Response to RC1:

We thank the reviewer for the critical eye and for the very detailed and constructive criticism.

General Comments

Here is our response to the four general comments (GC1 to GC4)

GC1: Physics

It is true that sea ice physics is different between the two polar regions. We agree with the reviewer on the 3 listed aspects, namely the MYI age and formation, the dynamic environment of the sea ice and the differences in the snow depth and composition. There are more aspects manifested in ice motion patterns, ice kinematics (particularly lead formation) and dynamics of marginal ice zones. However, characterization of those differences is beyond the subject of this study. Differences in ice physics between the 2 regions warrants a separate study, perhaps in the form of a review paper or a book chapter. In this regard, we would like to assert that little is known about sea ice physics in the Antarctic compared to the Arctic. Just a quick look at the literature is enough to prove this point. Most books on Arctic sea ice have an opening chapter on ice physics. This is not the case with the only book on Antarctic sea ice (published in 1998, edited by M Jeffries). Nevertheless, a summary of the major differences between gross properties of sea ice and snow in both polar regimes can be compiled and included in the Introduction section though this would increase the length of the manuscript, which we are hesitant to offer.

The objective of the study is stated as to utilize the ECICE algorithm to quantify the concentration of each ice type (YI, FYI and MYI) with a focus on MYI in the Antarctic region (similar to what has been done for the Arctic ice). It is true that the physics of ice and snow impacts the observed signal (from passive and active microwave sensors in this case) but we don't have to dig in to address the questions of how and why. All what we need to do is to provide the algorithm with a realistic probability distribution of each observation used in the processing for each ice type (without having to explain the physics behind them). This is already covered in Section 2.3. This is no different than the description of other commonly-used sea ice concentration algorithms that use tie points to represent the used radiometric parameter for a given ice type (e.g., NASA Team, Bootstrap, ASI, etc.). In the original papers of those algorithms there is no description of the physics that engender the tie points. Only sampling from homogeneous area of each ice type was needed. This is what we have done in this study (and in the application of ECICE in the Arctic), namely relying on samples from representative area of the given ice types. ECICE has the advantage of using the probability of occurrence of all possible radiometric values from a given ice type, not just a single tie point.

Finally, yes … the snow cover affects the observed signal in ways that we do not fully understand. That is why using the probability of occurrence of all possible values from a certain ice type when covered by snow under different conditions becomes necessary (the advantage of ECICE as mentioned above).

GC2: Previous work

The reviewer has raised a few points under this title. We can provide a few statements in the Introduction to point out the benefits of ice type mapping to the modelling community. However, we see no need to discuss the sea ice type mapping in the Arctic (we already provide reference to the ECICE application to the Arctic ice). Less attention has been paid to the Antarctic sea ice simply because there are no economic or geopolitical benefits attached to it. A case in point is the yet unclear impact of global warming on the Antarctic sea ice, when so much information (observation and modelling) is readily available about the impacts on the Arctic ice.

As for the current knowledge about Antarctic sea ice type distribution, we will provide more information in the Introduction but not about the typical emissivity, brightness temperature and radar backscattering as suggested because we see this as distraction from the objective of the manuscript. One final point in response to the comments under the GC2 section: ECICE was not modified to apply it to the Antarctic ice. Only the suitable input data had to be used.

GC3: Description of the methodology

We will complete the description. There is no difference in the application of the method between Arctic and Antarctic. The difference is in the input data. We shall further clarify how the sampling to compile the probability density functions of the radiometric and backscatter parameters from areas representing ice type was achieved.

GC4: Description and interpretation of the results:

The reviewer raised the point of background information about the difference in physical properties between Arctic and Antarctic sea ice. We addressed this point in GC1. The reviewer also suggested that it was mandatory to include an expert on Antarctic ice and snow to appropriately interpret the results.

We would like to offer the following arguments. The algorithm is about identifying and quantifying the 3 ice types in each resolution cell of the data based on the input probability density function of the used radiometric values for each ice type. If the input is wrong the output is wrong. Therefore, what is needed from an expert is help to identify authentic samples of a given ice type to construct the distributions. We will check this point, though not raised by the reviewer.

The actual suggestion of the reviewer is to use the knowledge of the expert to support the conclusions from the results. The reviewer presents 3 good themes as examples: the spurious MYI in the Weddell Sea and Ross Sea, the likelihood of leaking YI signature into FYI, and the handling of the iceberg signature, which confuses the identification of MYI. We will think about including an expert coauthor to help addressing these issues if we cannot address it fully.

As for the situations where the sum of the 3 ice types do not sum up to 100%: Note that the (uncorrected) concentrations of the three ice types *plus* the open water fraction add up to 100% (which is guaranteed by the equality constraint in ECICE (see L111-113) whereas the sum of the three ice type concentration adds up to the total ice concentration which can be 0% to 100%. See also the specific comment on that, item (11.), below.

Finally, yes, in this study we aimed at qualitative evaluation of the results because we thought that this would be appropriate for a first study to apply a new technique to the Antarctic sea ice. The purpose is to “prove the concept” rather than provide a comprehensive data set on ice type distribution for use in models. We will work to complete the description of the data used for the evaluation and try to incorporate some quantitative evaluation in certain areas where more information is readily available in the literature, e.g., MYI in the Weddell and Ross Seas.

Specific comments

Note that we quote the reviewer's specific comments and suggestions in red and have numbered them.

(1.) Line 33-34 "... but it is ... satellite data" is perhaps an a bit too general statement which i) could be specified better by telling the approaches used of doing so (aka: using instantaneous microwave observations) [note: using multi-annual time series of satellite data would work as well], perhaps by including the work of Comiso et al. (2011?) who figured out the differences in the signature of Arctic SYI vs. MYI, and which ii) could be amended by the fact that sea-ice age data retrieved for the Arctic (but not the Antarctic) are based on ice motion data which are in fact derived using satellite data. Hence it IS possible but nobody looked into it yet.

We will say "...using satellite data directly" and consider extending this a bit along the lines of the above comment.

(2.)L43-50: This paragraph is meant to provide the fundament for why ECICE needs some form of adaptation when applied to Antarctic sea ice. In that respect and given that this paper is the first attempting to derive partial ice-type concentrations for Antarctic sea ice, it would make a lot of sense to provide an adequate review of the difference in the sea ice AND snow properties year-round between the Antarctic and the Arctic that is back-up very well by a convincing set of references. This paragraph does not fulfil that role and should be re-written. --> GC1 / GC2

See our answers to GC1 and GC1

(3.) L47-49: "The ice cover ... The turbulent ..." --> I encourage you to provide 1-2 references each that underline these statements - particularly the notion that Antarctic sea ice is rougher - but also the evidence that the sea-ice structure is often different in the Antarctic compared to the Arctic.

We will elaborate a bit on that and include references.

(4.) L53/54: "Beside MYI ..." --> It would be very important to underline that in fact a substantial amount of the MYI along the East Antarctic coast is actually fast ice. This is often true (older than 2 years old) multi-year ice and is of even larger importance for the ecosystem and has effects on buttressing the ice shelves.

We will be more specific here.

(5.) L55-56: "pancake ice can form" --> Isn't this underestimating the fact that a lot of the Antarctic seasonal sea ice is actually formed via the so-called pancake ice cycle first published (in the 1990ties or late 1980ties?) by Lange et al. ?

We agree with the reviewer. Sea ice is formed in the Antarctic under turbulent atmospheric and weather conditions. Hence, pancake is common. Once again, it is possible to include pancake ice as a separate entity in ECICE using samples from authentic data. Bit this goes beyond the purpose of the manuscript. We admit, however, that pancake ice can be confused with MYI if radar data is used alone. But the combination with passive microwave can help. We have not done work to confirm this matter. We will emphasize the importance of pancake ice and refer to the Lange et al (1989) paper.

(6.) L60/61: Please check whether it is really the sea ice type that is required or whether these models wouldn't primarily be happy with using improved data of the sea-ice thickness (distribution), the degree of deformation and the snow load. Also, when it comes to validate a climate model I suspect that there are very few that already provide "sea-ice type" as a variable. They might provide ice age though.

We will check that.

(7.) L64++: "Recently, ..." --> While it is ok to already mention ECICE here, I ask you to provide a bit more background about algorithms that have been developed in the Arctic to separate FYI from MYI and to provide MYI concentration - first and foremost the NASA Team algorithm.

In addition to that, in order to put the value of your work into a wider context, I also ask the authors to provide more background about other attempts to discriminate between Arctic ice types. It is important that the reader understands that there is almost a full zoo of methods focusing on discriminating between different ice types in the Arctic - in addition to the NASA Team algorithm. To mention in addition to your and Ye's work is the work at met.no, at IFREMER, at BYU (David Long and his group) (and possibly others) that use coarse resolution satellite observations followed by the uncountable attempts to discriminate ice types using SAR. In contrast, activities in the Antarctic are very sparse.

You might argue that you are looking for ice type CONCENTRATION and not a simple discrimination. That is true, but even here your work is more upfront than any other work and this needs to be (implicitly) stressed.

You might also argue that ice type CONCENTRATION is the more important parameter, but if I understood your introduction so far correctly, then we are in need of ANY information about the ice-type distribution of Antarctic sea ice (other than land-fast sea ice), no matter whether this is a binary classification result or whether it is (already) an ice type concentration.

Because of this I ask you to one more time dig into the literature and try to find out what others did in this sector. If we omit polynyas / fast ice - for which a lot of studies exist - then there is not too many, perhaps add: Lythe et al., Classification of sea ice types in the Ross Sea, Antarctica from SAR and AVHRR imagery, International J. Remote Sensing, 20(15), 3073-3085, 1999, http://dx.doi.org/10.1080/014311699211624

and Ozsoy-Cicek et al., Intercomparisons of Antarctic sea ice types from visual ship, RADARSAT-1 SAR, Envisat ASAR, QuikSCAT, and AMSR-E satellite observations in the Bellingshausen Sea , Deep Sea Res. II, 58(9-10), 1092-1111, 2011, https://doi.org/10.1016/j.dsr2.2010.10.031 --> GC2

We will include a brief discussion on other attempts of ice type discrimination

(8.) L84/85: "It takes input ... any given ice types" --> So, I can input 6.9 GHz AMSR2 TB H-pol and 5 GHz ASCAT observations and can obtain the partial concentration of pancake ice? Or I can input 91.6 GHz SSMIS TB at H- and V-Pol and get the partial concentration of MYI ice and FYI ice? If this is not the case then I suggest to re-write this sentence according to the actual capabilities of ECICE which seems to be oversold a bit here.

This is only the necessary condition (number of input channels greater or equal number of surface types). The sufficient condition is, of course, that the distributions of the surface types in the input channels are independent/different.

(9.) L87/88: "to account for anomalies ... One anomaly causes ..." --> I suggest to re-phrase these statements. It is not clear what you mean by "observations". To me observations are the data you obtain from the satellites, i.e. brightness temperatures or backscatter coefficients. Hence, I ask myself what anomalies are in this regard? You possibly refer to those cases where ECICE fails to interprete the input satellite data into the correct total and/or partial ice concentrations, creating anomalous high or low concentrations and/or an anomalous misclassification of MYI as FYI and vice versa. Therefore it might be more correct to state that one set of satellite observations can be the result of several different combinations of physical parameters, causing ambiguous retrieval of total and/or partial concentrations when input into ECICE.

We rather meant “observations” in the sense of “anomalous ECICE results”. As this is unclear, we will rephrase it.

(10.) L100-102: Please provide 2-4 references for publications that could underline your statement for sea-ice concentration and sea-ice type concentration - for both passive and active microwave observations.

We will refer to the overview paper by Ivanova et al. (2014): With the exception of ECICE, all 11 algorithms presented there use tie points, including all ``standard" ones like NASA Team or Bootstrap algorithms

(11.) L111-113: What happens, during the retrieval, if fractions do not add up to 1 and/or for fractions below 0 or above 1? Are these set to 1 (or 0) before the median of all realizations is computed?

This is avoided by introducing the inequality constraint in the constrained optimisation approach of ECICE, as mentioned in L111-113.

(12.) L116: How is the spread around the median computed? How many valid values are required for a median and its spread to be computed (assuming that not all 1000 realizations provide a valid result)?

This is explained in the original paper of ECICE (Shokr et al. 2008) and in the book “Sea ice: physics and remote sensing” (Shokr and Sinha, 2015, chapter 10). We think we do not have to repeat the information in order to not to distract the reader. However, we can introduce a few lines to explain.

(13.) Lines 119-124: I am missing the physics and references in this paragraph. What are the physical properties of the ice types that cause the different radiometric and backscattering properties that allow us to discriminate between the three ice types? Which of these are influenced by which snow physical properties that make MYI to look like FYI? How about the ambiguities between YI and FYI?

You use snow metamorphism only in the context of "return of cold temperatures" albeit snow metamorphism encloses a wide variety of changes of the snows' crystal structure and composition under the action of temperature, humidity and wind. This should be re-phrased. In addition "warm spells" only cause "snow wetness" to develop if the temperatures are high enough; still, even with considerable below freezing (-5 degC) temperatures snow metamorphism (rounding of grains, etc.) is present. --> GC1

These five lines are only the very brief introduction to the correction schemes described in detail in the next two subsections. Stating here that melt-refreeze causes snow metamorphism does of course not imply that snow metamorphism can only be caused by melt-refreeze...

(14.) L126-134: This paragraph describes the temperature correction as developed for Arctic conditions. You appear to adopt it 1-to-1 to Antarctic conditions as is indicated by the last sentence in this paragraph. Without an adequate introduction and review of the physical, radiometric and backscattering properties of Antarctic sea ice and its snow cover compared to the Arctic, this raises my concerns. On the one hand differ Antarctic MYI and partly also FYI physical and microwave properties from Arctic ones. On the other hand differ Antarctic snow properties often fundamentally from those in the Arctic - not to speak of the frequency with which the weather influences the microwave signature of Antarctic sea ice compared to the Arctic. I am wondering whether a close collaboration with specialists in this field would not substantially improve both, set up of the algorithm and interpretation of the results.

Using the same setting as for the Arctic is a preliminary approach. As stated in our response to GC4 (see above) this study is a "proof of concept". As to the underlying physics, see also the response to GC1.

(15.) L150-152: "this correction scheme ... to MYI" --> I suggest to separate this correction from the drift correction because it has nothing in common with it. I further suggest that you make clear which form of snow metamorphism you are refering to here. In L154 you introduce "HR" as being related to the "onset of snow melt" which, at first glimpse, would suggest an increase in snow wetness and hence elevated brightness temperatures, making MYI to look like FYI rather than the other way round as is stated here. You are possibly refering to melt-refreeze cycles or the like and need to specify this here to avoid confusion.

The additional correction referred to here needs the "MYI domain" of the so-called "drift correction" and is done in the same step, therefore we do not consider it a separate correction. Note that the "drift correction" (in spite of the name) also corrects for the effect of snow/ice metamorphism as it eliminates ice that "looks" to the microwave instruments like MYI but is not. In other words, it uses the ice drift data to make sure that MYI does not appear very far from its expected domain of expansion. We will try to be more specific

(16.) L150: To me "Ex-MYI" implies that this sea ice once was MYI and now is a different ice type. How about you name it "artificial MYI"?

We will look for a more appropriate term like "erroneous" or "spurious" MYI.

(17.) L166-169: I suggest to add the actual resolutions and sampling interval of the AMSR2 channels used.

Please provide information about the native spatial resolution of the ASCAT data and how you gridded these into the NSIDC grid of 12.5 km grid resolution. It appears to me that the statistics is different for these data than for the AMSR2 data because of the different viewing geometry and swath width.

What would also be important to know is whether the sigma_nought values were corrected towards a certain common incidence angle (e.g. 40 degrees)? If this is not the case, please provide a comment why you deemed that as not being necessary.

Yes, we agree this information is missing and will insert the details here.

(18.) L170-180: Your description about the choice of sample areas and time periods is not specific enough to my opinion. I have the following questions:

18.1) Apparently you used ASI SIC maps to define your sample areas. What is the requirement regarding the SIC to have a grid cell contributing to the sample?

18.2) How did you define "beginning of the cold season"?

18.3) For which time period (just 1 day?) did you select grid cells from the Weddell Sea defining the MYI distribution?

18.4) For which time period(s) and region(s) did you select grid cells defining the FYI distribution?

18.5) From which time period(s) and region(s) did you define the YI distributions based on the PSSM data set?

18.6) How did you take into account the YI that develops during the ubiquitous pancake ice cycle in the MIZ that might cover several hundreds of kilometers? Isn't this, not the one growing in the polynyas, the far more relevant YI type in the Antarctic?

18.7) Where exactly, with respect to the ice edge, are your open water sample areas located?

18.8) How representative are the open water sample areas in the regions and months (March, Ross Sea; August, everywhere?) chosen for the weather influence?

One way to answer at least some of these questions would be to create a map in which you show the locations of the sample areas and, via color coding, the time-periods and/or frequency with which you used selected the data.

We will specify the details on the sample area selections and locations, trying to answer all of the above questions.

(19.) Table 2:

- I have concerns with two values in this table. Why do you define the END of the warm episode with a positive (2degC) air temperature? Is this a typo? If not it is absolutely not understandable and needs some justification.

There is actually a typo: the T1 must be -2°C and T2 +1°C. Note that the condition for the start of a warm spell is T>T1 *and* a reduction of MYI concentration in one day by 10%, and the condition for the end is T<1°C *and* a rise of MYI concentration by more then 10% in one day. Only in such a ase, the MYI concentration is corrected. T1 and T2 have been found empirically, as described in detail in Ye (2016a). Note also that T1 and T2 are 2-metre air temperatures, not ice surface temperatures.

- What is the motivation for the very long maximum duration of the warm episode? This does not sound overly reasonable to me - neither for the Arctic nor for the Antarctic actually. I can guess that the length of this period is chosen this way because the melt and melt-refreeze processes change the physical and therefore microwave signature of the snow / sea ice system for a considerable number of days; even after freezing conditions have returned the modified microwave signature might still last (e.g. Voss et al., 2003, in Polar Research and his work related to that).

Even though most warm episodes will last only few days, there is no harm in setting the *maximum* duration to 10 days in order to also catch rare longer events. It is computationally more expensive, but not in a significant way.

- Apart from that I am wondering how such a long maximum duration does match the comparably high frequency of warm events caused by cyclones passing over the sea ice. I'd say that such events can be quite short-lived. Therefore, depending on whether you aim for a monthly or a daily ice type product one could recommend to use a considerably shorter maximum duration of such events of just 5 or even 3 days.

Yes, a reduction of the maximum duration will reduce the latency times if we want to do NRT data, but that was not a concern of the present study.

- In case you comment on the choice of these parameters later in the paper, i.e. in the context of the discussion, please point the reader that already here to increase the credibility of your choices.

Yes.

(20.) Table 3:

- I note that the choice of the values for these parameters specifically for Antarctic conditions has not been discussed and/or movitated so far. You might want to do that, please.

- In case you comment on the choice of these parameters later in the paper, i.e. in the context of the discussion, please point the reader that already here to increase the credibility of your choices.

We will add a brief discussion on the parameters.

(21.) L187-191: Three comments here:

21.1) What kept you from using ERA5 data? Is there are credible argument to stick to ERA-Interim data for surface temperature data in the Antarctic?

When most of the presented work was done, ERA-5 was not available yet.

21.2) Tschudi et al. (2016) appears to be a bit outdated given the fact that there is a version 4.1 of the NSIDC sea-ice motion data set, referenced as Tschudi et al. (2019 or even 2020).

We will update the reference.

21.3) What is the motivation to use this rather low resolution OSI SAF sea-ice drift product? Doesn't it harmonize with the overall 12.5 km grid resolution you aim for much less than the NSIDC sea-ice drift product?

The source of the drift data was decided at the start of the study, several years ago. At that time, it seemed a good solution (probably,the temporal/seasonal coverage for Antarctic was better at that time). We are actually considering switching to NSIDC drift data for the future.

(22.) L192: In Table 3 you mention the TB at 37 GHz, not 19 GHz; please check.

This is a typo. Thanks for finding it, will be corrected.

(23.) Figure 1: I have a number of comments here; comment #1 and #2 are related directly to the figure content while comments #3 to #5 are related to the omission of relating the results shown to previous work.

23.1) What does the "Distr. set: AQ2" in the title of each panel refer to? Could it be removed?

Yes, this was a working title, will be removed or replaced by s.th. more meaningful.

23.2) What is the statistics behind the data? What is the time period? At how many data per surface type do we look?

We will list the details on that.

23.3) What explains, to your opinion the fact, that GR3719 is a bit lower than is classically observed for open water and, particularly, for FYI (compare the tie point triangle used in the NASA-Team algorithm).

This probably depends on the choice of open water samples. We will elaborate on the sample choices anyway (see item 18.)

23.4) How do your values compare in general to tie points used by ordinary sea-ice concentration retrieval algorithms?

23.5) How do your backscatter values compare to values for C-Band radar backscatter of Antarctic sea ice cited in the literature?

We will include a brief discussion of these two points.

(24.)L201-203: While I am fine with using AMSR-E instead of AMSR2, I have concerns to simply replace ASCAT (C-Band) with QuikSCAT (Ku-Band) as signal penetration into and interaction with the snow / sea ice system differ - in addition to incidence angle and resolution.

Yes, correct, one cannot simply replace the scatterometer data. The respective distribution functions for the surface types are needed as well. Retrieval with QuickSCAT instead of ASCAT and with AMSR-E instead of AMSR2 has already been done (see Ye, 2016a) – we will point this out in the manuscript.

(25.) Instead of working with piece-wise available sea-ice drift products it might be a very good idea to use one consistent data set, namely the NSIDC one - unless you find an alternative with year-round coverage (IFREMER?); yes, NSIDC is not an optimal choice but with that you avoid inconsistencies and jumps in your then much longer (by combining AMSR-E and AMSR2) time series.
You might want to consider to simply delete these three lines here.

Yes, of course. If there is a consistent time series of drift data for the whole period, all data should be processed and reprocessed using that one.

(26.) Section 3.1:

(26.1)- How many Sentinel-1 SAR data from which dates were used? Where were these located (provide a map with the frames)? What was the time difference between SAR image acquisition and ECICE product? What is the "time stamp" of the ECICE products [0 UTC, 12 UTC]?

We use daily gridded brightness temperatures (see L166ff.), interpolated from all swaths of one day (details on that will be included) – so there is no unique time stamp.

(26.2)- Where were the Sentinel-1 SAR images taken from. Which type of SAR images was used (Wide Swath, Extended Wide Swath, ...)? How were the SAR image (pre-)processed for the evaluation? Was any drift correction applied to the SAR images?

We will include the missing information.

(26.3)- You decided to provide a qualitative intercomparison without computing radar backscatter (sigma_nought) values. Why? Wouldn't your results be much more credible and useful if you would come up with fractions of MYI and/or FYI derived based on a rough (by means of sigma_nought value) classification from the SAR images and compare those to the ECICE MYI concentration maps?

See response to item 26.7

(26.4)- L214-215: "In SAR images, MYI ... sub-surface layer" --> This very qualitative and not overly scientifically formulated sentence applies to the Arctic. Melt processes during summer in the Antarctic differ considerably from the Arctic and I doubt that one can speak of a "bubbly sub-surface" layer here. Please revise your wording taking int account the specifics of seasonal changes in microwave signatures in the Antarctic compared to the Arctic.

We agree and will revise as suggested.

(26.5)- L221/222: "the iceberg ... as FYI" --> I don't agree. The SAR signatures inside that 70% FYI polygon encircling the iceberg are brighter than outside the polygon. The isoline does also not indicate at which side FYI concentrations are actually higher or lower. Given the fact that the area southwest of the iceberg is certainly dominated by FYI I suggest to re-phrase this statement along the lines that for that polygon both FYI and MYI concentrations are below 70% but that you don't know which is the dominant one. See also your Figure 5.

We will look into the details and revise.

(26.6)- For one grid cell, do partial concentrations sum up to 100%? I am asking because in the area indicated as > 50% YI fringing the Antarctic Peninsula there is evidence for MYI concentration > 50%. Did you actually check for maps like the one shown in Fig. 2 what the sum YI + FYI + MYI concentration is? It would interesting to see an example of this - perhaps in the appendix or in supplementary material.

See response to GC4 (4th paragraph). The sum is the total ice concentration (between 0% and 100%).

(26.7)- L222: This last sentence about the "quality" of this comparison I deem almost obsolete without information about how many SAR images of how many regions from which dates have actially been taken into account.

We will include more details about the number of scenes compared. We admit that the comparison is not very detailed, but rather an initial sanity check that focuses on the rather prominent MYI-FYI boundary usually found in the inner Weddell Sea – we will reword the sentence in this sense.

(27.) Section 3.2:

(27.1)- How are the weekly charts derived with respect to temporal availability of the input data? Is always the latest highest quality data set used for a respective grid cell (or pixel)? Or what is the compositing method used?

(27.2)- Does "microwave satellite imagers" include SAR? What is the dominant input data source for the charts you have used?

(27.3)- What does "analysis ... by experienced specialists" mean? Is this a manual analysis? Is the analysis done by one specialist or a team of specialists and what are the quality measures?

(27.4)- L228/229: What is the size of such a pixel? What is the grid that is used here? Is it the NSIDC polarstereographic one? Looking at Figures 3 and 4 I get the impression that in these ice charts the classification is not done pixel-by-pixel but rather in form of polygons that contain ice of similar characteristics and concentrations - such as done, e.g., by the Canadian and Danish Ice Services. Could you please check once again how the ice charts you show in your manuscript were generated, and if need be, re-phrase your description?

(27.5)- Were the input data projected into a common grid prior to ice chart generation?

(on 27.1 – 5) Ice charts are generated based on SAR image analysis as the prime data sources but combined with many sources of ancillary information. SAR images are analysed visually but ice analysis operators. One operator analyses each image, hence the analysis is subjective. However, those analysts are well trained and experienced. The ancillary data include climatic information about the area, the recent history of the ice filed, meteorological data, observations from ships and ice breakers Such charts consist of polygons, not pixels, thus we have to reword the sentence mentioning "pixels" which solves item (27.4)

(27.6)- L234/235: What is your "cold season"? What do you mean by "sporadic comparisons with data from other years"? How many, from where and which dates were these additional comparisons?

By "cold season" here we mean the period when the works because there is no widespread and sustained surface melt, i.e.n March to October. We are considering adding all comparisons as supplementary data.

(27.7)- L243-246: Your observations of the different labeling of ice as FYI or MYI between AARI and NIC charts could be the result of different definitions of when FYI is re-labelled MYI ice by the producing agencies? Did you check that? I note that a switch in March / April disagrees with the WMO recommendation you mentioned further up in your manuscript.

We think this rather shows that the two different teams of ice analysts came up with different assessments of their data, maybe also because they use different data. This is hard to find out.

(27.8)- I note in addition that there are more fundamental differences between the AARI and the NIC ice charts in the Eastern Weddell Sea regarding the location of YI and FYI.

See above...

(28.) Figure 3:

- I suggest to use a title for the MYI concentration that is consistent with the other two ECICE results.

Yes, the current titles are just working titles, we will use meaningful ones.

(28.2) - Putting the legend of the AARI ice chart into appendix is not a good solution. I suggest the following: You crop all maps to an area that excludes all the annotations in the AARI ice chart, put all ECICE results into the second row of panels and put the ice chart legend in the first row of panels next to the ice chart. In the second row of panels you could then also follow your approach from Fig. 5 and provide one legend with the title "Ice-type concentration", marking the ice type itself in the map (actually it is in the panels' titles but perhaps you consider to remove these anyways.).

We will consider the suggestion.

(28.3) - Finally, I note that you seem to use an old land mask to mask out Antarctica, still containing an overly long "Trolltunga" of the Fimbul Ice Shelf and the Mertz Ice Shelf. Given the fact that you focus here on AMSR2 data it might be a very good idea to use an more recent and hence more accurate land mask.

We have used the land mask that comes with the AMSR2 data (!), but consider improving on that.

(28.4) - L258 / Fig. 5: "with the summation of their fractions equal to 100%" --> When I look at Figure 5 I doubt that this statement holds. I guess it needs to be replaced by the actual total sea ice concentration that is obtained with the ECICE algorithm because there are quite some areas downstream of the Ronne-Filcher Ice Shelf polynya where YI conc. + FYI conc. + MYI conc. add up to something between 80 and 90%. I am sure you will get back to this in the discussion section. But it certainly does not hurt to either state that "theoretically" the partial concentration should add up to 100% but that this is not always the case, or correct your writing accordingly towards that the sum of the partial ice concentrations adds (of course) only up to the actually existing amount of sea ice. --> GC4

As stated before, the sum is 100% only if the open water fraction is added as well. We will make sure this is stated clearly in the manuscript to avoid misunderstandings

(28.5) -L261/262 / Figure 5: Your maps do also reveal that ECICE seems to have a problem discriminating between YI and MYI because the area just next to the Ronne-Filcher Ice Shelf appears to be characterised by some YI, no FYI and some MYI as one can observe a fringe of non-zero MYI concentration in that area.

This might also hint at a problem of the coastline – we will look into that.

(29.) Section 3.3:

(29.1) - Please provide more information about the PSSM maps. What are the grey areas masking parts of the maps shown? Where did you get the data from? What is their temporal and spatial resolution? How many of these maps did you use for which regions? The scope of this part of your intercomparison remains vague.

The grey areas are area not covered by the data, dependent on the definition of the region covered. The source of the data was already given above, section 2.3, L175-180, but we will include some additional information about the data provider there

(29.2) - Given the fact that you look at years 2017 and 2018 I assume it is SSMIS data and not SSM/I anymore, am I correct? Please correct your writing accordingly.

Yes, we will correct that.

(29.3) - I would appreciate if you could comment on the quality and limitations of the PSSM based ice type maps. What the approximate thickness limit between thin ice and "other ice"? Does "thin ice" mean that there is 100% thin ice or could this potentially also be 50% thicker ice interspersed with open water?

We will insert this information. The thickness limit is , according to the data provider (ICDC, CEN, Univ. of Hamburg), 10–20 cm

(30.) Figures 6 through 9:

(30.1) - I suggest to reduce the size of the panels considerably. In particular I recommend to make the PSSM map the same size as the white box shown in the left panel denoting its location. Even better would be if you'd crop the maps in the left panels to the size of the PSSM map. That way would would be able to reduce the number of these figures from 4 to 2 or perhaps even 1. Did you try, in this context, to combine the information from both panels into one? Perhaps by extracting isolines from the PSSM maps and superpose these onto the YI concentration maps?

We will work on the figures as suggested.

(30.2) - Did you chose the dates shown in the manuscript arbitrarily? If so, make a note.

(30.3) - Were these the only PSSM maps you considered in your comparison? If not how did the comparison go for all the other maps? Did you derive any quantitative information?

We just show a few representative examples, but consider giving some summary information on a more comprehensive comparison.

(30.4) - Looking at these YI concentration maps reminds me one more time the issue of how the ECICE ice-type concentration maps deal with cases of considerably less than 100% total sea-ice concentration because I note that all the YI concentration maps shown in Fig. 6-9 reveal lower YI concentrations in the core of the polynyas.

As mentioned, the partial ice concentrations of the three ice types sum up to the total ice concentration which is between 0% and 100%. We do not see a problem here but will look into that.

(31.) L280: "given the limitation that a rigorous validation ..." --> Given the fact that you kind of advertise the data set obtained with this paper and given the fact that this is first attempt to provide such a data set, I don't take it as a positive sign of credibility of the data set produced, when this paper only deals with a very general, little quantitative evaluation. The results presented are partly very vague and the description of the physical background being the foundation for the approach used and the data set is not overly exhaustive and - at least for me - not convincing.

See response to GC4 (aim: proof of concept)

(32.) L281: "Large icebergs are often erroneously retrieved as FYI" --> Is this your result? You could state this more clearly. But, when doing so, please take into account my comment made to Fig. 2 with respect to this issue.

Icebergs and their discriminations were not our focus here, but we will check and revise if necessary (see item 26.5)

L288-300: I was kind of expecting that you would run into problems with weather-induced variations in the snow physical properties and resulting microwave signatures. Since in your manuscript the physical foundation and description of the processes and properties resultung in specific microwave signatures is not overly detailed and mature, it is of course difficult to discuss these observations. I find that your attempt to explain your observations go into the correct direction but are far from being conclusive and is too vague. I'd say you could delineate the reasons that caused the MYI concentration over-estimation much better and much more specifically by means of checking the input data values and compare these with what is known from literature. It might make sense to take into account ERA-Interim and/or ERA5 data (you use them anyways) to discuss you observations also in the context of melt-refreeze, ice-snow interface flooding, slush refreezing, snow-ice formation and the like. I again recommend to take a look at the work of Voss et al. (2003) and the related doctoral thesis.

As to the physical foundations, see our response to GC1. However, we admit that we can improve on the interpretation of the results and thank the reviewer for the suggestions.

(33.) Figure 11: Looking at that figure again makes me to think whether you ever tried to look at maps of YI conc + FYI conc + MYI conc? It appears to me that there are patches of spuriously large MYI concentration that coincide with a total sum of partial concentrations above 100%.

Well, when only modify the MYI concentration in the correction schemes. which breaks the sum rule that the partial concentration *plus the open water fraction* add up to 100%. We will mention this at an appropriate place. The problem is that we have more than 2 ice types, so if we increase/reduce MYI concentration, it is not clear how to distribute the corresponding reduction/increase to FYI and YI, so we refrain from it. If we just had one ice type apart from MYI, it would be easy to preserve the sum by just mirroring changes of the MYI concentration.

(34.) L299: None of the references listed in this line deal with pancake ice and its backscatter. These are all references dealing with the snow cover and should be put into L298 behind"... MYI in that respect."

Yes.

(35.) L292: How credible are - to your opinion - these MYI occurrences "far offshore in the outer Ross Sea"? Which process can cause these?

Final question to L288-300: How did you compute the total MYI area shown in Fig. 10? Did you apply a threshold MYI concentration or did you count from 1 % onwards? What did you use as gridcell area to compute the total area?

As stated, we calculated the area which means we take into account the MYI concentration from 0% to 100% and the (space-dependent!) grid cell sizes of the polar stereographic (NSIDC) grid.

(36.) L302-303: You can look yourself into the likelihood of (1) by checking the drift data you used. How did you cope with data gaps in the drift product? Did you include the quality flags?

Good point. We will check that once more

(37.) L306-307: "such seeding points" --> please explain this in more detail or delete it. Questions I would have is how this happens and why this should have an influence on the MYI concentration in particular and not on the other partial concentrations.

We will improve the explanation of that

(38.) L309/310: I don't understand why you refer to an observation of Ted Makysm when you yourself used the data for the drift correction. Didn't you yourself take a look at the data once you suspected that these could include spurious drift estimates? This is inconclusive.

We will have to look into that.

(39.) Line 311-314: I suggest to not look into the data used but first try to understand which sea ice and snow physical properties you encounter during the course of one cold season and to further understand how the microwave signature looks like. This might require to look into 1-dimensional numerical modelling of microwave emissivities and of microwave backscatter as a function of sea ice and snow properties. There is a paper by Willmes et al. (2014) in the Cryosphere and there is work by Tonboe et al. that might help here.

We well look that the matter, but do not think a dedicated modelling study would make sense for this manuscript (but is worth doing in the near future).

(40.) L319-322: "The most likely reason ..." --> While your observation from Fig. 10 seems to be credible, I am wondering whether this isn't an over-simplification of the situation. I agree, wettening of the snow cover can mask MYI so that it looks like FYI. But at the same time the re-freezing of the slush at the ice-snow interface, ice lenses, whatsoever causing larger grain sizes can have the adverse effect and making FYI looking like MYI. A deep snow pack and/or substantial deformation of FYI has the same effect as demonstrated by one of the co-authors for the Arctic ocean. In addition, and here the authors were right earlier of course, pancake ice is a nasty fellow and could possibly also likely to be misclassified as either of the two thicker ice types - adding to their partial concentration.

We agree with the reviewer's points. The physical processes of the snow and ice modulates the radiometric and the scattering data. However, the advantages of using the probability distributions of all possible values of a given observation from a given ice type warrants the inclusions of all possible conditions. Sure, the now wetness and refreezing changes the observations but if the input distributions encompass all the possible changes, then the correct classification is warranted. As for the point of possible misclassification of pancake ice as MYI, we have not considered it. The two entities may not be misclassified using the present data set because while they have nearly same backscatter, their radiometric emission is different. This is an advantage of using the combination of passive and active microwave. Pancake ice is not part of the purpose of the study but its confusion with MYI should be considered. Here, ancillary information is required to avoid the inclusion of pancake ice areas.

(41.) L323-326: Two more thoughts on this: Beginning in October the expansion of the Antarctic ice cover stops and the lateral movements switches to a retreat / compaction type. In addition, due to the dispersion the fraction of MYI per grid cell has decreased to a value that is likely not large enough anymore to be adequately detected by ECICE. I am sure this is something you can check in your data. One could hypothesize that computing the total ECICE MYI area is reasonable as long as ECICE is capable to derive the MYI concentration with high accuracy ... which I doubt is the case when the partial concentration has fallen below 30% and when the MYI coverage has dispersed in many small floes embedded in a mixture of YI and FYI.

Also any MYI that has arrived in the MIZ (in the Weddell Sea) is now likely to melt as air temperatures are not cold enough anymore out there to keep it alive. From that point of view I find a rather decay of the MYI area in the September / October time frame not overly surprizing.

The accuracy of the results from ECICE has not been estimated quantitatively because this requires in situ observations. This statement applies to other algorithms too (e.g., MYI concentration from NASA Team algorithm). We do not know the minimum concentration that can be estimated but the manuscript provides data about the entire range of concentrations. We will include statements to clarify the lack of the quantitative accuracy. Operational ice charts cannot be used for this purpose (in our opinion) because of their coarse resolution, subjective method and most importantly the conservative estimates in these charts. Given all this, we don't think that the partial concentration is inaccurate if it falls below 30%. The limit should be lower than that.

(42.) L337: "and melt" --> Where did I find examples of these in your manuscript?

Maybe the figure in the manuscript do not show this very well. When looking at the daily evolution of MYI over an entire season, this is quite obvious. In view of this comment, we consider including an animation of one season in the supplementary material.

(43.) L343-344: "The new time series ... outweighs the shortcomings that still exist." --> I do not agree to this statement because of 1) the unmature physical foundation, 2) the vague interpretation of spurious ice type concentrations and 3) the very qualitative evaluation.

We still think that this data set, which we will term "preliminary", has its merits (it has actually already been used by Antarctic Cruises of the University of Cape Town) and serves as a proof of concept (see also response to GC4)., We will rephrase accordingly.

Typos / Editoral Comments:

L27: I suggest to look for a more recent paper making this statement, e.g. Kwok 2018 in Environmental Research Letters.
L30: There should be another reference from Parkinson and DiGirolamo from 2022 in Remote Sensing of Environment.
L49: Typo: "...sea ice For ..." --> "... sea ice. For ..."
L79: "existing a ice chart" --> "existing ice chart"
L114: Typo: "coast" --> "cost"
L152/153: Please explain all the mathematical expressions that are used here for the first time.
L159-164: You have introduced the sensors' acronyms further up and can omit that here.
L207: I am sure a reader would appreciate to see 1-2 references here.
L210/211: It might make sense to mention already here that the polynya maps used in the comparison are from a different year than those used for algorithm tuning.
L222: "Sentinel-1 scenes" --> "Sentinel-1 SAR scenes"
L256: Typo: "forth" --> "fourth"; see also L262
L266: Typo: "where" --> "were"
L286: "often" --> since you deal with a limited number of years in this paper you could perhaps mention all years during which you observe this increase.
L332: "outside the melt season" --> "during the freezing season"

"spatial" --> "grid"
L334/335: "... is well captured" --> You could add "by our ECICE results" to make clear that this is your result.
L334: "... in the Antarctic" --> add: "in addition to ship-based observations of the ice conditions."

We will correct the listed typos and errors