<u>Response to reviewer 2: 'Automated ArcticDEM iceberg detection tool: insights into area and volume distributions, and their potential application to satellite imagery and modelling of glacier-iceberg-ocean systems' by Shiggins *et al*.</u>

We would like to thank Till Wagner for their comments which will help to improve the manuscript. Our responses to each of the major and minor comments raised and how we intend to address them for the revised version of the manuscript are outlined below. For this, reviewer comments are copied verbatim in blue, and our response to each is given in black. All line numbers quoted with the prefix L (e.g. L123) refer to those in the original submitted manuscript. All line numbers quoted with the prefix RL (e.g. RL123) refer to those in this response document.

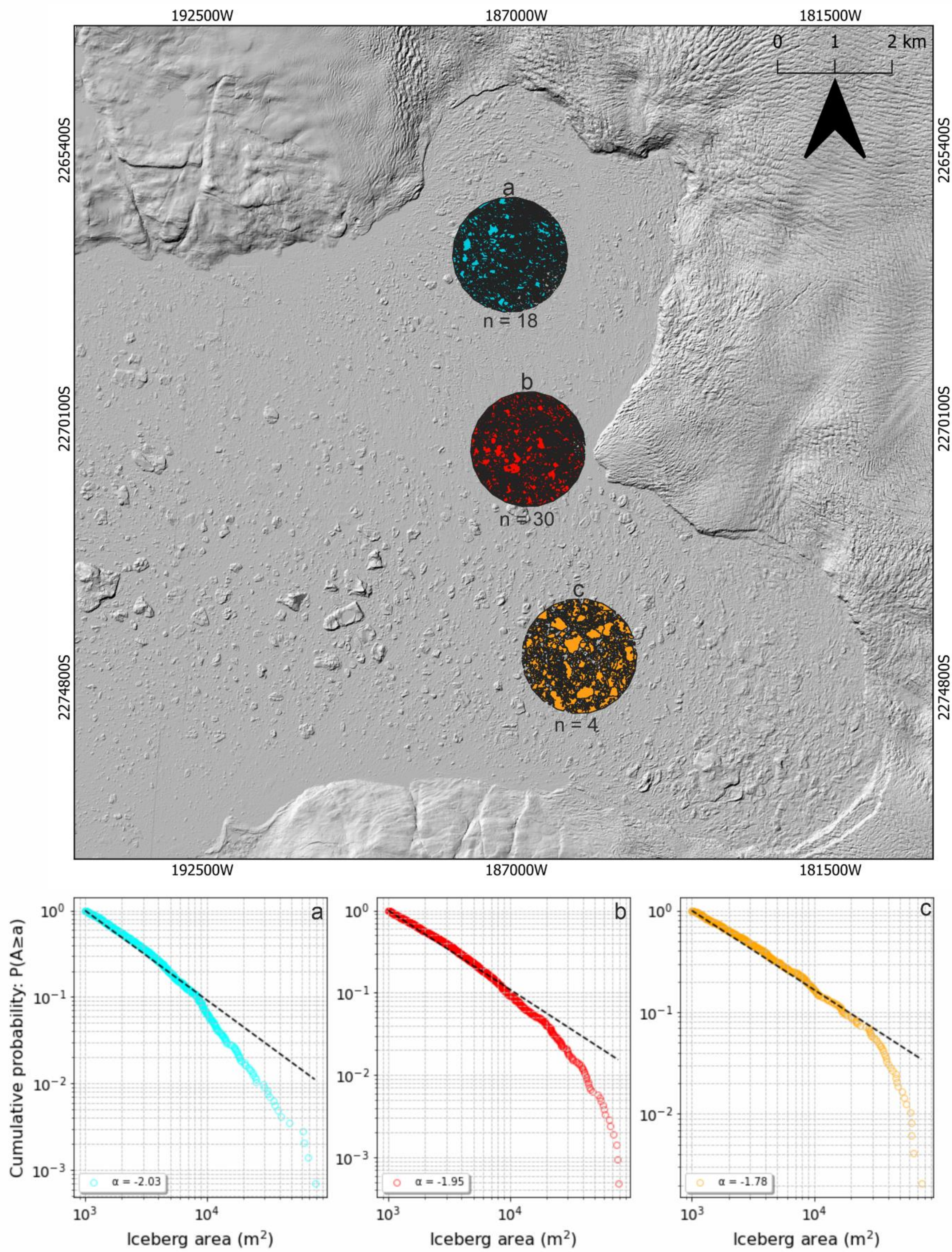
<u>Main comments:</u>

1. Availability of ArcticDEMs and picking the right ROI. I was able to run the code on SKJI without much trouble and could also approximately reproduce some of the distributions in the paper (e.g. something similar to those in Fig 6). However, when I tried to explore other random glaciers around Greenland I struggled to find ones with any available ArcticDEM scenes. I randomly tried ~10 or so glaciers in different regions and only 2 identified any ArcticDEM scenes (2 scenes each) for the ROIs that I picked. It was not clear to me from the manuscript how exactly to pick the ROIs and I tried to emulate the shapes provided in Fig 1 but realized I had no further knowledge how these were determined. This may be part of the reason why I couldn't detect more DEM scenes. I was also struck by the fact that picking slightly different ROIs in front of SKJI resulted in detecting a different number of scenes and also in somewhat different slopes for the area distributions. I appreciate that a comprehensive account of where ArcticDEM scenes are available may be beyond the scope of this study, but the lack of information in this regard limits the utility of the product. Relatedly, it would be helpful to have some practical guidance of how to draw the ROI polygons to best harness the strengths of the algorithm. Finally, a discussion of how much the results depend on the number of scenes available would be helpful. This could for example be explored by running the analysis on subsets of the SKJI scenes and showing the resulting spread in power law slopes, or similar?

- We appreciate the reviewer undertaking comprehensive testing of the tool as it is extremely useful to gain feedback on its usability for those encountering it for the first time. With regards to the definition of the polygon for the ROI, there is a GitHub read.me (https://github.com/ConnorShiggins/Google-Earth-Engine-and-icebergs) available which is included in the text of the manuscript. This includes a walkthrough on how to define the ROI. However, after these comments, we will clarify in the text of the revised manuscript that this tutorial exists for users wanting to obtain a dataset (at L169).
- With respect to the point raised regarding how slightly different ROI definition impacts results, as suggested by the reviewer we have conducted analysis on 3 ROI subsets for SKJI in front of the northern, central and southern regions

of the glacier (see Draft Figure 3). For the northern branch, results from 18 DEMs available returned a power law slope of -2.03, the middle ROI with 30 DEMs returned a power law slope of -1.95, and the southern branch subset had 4 DEMs available returning a power law slope of -1.78. Consequently, there is some variation in both image availability and the α value for each section of the fjord. Understanding what is driving this localised variability is poorly understood and certainly deserves detailed study in and of itself. However, given the potential for changing calving styles through time and variation in space of calving dynamics in front of each terminus region, it is not possible to say here whether these differences arise from data availability or differences in fjord/glacier dynamics. Such analysis would require detailed understanding and analysis of individual glacier dynamics and their spatial and temporal variability, which we suggest is beyond the scope of this study.

- As commented upon by the reviewer and highlighted in the response above, choosing different ROIs can lead to variation in the number of DEMs available for analysis. This is especially noticeable at SJKI as the terminus is over 40 km long and ArcticDEM strips rarely cover the entire fjord region. We also note that this is likely to have most significant impact on glaciers with long margins (e.g. SKJI, Humboldt, 79N), and will have less of an impact on termini in narrower (e.g. ~2-10 km wide) fjords. To potentially increase data availability across an ice front, the filter threshold defining the lower limit of ROI coverage can be lowered to allow more DEMs to be taken forward for subsequent analysis (default is 80% or 0.8 in the workflow). This is defined in line 220 of the code with a variable name '*imageAreaCoverage*'. However, doing so may lead to less accurate definition of sea level for each image. To clarify this, we will add discussion of Draft Figure 3 in the main text at L503, and will provide full instructions and caveats as part of the GitHub readme. It should also be noted that the new GUI functionality included in response to Reviewer 1's main comment 2 (RL54) will allow users to get an indication of how much ArcticDEM data may be available for different glaciers across the Arctic region.



Draft figure 3. Subset sampling across the ice front at SKJI to determine distributional changes depending on the data available. The power laws are below and respective to their position in the fjord by letter and colour. The 'n' is the number of ArcticDEM scenes in the image collection of the detected icebergs.

2. Degree of automation. There are a couple of user inputs which are not straight-forward to set, namely the ROIs (see comment above) and the elevation threshold. The elevation threshold seems to be somewhat of a complex issue (see also the other reviewer's comments about distinguishing rafted vs non-rafted iceberg clusters). However, from looking at Fig 5 it looks like key statistics such as iceberg frequencies and the power law slope are not overly sensitive to this threshold, and I was wondering whether a 1.5m cutoff could simply be applied to all glaciers (including SKJI) at least in the paper, with a discussion that one may want to adjust this for certain purposes (such as focusing on the specific distribution of small icebergs); I am such mostly suggesting a minor reframing of the language here. As an alternative (and more involved) approach, one could come up with an optimization scheme that picks the threshold for each glacier depending on specific output statistics? Relatedly, it would help clean up the presentation if a single x_min could be picked for the glaciers in the paper (with an accompanying discussion analogous to the one for the elevation threshold)? As a minor point I would suggest removing the word "fully" from l.12.

- Using different thresholds above sea level for different glaciers illustrates the flexibility of the workflow and allows users to change it depending on their research question. In the manuscript we aimed to show examples of this by varying the detection threshold and expanding on the circumstances in which it is appropriate to do so (at L481-483). We will add to this discussion a small paragraph (at L496) which outlines how changing the detection threshold may alter the icebergs detected (e.g. a higher detection threshold will result in fewer small icebergs being delineated), and highlight that instructions on how to do this are in the GitHub read.me. An optimisation scheme for setting a detection threshold was something we did consider, however it would require multiple iterations of computationally intensive parts of the code across all available ArcticDEM strips in order to maintain consistency of data output. We therefore decided against implementing this option in order to retain code efficiency, data consistency, and the speed with which users can obtain outputs.
- While we agree that being able to define a single x_{min} for all the glaciers would be ideal, doing so would risk severely limiting data available for analysis. For example, setting an x_{min} at UI and KNS equal to that at SKJI would result in KNS and UI losing approximately 30% of iceberg observations. This would lead to potential over-estimation of how large the iceberg distributions are for these glaciers. The difference in calving styles and overall iceberg size distributions at each glacier also raise questions as to whether applying similar x_{min} values at each site is appropriate. Again, such a choice can be made by the user during post-processing depending on the research question under investigation (i.e. what range of iceberg sizes are users interested in). In light of this comment, we will clarify in the text that these are the specific reasons different x_{min} values are defined (at L506).
- We will also replace the word ‘fully’ on L12 with “highly”.

Specific comments:

1. (L.59): is solar illumination also a limiting factor for the DEMs?

- Solar illumination does not impact the ArcticDEM data itself, though will have impacted whether the WorldView data used to construct the DEMs could be used for DEM generation. Given that this manuscript does not generate the DEM data from WorldView imagery, but instead uses the ArcticDEM strips that are available (and which retain no solar illumination related metadata), we do not include solar illumination as a limiting factor for analysis of the DEMs.

2. (L.69): "iceberg area distribution" vs l.70 "area-size distributions" I presume this refers to the same thing, so maybe pick one?

- We will choose ‘iceberg area distribution’, and endeavour to ensure that we make use of consistent language in this and other cases throughout the revised manuscript.

3. (L.70): Just a side note: we also used such size distributions to look at iceberg decay in Antarctica in England et al "Modeling the breakup of tabular icebergs". Science Advances 6.51 (2020): eabd1273. This was based on the Antarctic size distributions in Tournadre et al "Antarctic icebergs distributions 1992-2014". J. Geophys. Res. Oceans 121, 327–349 (2016). You may not want to bring in Antarctic references here, so feel free disregard this comment.

- Thank you for highlighting these works and we appreciate the comment. Research in Antarctica is of course relevant and we will add these references to the revised manuscript (at L73).

4. (L.75): I suggest explicitly stating what "x" represents (surface area in m^2 (?)). I was also wondering whether "a" or "A" may be better since "x" often refers to a distance and since in the vert. axis label of Fig 6 you write "P(A>a)", so if you stick with "x" you may want to adjust this label.

- We will adjust this label as suggested in the revised manuscript.

5. (L.88): maybe add "([as discussed] in Scheick et al., 2019)", otherwise it reads as if Scheick et al were misrepresenting the data

- This will be changed in the revised manuscript.

6. (L.91): "determine" instead of "interrogate" (?)

- This will be changed in the revised manuscript.

7. (L.102): Similar to the comment on Scheick et al.: it is not quite clear whether Sulak et al were among the few studies to directly estimate iceberg volume (maybe just move the reference to right after "few studies"?)

- We will move the reference to the suggested position in the revised manuscript.

8. (L.140): maybe clarify over which time period this retreat happened?

- The time period of the retreats (2000-2002 and 2013-2015 respectively) will be added in the revised manuscript (at L142).

9. (L184-194) (see also general comment 2): this reads a little like picking the right threshold is more of an art than a science. I'd suggest reframing this a bit.

- Yes, it is correct that choosing the “correct” threshold is somewhat of a subjective choice on the part of the user. However, this can be informed by prior knowledge of iceberg density. For example, if glaciers are known to have particularly dense melange cover dominated by large icebergs (e.g. SKJI, Helheim, Kangerlussuaq), a higher threshold may be more appropriate. Where there is dense melange cover with smaller icebergs (e.g. KNS), or where there is typically open water, then lower thresholds will produce more comprehensive data (i.e. more likely to include small icebergs and/or iceberg rafts). To address this comment we will explicitly flag on L496 that discussion of this point is raised later in the paper, as mentioned in our response to main comment 2 (RL82).

10. (L.210): 5.3 "km^2" to 41 "km^2"

- Thank you for spotting this and we will update.

11. (Table 1): How are the uncertainties in the power slopes calculated? There also seems to be a rogue "-" after 8.629 (and the misplaced line number 225). Out of mere curiosity I was wondering whether there is much of a seasonal fluctuation in any of these statistics? I guess you only have summer DEMs?

- The uncertainties are generated using a Python power law package (Alsott et al., 2014), and the uncertainty is calculated as one standard deviation of the residuals of the relationship between iceberg area or volume versus frequency. This will be clarified in the text on L217.
- We will remove the rogue “-“, as well as the misplaced line number.
- A very interesting point regarding seasonal fluctuations, but as correctly noted, we only use DEMs between July and October to avoid rigid melange and seasonal ice tongues where the workflow has higher risk of returning erroneous data. This is already flagged to the reader on L164-165, though the implications of this will also be reiterated in the discussion at L477 in the revised manuscript.

12. (Fig 3): The automated and manually detected volume sums for KNS are almost identical, much closer than for the other two - yet their power law slopes (Fig 6c) are more divergent than for the other two glaciers. Could you comment on that? I also noted that SKJI has a rather large % difference in manually and automatically detected iceberg volume. Could you comment on why that is and why we need not be concerned about that (or should we)?

- The percentage difference at SKJI between the automated and manual methods arise as a result of the manual user not identifying smaller icebergs in the DEM (discussed L484-485). Also, given that the automated approach performs analysis on a per pixel basis, whereas a manual delineation is almost certain to cross pixels, the automated approach is more likely to provide a more accurate characterisation of iceberg areas and be unaffected by manual user digitisation error (either through user under-estimation of extent, or over-estimation through failure to separate out adjoining icebergs). User digitisation error will also have a proportionately greater impact on smaller icebergs and is most likely to account for the mismatches in power law slope values observed (e.g. Figures 3 and 6). It is challenging to disentangle whether these small differences arise from user digitisation error or workflow error given that the definition of an iceberg margin is somewhat subjective and will vary between users. To avoid potential for bias in manual digitisation, we will also note that these were performed by a single operator (at L484). To clarify each of the points above, we will add to the discussion at L487.

13. (Fig 5 and Fig 6.): The given value for alpha (KNS) in Fig 6c is -2.38, while the KNS alpha values range from -2.1 to -2.3, and close to -2.25 for threshold = 1.5m. Why is there this discrepancy?

- The data presented in figure 5 includes all the data for KNS (i.e. all icebergs from 16 images), while the data presented in Figure 6c for validation is based on data from a subset region of a single image. It was necessary to use a subset of an image for validation as comprehensive manual digitisation of entire scenes is impractical. The differences in alpha values for KNS between Figure 5 and Figure 6c therefore arises from the latter representing the iceberg distribution of KNS at a single point in time for only part of its fjord. This will be noted in the text (at L506) as a point alongside discussion of main point 1 (RL43-58) regarding how subset areas of ROIs can influence the values of power law slopes (Draft figure 3).

14. (Fig 6): I was initially confused that the slopes on the log-log plots of figure 6 have are approx 1, whereas alpha =~ 2. I then realized that you are plotting CDF and the slope for a CDF = alpha -1. Maybe this could be noted in the text or caption?

- We will note this in the caption in the revised manuscript.

15. (Fig 7): The 5th and 95th percentile are given as power law relationships, for which I would have expected straight (dashed) lines in the figure, but the lines are somewhat wiggly. Why is that?

- The 5^{th} and 95^{th} percentile lines are not straight because the data values are derived from the binned ranges of $\log_{10}$ (x+0.1) increments. Adding on lines of best fit for the percentiles made the plots too crowded, and obscured the data. The area-volume relationships given in the text for the 5^{th} and 95^{th} percentiles (Equations 5 and 6) are based on the lines of best fit that have been derived for these binned mean values. We will clarify this point in the text (at L341), and in the figure caption in the revised manuscript.

16. (Fig 8): The resolution of this fig is somewhat low (also the horizontal label of panel a is cut off?)

- We can remake this figure ensuring a better resolution and thank you for noticing the x-axis on the subplot being cut off.

17. (Fig 9): horizontal axis label: "iceberg area (m^2)" (not increments)

- This will be changed in the revised manuscript.

18. (L.487): delete "is achievable" (or "it is able")

- This will be changed in the revised manuscript.

19. (L.542): I would suggest replacing "excellent" with "good" (?)

- This will be changed in the revised manuscript.