

The manuscript would benefit from some further discussion for the general reader that better situates the findings in the context of the published literature and suggests implications for the broader scientific community, rather than simply for the development of neXtSIM. For example, why is it the mono/multifractal scaling interesting? Does the work suggest parametrizations that could be used in traditional climate/sea ice models? This comment is merely a suggestion and the authors are not required to address it.

This is a very insightful comment as it goes to the heart of the motivation of the paper. Both this comment and comments by the other reviewer show that the paper's motivation needs to be stated more clearly - something we have tried to do in the modified manuscript.

#### Major comments

I find the exclusion of the Beaufort Sea from the model-observation comparison to be poorly motivated. The model-observation comparison does not seem fair if the authors exclude regions where the comparison is poor. The authors need to motivate the exclusion of this region better and quantify to what extent the exclusion of the region affects their conclusions, or refrain from excluding it in their analysis.

Indeed, the motivation for excluding the Beaufort Sea was not clear in the original submission. Our main reason for doing so was an apparent inconsistency in the observations between the Beaufort Sea and the rest of the Central Arctic. Further work on this issue - following this review - has, however, made it clear to us that excluding the Beaufort Sea is not well justified. In the revised text all analysis and comparisons are made on the "Arctic" and "Central Arctic" regions only.

Data availability – According to The Cryosphere data policy, 'Authors are required to provide a statement on how their underlying research data can be accessed.' This is missing from the manuscript.

We've added this to the revised manuscript.

#### Line-by-line comments

L7 Wasn't the Central Arctic region chosen to avoid the presence of polynyas?

It should have said wider Arctic - this has been corrected in the revised text

L18 'In particular . . . ' – add citation(s)

Done

L19 'Leads . . . are a much more temporally and spatially clustered gateways' – this is unclear, please revise

We've rewritten the sentence as "Leads in the centre of the Arctic Basin, on the other hand, are much more difficult to study because they are much narrower and shorter-lived than polynyas, and at the same time can form anywhere in the Arctic Basin."

L25 'causes' .. 'causing' – repetitive, rephrase

Changed 'causes' to 'drives'

L32 Missing parentheses

Fixed

L33 What area did the satellite image cover?

It was a 60x66 km<sup>2</sup> image, we've added this to the text.

L36 'accurately reproduce the properties of lead fraction statistics' - add citation(s)

We've rewritten this statement to be more accurate and added references

L43 In what situations does the mixed layer deepen in response to brine rejection? My understanding is that this is what occurs in low resolution models. Barthélemy et al. (2015) (<https://doi.org/10.1016/j.ocemod.2014.12.009>) could be cited here.

The mixed layer shoals when (relative) ice velocity is low and the brine forms a plume that sinks to the bottom of the mixed layer. When the ice velocity is high the shear induces mixing of the brine which in turns causes a deepening of the mixed layer. In (most) low-resolution ocean models the brine is released uniformly into the first ocean layer causing mixing and deepening of the mixed layer - similar to the high-ice-velocity regime. Nguyen et al showed that releasing the brine at the bottom of the mixed layer improves the simulation. This is described in the text, although we have modified it slightly to make it clearer.

We were not aware of the Barthélemy paper, but it was very interesting to see the importance of including both the low- and high-velocity regimes. So we've added a reference to that as well. Thank you for pointing it out.

L48 'is actively being researched' – add citation(s)

Done

L49 'Lead formation is closely linked . . .' – add citation(s)

This sentence was poorly formulated. We've reformulated it to read "When sea ice deforms ridges and leads are formed." - which doesn't require a citation.

L54 Suggest defining multifractality here.

We have added a discussion of the relevance of (multi)fractality here.

L55 'This fundamental property . . .' This sentence needs more explanation for the general reader.

We've removed this sentence in favour of the more detailed discussion added.

L74 Please provide more details on the slab ocean. Does it include any representation of ocean currents? How might the simplicity of the modelling configuration affect results?

We intentionally didn't go into much detail about the model - including the slab ocean - since this is presented in more detail in Rampal et al (2019) and Rampal et al (2016). We did still add the sentence "Oceanic heat loss results in lowering of the slab ocean temperature, which may be compensated for by new-ice formation and nudging of the slab ocean layer temperature to reanalysis results." to the text and hope you find it sufficient to address your comment.

As for how using a slab ocean affects the results then we expect this to be minor. We expect a larger contribution from the atmosphere in this respect, and that is mentioned in the discussion. To address your comment we've added a few sentences to this effect in the discussion section.

L95 Define 'node' and 'cohesion'

'Nodal spacing' was left over from a previous version of the manuscript. The correct phrasing is simply '12.5 and 25 km resolution', and we have changed this in the revised manuscript. Mentioning the cohesion at this point borders on being too precise, in our opinion and we prefer to refer to the discussion in Bouillon and Rampal w.r.t. scaling of cohesion. We have slightly reformulated and rephrased to make this clearer.

L99 'The deficiencies of the linear viscous model are well known' – add citation(s)

Done

L109 How is it an improvement?

The new product fixes an overestimation of the lead fraction present in the original.  
We've added a sentence to that effect in the text.

L126 Why are the heat flux magnitudes provided as snapshots rather than daily means?

They are not - this was left over from an earlier version of the paper. The heat fluxes and lead fractions are both daily means.

L154 Reword 'for future works' to 'in future work'

Done

L167 'that gives good statistics.' What does this mean?

It just means that it gives the same slope of the PDF, as per the latter part of the sentence. It now reads "We, therefore, choose a threshold thickness for the model that has the same slope of the PDF as the observed one, as shown below."

L168 Why not simply use a threshold of 10 cm? How much does the choice of threshold affect the results?

We wanted to see if we could deduce the appropriate threshold from the model statistics. The fact that the two are very close shows that the processes in the model are a reasonable reproduction of those in reality. Having done this it doesn't matter much which value we use for the comparison, the results are essentially the same. The paper also says that variations of the threshold of about 1 cm give the same results.

Note, however, that the 10 cm threshold is more of an educated guess and probably not a fixed threshold anyway - there is also a dependence on emissivity, frost flowers, and probably other factors as well. Since the 10 cm threshold is a very rough estimate we see no reason to prefer that over the model deduced threshold of 9.1 cm.

L182 Why do you think the model does not capture this?

We don't exclude the Beaufort Sea anymore, so this comment is not relevant.

Fig. 1 caption – define the red dashed lines. 'read' -> 'red'

The dashed line was what we exclude as the Beaufort Sea, but this is not relevant anymore.

L191 'excellent agreement' The figure is in log space, so some of the model- observation differences seem not insubstantial.

Indeed this was maybe overly enthusiastic wording. We go into more detail below about the model-observation differences to show that the observation shortcomings are such that it's hard to say to what extent the differences are indeed substantial. We, therefore, replaced 'excellent' with 'good' and added a sentence highlighting that the agreement between model and reality is probably much better than the comparison in the figure seems to indicate at a first glance.

Fig. 2 caption – are these lines excluding the Beaufort Sea? Typo in 'Arctic'

Yes, but as we don't exclude the Beaufort Sea anymore, so this comment is not relevant.

L210 What is 'proper' spatial scaling?

We've removed the word 'proper' - it should not have been there.

L275 'after some algebra' – this wording is too casual for a journal article

We've removed the phrase 'after some algebra'

L315 Reword 'this model shortcomings' to 'these model shortcomings'

Done