Response to reviewer 1: '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*.

We would like to thank the reviewer 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.

Main comments:

1. Please can the authors comment on the two iceberg distributions found at two of their study sites? I thought the purpose of defining a threshold above sea level was to remove the chance of multiple bergs that are held together by melange being detected as single icebergs. But in your results (e.g. Figures 7 and 8) you present two distributions for SKJI and KNS. You suggest that Distribution 2 does in-fact represent bergs frozen together by melange. Should the threshold above sea-level therefore be increased, to remove this phenomenon? You would then only retrieve a single distribution per study site.

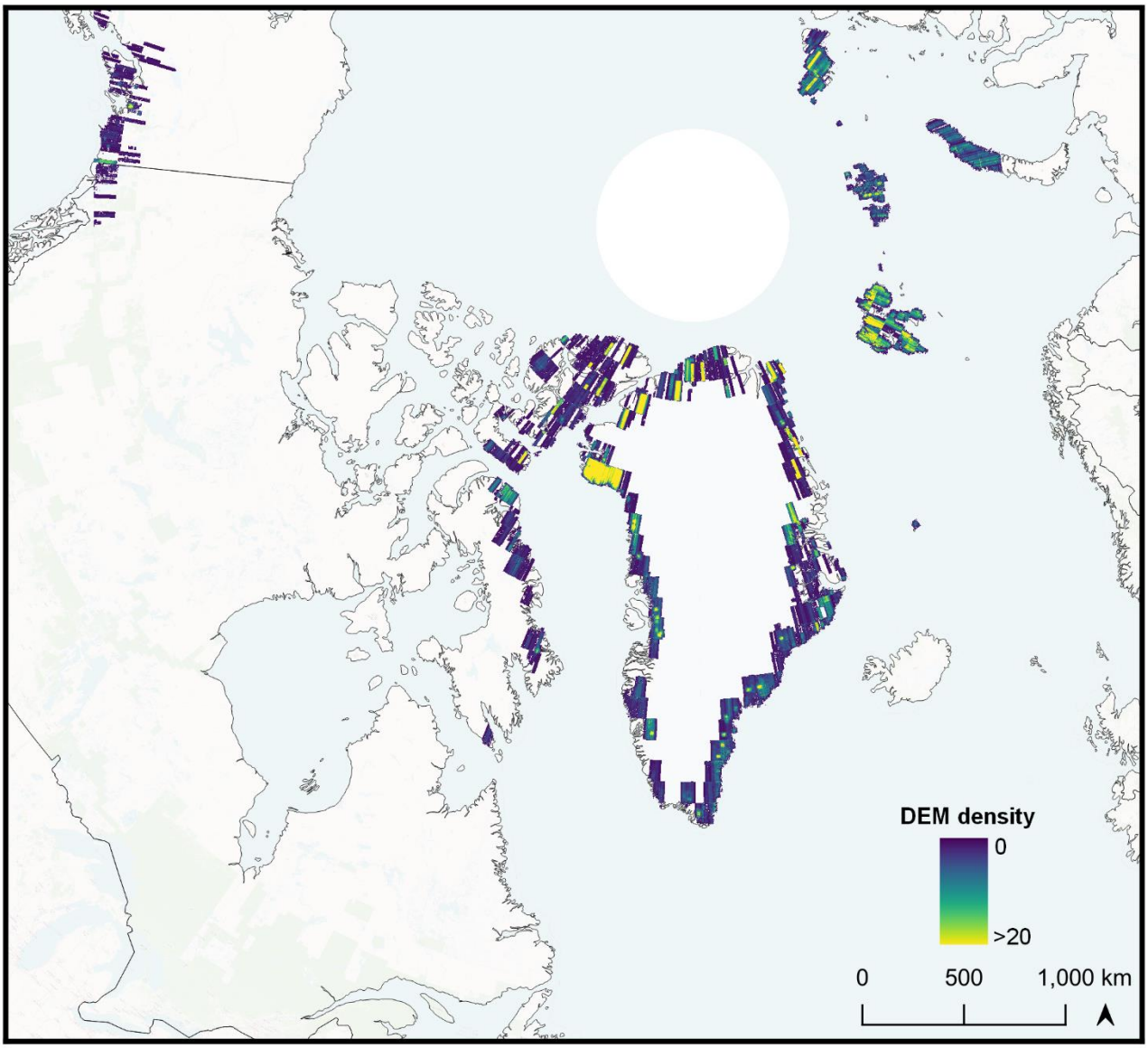
    - Whether a user wishes to obtain data including iceberg rafts or individual icebergs will be dependent on their definition of an iceberg within their research question. The approach presented in the manuscript allows users to choose whether iceberg raft data are retained or not through the definition of the threshold above sea level value for iceberg identification. For example, if only iceberg outlines are desired, a higher threshold above sea level could be defined by the user. By doing so, distribution 2 (iceberg rafts) would not be identified. However, a higher threshold would mean that smaller icebergs with lower freeboard heights may be missed. Conversely, if the user's research question requires all iceberg *and* iceberg raft cover from an ROI, results in the manuscript show that a lower threshold (e.g. 1.5m) will provide such data. A further alternative approach is that the iceberg raft distribution could be separated from the iceberg distribution as part of user post-processing (e.g. Figure 8 insets). The examples provided in the manuscript show the flexibility of the iceberg detection workflow depending on the type of iceberg data they wish to obtain (Figure 5).
    - To address this comment we will clarify that the research question being investigated is crucial for defining the iceberg detection threshold by inserting at L502: 'If a user's research question requires both iceberg and iceberg raft cover (distributions 1 and 2) within an ROI, the default threshold of 1.5 m above sea level is suitable, as is the 3.0 m threshold for more densely ice covered fjords such as SKJI. If only iceberg outlines are needed, a higher detection could be defined to remove iceberg rafts (distribution 2). It should be noted that setting a

higher detection threshold would result in the potential loss of data relating to smaller icebergs which have lower freeboard heights, and fractionally lower iceberg volumes obtained from larger icebergs. An alternative approach that would retain smaller icebergs and not result in the minor under-estimation of iceberg volume would be to use a lower threshold (e.g. 1.5 or 3 m), with data from distributions 1 and 2 separated as part of post-processing (e.g. Figure 8 insets).'.

2. The authors make a couple of references to this method having the potential to be upscaled across the full continent. However, they also suggest that there would need to be good enough data coverage for this. Please can the authors clarify whether there is enough data for pan-Arctic application or not?

- We have created a draft supplementary figure (see Draft Figure 1 below) which shows the coverage of ArcticDEM strip data on the pan-Arctic scale using locations identified as marine terminating glaciers for non-ice sheet and ice sheet glaciers (from the Randolph Glacier Inventory (RGI) v6 and Goliber et al. (2022) respectively). The map has been created by identifying the footprints of ArcticDEM strip data where there is overlap within 5 km of the point locations provided by Goliber et al. (2022) for Greenland, and having any overlap with RGI glacier outlines whose metadata show them as being either lake terminating, marine terminating or shelf terminating. This figure will be included as supplementary data in the revised manuscript. However, as RGI data use a benchmark of glacier outlines observed at near to 2000 as possible and some glaciers have now retreated into proglacial lakes (e.g. in Iceland) or changed their terminal environments, this map may not include ArcticDEM coverage of these glaciers. Consequently, we have also created summary maps showing all ArcticDEM data coverage irrespective of whether they cover glaciers or not.
- To allow users to get a quick impression of data availability for a given ROI we have now included new functionality within the GUI to view a series of summary maps showing ArcticDEM coverage. This includes:
    1. Map showing July-October coverage for known calving glaciers (Draft Figure 1; i.e. data least likely to be affected by solid melange/sea ice).
    2. Map showing all ArcticDEM coverage for known calving glaciers irrespective of acquisition time.
    3. Map showing all ArcticDEM coverage from the entire dataset irrespective of whether a glacier is thought to be there or not.
    4. As map 3, but for the months July-October.
- Further functionality to allow users to filter DEMs by month of acquisition has also been added to the GUI. The analysis workflow for this revised GUI has otherwise not been changed. The revised GUI for inclusion can be accessed at the following link: (https://code.earthengine.google.com/ad11c00c37b7ad88e28c4493ee6eec64).
- It is worth noting that these maps show where ArcticDEM data are available irrespective of the quality of the DEM data. Consequently, they do not indicate that all of the DEMs will be of sufficient quality/coverage to allow it to be used for analysis.

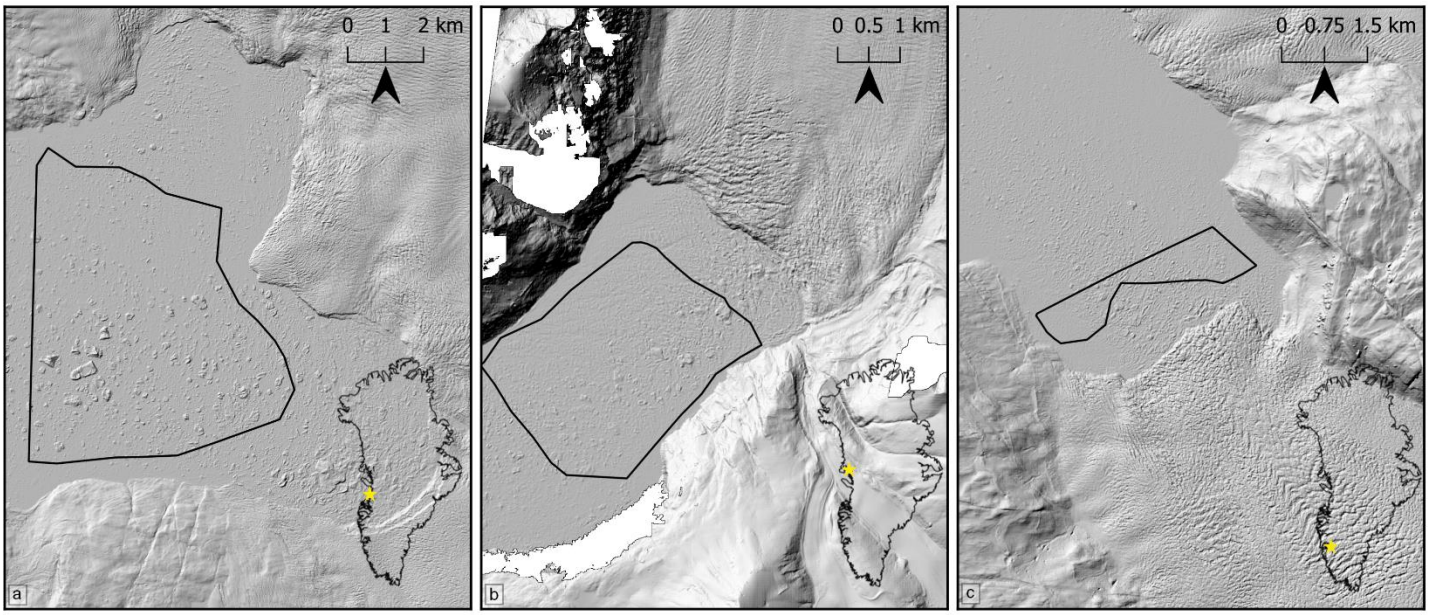
- The above will be clarified in the text (at L163) and as part of the GitHub read.me walkthrough (https://github.com/ConnorShiggins/Google-Earth-Engine-and-icebergs).



Draft figure 1. Google Earth Engine ArcticDEM v3 strip data availability (July-October) for Greenland's calving margins (Goliber et al., 2022) and the extent of all marine/lake/shelf terminating glaciers extent in the remainder of the Arctic (RGI v6; Pfeffer et al., 2014).

3. It would be good to see some figures showing what the DEM data looks like. You may have readers who have not worked with the Arctic DEM before, and it makes your workflow hard to understand without seeing some visualisations. Please can the authors add a figure (or two) where they deem it most appropriate.

We will replace our figure location maps that used Sentinel-2 imagery with hillshaded ArcticDEM data to provide readers with an indication as to what the DEM data look like. See below for the new draft location figure to replace Figure 1.



Draft figure 2. New location maps for the study sites with changed imagery (ArcticDEM from Sentinel-2) and ROI outlines in black.

4. Please can the authors double check that all results that they present have an equivalent section within the results section. Readers new to the topic need to fully understand (and even be able to recreate) how you take a 3-D DEM and produce area to volume conversions (for example).

    - After reviewing the manuscript in response to this comment, we believe all the data presented has a section within the results and discussion. With regards to reproducing our area-to-volume conversion relationships (Figure 7), all that is required is a power law relationship between the two variables (in this case area and volume) which was followed from previously published work (Sulak et al., 2017). We will provide the basic Python script which calculates the bin mean of each size class (area and volume) on the same GitHub read.me for users. All code produced by the authors that is used to post-process the output data, and the output data itself will be appended as supplementary data files in the revised manuscript. This will allow readers to both reproduce our results and workflow for other ROIs.

Specific comments:

1. (L15): Do you mean the GEE task run time is 6 minutes? Make this clearer.

    - The execution output time noted in the abstract is for the 3 glaciers which range from 6 minutes to 2 hours. We will clarify this in the revised manuscript (at L15 to L16).

2. (L25): Is there sufficient data coverage for a pan-Antarctic study? If not, I probably wouldn't say this.

- We assume the reviewer means the Arctic rather than Antarctic, though we provide responses relevant for each pole. For the Arctic, the new Supplementary Figure 1 (RL100) clarifies that pan-Arctic coverage is theoretically possible given the nominal availability of ArcticDEM data strips. However, a precise assessment of this would not be possible without performing the analysis itself, which is beyond the scope of this paper. For Antarctica, though similar strip data to ArcticDEM are available through REMA, the Antarctic is not the focus of this study. Anecdotally, from the experience of the authors using REMA versus ArcticDEM strip data available in Google Earth Engine, coverage and geolocation accuracy of the former tend to be poorer than those of ArcticDEM, posing challenges to pan-Antarctic application. The above will be clarified in the revised manuscript at L25.

3. (L30): Do you have a reference for shipping?

- The reference for shipping is Bigg (2015) (at L31).

4. (L33): add a 'that' after suggested

- This will be changed in the revised manuscript.

5. (L37): I don't think this sub-heading is necessary, especially as it captures all most all of your introductory material anyway.

- This will be removed from the introduction in the revised manuscript.

6. (L41): hyphenate 'Sentinel-2'

- This will be changed in the revised manuscript.

7. (L44): If CNN makes using optical imagery 'better', what is its disadvantage? Why do you need to use your method instead?

7. (L44): If the next paragraph is an attempt to address this, just make the link between paragraphs clearer.

- Convolution neural networks (CNNs) can be difficult to construct, requiring substantial training data that are often obtained from manual labelling of images. This can be computationally and user time intensive, while different training data used within the same CNN architecture will also provide different results. Though CNNs can produce high quality data (e.g. Rezvanbehbahani et al., 2020), the quality of data produced are highly contingent on the quality and range of their training data. The potential transferability of CNNs for iceberg detection beyond individual study locations and across different image illumination conditions remain relatively untested. Many CNNs are also not necessarily deterministic, so may also provide different results given identical

training data and CNN architectures. Additionally, CNNs using optical/radar satellite imagery will still be limited to only expressing a planform surface area, rather than a volume. Consequently, volumetric data can only be estimated through empirically derived area-volume conversions such as those presented in this manuscript (Equations 2 to 6).

- The approach presented in this manuscript using ArcticDEM data therefore offers advantages over CNNs in that our workflow is deterministic, applicable over wide areas, and can provide fully reproducible data of both iceberg areas *and* volumes. To address these comments, we will include mention in L44-45 regarding the difficulty of applying CNNs over large spatial scales.

8. (L53): replace 'are' with 'is'

- This will be changed in the revised manuscript.

9. (L99): hyphen needed between 'Sentinel' and '2'. Check elsewhere.

- This will be changed in the revised manuscript.

10. (L100): would there be a limit to this? If we kept using data with a finer spatial resolution I assume there would come a point where the xmin would stop decreasing?

- This is an interesting point which could be considered in future work using satellite imagery of different spatial scales and/or resampling individual high-resolution images to coarser resolutions. Though it would be possible to speculate that there may be a "minimum x_{min} value", we do not wish to do so here without data that explicitly supports this conclusion. As this would require substantial further analysis and is not an aim of the paper, we do not think it is possible to make such an assertion in this manuscript.

11. (L106): Do you want to identify ice bergs frozen together by melange though? I thought you wanted to avoid this and just wanted to identify individual icebergs?

- See response to main point 1 (RL13).

12. (L111): What makes the data suitable?

- Suitable data for constraining iceberg freshwater fluxes ideally require knowledge of an iceberg's volume and area (i.e. knowledge that could be parameterised within a fjord model to estimate how much freshwater could potentially be melted into the fjord and at what rate). Additionally, assumptions in numerical models are currently made regarding an iceberg distribution (e.g. power law slope = -1.8 to -2.0; Davison et al., 2020). To clarify what makes iceberg data observations suitable for inclusion in fjord models, we will add a sentence at L112 stating: "Models that include quantification of iceberg meltwater flux currently make assumptions regarding iceberg area/volume distributions within fjords, though direct observations of these from DEM or 2D satellite data are currently rarely available.".

13. (L129): Tidy up these figures where possible. The 'a' 'b' 'c' labels, north arrows, and scale bars would be better on a white background rather than a translucent background. Could you also make all the ROI outlines either green or red?

- Draft Figure 2 has been created in response to main comment 3 (RL04). This has been changed to show examples of ArcticDEM imagery and the ROI outlines have been changed to black for colour accessibility and consistency.

14. (L133): I would say this bounding box is green? Comment on the subset map also.

- Colour will be changed to black. See Draft Figure 2.

15. (L136): '-1' needs to be in superscript

- This will be changed in the revised manuscript.

16. (L145): what do you mean by this?

- This will be rephrased to: ‘The terminus depth of the glacier ranged from 230-500 m between 2013 and 2015 (Morlighem et al., 2017).

17. (L167): Is this enough to draw robust conclusions from?

- While 3 images at UI is less than at SKJI and KNS, the absolute number of observations and quality of data remains a substantial improvement on manual digitisation (e.g. 6,973 icebergs identified at UI for 3 images versus 712 icebergs manually delineated from 8 DEMs in Sulak et al. (2017)).

18. (L177): what if the ROI is dominated by sea ice, and there is little open water?

- The analysed DEMs are limited to between July and October of every year, minimising the likelihood that rigid melange and/or sea ice will be present at the glacier terminus. This means that the most frequent elevation in an individual DEM for these months is likely to be at or very near to the local sea level. Where continuous, solid sea/fjord ice cover dominates a scene the reviewer is correct that this may result in an over-estimation of sea level within the workflow. The value of the derived sea level is currently appended to observations exported from the workflow as metadata, allowing users to potentially filter data with anomalously high sea level values during post-processing. The requirement to do this will be contingent on a user’s research question. This will be clarified in the text at L177-178.

19. (L181): In the text (above) you state that the filters are replied in the opposite order. Correct either the figure or text.

- This will be changed in the revised manuscript.

20. (L182): Adding colours to this figure would help to differentiate between steps, rather than, or in addition to, different steps. However, at the moment, I cant work out why some steps are encased in different shapes?

- We did not use colour in the first version of the manuscript for accessibility (e.g. colour blindness). The workflow steps are encased by different shapes because they represent different elements of the code. In the revised manuscript, we will add the meaning of each shape in the figure caption at L183 as follows: 'Each step of the workflow is encased by different shapes representing different processes in the code, i.e. ovals = the beginning and end of the workflow; the inverse trapezoid = a manual requirement; italicised parallelograms = data inputs; rectangles with inset lines = predefined filter processes; and rectangles = code processes. We will also add a legend to the figure indicating what each shape indicates.

21. (L188): Would it not have still been better to have worked in 0.1m increments here too?

- The increments of 0.5 m at SJKI only resulted in a small variation of 0.04 across all values of the threshold (1 to 5 m) as shown in Figure 5d. Consequently, these increments resulted in small absolute variation in power law slopes, meaning that it would be unnecessary to use increments smaller than 0.5 m at SKJI. We will state on L188 in the revised manuscript that: 'There are small variations (~0.04) in the power law slopes at SKJI across all detection thresholds tested, demonstrating a relative lack of sensitivity of power law slope to threshold value used.'.

22. (L195): I would argue this information is implicit in binary, but I suppose you are stating which values represent what.

- Yes, we wanted to ensure readers who may not be aware of binary images understood the process behind the iceberg detection.

23. (L199): From your figure I can see that you export results to Google Drive, is there an option to export results as GEE assets?

- Yes, it is possible to export output to GEE assets within the workflow during the export stage. An explanation of how to do this will be added to the GitHub readme.

24. (L205): How did you get to these values, did you conduct any form of testing?

- These values fall within the known x_{min} values from previously published work (e.g. Scheick et al., 2019, Rezvanbehbahani et al., 2020). This will be clarified in the revised manuscript (at L205).

25. (L210): rather than this, just state the areas of the three ROIs.

- This will be changed in the revised manuscript.

26. (L213): This is vague. If they are quantitatively comparable, please provide the statistics.

- We can add the Pearson's r-values in brackets in the main text if required, though these values are also given in Figure 3.

27. (L217): Some of this info could probably be placed in supplementary info, then this table will be a bit less crowded.

- We think all the data presented in Table 1 is necessary and provides useful information for readers to refer to in the main manuscript without the need to access supplementary files. We therefore propose to retain data presented in the submitted manuscript for the revised version.

28. (L225): ???

- The 225-line number has entered table 1 accidently when formatting and will be corrected.

29. (L235): Increase size of axis font.

- This will be changed in the revised manuscript.

30. (L240): Please include a description of the statistics in this table in your methods section. I know it may seem obvious, but the methods for any result obtained should be provided.

- The Pearson's r-value is stated in Figure 3 and the respective caption. We will add to the methods that we used the Pearson's r-value to gauge the strength of relationship between the automated and manual delineations (at L207).

31. (L250): Please increase the size of the scale bars here so that they are legible.

- This will be changed in the revised manuscript.

32. (L266): This is a stylistic preference, but I would re-write this sentence so that you are always saying 'sea level ranged from' or X's 'range was' rather than mixing between the two.

- This will be changed for the revised manuscript, and we will endeavour to ensure consistency of language used throughout.

33. (L272): Please increase size of font on axis

- This will be changed in the revised manuscript.

34. (L296): Please re-write this sentence to make it clearer. At first I thought you were saying the y axis with their log scales were different, but they are not

- We will remove the word ‘normalised’ from the caption and clarify that the y-axis log scales are not different.

35. (L302): Do you know which of these scenarios is actually true from visual interpretation of data?

- In retrospect, we feel this point might be better suited in the discussion (at L485) and it will therefore be moved to expand on the comment made.

36. (L310): Please increase font sizes.

- This will be changed in the revised manuscript.

37. (L323): State what the black lines represent, and perhaps make them red/ blue so the reader can see whether they are linked to the manual or automated dataset.

- The black lines represent the lines of best fit for the icebergs in each distribution of the manual and automated approaches and we will clarify this in the figure legend and caption in the revised manuscript (at L329). On drafting a version of the figure where the colour of lines matched the data points we find that this reduces the clarity of the figure as we are unable to visually discriminate between data points and the lines of best fit. While admittedly not ideal, we propose to retain the lines of best fit as black.

38. (L331): Please can you better describe the methods used to achieve this in the methods section.

- In the methods we will add a sentence at L207 saying ‘New equations for the conversion of iceberg area to volume are derived from the resulting iceberg datasets. These are expressed as power laws to provide consistency with previously published work (e.g. Sulak et al., 2017).’.

39. (L341): how do you define small / large? Can this be quantified?

- We define the separation between small and large icebergs as 1000 m^2, as that is consistent with Rezvanbehbahani et al.’s (2020) definition. We mention this later in the manuscript (L457), but we will refer to this directly in the revised manuscript at L341.

40. (L376): Maybe place this figure after you have mentioned the two distributions, as currently I see this figure and the contents do not make sense until later in the text.

- If the manuscript reaches copy-editing stage, we will ensure that the figure is placed at an appropriate point within the paper to reflect this comment.

41. (L395): What are you trying to say here? It is unclear to me. Please re-write.

- This will be reframed as: ‘By calculating mean iceberg area and volume for binned increments of $\log_{10}(X+0.1)$, this reduced the potential for biasing the overall area-volume relationship towards smaller, more frequently observed icebergs.’.

42. (L424): Please increase font sizes

- This will be changed in the revised manuscript.

43. (L463): Given that the legend is the same for each of these subfigures, you could probably just put it on one subfigure. I would keep it in (c) and remove it from (a) and (b)

- To avoid any potential for ambiguity, we suggest that it is appropriate to retain the legends in each subplot.

44. (L465): The last bit of the caption here (stating the count vs volume of small icebergs) isn't really something that belongs in the figure caption, it should be in the text.

- We will insert this section of the figure caption into the text on L460.

45. (L474): Please can you comment on data availability? Does it allow for pan-Arctic application?

- See response to main comment 1 (RL13) and minor comment 2 (RL136).

46. (L475): rephrase to 'is quick to execute'

- This will be changed in the revised manuscript.

47. (L476): change to defining

- This will be changed in the revised manuscript.

48. (L483): I assume you mean a mismatch between manually delineated and automatically delineated icebergs? If so, please make this a bit clearer.

- We will clarify this point by stating the automatic approach only analyses whole pixels (L484). See also response to minor point 49 (RL398) in this review, and RL184 in response to Reviewer 2 minor point 12.

49. (L485): Please clarify what you are saying here. Do you mean to say that the manual classifications over estimate iceberg size relative to the automated classifications?

- Yes, we will clarify this in the revised manuscript by stating: ‘The automated approach identifies icebergs through analysis of whole pixels, rather than the

manual delineation which will have iceberg outlines digitised across pixels' (at L485).

50. (L486): Is figure 4 actually showing hillshaded DEMs? If so please state this in the caption and proximal text.

- Yes, the DEMs in figure 4 are hillshaded with the detected icebergs shaded with their respective outlines. In the revised manuscript we will add this to the Figure 4 caption that they are hillshaded, and we will also change the colour scheme to allow the hillshading to be seen more clearly by readers.

51. (L520): Please can the authors comment on this? I thought the purpose of the threshold set for height above sea level was used in order to prevent the detection of multiple icebergs 'stuck together' by melange? Surely at these two study sites you need to increase the threshold, and then you would only get one iceberg population?

- We wanted to highlight in the manuscript that it is possible for the workflow to identify different iceberg distributions present in the fjord. The user definition of the threshold above sea level allows flexibility for the user to obtain data most relevant for their research question (i.e. it is possible to derive separate relationships for rafted and non-rafted icebergs). The section as written illustrates that the workflow allows flexibility for this. Our response to the reviewer's main comment 1 (RL13) will also help to clarify the point raised here.