Dear Editors and Reviewers,

We would like to thank you all for the dedicated and insightful comments and advices. We have put a lot of effort on revising the manuscript regarding the structure, expression and readability. Comments from the two reviewers are responded on point-by-point basis.

The reviewers' comments appear in black. The responses are in blue and the proposed changes to manuscript are in ***bold italics***.

==

Report 1

The manuscript has been improved. Thanks for the author's effort to make a more strengthened paper. I recommend that this manuscript be published after some minor revisions of table and figures.

Table 1: it would be nice if it is specified whether the Satellite Input is based on the microwave radiation or scattering.
Reply: Thank you for the advice. The table is modified as suggested. In addition, a column of the grid is added.

Fig. 2: Please specify the imaging coverage of RS and S1.
Reply: Thanks. Notes of "RS-1" and "S-1" are added to Figure 2.

Fig. 3: figures are missing the geographic information such as latitude/longitude grid and scale bar. And please specify which satellite SAR images the sub-figures are from in the caption.
Reply: Thanks for the advice. The geographic information, scale bar and notes of "RS-1" "S-1" are added to Figure 3.

==

Report 2

Review of
Inter-comparison and evaluation of Arctic sea ice type products
By Ye, Yufang, et al.

I am not providing a summary of this manuscript because I reviewed a former version of it.
The authors have improved the manuscript considerably and have taken into account quite a number of the concerns that were brought up during the previous round of reviews.

The readability of manuscript and the credibility of the results presented do, however, still suffer from deficits in some parts of the description and from overrating of the mostly qualitative elements of the intercomparison carried out.

General comments:

GC1: The deficiencies of proper reference of the figures from inside the text and the numerous typos and strange formulations make this manuscript difficult to review and to read. Please next time when submitting a manuscript consider having it proof-read by a native English speaker plus check it for consistency. It cannot be the task of an editor or a reviewer to tick all these. It is an immense work load and distracts from the scientific content of the manuscript.

Reply: Thank you for the advice. We highly appreciate the thorough review from the reviewers and editor. We are more careful this time on double-checking the reference of figures and have more thorough review on the text and language. All the co-authors have proof-read it to avoid strange formulations and unclear descriptions in the manuscript.

GC2: While the consideration of the sea ice physics and its relation to microwave remote sensing of sea ice has been improved considerably, there are still elements that should be improved - please see my specific comments.

Reply: Thanks for the advice. The manuscript has improved substantially thanks to the reviewers' comments. Revisions and replies are made based on the respective comments.

GC3: Please step away from considering this as an evaluation / validation or assessment study. It is an inter-comparison study, involving qualitative data set for inter-comparison and products which quality you aim to understant and report upon in this manuscript. Rather than attempting to rank the products, I recommend to state clearly that we require more and better specified and evaluated data sets for a quantitative evaluation of the sea ice type products.

Reply: Thanks for the advice. We understand the concern of the reviewer. In the revised manuscript, we try to tone down the value of the work as a validation study. Meanwhile, as the reviewer said, there is no well-evaluated dataset at present. The data we use here is already one of the best we could achieve. We therefore think it is acceptable to use wording such as "evaluation" in the manuscript.

Specific comments:

L43: Please check the content of Boisvert et al., 2016, with respect to whether this is indeed the paper you wanted to cite in this context. Perhaps the two papers of her et al. from 2015 fit better: Increasing evaporation amounts seen in the Arctic between 2003 and 2013 from AIRS data" or "The Arctic is becoming warrner and wetter as revealed by the Atmospheric Infrared Sounder"?

Reply: Thanks for the advice. The reference "Boisvert et al., 2016" is replaced with the

following reference in the revised manuscript.

Boisvert, L. N., Wu, D. L., & Shie, C. L. (2015). Increasing evaporation amounts seen in the Arctic between 2003 and 2013 from AIRS data. Journal of Geophysical Research: Atmospheres, 120(14), 6865-6881.

L53-59: I suggest to add that these possibilities to discriminate MYI from FYI by means of the different microwave signatures work well (only?) in winter when the snow cover is dry. In summer / during melt events, MYI and FYI often reveal a similar microwave signature.

Reply: Thanks. We add the following sentence to the end of these sentences (see L60-L62 in the revised manuscript):

"Note that MYI and FYI have such different microwave characteristics in winter but not in summer or during melt events when snow is wet, which leads to similar microwave signatures of the different ice types."

L66: I suggest to remove Brath et al as this is about helicopter-borne scatterometer measurements and not about satellite data classification.
I Suggest to remove Hughes as this is grey literature (and relatively old as well).

Reply: Thanks. It is modified as suggested.

L75: Isn't the ECICE algorithm cited here using Shokr et al., 2008 belonging to the other type of algorithms (SITC)?

Reply: Yes. It was a wrong citation here, we have deleted it as suggested.

L83-85: "some areas ... of MYI in ice charts" --> I would be careful with this statement because after all, what is done in the binary assignment or classification of a grid cell as FYI or MYI bears the same potential for over- or under-estimation of the actual fraction of the respective ice type. Hence, in light of this retrieval uncertainty it is perhaps ok to use data for the inter-comparison that have a similar drawback?

Reply: Agree. We therefore delete the statement of "overestimation".

L85/86: The sentence mentioning SAR could i) include the detail that SAR - like scatterometers - is an active microwave instrument but with a spatial resolution several orders of magnitude finer and ii) back up the information that SAR images are used for this kind of evaluation with literature.

Reply: The sentence is modified as suggested (see below). References are added to back up such information.

"Synthetic aperture radar (SAR) is an active microwave sensor as scatterometers but with several orders of magnitude finer spatial resolution. SAR images are also used to evaluate ice type classification accuracy (Ye et al., 2019; Zhang et al., 2019)."

L91: "for winters" --> just a comment: If you include the information that the discrimination between MYI and FYI works well by means of their different microwave signatures during winter suggested further up, then this notion of "for

winter" is a logical consequence of what is physically possible.

Reply: Thanks for the comment. We add statements regarding the different microwave signatures in winter and summer further up, according to the reviewer's comments.

Table 1: Please check the dates for SSMIS (2000 as the starting year is wrong) and for AMSR-E (12 as the end month is wrong).

Reply: Thanks. We double-check the dates and modify them accordingly (see new Table 1 in the revised manuscript).

L106: SMMR did not use a conical scan. Please check the respective documentation. One paper where you find helpful information about the series of SMMR, SSM/I and SSMIS sensors is this one: https://essd.copernicus.org/articles/12/647/2020/

Reply: Thank you for the advice. We double-check the specifications and modify the sentence below:

"It provides five-frequency, dual-polarized (ten-channel) Tb observations with an average incidence angle of 50.3°."

L113: In view of the frequencies listed in Table A1, don't you think it makes sense to state in one sentence that for the generation of the SITY products introduced in section 2.2 merely the near 19 and near 37 GHz channels of these instruments are used? This would explain at the same time why you only list these two frequencies in Table A1.

Reply: Agree. We add one sentence to the end of the first paragraph in section 2.1.1.

"Specifications of the different sensors are shown in Table A1, where only the channels used in the SITY products in Section 2.2 are listed."

L119: Note that there have been two different ERS satellites ERS 1 and ERS 2, both equipped with that AMI instrument, but only ERS1 started operations in 1991. You might want to make this clear in your text.

Reply: Yes, there are two ERS satellites. We meant both of them. In the revised manuscript, "European Remote-sensing Satellite (ERS)" is modified to "European Remote Sensing (ERS) satellites (ERS-1 and ERS-2)".

L122: OSCAT: There have been more than one OSCAT sensor. Which one is this?

Reply: Thanks for the advice. It is the OceanSat-2 scatterometer, we therefore replaced "OceanSat Scatterometer (OSCAT)" with "OceanSat-2 Scatterometer (OSCAT)" in the manuscript.

L128: While I agree that there is a linear relation between TB and emissivity I doubt that from the physical / mathematical sense this statement is correct. I'd rather see it that way: TB is linearly proportional to the physical temperature of the object and the emissivitiy is the proportionality factor which magnitude is determined by the physical / chemical properties of the material that are relevant for its microwave emissive behavior.

Reply: Thanks. The sentence is modified as below:

"…, which is linearly proportional to the physical temperature of the object, where the proportionality factor, the emissivity, is determined by the dielectric properties."

L130-138: This is a good starting point, containing already parts of the relevant information. I suggest to

i) separate FYI better from MYI in your explanation and correct the statements of how salinity changes [brine can only be "expelled" towards the surface of the ice, a process that if at all occurs during very cold conditions and for rather thin ice; most of the change in the brine content is via gravity drainage during winter and via flushing out the brine by meltwater during summer - which is the process where brine pockets become air pockets (or air bubbles as you write). These processes should be explained more correctly.]

ii) try to make clear what the role of snow might be.

iii) try to make clear which of the geophysical properties determine changes in the emissivity on the one hand and in the backscattering properties on the other hand - also taking into account frequency dependence and different polarizations.

Reply: Thank you for the advice. Modifications are made as suggested.

i) The word "expelled" is replaced with "rejected" and "bubbles" is replaced with "pockets". In addition, we rephrase the last sentence to better explain the parameter of GR3719.

ii) we add two sentences to explain the potential effect of snow (see L147-L149).

"The snow over sea ice also influences the emissivity. The addition of dry snow on the ice leads to reduced emissivity because of the increased scattering in the snow volume, while the moisture in a wet snow cover results in increased emissivity (Shokr and Sinha, 2015)".

iii) the backscattering properties are explained in the next paragraph, and we think the current updated information is sufficient to understand the principles of sea ice type classification, therefore did not add more sentences as mentioned above but rather add some references for more details.

The last sentence where you cite Vant et al. is not clear. What is meant by "demonstrated" in this context and what are "high frequencies" in this context?

A very good review of such properties is given in the book by Carsey, F.D., "Microwave Remote Sensing of Sea Ice" from 1992 which is available online.

Reply: The suggestions are highly appreciated. We have rephrased the sentence including Vant reference and moved it up in the paragraph. As suggested, we have included a reference to the relevant chapter 4 (Eppler et al.) in the Caesey book.

L140: Certainly GR is used in SITY products (I am not sure about PR but fine) ... but why? What is the advantage of using a GR over a TB? Ideally this is going to result from the revised section about emissivities and TBs before line 139.

Reply: Thanks for the advice. We add descriptions regarding the advantage of GR, and modify the original sentence as below:

"The emissivity is an intrinsic radiometric property of the material, but brightness

temperature is not (Shokr and Sinha, 2015). For this reason, polarization ratio (PR) and gradient ratio (GR) are usually used instead of Tb because they are independent of the physical temperature."

L143/144: You need to state that the subscript "p" in Eq. 2.2 stands for polarization and can either be H or V. You also need to state clearly which of the two frequencies f1 and f2 given references to the higher frequency - unless this is not (why not?) important.
Reply: Thanks for the advice. We have included the definition of the variables {h, v, f1, f2, and p} at the beginning and specified that p can be either h or v.
"PR is the normalized difference between the horizontally (h) and vertically (v) polarized Tbs for the same frequency (f), whereas GR is the normalized difference between Tbs at two frequencies (f1, f2) at the same polarization (p) which can be either h or v."
As the order for f1 and f2 only cause a difference in the sign, this is not important for the discrimination ability. Difference products define GR differently but the outcome is the same. We have added a sentence on this in the end of the paragraph.

L147-152: "Meanwhile ... SITY product" --> Also this I rate as a good start but the description needs to rephrased still to be clear enough. I recommend to
i) clearly state what the normal radar backscatter behavior is for FYI and MYI and why and only then ii) point out in which way deformation would change this general view. Please work with expressions such as surface scattering (dominant for FYI) and volume scattering (dominant for MYI).
When it comes to differences between frequencies, polarizations, etc. it makes sense to introduce (perhaps again) penetration depth of microwave radiation into sea ice of different types. I again warmly recommend to take a look at the various chapters of the Carsey book mentioned above.
Reply: Thank you for the advice. To explain the different backscatter of MYI and FYI more clearly, we rephrase the paragraph and added descriptions as suggested. See L161-L172 in the revised manuscript, e.g.
- *"… surface scattering is therefore the dominant scattering mechanism of FYI"*
- *"… air pockets within the subsurface layer of sea ice contribute to a higher volume scattering, which is dominant for MYI."*
Since the penetration depth can only explain the effect on frequency but not for polarization and observation angle, we add such descriptions in the previous paragraph, where descriptions of Tbs are given (see L143-144 in the revised manuscript).

L153: I might be wrong, but your description above is rather qualitative and I did not really get which specific "signatures" characterize MYI and which FYI. Does FYI exhibit a higher or lower emissivity than MYI? What about the PR? What about the GR? What about the radar backscatter coefficient? I would say that there is enough literature out to - for this first ever inter-comparison of SITY products based in algorithms that use these specific signatures - provide a summary table of relevant signatures at C- and Ku-Band for scatterometry and at near 19- and 37 GHz for

microwave radiometry.

Reply: Thank you for the advice. The word "signature" is not appropriate here thus cause misunderstanding. In the revision, we use "differences" instead. Otherwise, we hope that sufficient clarifications have been included in the revised manuscript.

L154: "top layer" --> Do you have an estimate of the vertical dimension you are talking about? Are these millimeters? Centimeters? Does this behaviour depend on the frequency?

Reply: Thanks for the advice. It is millimeter's scale, which depends on frequencies, but the difference is very small. To give clearer description, the sentence is modified to: *"…, when microwave radiation can only reach the top layer (from several to tens of millimeters) of melting snow".*

L176-178: Could it be that the Maaß and Kaleschke 2010 reference solely applies to the land-spill over correction and the Wentz 1997 one solely applies to the RTM based correction of the atmospheric noise? Please check and revise. Also "land spill-over due to the influence of land" should perhaps simply read "land spill-over effects on the measured TBs"

Reply: Yes, it was a typo here. We revise the citation here as suggested, and modify the "land spill-over due to the influence of land" to "land spill-over effects" in the revised manuscript.

L184: "Tb observations" <----> "using the classification parameter GR" --> It is not entirely clear whether and if so why the training dataset comprises TB values or GR values.

Comment: The above-asked-for table with typical values of the various input parameters taken from literature would assist very well here in understanding that this is obviously a rather simple approach involving the GR and a GR-threshold value including its typical variability for MYI and FYI.

Reply: "Tb observation" here meant "GR observations". To avoid misunderstanding, we replace it with "GR_37v19v observations" in the sentence.

Regarding the comment (and the above-asked-for table), we still think it is unnecessary to include such details in this manuscript. The aim of this study is to inter-compare the eight SITY products, and analyze their performance based on the general principles. Documenting all the detailed info is out of the main scope of our manuscript.

L201: "reassign misclassified FYI" --> Sorry, not entirely clear. Is this filter taking care of pixels that are erroneously classified as MYI but are in fact FYI? Or is this filter taking care of pixels that are erroneously classified as FYI but are in fact MYI?

Reply: We meant to take care of the pixels are erroneously classified as FYI but are in fact MYI. To avoid misunderstanding, "reassign misclassified FYI" is modified to "misclassified FYI back to MYI".

L204: Which polar stereographic projection?

Reply: It is the NSIDC Sea Ice Polar Stereographic North Projection. We modify it accordingly in the revised manuscript.

L214: The reader might like a hint why OSISAF-SITY uses a slightly different GR than most other products based on the GR at 37 and 19 GHz channels.
Reply: GR_19v37v = minus one * GR_37v19v. These two parameters are in fact the same for any of the classification algorithm. That OSISAF-SITY uses GR_19v37v is for historical reasons and has not been changed since the beginning of this operational production. We add a footnote comment on this difference in the manuscript:
"The parameter GR19v37v is identical to -GR37v19v. But the different definition of GR does not affect the final classification outcome."

L220: Given the fact that C3S-SITY used quite a number of filters and corrections it would be good to confirm here in the text that OSISAF-SITY only uses the geographical mask to correct for eventual misclassifications of pixels as MYI and does not perform any other filtering.
Reply: Thanks for the comment. Yes, we confirm that OSISAF-SITY uses the same filters and masks as for C3S-SITYs, except the air temperature correction scheme in C3S-2. A sentence has been added to clarify this:
"In the post-processing stage, OSISAF-SITY uses the same OW filters and masks as those in C3S-SITY, except the final air-temperature correction scheme introduced for C3S-2 to correct for misclassified FYI (Aaboe et al., 2021b)."

L227+: I note that here and further down you do not specify the form in which the scatterometer data are input into the pre-processing stage - unlike for the radiometer data where it is clear whether these are swath data or not and at which stage of the processing you compute a daily map. Please therefore consider to also here, for the scatterometer products, specify whether these products are based on swath data or daily gridded data or whatsoever.
Reply: These data are based on swath data. This information is added to the revised manuscript (see L249-L251).

L229: I suggest to back up this statement about the incidence angle dependency with a reference.
Reply: Thanks. Reference is added as suggested.

L233: "data of March of each year" --> Would you mind to also tell us about the geographic region from which these thresholds are determined?
Reply: Thanks for the comment. To specify the region, we add "in the Arctic" to the end of this sentence.

L234: I recommend to not speak about "MYI signatures" but of "grid cells" or "pixels erroneously classified as MYI". In addition I am wondering whether these pixels are really removed or whether they are set to ice type SYI or FYI. If not, then I assume the

resulting ice type maps may have gaps?

Reply: Thanks. For clearer description, "to remove the unphysical MYI signatures" is modified to "to set the erroneously classified MYI pixels back to FYI".

L247: Is it correct to assume that all radiometer data come as daily gridded maps? Please mention this accordingly in the text.

Reply: Regarding the question here, yes, it is ok to use daily gridded Tbs in sea ice type classification. We therefore do not understand what the reviewer is asking for. This sentence explains the priority of the different radiometer data in the product, whereas the next sentence introduces format (daily gridded) of the respective radiometer data.

L254: How is the re-gridding to the finest spatial resolution among the input data realized? Bilinear? Nearest-Neighbour? Others methods?

Reply: Nearest neighbour method is used in the re-gridding. We add "using the nearest neighbour method" to the end of the sentence.

L267: I don't find the description of the data set sufficient enough. It is in particular not clear how a discrimination into FYI (SYI?) and MYI is made? This information should be given as a minimum - together with the granularity (in time). About what temporal resolution are we talking here? Weekly? Monthly? Annually? It should become more clear that this data set offers the spatial distribution of different ice age classes (younger than 1 year, 1-2 years, 2-3 years ...).

Reply: Thanks for the advice. We add the following sentences to explain the discrimination of the ice age and the respective temporal resolution.
"Ice age (i.e. 1 year, 2 year, … and 5+ years) is assigned according to the number of winters the ice parcels have survived. The age of the oldest ice within the grid cell of each week is regarded as the weekly ice age."

L271/272: Not clear what you mean with "middle-of-the-road scheme". Also, the fact that satellite data are combined with buoy data (and data from atmospheric re-analysis by the way) does not apply to the SIA data set but applies to the sea-ice motion product that is used to derive the SIA data set. I suggest to rewrite this statement.

Reply: Thank you for the advice. We revise the sentence as below for better clarification.
"Due to the scheme of using ice motion data derived from combined satellite and buoy data, NSIDC-SIA…"

L278: I suggest to delete this sentence because this is based on observations in the Weddell Sea, Antarctica. If you want to provide information about the uncertainty of the kind of ice motion product used here I recommend to search for publications by Sumata et al., and take a look at the two publications by Lavergne et al., one in 2010 and one in 2021.

Reply: Thanks. We replace it with a reference from Sumata et al. 2014 instead.

L279-282: While the statements provided here are certainly correct they apply to an

older version of the ice motion product (v3) which in fact contained artifacts. The data you are using are based on version 4 and therefore I consider the discussion given in these lines as not relevant. These are also not connected to the issues Korosov et al. found; these point to methodological shortcomings of the derivation of the SIA product from whatever ice motion product.

What is lacking for the SIA product is an evaluation beyond the published inter-comparison study results. The reasons for this is clear - a lack of proper data sets that could be used as a source for evaluation - but it should nevertheless be made clear that we are still waiting for an adequate evaluation of the NSIDC-SIA data set. This is something you could (and should?) mention here because it explains well why you do not consider the SIA data set as an evaluation data set, i.e. kind of a ground truth.

Reply: Thanks for the advice. We acknowledge that the artifacts apply to the older version products. This is already specified in this paragraph,

"… and has been largely mitigated by tuning the interpolation approach in the current version (Tschudi et al., 2020)".

For better clarification of the NSIDC-SIA evaluation, we add the following sentences:

"Although an adequate evaluation is still needed for the current NSIDC-SIA product, the good consistency and recent upgrades of the interpolation approach make it a useful dataset for SITY comparison.".

Section 2.4: In which form did you get and use the SAR data?

Which (pre-)processing was already applied before / has been applied by you?

What is the spatial resolution in these SAR images?

From the description in L284 it seems clear that the RS-1 images are at HH-pol; it is however not clear what the polarization of the used S-1 SAR images is.

Unless you refer to a much more detailed description of how the SAR images were "visually interpreted" further down in your paper, I strongly recommend to also here not use the term "validation". So far it again simply seems to be an inter-comparison of the SITY products with another unvalidated product which accuracy is not clearly specified. Hence, in L285 and in L290 "validation" needs to be replaced with "intercomparison".

Reply: Thank you for the advice.

To answer the questions raised by the reviewer, we revised sentences in the first paragraph of section 2.4 to provide information regarding the form of data, spatial resolution and the polarizations (See L312-L314).

"The three RS-1 images are in ScanSAR Wide (SCW) beam mode with nominal resolution of 100 m, whereas those from S-1 are in Extra Wide (EW) swath mode at HH and HV polarizations with nominal resolution of 40 m. The RS-1 SCW products and the Level 1 Ground Range Detected (GRD) S-1 product are both obtained from the Alaska Satellite Facility."

Regarding the "validation" in L285 and L290, it is modified to "accuracy assessment in case studies" and "case studies", respectively.

L290: While one can guess it - would you mind to tell the reader the dates for which you used these additional data? Is it the same days as you have SAR images of?

Reply: Yes, they are from the same days as the SAR images. In the revised manuscript, the first sentence is modified to "Auxiliary data from atmospheric reanalysis is used in addition to the SAR images in the case studies.". Since it specifies "used in the case studies", we therefore do not give a list of the dates in the text.

L307-309: "To account ... " --> While this is an ok solution it introduces uncertainties from the re-projection. A more easy and straight-forward way would have been to use the grid cell sizes published by NSIDC for several derivatives of this polarstereographic grid

Reply: Thank you for the advice. Since the uncertainty is small from the re-projection and in theory does not affect the inter-comparison of this study, we decide to keep the processing as it is.

L323-330:
- This reads like a sufficiently good qualitative recipe to interpret SAR images visually. The only part that I find is missing is the discrimination between sea ice and open water - the latter also having a mixture of brightness levels in SAR images depending on i) the size of the openings with open water and ii) the wind speed. You may want to add this here to complete this description. Please take a look at my comments directly to Figure 3.
- Also, please add one sentence informing the reader that you are performing a binary classification here (in the next paragraph you write about Kappa coefficient and accuracy, so I assume to classify the SAR images in FYI, MYI and, if need be, open water).

Reply: Thank you for the advice.
- regarding the discrimination between ice and water, we add sentences as below for further clarification (See L355-L358).
"(4) Backscatter of OW is dependent on the surface wind. It is low under calm conditions and could be high when the wind speed is high (Area D in Fig. 9). The more homogenous texture and lower auto-correlation of OW backscatter could be used to discriminate water from ice in SAR image (Berg and Eriksson, 2012; Aldenhoff et al., 2018)."
- Regarding the binary classification, we add explanations to the end of this paragraph, the revised sentence reads as below:
"… used as additional information for the ice type interpretation from SAR imagery (i.e. classification of OW, FYI and MYI)."

L333-335: While the advantage of using HV over HH polarization is clear from the old literature cited, I strongly recommend to also take a look into the more recent literature and cite at least one paper where this has been applied to either RADARSAT SAR or Sentinel-1 SAR imagery.

Reply: Thanks for the advice. We added the following references in the revised

manuscript.

Dabboor, M., & Geldsetzer, T. (2014). Towards sea ice classification using simulated RADARSAT Constellation Mission compact polarimetric SAR imagery. Remote Sensing of Environment, 140, 189-195.

Song, W., Li, M., Gao, W., Huang, D., Ma, Z., Liotta, A., & Perra, C. (2021). Automatic sea-ice classification of SAR images based on spatial and temporal features learning. IEEE Transactions on Geoscience and Remote Sensing, 59(12), 9887-9901.

L337: Please add 2-3 sentences explaining what the Kappa coefficient is and how you compute an "accuracy" from two binary classified images. What is the credibility of these two parameters when computed for a pair of non-evaluated data sets?
Reply: Thanks. We add sentences and equations to explain the definition of overall accuracy and kappa coefficient (See L368-L376). Although the SAR interpretation results are not standard or perfect validation datasets, it is used for SAR ice classification results thus should be sufficient to evaluate coarse resolution ice type products.

L339-342: Given the nature of the two kinds of data products you use for the intercomparison of the SITY products I strongly recommend to change the wording away from validation, evaluation, and reference towards inter-comparison.
Reply: Thanks for the advice. We totally understand that it is not appropriate to call it "validation". However, the wording such as "evaluation, assessment, reference" is okay to use when we are using data from multi-sources to analyze its pros and cons.

L347: As noted in the context of Figure 4: I would get rid of the dashed lines and the daily data in this figure. It adds noise instead of value.
Reply: Thanks. The figure is modified as suggested.

L351: "due to ice divergence" --> I am wondering whether it would be worth to include a schematic illustration which shows what you mean by this.
Reply: Thanks for the advice. Considering the length of the manuscript and the relatively clear description, we think it is not necessary to include additional schematic illustration.

L356: "of FYI similar as that of MYI" --> It is exactly the other way round: the MYI signature becomes similar to the FYI signature. Melting snow, e.g., results in an increase of the microwave emissivity and hence an increase in the TB. Melting snow also results in a decrease of the penetration depth of microwave radiation so that a scatterometer does not sense the ice type underneath the snow anymore but only the snow.
Reply: Yes, the reviewer is correct. We modify the expression to "which leads to noise in the radiometric and scattering signatures therefore unsatisfactory performances of the SITY algorithm" in the revision.

L382-387: This paragraph targets the daily MYI extent values. I suggest to either delete it ... or keep it with the notion that these daily extents are not shown (see my comments with respect to Figure 4). Alternatively, you could show an example figure where you compare day-to-day MYI extent variations independent of the NSIDC SIA MYI extent.
Reply: Thank you for the advice. Figure 4 is modified as suggested. In addition, we add the notion that the daily extents are not shown in the manuscript.

L388/389: This sentence is reporting the maximum deviation - which is fine. You might want to include a statement about the average difference between the products during those periods of winter where all run comparably stable in their performance, i.e. November through March.
Reply: Thank you for the advice. We add one sentence to the beginning of this paragraph, which reads as below:
"Between any two SITY products, the average difference in weekly MYI extent varies between 0.02×10^6 and 1.92×10^6 km^2 in winter, with values below 1.11×10^6 km^2 during the periods from December to March. The largest difference in weekly MYI extent reaches…"

L396 / L402 / L413: Number of the figure is missing. --> GC1
Reply: Thanks. All the numbering of figures is added.

L414/415: "The former" refers to which region? As written it refers to CAO and ESL and then your statement is wrong because there is no southward flow of MYI from region ESL.
Reply: Thanks for pointing it out. We meant the CAO region here. In the revised manuscript, we replace "The former" with "The negative MYI trend in CAO".

L416: I suggest to delete the "following the Transpolar Drift Stream" because it is confusingly used in the context of what is stated here. For sure the TDS has nothing to do with MYI export through Nares Strait and also whether MYI is exported into the Barents Sea has in first instance to do with whether sea ice survived summer melt just north of the Barents Sea and is then pushed into it by northerly winds.
Reply: Agree. We delete it in the revised manuscript.

L422/423: "it is less pronouced in the CAO region." --> What is less pronounced? The MYI extent? Or the difference between NSIDC SIA MYI extent and the MYI extent of the SITY products investigated? Please be more specific in your writing.
Reply: Thanks. This sentence is modified as below:
"In comparison, such difference is overall smaller in the CAO region."

L423-425: I am not sure your argumentation holds the way written. I recommend that you are more specific about the process that actually leads to a potential (...) overestimation of MYI extent in the NSIDC SIA product. And I also recommend that you try to relate this potential overestimation also to the actual MYI fraction instead of

generally writing about a "mixture of MYI and FYI". You might ask yourself the question whether this potential overestimation is the same at 80%, 50% or 20% MYI concentration.

Reply: Thanks for the advice. To give more specific descriptions, we modify the sentence as below:

"This indicates that the mixture of MYI and FYI (and the medium MYI fraction), which leads to the "overestimated" NSIDC-SIA MYI extent because of the oldest ice age assignment, …"

Subsection 4.2.2: Because of my inability to link the text of this subsection adequately to the figures because numbers are missing for the latter, I do not comment on this subsection. I am sorry. --> GC1

Reply: We apologize for the missing numberings. We have included all the numbering in the revised manuscript and double check all the citation carefully.

L468+: Just again the comment that the way the SAR images are interpreted does - in my eyes - not warrant to use the term validation or evaluation. It is an intercomparison between two kinds of ice type maps. I therefore once again recommend to get rid of the terms evaluation and validation and switch to intercomparison.

Reply: As mentioned in previous replies, although the SAR interpretation results are not standard or perfect validation datasets, it is sufficient to evaluate coarse resolution ice type products, which is also used for evaluating SAR ice classification results.

In order to tone down the value, we avoid using the word "validation" in the manuscript. On the other hand, expressions of "evaluation" and "inter-comparison" are kept.

L478: It is a particular day in November. Therefore please mention it to avoid the impression that we are looking at a monthly average. The year given in the text does not match with the year given in the figure.

Reply: Thanks. We have corrected the year and add the date as suggested for all the cases in the manuscript.

L478-483: The way you interpreted this SAR image makes a lot of sense - still bears the chance that

i) the area you classified as FYI in the eastern parts of the image is actually indeed MYI or that ii) a substantially wider fringe of the area north of the open water is actually compressed FYI in form of brash ice, possibly undistinguishable from MYI under the conditions shown.

What I would like to state here is that the intercomparison between your SAR image analysis results and the SITY products is very difficult to put into a reliable or credible quantitative measure as you try to do with the Kappa and OA values. I recommend that you at least mention that because of the lack of contrast in backscatter in the SAR image shown the actual border between FYI and MYI might vary considerably. While currently KNMI-Q looks ideal in comparison to your SAR ice type map, another interpretation of the SAR map might have resulted in a FYI / MYI distribution that

resembles the C3S or OSISAF products closer. There is a lot of ambiguity - not just in the products but also in your SAR image interpretation which I recommend not to hide.

Reply: We understand the concern regarding the reliability of the SAR interpretation results. However, we would like to mention that the interpretation is performed on the high-resolution SAR image. We should not expect to see all the details in the figures presented in the manuscript. The contrast between the FYI and MYI may not be that high in this case, but it is sufficiently good for the discrimination of FYI and MYI. Therefore, instead of including statements such as "the actual border may vary considerably", we add notions which read as below

"Note that quality of the SAR visual interpretation could vary with images. The identified border between FYI and MYI could deviate more from the actual border when the contrast in the backscatter is lower for the different ice types (e.g. Case 1).

L508: Kappa and OA values for NSIDC-SIA are similar to Case 1 where you did not comment about the mobility of the ice even though that case is at the ice edge AND you have substantially higher winds. Here, in case 2 the MYI tongue is embedded into a matrix of growing FYI and I doubt that within the time frame of one week there was too much movement.

Reply: Thanks for the advice. We agree that the mobility of ice could partly explain NSIDC-SIA in case 1. We therefore add one sentence to the paragraph regarding case 1 (See L552-L553 in the revised manuscript).

"Yet such difference is nearly negligible considering their different temporal resolutions and the mobility features of sea ice."

As for case 2, although the wind speed is low for the day when the SAR image is acquired, it could be high in other days of the week. We therefore decide to keep the statement here for case 2.

L518-519: The failure of these products to detect MYI is really strange and difficult to understand. It involves both, a product only based on QuikSCAT data and products based purely on radiometer data. It might be a very stupid question from my side but did you double-check whether the re-projections that were involved in one or the other product to do the intercomparison did not jeopardize your results? You know, if it would only be the radiometer based products or only the QuikSCAT based products I would understand this failure ... but we are talking about old ice (according to the NSIDC SIA map) which has a clear signature in passive microwave imagery during winter.

Reply: We agree that the failure of these products is quite strange here. We double-check the re-projection (and other processing) during the analysis and confirms that it was correctly performed and should not be the "reason" for the "failure". Since the failure occurs not only for purely radiometer or scatterometer based products, the unusual radiometric and scattering signatures might result from changes of snow properties, which lead to radiometric and scattering signatures between those of MYI and FYI thus are sensitive to the thresholds in different SITY products and exhibit failure in different products. Despite being out of the scope of this manuscript, it is worth mentioning that the soon-to-be-released version 3 of C3S-SITY seems to capture

this MYI pattern, at least to some extent (being either MYI or Amb)

L521-523: Not sure whether inside the pack ice in winter the statement about moblity as a means to explain discrepancies holds.
Reply: Thanks for the comment. To tone down the statement, we replace "… are mainly attributed to…" with "… could be attributed to …"

L531-533: This failure to classify a large part of the ice as ice in the middle of winter is a no-go for such a product. Strange.
Reply: We agree with the reviewer. Reasons behind the strange performance in the IFREMER SITY product are unknown to us as well, which needs further investigation.

L536/537: "westward shift" --> This might be in fact one of the cases where the NSIDC SIA product "classifies" a grid cell as MYI ... even though the MYI fraction in some of the grid cells is certainly barely above the 15% threshold. Just a comment, no action required.
Reply: Thanks for the comment.

L537/538: Why do you highlight the ice age here but not for case 3?
Reply: Thanks for the comment. For case 3, the main performance of NSIDC-SIA is over- and underestimation in different areas. The SITY distribution in NSIDC-SIA is not that similar as the SAR image. We therefore did not highlight the ice age there.

L544/545: I am not overly convinced that the sea ice with this bright signature is associated with land-fast sea ice. This kind of ice is usually level ice with little deformation. In SAR images it often is represented as rather dark and homogeneous patches along the coasts / around Islands; actually the HH-SAR image of Figure 13 shows land-fast sea ice in the immediate vicinity of Severnaya Zemlya islands.
Reply: Thanks. We agree that it could be deformed FYI in area C. The sentence is therefore modified to "it is more likely to be land-fast ice or deformed FYI …"

L550/551: Not sure this statement about that this "lasts for the whole winter" is appropriate given that the SAR image is from April 27 and therefore at the verge to spring.
Reply: Thanks for the advice. We meant to say that the error occurs in early winter months and lasts for the whole winter. The sentence is modified as below in the revised manuscript.
"This error starts to occur in November and lasts for the whole winter."

L599: When does precipitation appear "suddenly"? What is "thick precipitation"? Please revise your wording; it is not clear what you want to state here.
Reply: Thanks for the question. We meant to say that wet or high snow precipitation could lead to significant changes of GR37v19v. The sentence is modified as below:
"…, or when wet or high snow precipitation appears"

L605/605: "As a results ... is used." --> How about the other way round? Isn't is as likely that MYI is misclassified as deformed FYI?

Reply: The discrimination between MYI and FYI is the relatively high and low backscatter. However, deformed FYI could have as high backscatter as MYI thus could lead to misclassification of the deformed FYI as MYI. Take case 4 as an example, we would say, it is more likely that deformed FYI is misclassified as MYI. We therefore keep the expression here.

L610: Have you looked at MYI ice cores? Did you find air bubbles of 2 cm diameter? I think this is an order of magnitude too large. I suggest that you dive into papers / book chapters describing ground- or air-borne active microwave measurements over different ice types and attempts to understand the reasons for the observed differences. The book Microwave Remote Sensing of Sea Ice by F. D. Carsey might be a good source - as well as papers dating back to the 1980s.

Reply: Thank you for the question. We read the books and references as suggested. It shows as the reviewer comment, that it is an order of magnitude too large. In the revised manuscript, we modified the sentence as below:

"This could be due to the fact that Ku-band scatterometer is more sensitive to the volume scattering in MYI."

L612: While in the context of brightness temperatures you refered to melt-refreeze cycles and the role of snow wetness this is missing completely here in this discussion. What happens to the radar backscatter signal if the snow gets wet?

Reply: Thank you for the question.

The dielectric constant increases as the wetness increases. The increase of the real part of dielectric constant (permittivity) causes increasing surface scattering, whereas the increase of the imaginary part (loss) causes decreasing volume scattering. The total backscattering decreases when the wetness increases from 0 to around 2% and stabilizes when the wetness continues increasing (Koskinen et al., 2000; Shokr and Sinha, 2015).
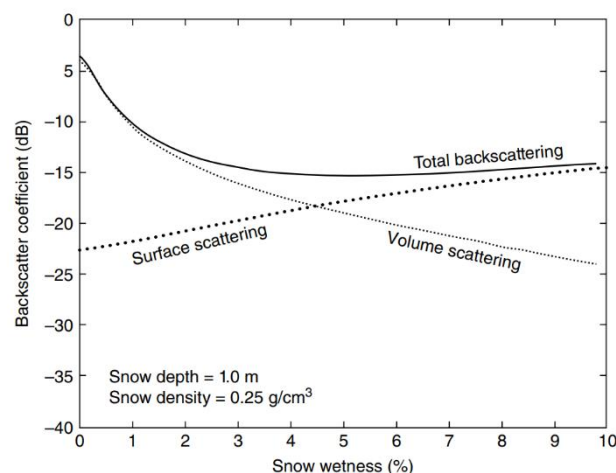
In the revised manuscript, we add the following sentences to describe the effect of snow wetness on backscatter.

"Factors such as snow wetness could also influence the backscatter of sea ice thus the efficacy. An example is given in Shokr and Agnew (2013), where the increase of snow wetness causes attenuated (decreased) backscatter of MYI and eventually leads to misclassification of MYI as FYI."

L617:
- Why this add-on "that is ice-free during summer"? What is the problem here?
- Also: "does not help" appears to be too global a statement. For the case you chose, the combination does not reveal advantages but there are enough other cases where it works very well. I recommend to tone this statement down by writing, e.g., "does not always help" or similar.
Reply: Thanks for the advice.
- It is unnecessary description here. We therefore delete it in the revision.
- It is modified as suggested.

L619: "the worst SITY classification" --> I can buy this statement for Zhang-SITY but while agreeably OSISAF-SITY is not correct IFREMER-A is clearly worse. Consider rephrasing please.
Note that April 27 is already a time of the year where solar radiation induced snow metamorphism can play a large role in shaping both passive and active microwave signatures of sea ice - even though 2m-air temperatures are still below zero.
Reply: Thanks for the advice. We agree with reviewer and delete "OSISAF-SITY" in the revision.

L645/646: "thus smaller microwave signature differences between MYI and FYI" --> I don't think that the Tschudi et al paper is an appropriate reference for this change in the physical properties of the multiyear ice when becoming younger. You might look into earlier work published by Comiso et al. or others in this regard.
Reply: Thanks for the advice. The citation here is not appropriate. We put the Tschudi et al., paper to the previous sentence, and replace it with the reference of Rivas et al., 2018.

L675/676: "surface processes such as wet snow attenuation and changes in brine salinity" --> You mix to kinds of processes here, it seems. One is the inter-action between surface properties and microwave radiation (the attenuation part) and one is a physical process (salinity change). I recommend to first give examples of the changes in physical properties and then write about the impact these changes could have on the microwave signature.
Are you sure that the salinity of the brine changes? Isn't it rather the bulk sea ice salinity or the brine volume?
Reply: Thanks for the advice. We meant the processes of snow metamorphosis and changes in bulk salinity of sea ice. The sentence is modified as follows:
"… due to processes such as snow metamorphosis and changes in bulk salinity of sea ice."

We understand that the reviewer would like to see more examples in the manuscript. However, considering the length and given references, we think it would be okay to make the statement in this way.

L682: "finer resolution does not ..." --> You might want to bear in mind that the 4.45 km grid resolution of the Zhang SITY product is based on using satellite data that were resolution-enhanced using the SIRF technology. Such a procedure does not improve the information content ... hence it does not matter how fine the grid resolution is, if the relevant information is blurry - not resolved properly at the original coarse spatial resolution - then it will remain blurry.

Reply: Thanks for the advice. It is not that appropriate to say finer resolution. We use "finer grid spacing" instead.

L683-684: Why mentioning the near-90 GHz channels here? Is it relevant? Are these used somewhere for ice type discrimination? If not then I suggest to delete this sentence.

Reply: It is used for MYI identification in some algorithms, e.g. Lomax et al., 1995. Since it is not exactly for ice type discrimination, we decide to delete the sentence.

Lomax, A. S., Lubin, D. and Whritner, R. H. (1995) The potential for interpreting total and multiyear ice concentrations in SSM/I 85.5 GHz imagery, Remote Sensing of Environment, 54(1), pp. 13–26

L690: "quantitatively" --> for me the inter-comparison of visually interpreted SAR images with the SITY maps is not a quantitative evaluation, also the computation of a Kappa coefficient does not help in this regard. I therefore again recommend to tone down your statement towards inter-comparison.

Reply: As mentioned in previous replies, although the SAR interpretation results are not standard or perfect validation datasets, it is sufficient to evaluate coarse resolution ice type products, which is also used for evaluating SAR ice classification results. We therefore would like to keep the statement here.

L693-695: Please make clear whether the ranges of years denoted together with the two products selected refer to the period for which these products are available or to the time period over which you have averaged the differences.

Reply: It refers to the respective available period of the products within the range of 1999-2019. To make clear about the meaning of years, we modify the sentence as below in the revision.

"The bias of MYI extent between the SITY products (during the different periods) and NSIDC-SIA varies from -1.32×10^6 km^2 (OSISAF-SITY, during the SSM/I-only period, 2006–2009) to 0.49×10^6 km^2 (KNMI-A, 2009–2019)."

L701: I find the numbered list of paragraphs that follows ok in terms of where which products seems to perform good or not so good. I don't think, however, that you should mention your attempts to explain the differences in the performance because i) your inter-comparison is based on a small set (5) of qualitatively interpreted SAR images

(which would be classified differently in an independent follow-on study), resulting in qualitative statements and because ii) you did not investigate / show evidence of misclassification of ice types due to the three mentioned main influencing factors. Also here you remain rather descriptive and do not go into depth. Hence most of the "explanation" given here is rather of hypothetical nature which proof requires further work.

Reply: Thank you for the comment.

We double-check these paragraphs. Most of the sentences are descriptions of their performances (good or not so good during certain period or under certain conditions), which the reviewer (and we) regarded as okay and are in fact based on the "qualitative" comparison with the SAR interpretation results (along with NSIDC-SIA).

When explaining these differences, we use words such as "are likely to" "could be resulted from" "could be caused" considering the uncertainties as mentioned above. This study aims to inter-compare the products and provide sensible explanations. We therefore do not quite understand the comment here and decide to keep as it is.

L723+ --> A lot of what follows here is a repetition of what is written in the discussion. Please condense and only provide 2-3 key points which you can also very well back up with the results you obtained. Make sure to highlight the nature of the results (qualitative / intercomparison) and give the outlook towards what would be needed to carry out a quantitative intercomparison or even evaluation. This is in my eyes much more important than, as you do at the end, to highlight which future satellites might be available. First we need to find procedures and well-evaluated data sets for a quantitative evaluation of the ice type products. Without these any novel ice type products from new satellites will be as useful (useless?) as the existing ones.

One obvious step would be to improve sea-ice motion estimates to improve the sea-ice age product - ideally with a smaller temporal resolution. And then: evaluate it adequately.

Another step would be to use well-evaluated ice type information from SAR images that underwent an unsupervised classification. Such information would make that part of the inter-comparison work involving SAR much more credible and the results could potentially even be interpreted quantitatively.

Reply: Thanks for the advice. We have condensed the texts as suggested and kept the sentences highlighting the results only. In addition, we add sentences regarding the well-evaluated dataset at the end of the manuscript. See L795-798 in the revised manuscript.

Typos / Editoral comments:
Abtract:
L17 / L18: Specify the QSCAT and the ASCAT periods as many readers will not be aware of these.
Reply: Done.

L17 / L18: "agree best" and "perform the best" --> Consider rephrasing such that you provide actual numbers of over- or underestimation.
Reply: the respective biases are added here, ", with smallest bias of -0.001×10^6 km^2 in FYI extent and -0.02×10^6 km^2 in MYI extent, respectively".

L20: "their performances" --> What is "their" referring to? Products? Sensors? Algorithms?
Reply: "their performances" is replaced with "the performances of the SITY products".

L38: I suggest to add the reference to Tschudi et al., 2020 to the one of Maslanik et al.
Reply: Done.

L42: I suggest to delete "such as the Atlantic Meridional Overturning" because this is a circulation in the ocean.
Reply: Done.

L51: "forecasting(Jung" --> "forecasting (Jung"
Reply: Done.

L64: "... other on input ..." --> better "... other in terms of input ..." or "... other regarding input ..." or "... other with respect to input ..."
Reply: Done.

L73: "It is found the ..." --> "The ..."
Reply: Done.

L77/78: "there is rarely ..." --> suggest to delete this part of the sentence since it is in contradiction to the previous one and begin this sentence with "The performances ..."
Reply: Done.

L90: "among the SITY" --> "among some existing SITY"
and further : "and give comprehensive evaluations on the ... " --> "and to assess the quality of the"
Reply: Done.

L113: "SSMR" --> "SMMR"
Reply: Done.

L119: "measures" --> "measured" because it is not operating anymore.
Reply: Done.

L123: "for inner ... from ..." --> "of 48.9 degree and 57.6 degree for the inner HH-polarized beam and the outer VV-polarized beam, respectively, from ..."
Reply: Done.

L124: "antennas, whose incidence angles varies between 25 degrees and 65 degrees" --> "antennas, each measuring backscatter over the incidence angle range of 25 degrees to 65 degrees."
Reply: Done.

For consistency with the previous subsection you could add the periods of operation for QuikSCAT and ASCAT.
Reply: Thanks. The periods of QSCAT and ASCAT are specified in the revised manuscript. See L128, ", which operates from July 1999 to November 2009", and L130, *", the data of which is available from May 2007 to present"*

L146: "microstructure" --> such as density, grain size and orientation ...? Consider adding.
Reply: We replace "and microstructure" with "density and grain size"

L153: "it" --> "these ice types"
Reply: Done.

L170/171: "and the climate record covered the period 1979-2020." --> "and was updated until 2021, covering the period 1979-2020."
Reply: Done.

L172: "is not be" --> "is not"
Reply: Done.

L180/181: "The swath data ..." Since this seems to be the final step of the pre-processing I would formulate it that way. For example: "As the last step of the pre-processing the corrected TB swath data are gridded into daily 25 km EASE2 grid TB maps." --> using which kind of gridding?
Reply: the sentence is modified to
"As the last step of the pre-processing, the corrected Tbs swath data are gridded into daily 25 km EASE2 grid Tbs maps using an equal-weighted average (also called a circular top-hat averaging window) of data within a radius from the grid centre(Lavergne et al., 2022)."

L194: "Water" --> "water"
Reply: Done.

L209: "is switched to AMSR-2" --> "was switched to AMSR2"
Reply: Done.

L210/211: Since the pre-processing of the C3S-SITY product is also done on the swath data I suggest to reformulate accordingly: "Unlike C3S-SITY, the core Bayesian ..."

Reply: Done.

L224: "KNMI-A respectively, available during" --> "KNMI-A, respectively, available for"
Reply: Done.

L268: "and active" --> delete, this is wrong. Such data are not used in that product.
Reply: Done.

L270: I would not call the work cited here as "assessments" and suggest to rephrase along the lines: "has been shown to provide very useful additional information about the changing Arctic sea ice cover because ..."
Reply: Done.

L274: "and ice motion data" --> "and the quality of the ice motion data."
Reply: Done.

L284: "Two" --> "two"
Reply: Done.

L288: "acquiring" --> "acquisition"
Reply: Done.

L297-299: "... Basin and limited ... . Note ... and analysis." --> I suggest to shorten this by writing: "... Basin excluding the area north of 87 degrees North with its observation data gap due to the satellites' inclinations (see Belmonte Rivas et al., 2018 and Fig. 2)."
Reply: Done. The sentence is replaced with
"…Basin excluding the area north of 87 degrees North with its observation data gap due to the inclination of satellites (see Belmonte Rivas et al., 2018 and Figure 2)."

L299: "as the integral extent of pixels" --> "as the sum of the area of all grid cells"
Reply: Done.

L303/304: "by the integral extent of pixels" --> "as the sum of the area of all grid cells"
Reply: Done.

L309-311: "Besides ... " --> "In order to compare the MYI extents at the same temporal resolution, SITY product MYI extents are averaged weekly to match the temporal resolution of the NSIDC SIA MYI extent."
Reply: Done.

L331: "to the UTM projection" --> "to the respective UTM projection" as this varies with longitude and latitude.
Reply: Done.

L364: "with bias" --> "with a bias"
Reply: Done.

L366 and L367: "bias" --> "biases"
Reply: Done.

L373: "scatteromter" --> "scatterometer"
Reply: Done.

L375: "product are" --> "products are"
Reply: Done.

L393: "difference" --> either "the .... difference" or "differences". Also note the wrong superscript "e" of the "10" in this line.
Reply: Done.

L398/399 and L400: "deviation of" --> Either it is a deviation of a quantity from another quantity ... which is not the case here, or it is the "deviation between" different products ... which is the case here. Please rephrase accordingly, because what I guess you want to tell is that, e.g., in mid-winter the difference of the MYI extent derived from the different SITY products is small - or in other words: The deviation of the MYI extent between the different SITY products is small.
Reply: Done.

L404: "For the" --> "Regarding the" or "With respect to the"
Reply: Done.

L404: What do you refer to by "the latter"? Is it "the other products"? Or are you refering to OSISAF-SITY?
Reply: "the latter" is used to refer to OSISAF-SITY, to avoid misunderstanding, we replace "the latter" with OSISAF-SITY.

L405: "mild ... trend" --> weather can be mild ... how about "weak ... trend" or "small ... trend"; furthermore: "rapid ... trend" --> rapid has something to do with speed and time, hasn't it? How about "large ... trend"?
Reply: Done.

L414: "while that ... increasing". --> either: "while the trend in the BCS regions is either zero or positive" or: "while the MYI extent in the BCS region remains constant or is increasing."
Reply: Done.

L418: "torwards south" --> either: "towards the South" or "south"

Reply: Done. It is replaced with "south".

L419/420: "In the BCS region, ... out of this region ... from the CAO region." --> consider rewriting this sentence. Certainly MYI drifts in the BCS region following the Beaufort Gyre. It enters the BCS from the North along the CAA and it eventually exits the BCS westward into ESL or back northward into CAO at the western borders of the BCS region - eventually entering the TDS. It also simply melts there (in summer).
Reply: Done. This sentence is revised as below:
"In the BCS region, large quantities of MYI enters this region from the north along the CAA and eventually exits BCS westward into ESL or back northward into CAO at the western borders of the BCS region."

L421: The statement of an increasing MYI extent in the BCS region should be supported by a notion of seasonality. Most likely MYI extent in this region is at a minimum in September and increases towards winter by MYI drifting into it from the North.
Reply: Thanks. The sentence is modified as below:
"The nearly constant or increasing MYI extent in the BCS region could be caused by the fact that the MYI extent in BCS reaches a minimum in September and increases toward winter by MYI drifting into it from the north."

L493: Again, please enclose the full date.
Reply: Done.

L498: "As shown in ..." --> I guess that fact that Table 3 shows up here was not planned?
Reply: Thanks. It was a wrong format here. We have revised it accordingly.

L505: "of SAR image" --> "of the SAR image"
Reply: Done.

L506: "to the case in" ... something missing here?
Reply: Thanks. We meant the case in Figure 8. It is modified accordingly.

L525: "as" --> "is"
Reply: Done.

L549: "nearly" --> "near"
Reply: Done.

L552: "northeast of the image" --> Did you mean "in the northeastern part of the image"?
Reply: Yes. It is modified it accordingly.

L559/560: What do you mean by "distinct"? Do you perhaps mean "different"?
Reply: Yes. "distinct" is replaced with "different" in the revision.

L561: "as a cross-validation dataset." --> "as an inter-comparison data set."
Reply: Done.

L564: "o of" --> "of"
Reply: Done.

L594: "...capability to separate and physical ..." --> Please check this sentence; its meaning is not clear.
Reply: Thanks, the sentence is modified as below:
"The efficacy of input parameters depends on their separability of sea ice types and the relevant sea ice physical properties."

L598: Over sea ice I would speak of melt-refreeze cycles only and hence remove the "wet-dry cycles".
Reply: We would like to keep the statement of "wet-dry cycles" since the change of snow wetness does not necessarily mean "melt-refreeze".

L602: "in C3S-2" can be deleted here.
Reply: Done.

L604: "disparate" --> I know, this comment comes somewhat late in my review but I recommend to check the meaning and usage of "disparate" versus the meaning of "different". I doubt that the microwave and/or scattering properties of MYI and FYI can be termed "disparate". They are different but they share similar (but different) basic scattering mechanisms. Unless you can be sure that the differences are so substantial that they exclude each other, i.e. absolutely no volume scatter in case of FYI or absolutely no surface scatter for MYI, or the like, I recommend to always rather speak of "differences" - here and everywhere (...) else in the paper.
Reply: The word "disparate" is replaced with "different" or throughout the manuscript.

L623: Either "on an a priori training dataset" or "on a priori training datasets"
Reply: Done. It is revised to "on a priori training datasets".

L625: "dataset" --> "datasets"
Reply: Done.

L639: "variabilities in the" --> I guess "variabilities as in the" is better.
Reply: Done.

L642: "fails to identify narrow MYI tongue in peripheral seas" --> "fails to identify features such as a narrow MYI tongue often observed in the Arctic peripheral seas ..."
Reply: Done.

L662: "aims to reassign the erroneously classified FYI" --> Not clear, better: "aims to re-assign the ice type MYI to grid cells where MYI was erroneously classified as FYI"
Reply: Done.

L673: "the five series SITY products" --> ?? What series?
Reply: "five series" is deleted to avoid misunderstanding.

L686: "... especially the fraction of MYI. The change of the SITY ..."
Reply: Done.

L687: "... inter-comparisons and analyses of SITY products ..."
Reply: Done.

L689: Please state that the NSIDC-SIA product is a weekly one and that you averaged (?) all SITY products to the same temporal resolution before the comparison.
Reply: We add "daily" and "weekly" to describe the SITY products and NSIDC-SIA product, respectively.

L719 "disparate" = completely different? Really? See my previous comment about usage of "disparate".
Reply: It is replaced with "different" as mentioned above.

Table 1:
"SSMI/I" --> "SSM/I"; "AMSR-2" --> "AMSR2"; "ASMR-E" --> "AMSR-E"
In the text you describe OSCAT but I cannot see it used in any of the products listed here. Consider removing it in the text?
You state in the text that all SITY products provide daily estimates. You can therefore delete the column denoted "Frequency".
I suggest to change "grid size" into "grid resolution". Is the type of all grids the same (all EASE or all polarstereographic)? If not it might make sense to add a column where this is specified.
Reply: Thanks. The table is modified as suggested.

Table 2 footnote: "to verify ... open water" --> better "to assess the correct discrimination of sea ice from open water."
You might want to explain also what "ice motion confining" means.
Reply: The footnote is modified to "Filters based on gradient ratio and brightness temperatures are used to filter out the open water pixels".
Notes are added to explain the "ice motion confining and spatial filtering".

Fig. 3:
- Please for all scenes add a date.
- Do all scenes have the same spatial scale? If not please provide a scale along with every scene, if yes provide it once.

- I suggest to include in (b) that the bright features may also be due to openings in the ice cover under high wind speed conditions.
- I suggest to add in (f) that these are MYI floes in a matrix of younger, presumably FYI.
- Whether (c) indeed shows brash ice between ice floes depends a lot on the location and the season (which are unknown?).
Reply: The Figure and notes are modified as suggested.

Fig. 4:
- I suggest to delete the dashed line with the daily MYI extent. It does not add value to the figure now that you have the weekly average - rather it adds noise.
- "the shaded area represents" --> "the shaded area in the same color as the respective solid line represents"
Reply: The figure and notes are modified as suggested.