The Cryosphere Discussions

The Cryosphere Discuss., referee comment RC2
https://doi.org/10.5194/tc-2021-368-RC2, 2022

Comment on tc-2021-368

Christopher Horvat (Referee)

Referee comment on "Characterizing the sea-ice floe size distribution in the Canada Basin from high-resolution optical satellite imagery" by Alexis Anne Denton and Mary-Louise Timmermans, The Cryosphere Discuss., https://doi.org/10.5194/tc-2021-368-RC2, 2022

This is my review of Denton and Timmermans (2021).

Here the authors use image processing to obtain floe size distribution (FSD) measurements spanning a scale range from 50 m^2 -5 km^2, and over a wide time period in the Canada Basin. They find both that FSDs obey a power law relationship and that there is a connection between sea ice concentration and power law coefficient. Generally I find this paper to be an interesting contribution to the literature. It is well written and has a comprehensive review of existing work and does a good job of contextualizing the work here. I am also particularly excited that someone has characterized the GFL dataset! I have a few minor comments, and one important one about statistics that needs to be taken into account. The latter is a straightforward application of a method to data already presented in the MS, so it really is likely to be a minor revision, though it may have an impact on the results. So for the authors I’m recommending “major” revisions only to give more time to include that bit of analysis and discussion. Up to these changes, I’m happy to recommend this paper for publication.

Best,

Chris Horvat

——

Main comment.

L30, and generally - It is worth describing the power-law behavior of the FSD as a "hypothesis". As was pointed out by Herman (2010), there actually isn't strong evidence of a power law in most data, but the use of this to describe the distribution is because it is handy for multi-scale distributions. As was pointed out in Stern's two 2018 papers, the appearance of power-law behavior can appear spuriously as a result of choosing a particular way of plotting the distribution. This hypothesis can be formally tested, and the mechanism for doing so is described in papers by Clauset (2007,2009,2014), who maintains code to perform fitting, goodness-of-fit tests, and p-value computations at https://aaronclauset.github.io/powerlaws/. When examining multi-scale distributions, one should report the p-value for fits, and the start of a power-law tail. There are many reasons for this!

First, often the intuitive "eye test" for power-law behavior is flawed, and alternative distributions fit better (as examined by Herman (2010)). In studies where this was adamantly required by certain unnamed co-authors, (e.g. Hwang, 2017), few of the obtained distributions actually had power law tails, or only could not reject a power law hypothesis over a small range of size scales. It is worth knowing this!

Second, power-law behavior is fundamentally about the tail of a distribution. Thus issues at the "artifact scale" are not important when performing such tests. This allows you to expand the range of floes into the less certain smaller sizes without actually losing information - if the fit doesn't extend there, it won't be included.

Third, there is a mechanism for looking at binned (Virkar and Clauset, 2014) and un-binned data (Clauset et al, 2007,2009). The binning employed here seems to be done to facilitate fitting, but this is not necessary when using the Clauset-style tests. You have the full information on floe areas, so you can directly use that to obtain fits, goodness of fits, and p-values.

I would suggest that these tools be applied to the distributions obtained here, and p-values and tail beginnings (x_{min} in the power law toolbox) reported for power-law fits before making statements about the distributional fitting. Without them, it is hard to be convinced rigorously that this hypothesis is not "not true". Including these makes the reporting of PL behavior robust.

Finally, it would be good for the community if your segmentation algorithm was posted publicly and DOId, through e.g. Github/Zenodo. Almost all existing algorithms used in FSD studies are similar in structure, but not able to be used by others. Thus each time a new group wants to compute FSDs they have to reinvent the wheel. Having an available FSD code would benefit many, and ensure your work was properly credited.

Specific Comments:

L26 - it is worth looking at the papers of Roach (2018,2019), Zhang et al (2017) who also discussed and theorized the impact of the FSD on both local and model-scale melt partitioning.

L95 - can you clarify how your method differs from the methodology used by other authors? Is it a similar concept?

L198 - I appreciate that you are using floe area to define the FSD - not sure why more authors don’t do this. It makes for more accurate comparison to models (which use a fixed radius-area relationship) and is much easier to understand than the MCD. I would even add a statement here or in the discussion to make this point as it is quite helpful from the modeling side.

L180 - Have a look at the main comment here. I do not think it is a good idea to compute power law slopes from binned data, as you are sensitive to your bin choices. Still, there is a methodology for computing such slopes (Virkar and Clauset, 2014), which I would argue to employ here if you want to compute them using the binned data. However you have the raw area data, and so don’t need to resort to binning. In that case you can use the methodology of Clauset et al (2009) to obtain slopes and p-values for fitting. You will find that the binned PL slopes and the unbanned PL slopes *will* differ.

L223 - Again, I would strongly caution to use the Clauset method to obtain slopes. This is discussed in the series of Clauset papers, and later by Stern et al (2018), that fitting straight lines to log-log plots can often fail you in unexpected ways.

L279 - please clarify how you are testing for significance. You are highly sampled in the high SIC range, and weakly so in the low SIC range. What went into the choice of such a significance test, and why? For example, you bin floe areas and then fit them - what guides that choice, but not a similar choice for SIC? Hopefully that points a bit towards why it might be preferable to avoid binning.

L360 - An extensive discussion of how multiple FSD size regimes can emerge was performed in Horvat and Tziperman (2017) which I will shamelessly plug. It is worth noting the difference in locations may impact the processes that give rise to PL behavior (i.e. the impact of waves leads to more small floes, etc).

References.

Clauset, Young, and Gleditch. On the Frequency of Severe Terrorist Events. 2007

Clauset et al. Power-Law Distributions in Empirical Data. 2009.

Virkar and Clauset. Power-law distributions in binned empirical data. 2014.

Herman. Sea-ice floe-size distribution in the context of spontaneous scaling emergence in stochastic systems. 2010.

Horvat and Tziperman. The evolution of scaling laws in the sea ice floe size distribution. 2017.

Roach et al. An emergent sea ice floe size distribution in a global coupled ocean--sea ice model. 2018.

Roach et al. Advances in Modeling Interactions Between Sea Ice and Ocean Surface Waves. 2019.

Zhang et al. Sea ice floe size distribution in the marginal ice zone: Theory and numerical experiments. 2016.