Dear Dr Fabien Montiel,

We would like to thank you for providing valuable comments and suggestions. Please find our responses to your comments in blue text and revised manuscript in red text below.

Best regards,

Yanan Wang and co-authors

Response to Referee #2

First, I would like to apologise to the authors and editor for the delay in submitting my review. The manuscript attempts to compare floe perimeter density data obtained via satellite imagery at 2 locations in the Arctic Ocean against the predicted density of 3 floe size distribution (FSD) models, which all include a parametrisation of wave-induced fracture. The goal is to evaluate the performance of the models. Results show that the discrepancy is quite significant in a number of ways. In particular, the models generally predict much larger perimeter densities than the observations, meaning an overestimation of small floes. The authors then attempt to explain the discrepancy by discussing potential issues with specific parameterised components of the models considered.

The manuscript presents original work, is overall well written and the topic is highly relevant to improving the modelling capabilities of FSD resolving sea ice models. That said I have a number of important issues with the way the study has been designed, which may explain the poor agreement between models and data. These are detailed below as well as other comments. I therefore recommend the manuscript undergoes major revisions before it can be further considered for publication.

Main comments

1. My main concern relates to the way the study regions for both models and observations were selected in relation to one another. If I understand correctly, the satellite images were chosen at 2 specific locations, one in the Chukchi Sea and one in the Fram Strait. In contrast, the regions selected for analysing model outputs are much larger. They do include the specific observational locations but extend over much wider regions, selected to include the ice edge. My issue is that the variability in FSD in these regions is likely to be much larger than that at the locations of the data. If the data is collected far from the ice edge (which could have been estimated), floe sizes are likely much larger than closer to ice edge. Therefore, I am not sure the comparison between model outputs and data is fair with respect to the models. In fact, when the comparison is refined to a subdivision of the initial study regions, the agreement is improved. I am not an expert in analysing satellite imagery, so my following suggestion could be naïve and/or uninformed, but this is what I would have done: (i) select all the imagery available in the model study regions, not just those at the specific locations selected, OR, if this is too much data to analyse, (ii) take a random sample of images spanning the study regions. Either

way, the variability in the data would be a lot more representative of that predicted by the models. I am not suggesting that the authors redo the entire analysis for this paper, but I feel like this limitations needs to be given a lot more emphasis in the manuscript. At the very least, bring this up in the Discussion section, but what would be even better is to add a sub-section looking at the comparison models/data when the model outputs are only selected in a smaller region around the data locations, even more localised than the subdivision shown in figure 6.

Thanks for your comments and suggestions. For models, we have tried to select a small region (5×5 grid cells) for comparison. However, for the comparison between observations and several models that have different atmospheric forcing, there are very different ice conditions within this small region, e.g., one model shows MIZ but the other two show interior ice pack. This situation will cause the unfair in the model-observation comparison. By choosing large regions including the ice edge, we can make sure the perimeter density simulated by all models indicate the average perimeter density in the region far from the ice edge and close to the ice region.

For observations, all MEDEA images used in this study in the Chukchi Sea and the Fram Strait are in the same fixed location, 70°N, 170°W for the Chukchi Sea and 84.9°N, 0.1°E for the Fram Strait. However, although the images applied in this study are in the fixed site, different imagery data can also indicate the perimeter density at different distances to the ice edge. For example, the sea ice concentration for 7 out of 13 cases in June in the Chukchi Sea is below 80%, indicating MIZ in the images. And the left 6 June cases indicate interior ice pack. Hence, when we compare the monthly mean (as the solid lines in Figure 2e-2k) rather than the case-by-case comparison, it is the comparison of the mean state of perimeter density both near the ice edge and far from the ice edge for both models and observations.

Thanks for pointing out this limitation. There are some extra MEDEA images for sea ice buoys available within the study regions, which is not in the fixed locations of "fiducial sites". However, for this study, we believe the variation of observations is enough for the model-observation comparison study. Thanks for your helpful suggestions. In future study, these images could be used to expand the spatial variability in the observations. We will add more explanations about these limitations in the Discussion and Conclusion section.

2. The title of the manuscript is misleading and should be changed, as the authors do not analyse the FSD but the floe perimeter density, which is different. The abstract also needs some work as it starts with statements on the FSD and then switches to perimeter density without making a link between them. I am not saying the perimeter density is a bad metric, but it is not the one advertised! It took some time to fully appreciate the meaning of Pi (again not an expert!). It would have helped me to show the Pi results with units km per km^2. I would have liked the relationship between these 2 quantities discussed in more details in relation to the results. For instance, do we expect FSDs to have similar qualitative and quantitative properties as the perimeter density distributions shown in figs 2e-k?

Thanks for your comments. In previous observational studies (Perovich, 2002; Perovich and Jones, 2014 and Arntsen et al., 2015), perimeter per unit area has been applied as a proxy for floe size distribution. As discussed by Perovich (2002), the use of the perimeter of the sea ice floe reduces the impacts of partially captured floes at the edge of the image for FSD retrieval. Besides, perimeter density per unit ice area is widely used as a metric to show the evolution of the FSD, e.g., Roach et al. (2019) and Bateson et al. (2022). As explained by Roach et al. (2019), "P_{ice} is weighted more heavily by smaller sizes, so Pi is more relevant for thermodynamic melting and freezing of floes", which is important in evaluating the FSD models. For more details on the discussion of the definition of floe number density N(r), floe area fraction F(r) and perimeter density P, please see our response to referee #1.

Thanks for your suggestions. We will revise the title to " Summer sea ice floe perimeter density in the Arctic: High-resolution optical satellite imagery and model evaluation" and include the relationship between FSD and perimeter density in the abstract. We also changed the units of Pi to km km^{-2} to make it easier to understand the meaning of Pi.

Other comments

3. L28-30: This statement is confusing, especially for the uninformed reader. Quantifying the MIZ is still an active an area of research. Mentioning the 2 definitions (i.e. wave-based and SIC-based) in this way makes it seem that they are equivalent, but that’s not true (see Brouwer et al paper). I suggest you pick one definition or expand the discussion.

   Thanks for your suggestions. We expanded the discussion as below.

   This results in the changing marginal ice zone (MIZ), defined as the ice-covered region affected by waves and swell by the World Meteorological Organization (WMO, 2014). However, due to the difficulty in detecting wave-ice interaction data from satellite imagery (Horvat et al., 2020), another optional definition of MIZ widely applied in previous studies is a sea ice-covered area with sea ice concentration (SIC) of 15%–80% (e.g., Strong and Rigor, 2013; Aksenov et al., 2017; Rolph et al., 2020; Bateson et al., 2020; Horvat, 2021).

4. L43: “viscous dissipation” is not the process governing attenuation, it refers effective/homogenised dissipative rheology of continuum viscous layer models often used to approximate attenuation caused by a non-homogeneous ice cover. In any cases, “dissipative processes” would be a better choice of wording here.

   Thanks for your suggestions. We changed the wording in the revised manuscript.

   The FSD affects the ocean surface waves propagation and attenuation through the ice, i.e., floes smaller than a characteristic wavelength of swells attenuate wave energy through dissipative processes

5. L53-55: The limitations of the power-law should be discussed. See Montiel & Mokus (2022, Philosophical Transactions A) for an overview.

   Thanks for your suggestions. We added more discussion about the limitation of the power-law hypothesis.

   Some other modelling studies assumed a particular shape of FSD (e.g., Bennetts et al., 2017; Rynders, 2017; Bateson et al., 2020). Previous observational studies reported a power-law behaviour existing in the tail of FSD, i.e., a straight line in logarithmic axes, leading to the parametrisations of fixing the FSD as following a truncated power law (Burroughs and Tebbens, 2001). However, the floe size power-law hypothesis has been contested by recent observations (Steer et al., 2008; Herman, 2010; Herman et al., 2021), which suggest different shapes or functions for describing FSDs. Results from laboratory experiments (Herman et al., 2017; Passerotti, 2022) and models (Herman, 2017; Montiel and Squire, 2017; Mokus and Montiel, 2021; Montiel and Mokus, 2021) also express the concerns that a power-law FSD is not appropriate for wave-induced sea ice breakup.

6. L77: I'm not sure I follow the argument that a low-resolution model output justifies the choice of a large model study area.

   Thanks for your comments. For the WIPoFSD model and CPOM-FSD model, the grid cell area ranges from about 2468 km2 to 3480 km2 in the study region. For FSDv2-WAVE, the grid cell area is between 541 km2 and 3802 km2 in the study region. Within every grid cell, the atmosphere forcing, wave forcing, etc. are homogeneous. The accuracy of FSDs in every grid cell relies heavily on the accuracy of external forcing. Applying a large region for models could better reveal the inhomogeneous effects of external forcing on the evolution of FSD and decrease the effects of external forcing bias in a few grid cells on FSDs.

   We will add more details of this argument to the revised manuscript.

7. L90: What's the area of the uncropped images? 250 km^2? It would be interesting to know.

   The size of the raw MEDEA image is approximately more than 15 by 15 km, i.e., covering an area larger than 225 km^2 (Kwok, 2014). For the World-View image applied in this study, the size of raw images is more than 200 km^2, but we only purchase part of it (~40 km^2). The image size information will be added in the revised supplement.

8. L91-93: A reference for the WV images and more details about the sensor should be given.

   Thanks. A reference for WV images will be added to the revised manuscript. A more information is added as shown below.

We also collected one WorldView-1 (WV1) and four WorldView-2 (WV2) images at the Chukchi Sea and Fram Strait sites (Fig. 1), which are collected in a 50 cm panchromatic band (WV1) at 0.52 m spatial resolution and 50 cm panchromatic band and 2 m 4-band Multispectral Bundle (WV2) at 0.47 m and 0.58 m spatial resolution. The size of the WorldView (WV) images is ~40 km^2.

9. L98-102: how is "floe size" defined? There are multiple definitions out there.

   Thanks for pointing out this. In this study, we define floe size as floe effective radius, the radius of a circle that has the same area as a floe. We have added this information to the revised manuscript.

10. Section 3.2 is confusing. The intro paragraph discusses differences between the different models, before describing the models themselves in subsequent sub sub-sections. It would make more sense to describe the models first and then discuss their differences/similarities.

    Thanks for your comments. The introduction paragraph gives a broad description and review of the recent FSD models including but not limited to models evaluated in this study. Section 3.2 introduces three specific models within them. In our opinion, it may be better to keep them, i.e., describing the differences between recent FSD models first and then introducing the details of the three specific models.

11. Section 3.2.1: More details about the WW3 configuration are needed. Is it a global or regional run? What ice attenuation parametrisation did you choose? Also, what do you mean "attenuation in the open ocean"?

    Thanks for your comments. It is a global run. As in Roach et al. (2019), attenuation is the sum of two parts. "The first is an empirical fit to a multiple-floe scattering theory by Meylan et al. (2021) that accounts for floe size, thickness, and sea ice concentration. This theory finds that the scattering resembles the linear Boltzmann equation derived by Masson and Leblond (1989) with an error correction that accounts for multiple scattering. Attenuation from this theory becomes negligible when wavelengths are much greater than floe sizes; therefore, Roach et al. (2019) include an additional term based on observations to attenuate waves like the inverse of the wave period squared".

    More details will be given as a brief introduction to WW3 in the revised manuscript.

12. L149: "process" is the wrong word. Maybe "theory" or "model"?

    Thanks. The wording is changed as follows.

    The WIPoFSD model implements the wave-in-ice model (WIM), originally based on the ice-wave interaction theory described by Williams et al. (2013a, 2013b)

13. L153-154: I feel more details are needed about how the fixed power was determined. Was it the same data as those used later for comparisons with model outputs. Were all the floe size data from all the images collated into a single dataset and then a power law fit was performed or were power law fit done for each image and then averaged? What was the range of floe sizes considered for the fit(s)? Also is 3 significant figures appropriate? What's the uncertainty on alpha? Was a goodness-of-fit test evaluation conducted?

    Thanks for your comments. It is not completely the same dataset. The exponent of -2.56 was determined for the dataset considered in the Bateson et al. (2022) study rather than the more complete dataset considered in our study. To describe the differences between the dataset used by Bateson et al. (2022) and our study, hereafter, we refer to the observations used for calculating the power-law fit used in WIPoFSD (Bateson et al., 2022) as Obs1 and the observations used for mode-observation comparison in this study as Obs2. The difference between Obs1 and Obs2 are:

    1) The different resolution: The resolution of the images used for deriving Obs1 was reduced from 2 m, while for Obs2, we keep its 1-m resolution.

    2) The different number of sites: Obs1 is calculated from the data derived from three sites: the Chukchi Sea, East Siberian Sea and Fram Strait. Obs2 only includes 2 sites: the Chukchi Sea and Fram Strait.

    3) The different number of cases in every site: For Obs1, there are 14 cases in the Chukchi Sea, 9 cases in the East Siberian Sea and 12 cases in the Fram Strait. For Obs2, there are 24 cases in the Chukchi Sea and 32 cases in the Fram Strait.

    We will clearly explain the WIPoFSD model output and any choices about model parameters from the Bateson et al. (2022) study in the revised manuscript.

    We applied power law fit for the floe size in each image first and then averaged. The range of floe sizes considered for the fits is 0 m-3000 m for Obs1.

    Yes, 3 significant figures are appropriate. If we keep 5 significant figures and set $\alpha = 2.5574$, $[P(\alpha = 2.5574) - P(\alpha = 2.56)] / P(\alpha = 2.56) \approx -1\%$, which only shows a 1% smaller perimeter density.

    Half cases pass the goodness-of-fit test.

    More details are added in the main text. A table of the power-law fit, uncertainty on alpha and p-value for the goodness-of-fit test of every case will be given in the revised supplement.

14. Eq (2): I don't n(r) has been defined.

    Thanks. The definition is added as follows.

    $n(r)$ is defined such that $n(r)dr$ is the number of floes per unit area for floe sizes between r and $r+dr$.

15. L175-177: Rephrase these 2 sentences as they appear to contradict each other.

    Thanks. We will move this sentence, "More details on the calculation of P_i is provided in the supporting information Sect. S1", to the end of Section 3.4.

16. Eq (4): gamma needs to be defined.

    Thanks for your comments. We added a clear definition of gamma. When adding this definition, we realize that there was a mistake in the value of gamma for the calculation of Pi. As a result, we have revised all equations in both the supplement and main text.

    As r applied here is effective floe radius $r_{eff} = \sqrt{a/\pi}$, the radius of a circle that has the same area a as a floe. We assume $p = 2\gamma r_{eff}$. Here the gamma is defined as $\gamma = \frac{p}{2r_{eff}} = \frac{p}{2\sqrt{a/\pi}} = \frac{p\sqrt{\pi}}{2\sqrt{a}}$, where p and a are the perimeter and area of a floe respectively.

    The gamma value we give in L23-L24 (the mean floe shape parameter γ is 2.27 in the Chukchi Sea region and 2.23 in the Fram Strait region) is $\frac{4a}{d^2}$. Now we have corrected both the equations and the Pi value shown in the results sections as follows.

    Here γ is a floe shape parameter, $\gamma = \frac{p}{2r_{eff}}$. From the analysis of MEDEA-derived FSD results, the mean floe shape parameter γ is 2.17 in the Chukchi Sea region and 1.96 in the Fram Strait region, respectively.

    The relationship between the areal FSD f(r) and CFND N(r) is revised to $\int_{r_0}^{\infty} f(r)dr = \int_{r_0}^{\infty} \pi r^2 dN(r)$.

    P_{ice} for prognostic model P_{i_prog} is revised to $P_{prog} = 2\sum_{i=1}^{12} \frac{\gamma f_i(r_{i_{max}} - r_{i_{min}})}{\pi r_i c_{ice}}$.

    P_{ice} for WIPoFSD is revised to $P_{wipofsd} = \frac{\int_{r_{min}}^{r_{var}} 2\gamma r n(r)dr}{c_{ice}} = \frac{2\gamma(3-\alpha)({r_{var}}^{2-\alpha} - {r_{min}}^{2-\alpha})}{\pi(2-\alpha)({r_{var}}^{3-\alpha} - {r_{min}}^{3-\alpha})}$.

    P_{ice} for observations is revised to $P_{obs} = \sum_{i=1}^{12} \frac{2\gamma A_{floe_i}}{\pi r_i A_{ice}}$.

17. L197: How do you estimate central tendency and spread in the figures provided for Pi?

    By applying equation $P_{obs} = \sum_{i=1}^{12} \frac{2\gamma A_{floe_i}}{\pi r_i A_{ice}}$, we calculate the P_{obs} of every images. Then calculate the average of Pi and the standard deviation.

18. Fig 2e-k: there appears to be a plateauing in the observations for small floes. It could be due to a resolution issue as discussed in section 4.2, but it could also be a

signature of the limitation of the power law. I know you are not trying to fit a power law here, but I believe this is an important point to make, especially when comparing to

Thanks for your comments. This is a very interesting point. We will discuss the possibility of the limitation of the power law in the Discussion and Conclusion section.

19. L212: use big O notation for order of magnitude.

    Thanks. Revised.

    The CPOM-FSD results also show a similar pattern, yet the model P_i values are in much better agreement with the observations for large floes ($r > O(10)$ m), especially during July and August in the Fram Strait region (Figs. 2g and 2h).

20. L215-216: I believe an explanation was provided for the presence of the “uptick”. It would be good to mention it here.

    Thanks. The explanation is added (see below).

    The “uptick” is an artificial feature derived from the model setup of an upper limit of floe sizes and a lack of fragmentation processes for large floes (Roach et al., 2018; Bateson et al., 2022).

21. L254-255: That statement seems to be quite a stretch at this stage and sort of comes out of nowhere. Given the seasons considered, I would expect lateral melting to be a more important contributor to underestimated SIC.

    Thanks. In the prognostic models, the floe welding rate is set to be proportional to the square of SIC in the two prognostic models. What we would like to explain here is that the underestimated SIC is a contributor to the underpredicted floe welding rate. Will modify the sentences to make them clear.

    The underestimated SIC from the two prognostic models may be a contributor to the underpredicted floe welding rate during spring and early summer.

22. L270: “floes welding are negligible during this season” seems to contradict the statement in the previous subsection (see comment 21).

    Yes, the contribution of floe welding to the change of Pi in summer is negligible, i.e., floe welding does not happen in summer. But Bateson (2021) has assessed the effects of floes welding on the FSD in the CPOM-FSD model, suggesting that floe welding occurring in spring can influence the perimeter density in summer (Section 7.2.3 and Figure 7.9 in Bateson (2021)).

23. L334-339: again, I don’t find the argument about underpredicting welding very convincing. What about potential overestimation in the data? See Roach et al (2018, The Cryosphere)

    Thanks. Roach et al (2018) discuss the reason for the overestimation of sea ice concentration (SIC) as lateral melt. Similar to the response to comment 22, in this study, we discussed the underestimation of SIC will lead to the underestimated floe welding and thus the overestimation of perimeter density of small floes in summer. Our explanation may cause a misunderstanding.

    Thanks for your comments. We will modify the sentences to make these causal relationships clear.

24. L243: Again, how was the fit performed? Across all floe sizes or a limited range? What about goodness-of-fit?

    Thanks. I guess that you may mean L343.

    This exponent is calculated across all floe sizes. By setting the floe size range of 0-3000m, 50% data pass the goodness-of-fit test. We will add this information in the main text and a table including the power-law fit, uncertainty on alpha and p-value for the goodness-of-fit test of every case in the revised supplement.

25. L245-247: What about the possibility that the power law is not appropriate? I don’t see why a power fitted through historical data should predict a power law in the current dataset, even at the same location.

    Thanks. I guess that you may mean L345-347.

    As you mentioned in comment 18, almost flat FSD in the observations for small floes could result from the limitation of image resolution (section 4.2) or a signature of the limitation of the power law. Besides, as in Hwang and Wang (2022), this behaviour may be consistent with the fact that small floes are more vulnerable to lateral melt relative to large floes, which causes the deviation of the small-floe distribution prior to the tail. A discussion of all these possibilities is added in this section in the revised manuscript.

    Through historical data, we would like to indicate that the spatial and seasonal variation of the power-law exponent does not only exist in our datasets but also in previous studies. This could emphasize the view that a constant exponent is not appropriate.

Typographical errors

26. L58: delete “the”.

    Thanks. Corrected.

27. “welding” is misspelled a couple of times.

Thanks. Corrected.

Citation in the response to Referee #2:

Arntsen, A. E., Song, A. J., Perovich, D. K., and Richter-Menge, J. A.: Observations of the summer breakup of an Arctic sea ice cover, Geophys. Res. Lett., 42, 8057–8063, https://doi.org/10.1002/2015GL065224, 2015.

Bateson, A. W.: Fragmentation and melting of the seasonal sea ice cover, Ph.D. thesis, Department of Meteorology, University of Reading, United Kingdom, 293 pp., 2021.

Bateson, A. W., Feltham, D. L., Schröder, D., Wang, Y., Hwang, B., Ridley, J. K., and Aksenov, Y.: Sea ice floe size: its impact on pan-Arctic and local ice mass and required model complexity, 16, 2565–2593, https://doi.org/10.5194/tc-16-2565-2022, 2022.

Horvat, C., Blanchard-Wrigglesworth, E., and Petty, A.: Observing Waves in Sea Ice With ICESat-2, Geophys. Res. Lett., 47, https://doi.org/10.1029/2020GL087629, 2020.

Hwang, B. and Wang, Y.: Multi-scale satellite observations of Arctic sea ice: new insight into the life cycle of the floe size distribution, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., 380, https://doi.org/10.1098/rsta.2021.0259, 2022.

Kwok, R.: Declassified high-resolution visible imagery for Arctic sea ice investigations: An overview, Remote Sens. Environ., 142, 44–56, https://doi.org/10.1016/j.rse.2013.11.015, 2014.

Masson, D., & Leblond, P. H.: Spectral evolution of wind-generated surface gravity waves in a dispersed ice field. JournalofFluid Mechanics, 202, 43–81, https://doi.org/10.1017/S0022112089001096, 1989.

Meylan, M. H., Horvat, C., Bitz, C. M., and Bennetts, L. G.: A floe size dependent scattering model in two- and three-dimensions for wave attenuation by ice floes, Ocean Model., 161, 101779, https://doi.org/10.1016/j.ocemod.2021.101779, 2021.

Perovich, D. K.: Aerial observations of the evolution of ice surface conditions during summer, J. Geophys. Res., 107, 8048, https://doi.org/10.1029/2000JC000449, 2002.

Perovich, D. K. and Jones, K. F.: The seasonal evolution of sea ice floe size distribution, J. Geophys. Res. Ocean., 119, 8767–8777, https://doi.org/10.1002/2014JC010136, 2014.

Roach, L. A., Horvat, C., Dean, S. M., and Bitz, C. M.: An Emergent Sea Ice Floe Size Distribution in a Global Coupled Ocean-Sea Ice Model, J. Geophys. Res. Ocean., 123, 4322–4337, https://doi.org/10.1029/2017JC013692, 2018.

Roach, L. A., Bitz, C. M., Horvat, C., and Dean, S. M.: Advances in Modeling Interactions Between Sea Ice and Ocean Surface Waves, J. Adv. Model. Earth Syst., 11, 4167–4181, https://doi.org/10.1029/2019MS001836, 2019.