

Reviewer 1:

The authors processed 4861 MODIS images of sea ice in the Beaufort Sea spanning the months of March through September in the years 2000 through 2022, identifying more than 9 million ice floes. They then constructed and analyzed the sea-ice floe size distribution (FSD). They found that the FSD follows a power law with an exponent that varies seasonally. They also calculated the mean floe size and the orientation of floes.

Previous studies of the FSD have typically been based on the analysis of a handful of images, or at most a few hundred. The amount of imagery analyzed in this work exceeds all previous studies by a factor of 15 or more, and for that reason alone it is worthy of publication. But regarding the seasonal evolution of the FSD, the authors did not make a careful comparison of their results with previous studies, and there are some mathematical issues to resolve.

Response: We are grateful to the reviewer for a thorough review of our manuscript. We have revised the manuscript per the reviewer's constructive suggestions, which we believe has greatly improved the clarity of the presentation. We have addressed the stated concerns and made the necessary changes to the manuscript (shown in blue font). Significant changes were made to Figure 5 (now Figure 6) per the reviewer's comments and added a new appendix to show the analytical derivation of the observed relationships in our dataset. The specific revisions are described in the point responses below:

Lines 89-90. "The dynamic threshold value is the weighted mean for the 400-pixel (100 km) neighborhood of a pixel subtracted by a constant." A few questions:

Thank you for your comment. The text has been revised to clarify:

"The dynamic threshold value is the weighted mean for the 399 pixel x 399 pixel neighborhood (~100 km x ~100 km)."

1. A neighborhood is an area. Is the neighborhood 400 x 400 pixels (100 x 100 km), or is it 20 x 20 pixels (400 total, 10 x 10 km)?

Correct. In this case, it is 399 pixels x 399 pixels (centered around the pixel of interest).

2. Is the neighborhood a square or some other shape?

Yes, square.

3. What does "subtracted by a constant" mean? What constant?

Thank you for catching this detail. We do not subtract by a constant, so this has now been deleted. This was the description of the python function we used. But the default constant is zero.

Lines 161-167. I do not understand the bootstrapping procedure. Samples of (what?) are drawn from (what?).

Thank you for your feedback. We are referring to sample of floes in an image, drawn from all the floes in an image with replacement. Text has been modified for clarity:

Finally, we apply a bootstrap approach to quantify uncertainty in the derived parameters. We randomly select 100 days of segmented MODIS imagery with at least 50,000 km² of identified floes, and examine the uncertainty in the derived FSD for each of those days. We create 1000 bootstrap samples (of floes in the image) with replacement. Each bootstrapped sample is the length of the

original dataset (the number of identified floes in the image), and we calculate the power law distribution of the sampled floe areas. We then calculate the standard deviation of the alpha values for each image. The standard deviation of alpha generated from the bootstrapping of an image is on average, 0.024, ranging from 0.007 to 0.08. The standard deviation increases as the alpha value increases, indicating more uncertainty.

Lines 178-182. These sentences discuss the mean and median floe areas as shown in Figure 5(a). It should be noted that for a power-law distribution, the mean and the median are directly proportional to the minimum floe area, A_{min} , which is 5 km^2 in the present work. The power-law pdf with exponent $-\alpha$ is:

$$f(A) = ((\alpha-1)/A_{min}) * (A/A_{min})^{-(\alpha)}$$

The integral of $f(A)$ from A_{min} to infinity is 1.

The mean floe area is:

$$A_{mean} = A_{min} * (\alpha-1)/(\alpha-2)$$

The median floe area is:

$$A_{med} = A_{min} * (2^{1/(\alpha-1)})$$

The point is this: The mean and median floe areas are completely dependent upon the resolution of the measuring device. If the resolution of the sensor were to allow the detection of floes 10 times smaller, the mean and median floe areas would be 10 times smaller (provided the data continued to follow the same power law). The mean and median floe areas reported here do not reflect any kind of objective reality but rather the limitations of the sensor. They cannot be meaningfully compared to the mean and median floe areas detected by sensors with different resolutions. It's fine to calculate and plot the mean and median floe areas, and to note their variation over time, but their direct proportionality to the resolution (A_{min}) of the detection mechanism should be noted, if in fact the floe areas are power-law distributed.

Thank you for the comment and the detailed derivation of these equations. The mean and median values, while dependent on the A_{min} value (as pointed out), are also a function of alpha, which we show has a seasonality. Therefore, showing the evolution of the mean and median floe size is an interesting result. Throughout our analysis, we have held the x_{min} and x_{max} values constant and changes in the mean and median are a result of changes in alpha. When comparing different sensors, we held x_{min} and x_{max} fixed to ensure comparability of the calculated FSDs. We have added a discussion on the relationship between these parameters (section 4.1 second paragraph) and have described the choice of x_{min} and x_{max} values in more detail (section 3.5). Thank you for pointing this out, we believe the discussion has improved with this suggestion.

The previous comment raises a further difficulty. For a power-law distribution with exponent $-\alpha$, the mean does not exist when $\alpha \leq 2$. In other words, the integral of $A*f(A)dA$ from A_{min} to infinity is infinite when $\alpha \leq 2$. Now look at Figure 5(b). It shows that alpha is ≤ 2 quite often, and yet Figure 5(a) shows a finite mean floe area. It is not mathematically possible for a power-law distribution with the reported values of alpha to have a finite mean floe area. Fortunately, this is easily fixed by postulating that a maximum floe area (A_{max}) exists such that the distribution of floe areas follows a power law only for the finite range $A_{min} \leq A \leq A_{max}$. If this approach is taken, the authors should estimate A_{max} . If this approach is not taken, the authors should explain how to reconcile the mathematical contradiction.

Thank you for raising this point. We have been using a constant A_{max} value throughout our analysis.

This value is now in the text and the choice of this value is now described in the manuscript (section 3.5).

Line 188. "The power law alpha is inversely correlated to the mean floe area". This finding is based on the data. It could also be checked analytically for a power-law distribution. As noted above, if $\alpha > 2$ then $A_{mean} = A_{min} * (\alpha - 1) / (\alpha - 2)$. As alpha becomes larger, A_{mean} does indeed become smaller, although not linearly. But most of the values of alpha in this paper are LE 2. If the power law is defined on the finite interval $[A_{min}, A_{max}]$ then A_{mean} is a more complicated function of A_{min} , A_{max} , and alpha, but A_{mean} still decreases as alpha increases, although not linearly. Whether the authors wish to include such an analytical comparison to the data is up to them, I am only pointing out the possibility to do so.

Thank you for detailing the mathematical functions behind this relationship. The authors agree that this derivation would complement the information presented in the manuscript. We have added an Appendix to derive the analytical mean and median which are now shown in Figure 6a and b.

Lines 210-218. This paragraph compares the power-law exponents of the present study to those found by Stern et al (2018a), but the comparison is not done correctly. The present study uses floe area as the measure of floe size, while Stern et al (2018a) used floe diameter. Assuming that floe area (A) is proportional to the square of floe diameter (x^2), a power-law distribution in A with exponent $-\alpha$ is equivalent to a power-law distribution in x with exponent $-2*\alpha + 1$. With this conversion, the authors should make a quantitative month-by-month comparison of their results with those of Perovich and Jones (2014), Hwang et al (2017), and Stern et al (2018a), all of which looked at the seasonal evolution of the FSD in the Beaufort Sea.

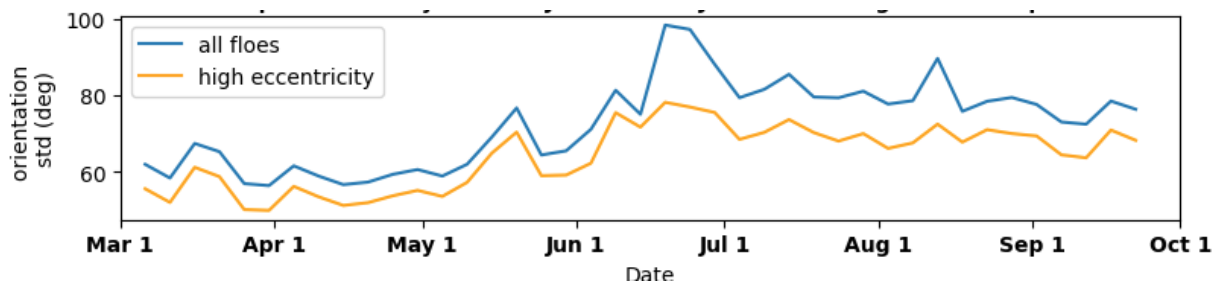
We agree. We have included a description of this conversion to the manuscript and added a quantitative comparison with the studies mentioned here (Section 5.2).

Regarding floe orientation (section 4.3), a couple of comments:

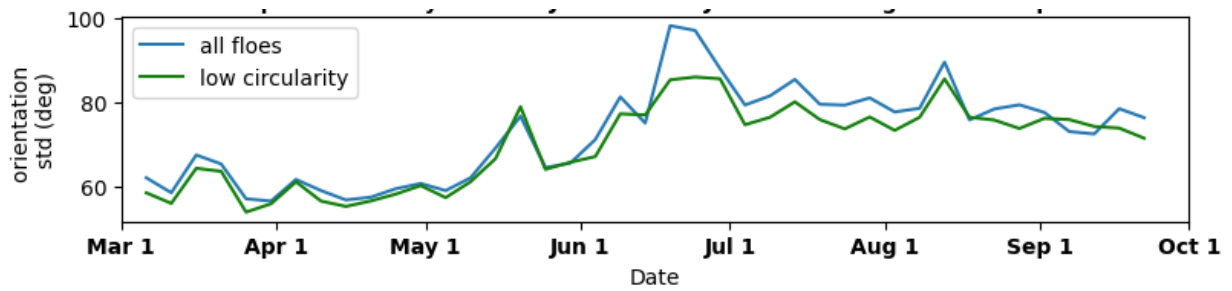
(1) It's possible that nearly circular floes could have a more variable orientation than elongated floes, because small changes in their perimeter could trigger large changes in their orientation. Therefore, as floe orientation becomes more variable (as in Figure 5(c) in June), how do we know whether it's due to floes becoming more randomly oriented or floes becoming more circular?

This is a good point. We looked at a subset of larger floes (>75th percentile), a subset of less circular floes (<25th percentile) and a subset of higher eccentricity (major axis/minor axis) floes (>75th percentile). We observed similar patterns in the evolution of the standard deviation of orientation to ensure that the pattern we were seeing is not just because floes are becoming smaller, more circular, and the eccentricity decreasing. We have added a sentence to the paper to address this comment.

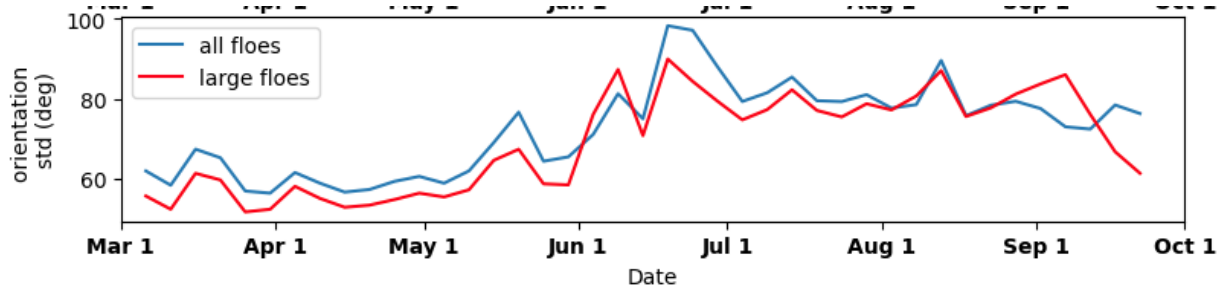
Here is an example of the evolution of all floes (blue) and the subset of just higher eccentricity (orange) - the trend is less extreme, but still present.



all floes (blue), low circularity (green):



all floes (blues) large floes (red):



(2) Lines 227-231 discusses the orientation of large rectilinear floes, concluding that their orientation is less variable in the early season compared to the full data set, but increasingly variable into the summer. It's not clear what the point of this comparison is. The reader is left hanging, with no figures or numbers or explanation of the significance of the result.

Thank you for your feedback. Floe orientation is related to the generation of fractures on the ice as a result of the stress being applied (e.g., from atmospheric features, like the Beaufort High and anticyclonic winds) and the effects of coastal geometry. By looking at the standard deviation of floe orientation, we can infer at what point the ice movement transitions to free rotation- and if floe rotation can be linked to ocean vorticity underneath the ice. We have added text discussing floe orientation in more depth:

“The orientation of the floes is an indication of the stresses that have caused ice breakup and can provide insight into the structural properties of the ice pack to inform how it may respond to future stresses. When ice floes are able to rotate without interaction with other floes, their rotation rates can be related to ocean vorticity (Manucharyan et al., 2022).”

Minor Comments

First line of the Introduction. Consider changing "from the atmosphere to the ocean" to "between the atmosphere and the ocean" because the fluxes go both ways -- atm to ocean and ocean to atm.

Thank you, changed.

Lines 52-53. "The older and thicker multiyear ice is now melting out" -- but the loss of older and thicker multiyear ice has been going on for a long time. It was first noted in 2007:

Maslanik, J. A., C. Fowler, J. Stroeve, S. Drobot, J. Zwally, D. Yi, and W. Emery (2007), A younger, thinner Arctic ice cover: Increased potential for rapid, extensive sea-ice loss, *Geophys. Res. Lett.*, 34, L24501, doi:10.1029/2007GL032043.

This phrase “now melting out” refers not to recent years, but to recent decades- as described in the

study referenced at the end of this sentence that is specific to the Beaufort Sea. We removed this confusing phrase and added the Maslanik reference to the end of the sentence. Thank you for mentioning this study.

Lines 95-102. This same segmentation procedure was also used by Stern et al (2018a) -- see their section 3.2.

We added text to describe that this algorithm is based on Denton and Timmermans (2022) and Stern et al, (2018).

Lines 95-96. "where where"

Corrected

Lines 110-111. "from from" should be "range from"

Corrected

Line 119. "This quality assurance steps" -- singular or plural?

Singular, changed to "step"

Line 130. "The standard error of the fitted distribution" -- is this supposed to say the standard error of alpha?

Yes, thank you for catching that, corrected.

Line 150. "a correlation of 0.99" -- is this supposed to be a squared correlation? Figure 4 says " $R^2 = 0.99$ "

Yes, corrected to "squared correlation"

Line 166. "The the"

Corrected.

Line 167. "values increases" -- singular or plural?

Changed to value increases

Figure 3 caption: Line 3. "imaeg" Line 4. Either write "captured" or delete the word "capture"

Corrected

Figure 5(b). The caption refers to "the slope anomaly" but the figure shows the actual slope, not the anomaly. Figure 5(c). The caption refers to "low circularity floes" in green but I don't see a green curve. Figure 5(d). The caption refers to "land (magenta)" but I don't see any magenta color. Figure 5(d). Is it really necessary to show all that gray area (clouds)? That makes it harder to see the green and blue areas, which are really the quantities of interest. Line 204. "Fig. 5b, shaded" -- I don't see any shaded part of Figure 5b.

Corrected. Cloud bars were removed, no shaded regions.

Lines 203-204. "the standard deviation calculated as in Equation 3." Is this supposed to say "standard error" or "standard deviation"? Does Equation 3 give the standard error or the standard deviation? See line 130.

Standard error. This has been revised.

Line 211. 59400 should be 594

Corrected.

Line 225. "that are can be"

Corrected.

Line 239. "increasing power law slopes" -- does this mean steeper or shallower? When alpha increases, the slope decreases (becomes more negative). When alpha decreases, the slope increases (becomes less negative). That's why it might be clearer to say steeper or shallower.

I see why this is confusing. We do not include the negative in the alpha- so increasing alpha means more negative and steepening slope. But this is not an increasing slope. Changed to increasing alpha (steepening power law slope)

Lines 241-242. This is the first mention of "discrete element models" in the paper. I would suggest that on line 37 the words "discrete element" be added just before the word "models".

Corrected.

Line 245. "new segmentation algorithm presented here" -- the segmentation algorithm is not new. See the comment above for lines 95-102.

Thank you for bringing this up. We changed this to "modified" and added a longer description in the introductory part of the methodology (Section 3) to describe how this has been adapted from previous work.

There are other minor typos not noted here.

The manuscript has been proofread to identify typos (see new version of the manuscript).

Reviewer 2:

The authors presented a study on sea ice floe size distribution in the Beaufort Sea based on the long-term MODIS dataset.

The method of ice floe segmentation is primarily based on the previously developed method by Paget et al. (2001). MODIS-derived ice floe segmentation was compared with the Sentinel-2 data based on three cases. They applied the method to 4,861 images of MODIS, identifying more than 9.4 million floes over 23 years. By analyzing the FSD in the Beaufort Sea over a long period, they unfolded the seasonal variability, including the decreasing mean floe area and increasing FSD power law slope as the summer progresses.

Some major comments and specific comments are below for reference.

Response: We sincerely thank the reviewer for a thorough assessment of the manuscript. We believe that these comments and suggestions have improved the clarity of the presentation. We respond to the reviewer and describe changes to the manuscript in blue font below.

Major Comments

1. Obviously, the title of the manuscript is too wide, and the authors only conducted analysis in the Beaufort Sea.

Thank you for bringing this up. We changed it to include “Beaufort Sea”.

2. The dataset is not well described. For instance, are all four bands of the S2 data used to segment ice floes?

Sentinel-2 has a total of 13 bands, four of which are at ten meter resolution. We apply the IFT to the Sentinel-2 imagery. The IFT uses the red channel for ice-water discrimination. We have added clarification to the description of the dataset.

3. The methodology is too brief to be understood. Is the method exactly the same as that described by Paget et al. (2001)? Are any improvements achieved?

We now provide further details in the text on the erosion-expansion methodology in the introductory part of the methodology section. The methodology is not the same as Paget et al. (2001). Paget et al. (2001) performs an expansion-erosion algorithm. Stern et al. (2018) and Denton and Timmermans (2022) methods are also based on erosion-expansion operations but add an iterative component to identify different sizes of floes each round. Our algorithm builds on these two algorithms but differs in that the structuring element shape and the ice-water discrimination techniques were redesigned for MODIS imagery.

The text has been edited to highlight the importance of iteration in the expansion-erosion methodology:

“Paget et al. (2001) describes an erosion-expansion algorithm that erodes the boundaries of floes to separate them, subsequently regrowing them to their original shape. Denton and Timmermans (2022) and Stern et al. (2018a) introduce an iterative procedure that cycles through erosion-expansion varying the amount of erosion to identify floes of different sizes. In this work, we build on the algorithm described in Denton and Timmermans (2022). We rewrite the algorithm in Python, and automate it to process thousands of images consecutively, introducing adaptive thresholds. The algorithm consists of four steps: pre-processing, ice-water discrimination, segmentation, and post-processing (Figure 2).”

Validation of the MODIS-derived sea ice floes is not convincing. On the one hand, only by presenting only three cases for nearly 5000 MODIS images seems to be inadequate. I would suggest that the authors add comparison experiments to provide a more robust argument.

Thank you for your comment. We need to keep in mind that the availability of validation

imagery is very limited. Sentinel-2 is only available 20 km from the coast. Finding pairs of near-coincident cloud-free images in two different data sets (MODIS and Sentinel-2) is challenging. We present three different ice conditions to demonstrate the ability of the segmentation algorithm to operate in different SIC scenarios. Also each Sentinel-2 image is 110 km x 110 km, so the analysis of three Sentinel-2 images and coincident MODIS imagery allows for comparison of 36,300 km², a sufficient area for validation. We have added additional text to the manuscript to address uncertainties.

On the other hand, the current comparison itself is not convincing. As the three cases shown in Fig.3, there are obvious distinctions (e.g., the N values in the figure) between the two datasets. The MODIS data have a spatial resolution of 250 m, which is much larger than the S2 data of 10 m. Therefore, a threshold should be set to exclude those floes that MODIS cannot observe at all due to its coarse spatial resolution.

Thank you for bringing up this point. Given the different spatial resolutions, we set the minimum floe area as 5 km² for both the Sentinel-2 and the MODIS dataset. Even though more floes are identified in the Sentinel-2 imagery, the floes picked up in both data sets display similar FSDs. This highlights the value of MODIS to analyze FSDs dating back to the early 2000s.

Moreover, the validation shown in subsection 3.6 fails to "perform equally well". These ice floes, which are not segmented by MODIS, not only include the clustered loose ice but also some pack ice. Such a discrepancy is too visible to be explained by the different data resolution. The proposed method probably has limitations in terms of accuracy (however, it is hard to judge because the method is described too briefly). It is possible that some steps, such as the use of adaptive thresholds, excessive morphological processing, or the direct discarding of low-intensity floes, lead to unreasonable results.

We now provide additional information on the methodology as outlined above. We added a sentence to address that the algorithm performs well on both types of imagery, but does identify more floes in Sentinel-2 due to the higher resolution:

“Although the algorithm is able to identify smaller floes in the Sentinel-2 imagery, the floe sizes retrieved and the floe size distribution agree well between the Sentinel-2 and MODIS imagery.”

We developed the algorithm based on the success of previous algorithms (Denton and Timmermans 2022, and Stern et al., 2018). The adaptive threshold is for accounting for variable but minimal low-level fog over the open water, as described in Section 3.2. Stern et al. (2018) also applies an adaptive threshold for ice-water discrimination. This adaptive threshold is different but similar in purpose. Further, as noted in Section 3.5, since the low intensity floes account for less than 1% of floes, removing them does not introduce additional uncertainty.

4. The title of section 4 should be narrowed down.

Thank you for your comment. We apologize if we are misunderstanding the comment, but we believe that the title of section 4 describes exactly what is included in the Section. The title is “Spring to Summer Transition of Floe Characteristics,” and we go on to describe how floe area, floe size distribution, and floe orientation evolve from March 1 to October 1.

I think that many relevant studies have been conducted with the FSD in the Arctic. Why not compare some previously derived FSD with the present results for further verification of its accuracy? Additionally, the statistical charts should be improved. For Fig. 5, it is fine to apply the "10-day running window", but one should also consider quantitative results such as scatters, boxplots, or upper and lower significance intervals to give the reader a clearer view of the author's raw statistical results. It would be better to try seasonal statistics, which might work better for the author's dataset (as too many MODIS data are excluded from analysis due to the cloud effect).

P2 L26 "see Stern et al. (2018b) for a comprehensive list of FSD studies": Even without a systematic review, a proper overview and summary of state of the art should be briefly presented here.

P2 L29-L31 "These studies have advanced our knowledge of seasonal evolution of the FSD": The authors' review does not serve to summarise this knowledge.

Here we grouped together a few comments regarding the suggestion for more comparison to existing studies.

We discuss the spring to summer evolution seen in our data as well as in other studies. We have added more quantitative comparison with the Stern paper as well. (Section 5.2). We added a figure to show how the cloud fraction evolves over the summer. This figure also includes the number of floes identified in each 10 day running window to address the reviewers comment on more statistical information. The floe area mean, median and 75th percentile are shown in Figure 6. The standard deviation of orientation and alpha value are computed with all the floes in the 10-day running window and thus boxplots and scatters are not applicable. The standard error associated with the alpha value is discussed in Section 5.2.

Specific Comments

P2 L25 "floe size distribution": The full name has already been presented. Similar issues occurred several times in the manuscript, including "SIC", "SAR", and "MODIS". Please revise them.

Thank you for catching this detail. We have revised the text.

P2 L27-28 "... from radar imagery ... high-resolution imagery ": The summary of research types is weird. Isn't a SAR image a high-resolution image? Please rewrite this sentence to provide a detailed review of these studies.

We appreciate your attention to detail. We changed to: high-resolution optical satellite imagery.

P2 L35 "(Lopez Acosta et al., 2019) demonstrate..." -> "Lopez Acosta et al. (2019) demonstrate...". Please also note other similar citations.

Thank you, these have been revised.

P3 Fig.1 (c): Rather than showing the annual average ice floe numbers here, I'd be more interested in first finding out the annual use of MODIS data. In particular, long-term statistics need to know the amount of cloud-contaminated data for each year.

Figure 1c shows the total floe count per year. We also include a figure (Figure 5) and an introduction to Section 4 to discuss these statistics further.

P4 Fig.2: The authors presented a very clear case of pack ice. In addition, I would also like to see the algorithm's adaptability to high SIC, melting ice, MIZ, ice-water mixing, etc. After all, the authors aimed to focus on the phenomena related to the transition, which implies a rather complex ice condition.

We appreciate this suggestion, but kindly disagree. The purpose of Figure 2 is to illustrate the methodology. We have included more examples of the algorithm applied to a range of ice conditions in Figure 3.

P5 L90 “400-pixel (100 km) neighbourhood of a pixel subtracted by a constant”: What is the constant?

Thank you for catching this typo. This has been revised for clarity. This was the description of the python function we used. But the default constant is zero.

The 400-pixel neighbourhood is a relatively large region. However, the masked MODIS contains many NAN values. How did the author choose the threshold at the edge of the NAN?

The threshold is applied to the masked image, so the NAN values are not included in the calculation for the threshold. We added a sentence: “At this point, the land and cloud pixels are masked and not considered in this step. “

Additionally, it would be better for the authors to argue for the rationality and specific benefits of adaptive thresholds where appropriate.

We agree with this comment. The justification for using adaptive thresholding is already described in the text: to consider variable pixel values of open water. We added a sentence to clarify this point: “In this way, we are able to account for varying brightness levels of open water given different ice concentrations and atmospheric conditions.”

P5 3.3 section: Since erosion-expansion is an important step, it would be better to show the effect before and after morphological processing using the Fig.2 case.

And, what is the necessity of performing multiple iterations? Actually, in my opinion, besides the fact that it does improve the visual effect, too much morphological processing may lead to losing the original sea ice features.

This is a good point. We adapted this iterative erosion-expansion algorithm from Denton and Timmermans 2020 and Stern et al., 2018. Here, we review the original algorithms and focus on the changes we have made to apply similar methodologies to a different dataset at larger scales.

We interpret the second part of this question as why there are so many erosions within a round or why there are so many iterations. Hopefully we can clarify both.

In the first round, the erosion is very aggressive, so that small floes are eliminated, and the large floes are heavily eroded. Then the large floes that remain are tagged, regrown (dilated) the same number of times they were eroded, and then removed from the image. This iterative processes is repeated but with fewer erosions each round, so the size of the floes identified are decreasing in successive rounds.

In each round, the floes can be eroded multiple times since the erosion structuring element is small enough so that the features are not deformed. By removing just the border pixels we are maintaining the shape. And then this process is done multiple times to separate floes.

These methods are consistent with the iterative expansion-erosion algorithm presented in Denton and Timmermans, 2022 and Stern et al., 2018. Figure 3 in Stern et al., 2018 demonstrates the process:

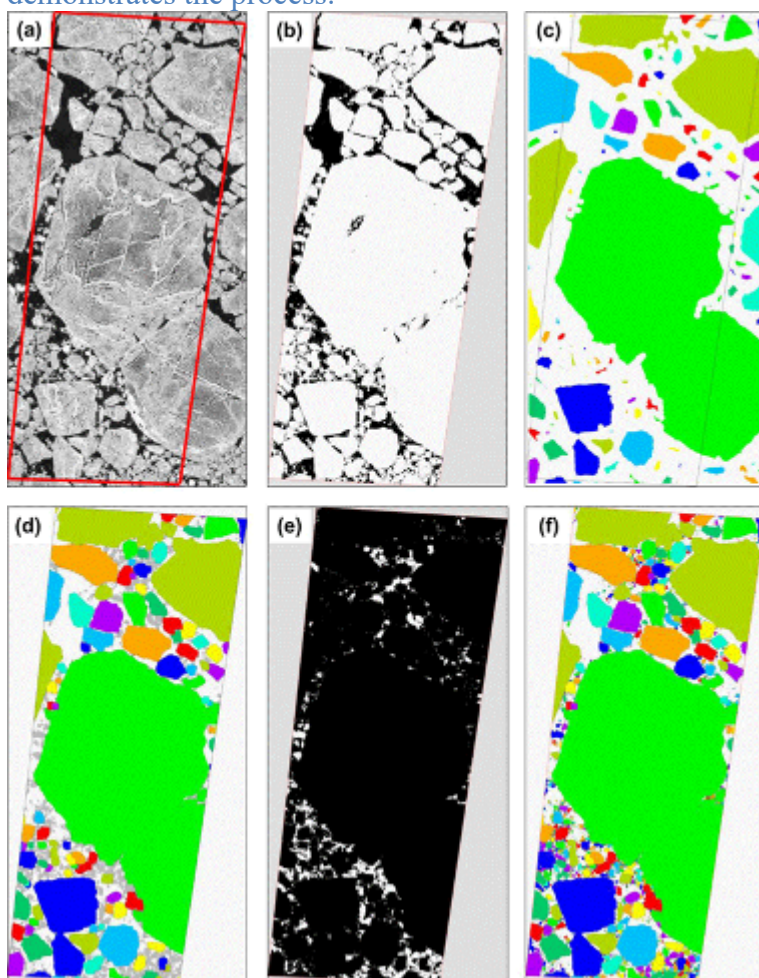


Figure Caption:

MEDEA image processing steps. (a) Subset #1 of July 8, 2014 (within the red outline). See Figure 9 for location. This subset measures 3294×6537 pixels (3.3×6.5 km). (b) Binary image, first iteration. (c) Eroded and labeled image with holes filled, showing individual ice floes. (d) Re-grown ice floes in color, with unidentified ice floes in gray. (e) Binary image, second iteration, after the largest floes have been tagged and removed. (f) Final labeled image showing all identified floes. DOI: <https://doi.org/10.1525/elementa.305.f3>

We have text throughout Section 3 that references the methodologies in these studies and includes more statistics of the erosion-expansion algorithm.

P5 L103-L105 “On average, 26% of the classified sea ice area is identified as individual floes”: Average of what?

Of all the images analyzed. This sentence is clarified: “Over the thousands of processed images, on average, 26% of the classified sea ice area is identified as individual floes, with remaining sections consisting of ice filaments, brash ice or pieces of ice smaller than the minimum detectable floe size.”

P5 L111 “...the variation of the floe orientation”: Which floes are the authors using circularity std to compare their variation? Are they ice floes that are tracked between time-series images, or are they all ice floes in the same image?

All ice floes in the 10-day running window of images. Added “within the 10-day window” to clarify

P6 L116-119: In fact, the low-intensity ice may be an important ice condition as well (especially in the transition, where it may represent a melting scenario). However, the author removed them outright, which would cause the subsequent results, especially for power law distributions, to be different from previous results. How can the authors justify this proposed step? Also, what are the units of 150? Can the authors prove <150 to be so-called brash ice rather than other types of sea ice (e.g. grey ice, melting ice)?

The pixel values are an 8-bit integer from 0-255. We look at the red channel as noted in the manuscript. 150 is determined empirically by testing various thresholds and examining the pixel brightness of floes, brash ice, and mixed pixels. Because of the size of each pixel (250 x 250 m), mixed pixels containing more than one surface type are present. The 150 intensity value ensures that low brightness ice floes are still included but mixed pixels over ambiguous surfaces are not included in this analysis. Also- as mentioned in the manuscript- this only removes $<1\%$ of the floes. The further analysis of how floe brightness intensity may represent the surface melt evolution is an interesting question but beyond the scope of this paper.

P6 L122-123 “that complements other parameters commonly used to describe ice floe fields, such as the sea ice concentration (SIC) and average ice thickness”: I would suggest deleting this sentence.

We choose to retain the information as we think it is important to acknowledge there are multiple ways to describe the sea-ice cover, but have made edits to the paragraph in question. We have reworded this: “Taken together with the floe geometry properties outlined in Section 3.4, the FSD and other parameters commonly used to describe floe fields, such as the sea ice concentration (SIC) and average ice thickness, allow us to study the physical processes that shape the structure and evolution of sea ice.”

P7 3.6 section: When using S2 images for segmentation, is there also a step to remove the <150 intensity? And, are the other parameter settings exactly the same (number of erosions and the calculation of adaptive thresholds)? What level of S2 data is used? These details may lead to a different adaptation of the ice floe segmentation algorithm to S2 data.

Thank you for bringing this up. We use Sentinel-2 Level 1C Top of Atmosphere Reflectance- this detail was added to the manuscript. The same algorithm was applied to both datasets, and there is a step to remove low intensity objects.

Fig 3&4: As I mentioned above, I don't think the comparison presents a good result. What is the "xmin"? Also, please present clearer quantitative results in the text.

We added the xmin value to Section 3.5. We have clarified the figure and added discussion on how the high sea ice concentration example exemplifies limitations of the algorithm. We hope that clarifications and additions elsewhere (including the addition of an Appendix deriving the analytical expressions) provide clearer quantitative results.

L149-151 "The areas of the matching 82 floes agree very well, with a correlation of 0.99, and an absolute mean area difference of 0.18 km² (Figure 4)": The authors only compared the 81 floes identified by both. I don't think this comparison is fair. The high-resolution S2 results should be treated as a reference, and all the S2-segmented floes should be compared to MODIS results. Obviously, the MODIS results have gross omissions.

The comparison of floe areas can only be done with the floes that are identified in both images. The power law distributions shown does include all floes identified in both the images that are in the MODIS range of detectable floe area. Here we are demonstrating that in the floe area range that MODIS can identify floes, there is strong agreement with the Sentinel-2 validation dataset. There is another question of the omission of smaller floes that I believe the reviewer may be referring to. We are carrying out a further error analysis to understand these limitations, but this is beyond the scope of this paper. Here we just assure that MODIS is acceptable for this range of floe sizes.

P7 L153-154 "...14% more floes are identified in the MODIS imagery compared to the Sentinel-2 imagery.": How can the author tell the 14%? I can only find that MODIS significantly underestimated the floe numbers.

We do not limit the floe size detected in the Sentinel-2 image. To compare likes, we sample the floes identified in the Sentinel-2 imagery that are in the MODIS size range. The 14% more floes are in the area range that MODIS can detect- so greater than 5 km². We have added a sentence to clarify that the Sentinel-2 identification includes floes less than 5 km². "as the identification of floes in the Sentinel-2 imagery was not limited to the MODIS range of floe sizes."

P7 L163 "We create 1000 bootstrap samples with replacement": I can't follow the author's point of this step.

This text has been modified for clarity. We include this step to show that the power law fit is a robust description of the dataset.

P7 L171 " SAR data is not as widely available": This sentence is misleading. If it refers specifically to long-term applications, SAR does have limitations. "other limitations or complications, such as speckle or granular noise". It is not correct. It is just the authors do not know how to deal with the valuable dataset.

Thank you for this point- we are referring to the availability of SAR data for this study which is a long term study that includes the summer time period. We have clarified the text: “However, SAR data is not as widely available for long-term applications, and has other limitations and complications, such as speckle, granular noise, and ambiguous returns when meltwater is present on the ice surface.”

P8 L181-182: “...median floe size exhibits a similar pattern as the mean, with maximum median value of 17 km² on 10 April...”: The variations in the median are not visible at all, and this image should be modified.

Thank you for the comment- this image has been modified. A new panel has been added to separate the mean and the median.

P11 L210 “... Stern et al. (2018a) ...”: Stern (2018a) included the analysis of small ice floes (<5km²), right? So, it can be different from the authors' results.

In fact, in section 3.6, the alpha estimated from the MODIS is also lower than the S2 result. It is suggested that the proposed algorithm is supposed to have some limitations. I would suggest that the authors explain such reasons in detail.

Yes, Stern et al 2018 study includes floes as small as ~ 2.6 km². We have added a line: “These values are similar and the small differences are likely due to location and time of the observed floes.” We acknowledge that this algorithm and dataset have limitations- this is already described in Section 3.6 Validation and Limitations.

P12 L223 “...we find a low variability in their orientation (Fig. 5c)”: Low variability? It seems to me that Fig. 5(c) displays a clear increasing trend. Doesn't this suggest that as the sea ice melts, the direction of the ice floes becomes more cluttered?

“this period” refers to the spring which was described in previous sentences. we replaced “this period” with “the spring” for clarity.

P12 L223-224 “This effect is especially noticeable in areas of high ice concentration, where the ice movement and readjustment to external forces is limited by the surrounding ice ”: It’s hard to follow. If the authors mean that there is a decreasing tendency for std as SIC increases, it should be expressed more clearly.

Otherwise, the most obvious phenomenon on this graph should be the very wide range of std variation for SIC > 0.7.

Added for clarity “lower standard deviation of orientation with greater SIC”

P12 L227 “large rectilinear floes”: This is less common usage, so please confirm.

Confirmed. Rectilinear means straight lines, and this is how we see the ice fracture in the spring.

See also:

Arntsen, A. E., Song, A. J., Perovich, D. K., & Richter-Menge, J. A. (2015). Observations of the summer breakup of an Arctic sea ice cover. *Geophysical Research Letters*, 42(19), 8057-

8063.

where in the abstract they describe the fracturing in the ice pack resulting in rectilinear floes.