Response to reviewer #2

"An improved sea ice detection algorithm using MODIS: application as a new European sea ice extent indicator", by Joan A. Parera-Portell, Raquel Ubach, and Charles Gignac

Dear reviewer #2,

Thank you for your constructive comments, they will certainly help us in improving the quality of our manuscript. In the first place I will answer your **general comments**:

Overall, I agree that the most important part of the study is the algorithm itself, not the monthly trends. The trends (and their agreement with previous studies that you point out) are just a demonstration of the algorithm's usefulness and reliability. However, we want to stress that the IceMap500 raw products are the swath map and the daily map, so our monthly maps are, essentially, a way of capturing the trends in the study area. Thus, there is averaging of the sea ice edge indeed, and here is where both the sea ice presence likelihood maps and the comparison to the Sea Ice Index (SSI) gain importance.

In the first case, the sea ice presence likelihood maps add valuable information to the monthly sea ice extent (SIE), such as how many times sea ice was detected in a certain area or pixel. This allows us to detect the places where sea ice (and also the ice edge) has been more unstable during that month, as the sea ice presence likelihood will drop (see Figure A). In fact, the likelihood maps allow even to detect cracks in the sea ice, and of course if sea ice has moved significantly the sea ice presence likelihood will be lower. Under a certain likelihood threshold, which is 10 % in our case, we set the pixels as NoData because they may not be reliable enough, and this usually leaves a small NoData buffer zone along the ice edge (see Figure B) where sea ice presence likelihood is >0 % and <10 % (0 % is water). It is within this buffer zone that we place the sea ice edge during the monthly SIE map creation, as we use an Euclidean allocation method to set those pixels either as water or sea ice. This generates a smoother sea ice edge, and should be considered as the maximum extent achieved during a given month.

On the other hand, the SII-IceMap500 monthly SIE comparison allows to assess where do our SIE maps agree or disagree with a commonly used index. The agreement analysis indicated that one of the main differences is the capability of IceMap500 of detecting small sea ice floes and fragmented ice and incorporating them into the monthly SIE map, especially in September. In March the SII and IceMap500 datasets are very similar, so even with the much greater spatial resolution of IceMap500 and the different SIE derivation approach they agree very well. Of course, IceMap500 was never intended to replace SII or other datasets, but instead to provide additional and higher resolution information which could be interesting for local or regional studies of sea ice conditions (the European sea regions, in our case). In our opinion, the agreement with the SII indicates that our maps, even the monthly SIE maps, are coherent and reliable. Some strong features of our monthly SIE maps are the increased classified area in comparison to MOD29 (which increases the sea ice presence likelihood values), and the more detailed SIE information along the shoreline, fjords and other areas that sensors with a coarser resolution might miss. Thus, we think that IceMap500 represents an improvement towards local and regional sea ice studies even at a monthly scale, especially taking into account the spatiotemporal information that the sea ice presence likelihood maps may provide, and is complementary

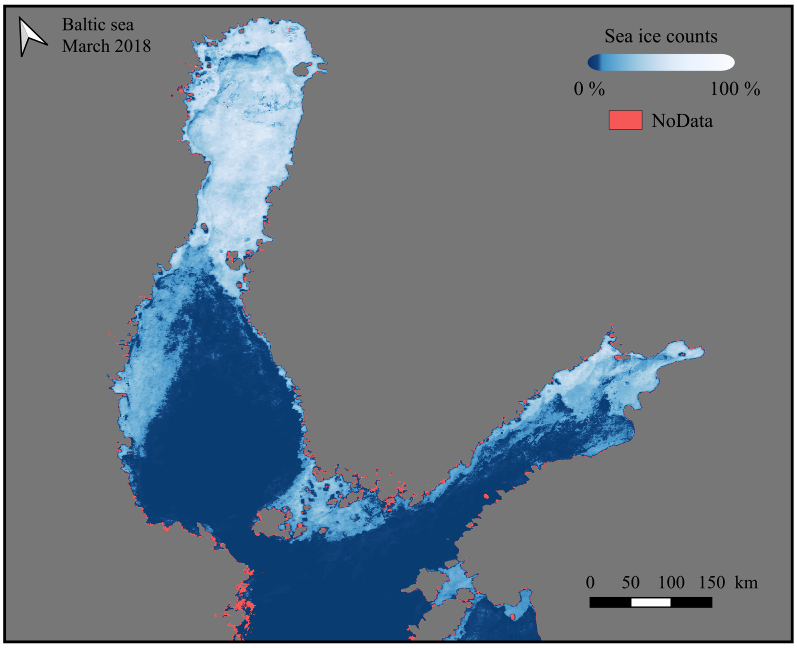

Figure A: IceMap500 monthly sea ice presence likelihood map of the Baltic sea.

to informations commonly gathered by national sea ice services for their operational ice conditions monitoring activities. Also, the likelihood maps can be generated within any given time period, so it might be a useful approach to analyze, for instance, weekly sea ice presence, and a more conservative threshold can be used to obtain SIE. Additionally, the moderate processing time of the algorithm should allow to produce datasets that could be of interest for sea ice services, marine infrastructure managers or for navigation, as a full daily map covering our study area (16 scenes) takes about 50 min to process in a machine under Linux Mint 20.1, with a 2.67GHz $\times$ 4 processor and 12 Gb RAM.

We think that, as you suggest, a revised manuscript would greatly benefit from a subsection focused on comparing daily maps from different sources, such as MOD29 and EUMETSAT Daily Sea ice Edge. Therefore, the effect of the NISE footprint could also be further discussed. Illustrating the different phases of the algorithm would be also an improvement in the Methodology section, as you say, and it would surely help readers understand how it works step by step. We plan to add this to the manuscript in the revision. We also plan to compare the SIE trends in the papers you have suggested, which we unfortunately missed. However, as the original goal of the algorithm was to provide complementary and higher resolution information for existing sea ice extent indicators used in the European Union, we plan to keep our current study area. In addition, the trends in our study are intended to proof the algorithm's fitness by comparing them to existing ones, but it is not intended to be an exhaustive trend analysis, as our goal is to demonstrate that IceMap500 can be used to monitor sea ice at a European scale level.

The applicability of IceMap500 to other sensors is also another topic which deserves further discussion and which would greatly enhance the manuscript. Most MODIS bands we use in IceMap500 have their equivalent in both VIIRS and Sentinel-3 SLSTR at almost the same wavelength ranges, with the only exception of MODIS band 7 (2.105-2.155 µm) which we use as a cloud detector in the MOD35 correction. The closest matches in VIIRS

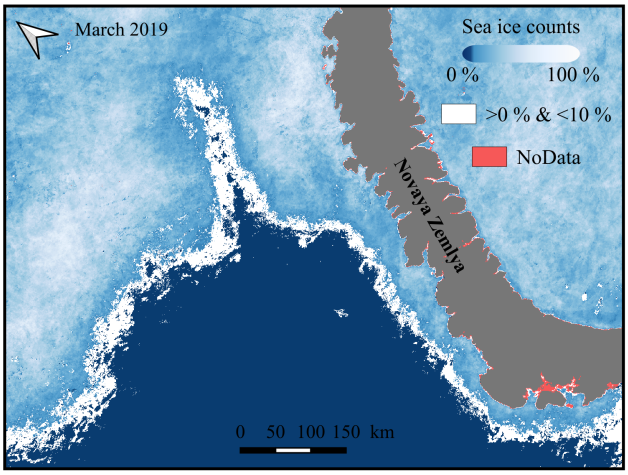

Figure B: Detail of IceMap500 monthly sea ice presence likelihood map, showing areas where likelihood is >0 % and <10 %.

and Sentinel-3 SLSTR have central wavelengths of approximately 2.250 µm, a region in which ice has a reflectance peak of about 20 %, while at 2.105-2.155 µm it has 5 %-10 % reflectance. It would be worth investigating 1) whether VIIRS and Sentinel-3 SLSTR require such artifact correction and 2) if they do, whether clouds and sea ice can still be distinguishable by their reflectance at 2.250 µm.

Finally, as for bare ice and thin ice, the thresholds in IceMap500 are designed to detect those surfaces as sea ice too. The 17 % ToA threshold using band 4 (545-565 nm) had been previously used in Riggs et al. (1999) and Gignac et al. (2017) and is intended to include most low-albedo sea ice. Validation in Gignac et al. (2017) clearly shows that the B4>=17 % threshold resides slightly into the upper standard deviation of the water class reflectance, so the risk of misclassifying melt ponds, new ice, leads and polynyas is low (see, for instance, Figure 6 in the Gignac et al. (2017) paper). The Normalised Snow and Ice Index 2 (NDSII-2) test also is shown to discriminate 96-100 % of the sea ice even during the melting periods in Gignac et al. (2017). The band 20 temperature threshold (1 °C) is intended to be a mask and not really a classification test, so it generates a buffer zone that performs pretty well including both cold water and new sea ice. Past studies such Zhang et al. (2017), already cited in our manuscript, show that melt ponds stay below 0.3 °C, so the threshold should be safe. However, both the NDSII-2 test and the band 4 reflectance test may still fail, as our validation results demonstrate. A strength of the MOD35 correction is that, if one of the threshold tests fail, those pixels are set as NoData and may be classified again during the correction. Then those pixels have a higher probability of being classified as sea ice, as the area in which the NDSII-2 test is applied is smaller and so the Jenks threshold tends to include more low-albedo ice. That is, always when ToA reflectance at band 4 is 17 % or above.

Specific comments

1) Abstract: I think we should drop the reference to IceMap250 and just explain what it is later in the manuscript. Also, we will mention the validation method we used to test the accuracy.

2) Introduction: We will take a look at the references you suggest and include them to update the sea ice information. References to the mentioned sea ice products will also be included, also with a brief description of each product.

3) Materials and Methods: all the references and text corrections you suggest will be added to the manuscript, including the NDSII-2 equation. Now, regarding your more technical questions:

l.89: This line is not clear enough and will be rewritten. It does not really mean that TOA reflectance does not depend on the physical properties of sea ice and water, but instead that TOA reflectance is not itself a physical property of sea ice and water because of the atmospheric contribution, although it is obviously related to the surface reflectance.

Section 2.2: We will explain step by step the workflow of the algorithm, including what you suggest here. For projection purposes we use NASA's HDF-EOS to GeoTIFF Conversion Tool. Also, we unfortunately did not mention the projection we use both in our figures and the algorithm itself, which is North Pole Lambert Azimuthal Equal Area.

l.107: What you suggest is an interesting approach, although we did not consider a solution like that because we wanted to use as few input data as possible for the sake of simplicity and efficiency. What we did is increasing the restrictiveness of the classification: even though a pixel passes the NDSII-2 Jenks optimization test, it still has to be confirmed as sea ice by the 17 % band 4 ToA reflectance threshold. Then, if water erroneously passes the NDSII-2 Jenks optimization test, it should be automatically discarded due to its low reflectance. This increase in restrictiveness does not suppose a significant change in the final classified sea ice area, because as we discussed earlier both tests are designed to include low-albedo sea ice.

l.139: Yes, it is intended to do so. This way the mapped area is considerably larger.

l.143: You are right, the mean and std are calculated every time from the swath data that is being processed.

l.167: We are not conducting atmospheric correction on the solar reflective data, nor on the thermal bands when calculating the VIS mask. However, band 20 is indeed atmospherically corrected, but only when performing the SST test. This way it is much easier to select a temperature threshold.

Section 2.3.3: Even though the area occupied by the NISE artefacts in the MOD35 mask is tagged as cloudy, the algorithm still has access to unmasked TOA reflectance so these areas can be analysed again. What IceMap500 does is to create a 25 km buffer zone around the areas classified as sea ice in the MOD35 mask. It then masks again the original TOA reflectance data using only the buffer zone, while a new cloud mask within this buffer is created using band 7. So, actually, the MOD35_L2 cloud mask is ignored within the areas that the MOD35 correction classifies. We plan to add a step-by-step graphic example in the revision, so I hope this process will become more clear to readers.

l.216: Unfortunately I cannot give you exact numbers, as the algorithm directly calculates sea ice presence likelihood in % without saving such information. However, we can estimate the maximum number of observations by assuming that the minimum is 1 (excluding water, that is, 0) and using the corresponding % to extrapolate the maximum. We get that the mean maximum number of sea ice observations is 50 in March and 62 in September.

l.255: NoData gaps tend to appear in the northernmost regions of our study area in the March monthly maps. This is a consequence of the poor lighting conditions during the winter months; remember that we only keep pixels tagged as day in the day/night flag of MOD35_L2. Obviously, this also makes the sea ice presence likelihood to drop. September has no such lighting limitations, so NoData gaps appear more randomly and are fundamentally linked to the cloud cover (see the examples in Figure C). Overall, the mean NoData area fraction of our monthly time series is 1.0 % in March and 0.7 % in September. However, March features a larger std (0.7 %) compared to the std of September (0.3 %).

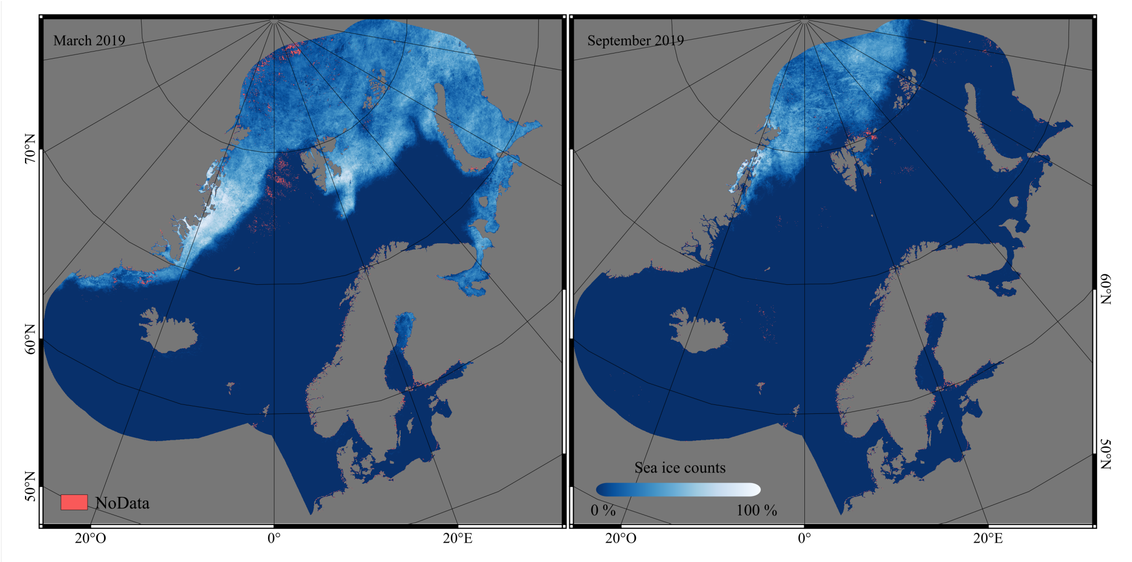

Figure C: IceMap500 monthly sea ice presence likelihood map of March 2019 (left) and September 2019 (right).

4) Results

l.240-242: With typical p-values of 0.05 and 0.01 our Arctic trend lines remain statistically significant. This information was accidentally omitted in the text.

l.244: You are right, both the visual and the quantitative analysis do not show any clear trend in the Baltic, so this line will be removed.

l.267: Indeed most scenes feature both surface classes. However, in March the extensive sea ice cover plus the presence of clouds may cause the classified water area to drop considerably in some scenes, especially to the east of Novaya Zemlya and Franz Josef Land. Even though there may still be some pixels classified as water, the area fraction compared to sea ice is very small and so random points used for accuracy assessment may not sample those water areas, causing kappa coefficients to drop.

Table 5: Right, we can give percentages with 0.1 % accuracy and reduce the decimals of the kappa coefficients from three to two.

Figure 9: Also right, we accidentally did not reference the figure in the text.

5) Discussion: We will include the reference you suggest and discuss on the effect of the Arctic Oscillation. As for the projection of the maps in the dataset (`https://doi.org/10.5565/ddd.uab.cat/196007`), you are right, there is no information about it in the readme file. We will amend this issue as soon as possible. All maps are in North Pole Lambert Azimuthal Equal Area.