General comment:
The authors choose rather complex statistical tools by analysing the heavy-tails of PDFs and the spatial scaling. These methods are appropriate to study the localisation of lead density and simulated heat flux, which is the main topic of the manuscript. However, a comparison of the spatial distribution of lead density as done in Wang et al. (2016) and Hutter & Losch (2020) is missing, although all data would be available for that. In Fig. 1 such a comparison is made for a snapshot of a single day. I recommend to add a comparison of spatial distribution of lead-density for the entire winter analysed in this paper (maybe replacing Fig. 1). In doing so, the model evaluation of this manuscript would be more comprehensive by showing that the model (might be)/is able to reproduce the large-scale spatial distribution and the strong localisation of lead-density.

Indeed, we have not made the kind of spatial distribution analysis that Wang et al and Hutter and Losch did. This would indeed be an interesting additional metric for a model-observations comparison, but we feel that adding it to the present paper would not be appropriate. The reason being that in this paper we are introducing the notion of using spatial scaling analysis to investigate lead fraction patterns. This is motivated by the spatial scaling we see in deformation rates, as we know that lead formation is closely linked to deformation.

In this paper, we show that both observed and modelled lead fraction patterns demonstrate spatial scaling and we then use the model to show that the (modelled but unobserved) heat-flux patterns also demonstrate spatial scaling. The model-observations comparison is therefore only there so that we can with some confidence say that there is likely also a spatial scaling of the heat fluxes in reality.

Adding the spatial distribution analysis, as suggested, would therefore not add to the central theme of the paper - which should be the spatial scaling of lead fraction and heat fluxes. We feel that adding an arguably parallel discussion to the paper in this manner would not be beneficial for the paper, but make it less focused and more difficult to follow. We therefore respectfully decline to follow the reviewer’s recommendation on this point. We do, however, feel that this comment, together with one of the other reviewer’s, makes it clear that the introduction and motivation of the paper needs to be improved - something that we have attempted to do in the revised manuscript.

Specific comments:
P2, line 32: “Andreas and Cash (1999); Esau (2007)” - wrong citation style

Fixed

P2, line 35: “including smaller leads increased by 55% the total estimated heat flux” - including smaller leads increased the total estimated heat flux by 55%.

Fixed

P2, line 35-37: I assume that the magnitude of the overall heat flux is adjusted by the tuning of thermodynamic parameters in coarse resolution climate models. However the spatial distribution and local magnitude might be off, if leads are not resolved in these models. Please clarify.

You are right. We have changed under-represented to misrepresented, which is more appropriate.

P2, line 48-49: “the statistical properties of leads in large-scale sea-ice models have not yet been shown to be robustly reproduced” - How about Wang et al. (2016) and Hutter & Losch (2020). Wang et al. (2016) shows agreement in the lead density in the Arctic between a model simulations and satellite observations. Hutter & Losch (2020) show that multiple spatial and temporal properties of LKFs, which are leads and pressure ridges, observed from satellite are matched by large-scale sea-ice simulations.

We have now included those references in the text. The point we wanted to make was that models that are normally used to study the Arctic have neither the resolution nor the numerics necessary to resolve these features. This is made clearer in the revised manuscript.

P2, line 61: “Section 2.1” - Section 2.1 presents only the model set-up. Please refer to Section 2.

Fixed

P3, line 61-70: This paragraph reads a bit wordy. Maybe consider to rephrase it.

We have rewritten parts of this and hope that it reads better now.

P3, line 83: “model mesh” - Model mesh or the mesh to which the model output is interpolated?

The model mesh. We have rewritten this sentence slightly to make this clearer.

P4, line 1-2: Not clear, from which data product concentration and from which product thickness is taken. Please clarify.

This is clarified in the revised text

P5, l 125 “order” - order -> orders

Fixed

P5, l 136: “2011” - 2011 vs model year 2007? In the model description it is written that the model is ran for winter 2007, later on in the paper you evaluate only the year 2011. Please clarify. Does this sentence anyways not rather belong to the results section?

It was supposed to say 2007 and we’ve fixed that (here and in several other places). We used 2011 in a previous version of the paper. We think that the sentence belongs here because it illustrates the method of using averages of JFM and does not address the results. The reference to figures 2 and 4 is for illustrative purposes and the figures are only discussed in the results section.

P5, l150: “PÌD âLij L−β(0)” - Supposing x_bar should represent the mean, it should be beta(q=1). For q=0 no scaling should be observable, if equation (2) is used (xˆ0=1 for all samples).

Yes, \bar{x} is the mean and it should be \beta(1), not \beta(0). We’ve fixed this.

P6, l152: “Stern et al. (2018) argue that this method provides a reasonably accurate estimate of the power-law fit.” - In addition, Stern et al. (2018) argue that no matter what method is used for estimate of the power-law exponents a goodness-of-the-fit test like in Clauset et al. (2009) should be performed. Please clarify, if you do such a test, or why it is not necessary in this case.

P6, l153: “might provide” - Replace by “provides”. Both Stern et al. (2018) and Clauset et al. (2009) say it provides better estimates. Given that the method is computationally not much more expensive, it is unclear to me, why you choose to use a more inaccurate method.

For the two points above: The second half of this paragraph was wrong. The MLE method of Clauset et al is to estimate if the PDF follows a power law, but this is not our concern here. We know that the PDF has a “fat” tail (doesn’t need to be a power-law) and then we can reasonably expect there to be a scaling relationship - which we do find. We’ve removed the erroneous portion of the paragraph and slightly rewritten the surrounding text without changing its contents or meaning.

P6, l171-172: “It is important to note that the simulated lead fraction is not strictly a lead fraction as it includes all open water areas, including polynyas (cf figure 1).” - How about using a smoothened concentration field to mask large open-water areas as around Svalbard?

If we’re doing that we would essentially be fabricating data since the smoothed field would have very different characteristics from the rest of the field.

P6, L 177: “showing a deviation from linearity at around 70%” - I can not see a clear deviation. Is it due to the dashed line style. An annotation to the plot could help to point the reader to what you mean.

This was not very clear in the original submission. But as we don’t exclude the Beaufort Sea anymore, so this comment is no longer relevant.

P6, l178: “When excluding this region, the observations also show a linear decrease (Fig. 2, solid blue line).” - This does not fit to the caption of the Figure2 (dashed for “Arctic” and solid for “Central Arctic”).

We’ve fixed the caption, which was wrong.

P6, l 182: “However, we suspect that the large number of small leads forming there may result in increased noise in the lead fraction product (see Fig.1) and an overestimation of the large lead fractions.” - Not clear to me. Please be more specific why more small leads lead to an overestimation of the lead fraction product. Or do you mean that by having many small leads the lead fraction increases, but the model does not resolve these small leads and therefore shows lower lead fractions?

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.

P8, l197: “than 20%” - Add (note shown) or reference to figure.

This refers to the discussion in the paragraph above, so we’ve added “(see above)” to the text.

P8, l 203: "strong indicator" - Be cautious, even if the scaling is right, the regional distributions could be off, i.e. high lead fractions found close to the coast or in Beaufort sea in observations could be reproduced by the model in different regions. To clarify this, please be more specific what you mean with lead-fraction patterns in the text.

It is right that our use of the words “lead-fraction patterns” is too broad and not supported by our results. We have therefore rephrased the sentence to read “... strong indicator that the model is simulating lead formation in a physical and realistic manner …”

P11, l 243: “In addition to these differences in the scaling, there also seems to be a difference in the nature of the structure function, depending on the model resolution” - Please also discuss the change in structure function for the lead fraction. It appears that the linear fit is not appropriate to fit the structure function of the coarse resolution models (The fit does not pass the uncertainty interval for q=1).

This is true and we have noted it now.

P13, l267-269: “We also assume that the closing is directly proportional to the area of the polynya since most of the heat loss and ice formation happens over open water.” - This assumption is not clear: I agree that ice formation is larger over open water, but if a polynya is formed instantaneously the entire area of the polynya starts to freeze at the same time. Please clarify.

This was indeed not clear enough in the initial submission. The point is that we can assume that the closing rate is directly proportional to the area of the polynya since the total heat loss and total ice formation can be assumed to be proportional to the area of open water. This leads to the differential equation and the rest of this simplistic model. We've modified the text to be more precise.

P14, l 219: “figure 4” - Please reference the subfigure for clarity.

Done

P15, l300-308: “This is partially due to the fact that neXtSIM . . . the lead-fraction and heat-flux scaling and structure functions across different model resolutions.” - This paragraph is not clear to me. It is difficult to follow your line of argumentation. Please revise and rephrase.

The idea here is that if the opening-rate scalings are not resolution independent then we should expect the lead fraction (and heat flux) scalings also not to be resolution independent. We've modified the text to try to make this clearer.

P16, l320: “Conclusions” - You provide rather a summary of the paper than a conclusion. So, please change the title of the section accordingly.

We've changed this to “Summary and conclusions”

Data and code availability: A statement is missing, where to find the code and data of this study.

Done

Figure 1: “Lead fraction larger than 0.05 is indicated in yellow.” - Why do you show the thresholded fields instead of using a colormap that highlights the 0.05 fraction about shows the entire range of lead fractions? I recommend to use a show the entire range of lead fractions.

We now use a colormap that highlights the important range of lead fractions.

Figure 2: “The dashed straight lines are linear fits discussed in the text.” - Could you use color to indicate which fit belongs to which data. Please use different linestyle for the fits and the “Arctic”. Please also add all lines to the legend to clarify. In the caption “Arcitc” should be “Arctic”.

We've reworked this figure after not excluding the Beaufort Sea anymore, and in this case there's no need to differentiate between the two fits. We have added the all the lines to the legend.

Figure 4. Please add (a) and (b) labelling to the subfigures.

Done

Figure 7: “he” to “the”. How do you choose the order of the polynomial fit of the structure functions here? For 12.5km and 25km the linear fit does not seem to be appropriate to fit the structure function, but rather a higher order (quadratic?) is required. This, however, would mean that the lead fraction gets multifractal for higher model resolution. Please elaborate on this difference when changing the model resolution in the text.

We've fixed the typo. We chose a linear fit for the structure function, same as in figure 3, and as you surmise. We chose this to be consistent with the analysis of the 6.5 km resolution run. Indeed the 25 km fit could be quadratic, but the uncertainty associated with the scaling at this resolution is so high that the significance of this is limited. This is now noted in the text.

References:

Wang, Q., Danilov, S., Jung, T., Kaleschke, L., and Wernecke, A.: Sea ice leads in the Arctic Ocean: Model assessment, interannual variability and trends, Geophys. Res. Lett., 43, 7019–7027, https://doi.org/10.1002/2016GL068696, 2016.
Hutter, N. and Losch, M.: Feature-based comparison of sea ice deformation in lead-permitting sea ice simulations, The Cryosphere, 14, 93–113, https://doi.org/10.5194/tc- 14-93-2020, 2020.