Respond to Reviewer #1

A) General comments:

1) There are a number of flaws in this study, with inconsistencies between its objectives, the way it is conducted, and its conclusions. Climate monitoring is clearly specified as the objective (l. 39 l. 388), but the method finally suggested cannot be used for long climate record. The data are only available starting in 2000. If the final goal is to produce long climate record, the selected method has to be applicable over long time series, otherwise the objective of the study is not met. The interest of this method with respect to the ones that are applicable to longer time series is not clear and needs to be discussed further.

Authors' answer:

Although climate monitoring has been mentioned in the manuscript, it is not our objective here. Our objective is to find the most appropriate SITC method for sea ice type distribution estimates in Arctic, which is linked to climate monitoring. To avoid misunderstandings, we have replaced 'climate studies' in l.388 with 'Arctic sea ice type distribution estimates'.

2) The comparisons are only applied to winter months. Why not summer months as well? The reviewer is well aware that the algorithms have strong limitations during summer months. However, the real problems have to be tackled. From a year to the next, the sea ice decline essentially occurs during the summer months, not during the winter months. Would it be interesting as well to evaluate the methods during that period? It is necessary here to better justify the use of winter data only, if summer data are not explored.

Authors' answer:

The following sentences have been added in Introduction:

In summer months, the snow on top of sea ice starts to melt, which limits the penetration of microwave signals and prevents the identification of different ice types. To avoid a comparison with large uncertainties, only the data in winter months were used.

3) Sea ice types are FYI or MYI. The type is thus directly related to the age of the ice. The paper uses two sea ice type algorithms. It also describes and uses sea ice age algorithms. From a naïve point of view, the two variables should be related, with the sea ice type derived from the sea ice age. Can you elaborate on the fundamental difference between the nature of the two variables and their relative merits? It is actually very surprising to see that the algorithms for sea ice types and sea ice ages are generated by partly common producers, but with very different methods for the two products and consequently possible inconsistencies. Consequently, the first step is to compare the two types of products (sea ice type and sea ice age), and if not consistent, they should be reconciled (especially if they are generated by the same groups).

Authors' answer:

Yes, the two variables, sea ice age and sea ice type, are related (FYI and MYI area the age-based ice types according to WMO). However, the methods used in the SIA and SIT products are based on different physics of sea ice. The SIA products use the time series of sea ice dynamics to track the sea ice thus determine its age, whereas the SIT products rely on the different radiometric signatures of sea ice (radiometer and scatterometer). We agree that the two types of products should be reconciled if not consistent. But this manuscript is about inter-comparison of SITC methods, not providing consistent MYI dataset. We therefore did not compare SIA and SIT products at the first step.

4) Snow impact on the signal is mentioned (l. 33), but not discussed in the rest of the paper.

Authors' answer:

We do not have further discussions because of the following two points. 1) This manuscript is to compare the 8 SITC methods but not to attempt to improve/modify them to account for the impact of snow. 2) None of the SITC methods, SIA and SIT products (here and at present) account for the impact of snow.

5) The study should focus on the analysis of a limited number of carefully selected products, based on the underlying physics of the observations involved in the algorithms, and on the method used in the algorithms, having in mind the improvement of consistent climate series.

Authors' answer:

This study focuses on 8 SITC methods, including 2 classic ones (BT and NT) and 6 non-traditional ones. To make it clearer, we added the following sentences on the selection of input parameters for ECICE (Section 3.1).

" Input parameters for ECICE were selected according to the following two principles. Firstly, PDFs from the three surface types (FYI, MYI and OW) should have good separations for that observation (e.g., PDFs of TB37H in Figure 3), which reflects the distinct radiometric signatures of the three surface types. Secondly, PDFs of the different observations have different patterns of separation (e.g., TB37H and and TB37V in Figure 3), so they complement each other on distinguishing ice types. In addition to the two principles, input parameters in the two pure radiometer-based methods, ECICE-SSMI-1 and ECICE-SSMI-2, were selected to correspond to the ones in the BT and NT method, respectively. ECICE-QSCAT-2 uses the same input parameters as that in Ye et al. (2016a, b), whereas ECICE-QSCAT-1 uses PR and GR instead of the brightness temperature channels. Similar as ECICE-QSCAT, input parameters of ECICE-ASCAT were selected accordingly. "

B) More specific comments:

1) The title of Section 3.1 mentions three sea ice type concentration algorithms. Finally, eight of them are compared, six of them from the same provider, and with common features between them and with the two others. Four are only derived from passive observations, with a maximum of four channels (19 and 37 GHz, possibly dual polarizations) at 25 km resolution. The four others are also using active microwaves, and they benefit from a twice better spatial resolution. Note that it is not clear if it is the spatial resolution of the product that is mentioned or its sampling. Any possibility to limit the number of products for comparisons, based on the physics of the observations? From the six products provided by ECICE, which ones are recommended by the provider? I would understand the comparisons of products with rather different passive microwave observations, with lower frequencies for instance. But here it is just the comparison of different ways of combining passive microwaves that have been around for several decades (starting in the 80s). In addition, the fundamental differences between the C and Ku bands of ASCAT and QSCAT is not discussed much, as well as their differences in terms of incidence angles or polarizations. These are real issues when building climate series. Similar comments apply to the sea ice age algorithms and sea ice type algorithms, with a lack of explanation of their differences in terms of physics.

Authors' answer:

The numbers in the "Resolution" column in Table 1 refer to the grid sampling. Since this easily can be mixed with the spatial resolution, we change the column title from "Resolution" to "Grid spacing".
We also add the following sentences, which relate the grid spacing to the true spatial resolution:
"The 25 km grid spacing of the pure radiometer products comes from the input radiometer data, which are provided at 25x25 km^2 grid. The footprint resolution of the used radiometer channels is 70x45 km^2 for TB19 and 38x30 km^2 for TB37. So, the true spatial resolution of the pure radiometer products is slightly coarser than the grid spacing. The 12.5 km grid spacing of the combined products come from the scatterometer data, which are provided at 12.5x12.5 km^2 grid. The radiometer data were regridded to the same size as that of scatterometer in the four combined methods."

We agree with the reviewer about the contribution of grid spacing. We have added the following sentences in Section 5, l. 394 in the revised manuscript.
"We explain this result with the additional sea ice information contained in the scatterometer signal. However, we cannot exclude that this result may also be influenced by the increased resolution of the scatterometers."

As for the 6 products from ECICE, since this algorithm (ECICE) is generic in terms of the input data, we examined it with different combinations of inputs as explained and justified physically in Section 3.1 (see answer to general comment 5). We considered the algorithm with a given input set as one method in itself (although the scheme of optimal solution is same) because each of the other methods is based on a unique set of input observations.

As mentioned by the reviewer, it would be interesting to use passive microwave observations at lower frequencies. We did not include such comparisons because of the low resolution. This has been mentioned at the end of this manuscript (l.423 in the revised manuscript) and could be done with the planned Copernicus Imaging Microwave Radiometer (CIMR). Besides, we have added a few sentences about how we select the input parameters for ECICE (see answer to general comment 5).

Once again, the objective of this manuscript is not to build climate series using the methods/products presented here. Details about the QSCAT and ASCAT data has been presented in Section 2.1. They are scatterometers at different frequencies (Ku and C-band), with different polarizations (VV and HH for QSCAT, VV for ASCAT) and incidence angles (54 deg and 46 deg for QSCAT, 25-65 deg for ASCAT normalized to 40 deg). Studies show that Ku-band scatterometer could be more sensitive to ice types than C-band (Ezraty and Cavanié, 1999; Rivas et al., 2018). This has not been confirmed in SITC methods, we therefore would like to investigate it in the study.

2) The tie points (section 3.2) are estimated on data covering almost all year. Why not extending them to the year for the derivation of a MYI product applicable to the full year? Would it be possible to analyze the seasonality of the responses on Figure 3? That could justify the use of winter months only. It would be also very informative to present similar results for the ASCAT and QSCQT backscattering parameters (same as Figure 3, with addition of information on the seasonal cycle).

Authors' answer:

As mentioned in the reply to general comment#2, we have added sentences in the revised manuscript to explain the snow melt effect in summer months, therefore did not extend the data to the whole year. In Figure 2, we show samples and the periods from which the tie points were obtained for MYI, FYI and OW. For the 2 ice types, no samples were obtained during summer because of the anomalous microwave signals. But for OW tie points, summer data were included because the signal is virtually not affected by the water temperature.

3) Figure 4 raises some questions. It seems that, regardless of the SIA product (SICCI or NSIDC), the extent of the multi-year ice steadily declines during the winter (ice with age > 1yr). I would expect it to decrease during the summer, but not during the winter months. If one makes the fair assumption that the sea ice does not melt during the first part of the winter, it means that the multi-year ice is transformed in FYI during the winter? That would be an interesting rejuvenating process! In the ECICE products, this is not observed: in general, MYI is stable or slightly increases during the first part of the winter, as expectation. The behavior of the SIAs seems very questionable and should be commented. Still on Figure 4, from April 2016 to October 2016, the MYI extent from SICCI-SIA is constant: this should mean that all the FYI melted during that summer. Is that confirmed by other observations? Note that the NSIDC-SIA product does not observe it and there is a ratio of _1.6 between the two SIA estimates for that month, raising doubts again on the SIA products. These key

aspects have to be clarified before any quantitative analysis of the differences between the ECICE and the SIA in 4.1.1.

Authors' answer:

MYI is the ice that survives for at least one summer melt. FYI can only become MYI after surviving the summer melt. It means that increase of Arctic MYI area at the end of summer (freeze-up months e.g. October 2016) is expected. On the other hand, FYI cannot turn into MYI in winter. The decrease of MYI area in winter is expected, mainly due to ice export to more southern areas (e.g., ice export through Fram Strait). To avoid misunderstandings of the expected changes of MYI area, we have added the following sentences in this section.

"According to the definition of MYI, sea ice that survives at least one summer melt, FYI can only turn into MYI after surviving the summer. It is expected that the Arctic MYI area declines within the winter. In Figure 4, "

4) As already said, comparison between the SIT and SIA have to be performed before other comparisons, as these products are both related to the ice age. Figures 6-7 actually provide an indirect comparison of the two SIT products. It would be much more efficient and clearer to directly compare the two SITs. The conclusion from the comparisons between the SIT and the SITC that the ECICE QSCAT are the best SITC product is very surprising and not convincing (l. 244 and following). The 'validation' with the tie point is not more convincing. The ECICE QSCAT 1 and QSCAT 2 products provide rather different results, especially at the beginning of the winter season (Figures 6-8-9). If it was the addition of the QSCAT that was beneficial to the retrieval, more stable results would be expected between QSCAT 1 and QSCAT 2. The spatial resolution issue is never discussed in the analysis, but all the ECICE products with QCSQT and ASCAT have twice the spatial resolution of the purely passive microwave products. That can clearly help the comparison with other products. Second the ECICE QSCAT 2 is the only product without the 19 GHz channel (19GHz has twice worse spatial resolution than the 37 GHz channels). That can help improve the final spatial resolution of the product for comparison with other parameters. All these factors can play a role in the differences between the algorithms, and should be discussed. Note that the notion of spatial resolution in this paper is very unclear, as it is likely closer to the spatial sampling of the data.

Authors' answer:

In the manuscript, we conclude that ECICE-QSCAT is the most suitable approach to reproduce the information in the SIT product, but we do NOT conclude that it is the best SITC product. This section (and Figures 6-8) is about comparison with SIT products but not using SIT products as the 'ground truth' data. The objective of this section is to find out which SITC method agrees better with the existing SIT products. The following sentences were added at the end of this section.

"Thirdly, among the three SIT products, the eight SITC methods have the best agreement with the KNMI-QSCAT-SIT product, with high MYI concentrations over MYI pixels and low values over FYI pixels." Besides, we have added a figure (Figure 9 in the revised manuscript) to show the seasonal cycle of MYI concentrations of the three SIT products. Texts in this section were modified accordingly (see Section 4.2 in Intercomparison_SITC_revised_track_changes.pdf).

To mention the contribution of spatial resolution, the following sentences were added to the manuscript in Section 5, l. 394 in the revised manuscript.

"We explain this result with the additional sea ice information contained in the scatterometer signal. However, we cannot exclude that this result may also be influenced by the increased resolution of the scatterometers."

As mentioned by the reviewer, it should be called spatial sampling but not spatial resolution. In the revised manuscript, we all use 'grid spacing' instead.

5) The validation with the SAR data could be very interesting. However, it is not quantified and only based on the qualitative examination of SAR images, without much information about the methodology. It is rather difficult to evaluate the different SITC and SIA products from these qualitative comments. The incidence angles of the SAR observations have to be mentioned. Does it change much from one side of the image to the other?

Authors' answer:

Quantitative validation would be ideal for the evaluation. However, qualitative interpretation from the SAR images with auxiliary information is already the best we can have. In fact, all the operational ice services are mainly based on visual analysis of SAR images. We have added sentences in Section 4.4 about how the identification of ice types is done in this study.

" In these case studies, identification of different ice types is mainly based on brightness, texture and spatial context of the SAR images. Auxiliary information such as ice charts (with stage of ice development) from the Arctic and Antarctic Research Institute (AARI) and Canadian Ice Services (CIS) were also used to help with the interpretation but was not shown here."

A sentence about incidence angle range was added in Section 2.4. "The incidence angle covers a range from 20 deg to 49 deg." Backscatter of the SAR images is incidence angle dependent. We did not take this into consideration, because identification of ice types uses other features of SAR images as well (e.g. texture and spatial context of SAR image), which are hardly influenced by incidence angle.

6) No SIA product is shown in these comparisons. Why? Seems like there is a systematic problem in the color bar of the SICCI-SIA problem (not age but percentage).

Authors' answer: (We guess the reviewer meant no SIT product is shown in the comparisons)

No SIT product is shown in the comparison because 1) the SIT products cannot be used as 'ground truth' data for SITC methods; 2) evaluation of SIT products is out of the scope of this study.

The SICCI-SIA product is correctly shown in the figures. Although the SICCI-SIA product is a SIA product, it produces concentrations of different ice types (age of 1yr, 2yr, 3yr, 3yr+). When plotting the figures, we merge the concentrations for all the sea ice that is older than 1yr.

7) The SAR and the scatterometer data are based on the same physics (backscattering from a surface), and ASCAT and the SAR share the same frequency. It is expected that the SAR and the scatterometer images show some similarities, at least qualitatively and at large scale. These comparisons between the SAR data should again concentrate on a few carefully selected parameters, with discussion on the underlying physics and algorithm principles. If used in climate data records, the transition from the ECICE-QSCAT and the ECICE-ASCAT algorithms have not been analyzed. Is it smooth enough for climate studies? Jumping from C to Ku band is very likely problematic. In addition, nothing is said about the possible use of ERS data before, to extend the data set.

Authors' answer:

It is correct that RADARSAT and ASCAT share the same frequency and are based on the same physics. However, backscatter from ASCAT is only one of the 4 input parameters. Similarities between the retrieved SITC and SAR image could be limited. In the revised manuscript, sentences about selection of input parameters for ECICE have been added in Section 3.1, which reflects the underlying physics and algorithm principles for the SITC methods' comparison.

As for climate series, since it is not the objective of this study, we did not analyze the transition between ECICE-QSCAT and ECICE-ASCAT methods. Usually, it can be done by normalizing the sea ice area (or MYI area) for the overlapping periods. This is out of the scope of this study, therefore is not further discussed.