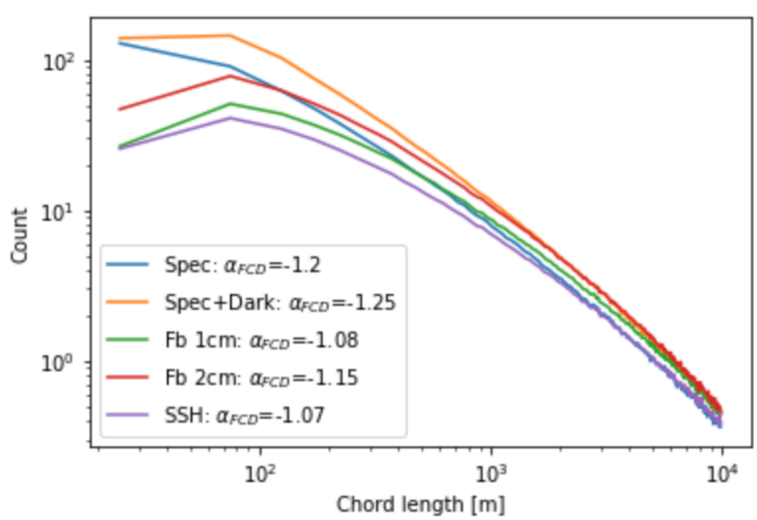
Reviewer 3

This is a review of Gupta et al (2024).

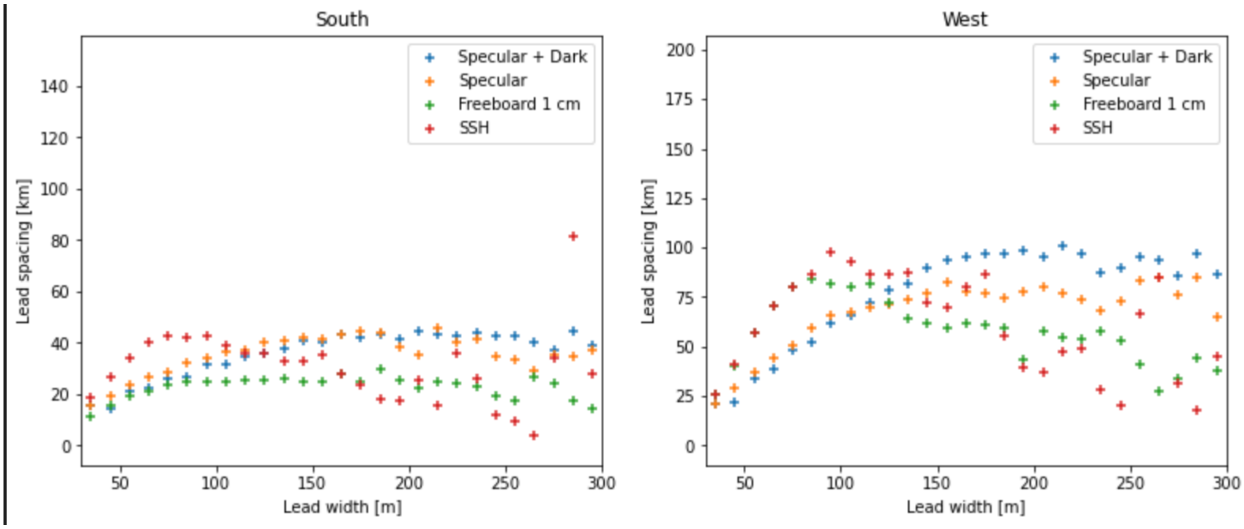
The work uses ICESat-2 altimetry to understand properties of the shape and size of sea ice floes in the Weddell Sea. Generally the paper is well-written and interesting - as a case study in how laser altimetry can be used to explore the properties of the sea ice surface. I think there are some methodological questions I am hoping to have addressed in a revised manuscript.

The authors do discuss this, but it is worth emphasizing. The use of radiometrically-defined "dark lead"s, remains a challenge with ICESat-2, because of a lack of confidence in their meaning. I would like to see your Fig. A3 as an anomaly rather than side-by-side plots to show the differences. Perhaps in addition, showing the histograms next to one another of all lead spacings and chord lengths. Chord/spacing measurements with IS-2 have often been confusing, as with the high along-track resolution it is very easy to chop a floe in half with a misclassification. It might be in the Weddell this is not an issue, so that's exciting! But more details here might be helpful.

We certainly notice differences in lead locations and widths when considering different lead definitions for a given track. When averaging over the large areas of the south and west Weddell Sea, these differences translate into some differences in the FCD. The plot below shows the annual-mean FCD curves for different lead definitions, as requested. We will include this in the Appendix. Despite these differences, we find that the seasonality of the FCD slope anomalies is mostly robust across them (Fig. 4a). We cannot comment on whether this is true only for the Weddell Sea or elsewhere.



The plot below can allow you to more easily compare the lead spacings of Fig. A3 for the annual-mean fields. We will include this in the Appendix.



More details are needed on the power law slopes. For example, how are you fitting alpha to the FCD? A treatment of this mathematically is in several places, including Virkar and Clauset (2014), but has to be done carefully. The use of binning can introduce spurious errors in the slope of such a distribution - see Stern et al 2018. The VC method is simple to apply and doesn't rely on the binning. The statistical tests examined there should be applied before discussing power-law fits (if they are not already used).

All the slopes alpha_XX are obtained by linear regression (ordinary least squares) over the binned data. We recognise that this method has been shown to cause biases in the estimation of the exponent in a power law model. Alternative methods, such as suggested in the comment above, assume that the distribution follows a power law and use this assumption to reduce biases, notably those caused by the tail end of the distribution, where data is usually more sparse.

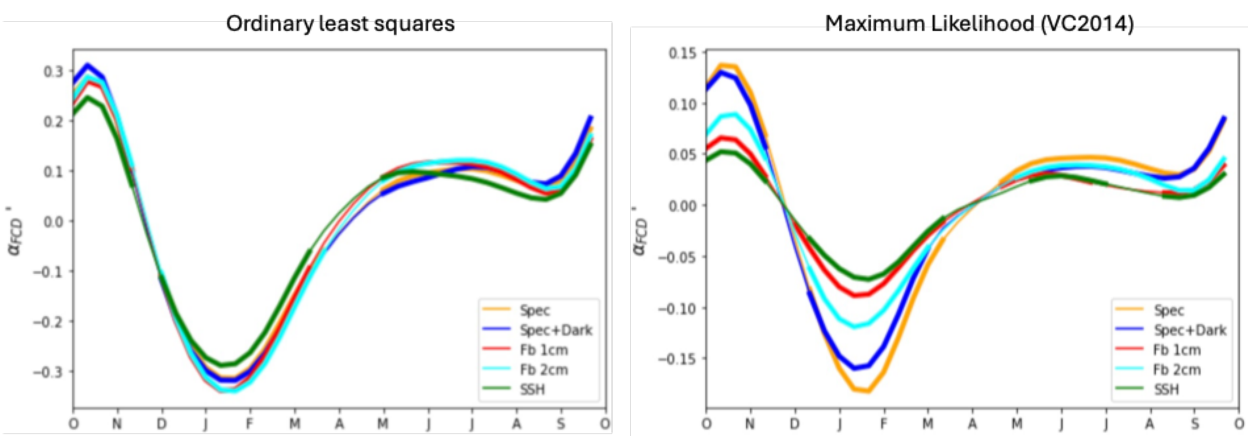
Here, we can clearly see that the FCD is not distributed like a power law over the range of floe chords that we considered - ie. the curves are visibly not straight lines in log-log space (Fig. 3a). Given this result, we do not believe that a power-law is necessarily the best model to fit this distribution, and others may indeed provide a better fit. However, our objective here is to characterise the broad-scale changes in the distribution over a seasonal cycle. We thus seek a simple model that is easily interpretable (in terms of slope and intercept) and can fit the curve reasonably well. The power law model meets these characteristics, with R2 typically varying between 0.9 and 0.98. Since we do not assume that the original curve follows a perfect power law, we do not believe it would be appropriate to use more sophisticated methods for inferring alpha_FCD.

We acknowledge that this approach comes with caveats:

- We are reducing a curve to a straight line. This prevents us from distinguishing over which range of floe chords the changes are occurring throughout the seasonal

- cycle. More detailed investigations could be possible with a more complex model, but this would be more difficult to interpret than a single slope, and is beyond the scope of our study.
- Biases at the tail end may still exist, but this is difficult to eliminate without making an assumption about the nature of the distribution, which we avoid.
- Biases due to bin spacing may remain. We do not find a notable difference between them in the range 75 to 300 m for the annual mean value. The default was 75m.

For completeness, we calculate power law fits of the FSD using the Maximum Likelihood method outlined in Virkar and Clauset (2014). The plot below shows these fits over the seasonal cycle, for the different lead detection methods, compared with the ordinary least squares method. The Maximum Likelihood method does make a substantial difference to the fit, but the overall seasonality remains robust.



On Fig. 3a we will show the power law fit lines. We will add a condensed version of the above to the discussion, which applies to other alpha values as well.

I would suggest renaming your alphas (one is a power law, others are not) because they correspond to different distributions.

This could be possible indeed, but we feel it would impede readability slightly. We will instead indicate more frequently which type of distribution we are referring to.

It is also important to discuss that there is fITD uncertainty, much like FCD uncertainty, because of the unknown misalignment of the laser with sea ice structure (in this case, ridges, say).

That is a good point. We will add a sentence to the discussion.

The extremely high correlation ($R^2 = 0.98$) between mean freeboard heights and mean chord length is striking, but somewhat concerning - do you think this is because thicker ice is easier to separate from open water? In that case, thicker ice means fewer missed classifications means wider FCD. Have you examined this possibility? Can you explain why this correlation is so strong?

The explanation you provide could indeed be part of the correlation we identify. The other physical processes we discussed may also play a role. However, we cannot disentangle their individual contributions. We will add the mechanism you propose to the discussion.

Regarding the interpretation of the high correlation, we note that the R^2 metric was applied to the 'mean' freeboard values when binned across chord lengths, which is well represented with a linear fit across seasons and regions (blue points and red line in Fig. A4), giving a high R^2 value. However, there is a substantial number of floes that deviate from the mean freeboard (grey clouds of points in Fig. A4), as quantified by the standard deviation (blue bars in Fig. A4).

We will add the standard deviation to Fig. 3c to avoid any misinterpretation of the R^2 value.

For the floe roundness elements - an important question here is what is the intended usage of floe freeboard roundness information? I assume this is to infer something about the dynamical behavior of the floes contacting on another? This could be better justified because this metric, especially with a laser altimeter that has uneven sampling and only measures freeboard, is somewhat unclear. It would be useful to see exactly how ICESat-2 "sees" an individual floe rather than the composite normalized image of Figure 7. - because floe surfaces are very non-round (see your Figure 1!) Freeboard variability can be from changes in sea ice density, and "floe roundness" is realistically more closely compared to the surface roundness alone. How round the floe is depends on factors under the ocean surface that ICESat-2 can't examine here.
It would be helpful to discuss averaging here - I presume in the paragraph beginning on L264 you mean you average over all floes, not average along the floe. How much variance is there in the resulting plots shown in Figure 7? This could be a good visualization - showing the variability associated with your compositing.

The roundness we calculate indeed only informs us about the surface above sea level, and not about what happens under water. We will remove 'lateral melt' as a possible explanation.

Yes, we do mean average over all floes.

The variability in the mean profiles in Fig 7 is indeed very high and reflects many asperities seen on floes in Fig. 1. The shading in the plot below reflects 2 sigma standard deviation. This means that the roundness metric is unlikely to be useful when only considering a small number of them. Nevertheless, we find it interesting that the shape collapses to such clearly defined half domes when averaging over a sufficient number of floes. It remains to be determined what processes may be responsible for this. As our knowledge of floe-level processes continues to increase, floe roundness may be a useful metric for differentiating the variability of these processes over time and spatial scales, and for model calibration.

Note that we also noticed a bug in our calculation for this plot where the curves were being scaled down artificially. We have now corrected this. Therefore, the d values shown below are slightly different to the submitted manuscript. We will add the corrected version of the plot to the manuscript, including the variability of the profiles.