "Adaptive clustering algorithm is used" is 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 is 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 ..." is modified to "narrow MYI ..."

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

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

Reply: Yes. "small FYI in MYI pack" is 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 your thorough review. We will double-check all the references and be more careful on the detailed information in the revision.