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

Connor J. Shiggins[1], James M. Lea[1], Stephen Brough[1]

[1]Department of Geography and Planning, School of Environmental Sciences, University of Liverpool, Liverpool, L69 7ZT, United Kingdom.

Correspondence to: Connor J. Shiggins (Connor.Shiggins@Liverpool.ac.uk)

Abstract. Iceberg calving accounts for up to half of mass loss from the Greenland Ice Sheet (GrIS), with their size distributions providing insights into glacier calving dynamics, and impacting fjord environments through their melting and subsequent freshwater release. Iceberg area and volume data for the GrIS are currently limited to a handful of fjord locations, while existing approaches to iceberg detection are often time consuming and are not always suited for long time series analysis over large spatial scales. This study presents a fully automated workflow for the detection of icebergs within Google Earth Engine using high spatial resolution timestamped ArcticDEM (Arctic Digital Elevation Model) strip data. This is applied to three glaciers that exhibit a range of different iceberg densities and size distributions: Sermeq Kujalleq (Jakobshavn Isbræ), Umiammakku Isbræ and Kangiata Nunaata Sermia. A total of 39 ArcticDEM scenes are analysed, detecting a total of 163,738 icebergs in 6 minutes to 2 hours for each glacier depending on the number of DEMs available and total area analysed, comparing well with manually digitised outlines. Results reveal two distinct iceberg distributions at Sermeq Kujalleq and Kangiata Nunaata Sermia where iceberg density is high, and one distribution at Umiammakku Isbræ where iceberg density is low. Small icebergs are found to account for over 80% of each glacier's icebergs however, they only contribute to 10-37% of total iceberg volume suggesting that large icebergs are proportionally more important for glacier mass loss and as fjord freshwater reservoirs. The overall dataset is used to construct new area to volume conversions (with associated uncertainties) that can be applied to two-dimensional iceberg outlines derived from optical or synthetic aperture radar imagery. When data are expressed in terms of total iceberg count and volume, insight is provided into iceberg distributions that have potential applicability to observations and modelling of iceberg calving behaviour and fjord freshwater fluxes. Due to the speed and automated nature of our approach, this workflow offers the potential to interrogate iceberg data on a pan-Arctic scale where there is sufficient ArcticDEM coverage.

1.0 Introduction

Iceberg production is of critical importance when considering the mass balance of ice sheets and glaciers (Bigg et al., 2014),
30 freshwater fluxes (Enderlin et al., 2016; Davison et al., 2020a), offshore infrastructure (Eik and Gudmestad, 2010), shipping,
tourism (Bigg, 2015) and ecological habitats (Laidre and Stirling, 2020). Their area size distributions can be used to infer
glacier calving dynamics (Sulak et al., 2017; Scheick et al., 2019; Åström et al., 2021, Cook et al., 2021) and also estimate
freshwater fluxes (Enderlin et al., 2016; Moon et al., 2018; Moyer et al., 2019; Davison et al., 2020a). It has been suggested
icebergs could account for up to 22-70% of the total mass loss by 2100 from the Greenland Ice Sheet (GrIS) (Choi et al., 2021),
35 though how future changes in glacier dynamics will influence iceberg size distributions (and vice versa) is currently poorly
constrained.

1.1 Background to iceberg detection

Multiple different approaches have been taken to iceberg detection, including analysis of optical imagery, synthetic aperture
radar (SAR) imagery and digital elevation models (DEMs). Semi-automated and/or automated iceberg detection utilising
40 optical imagery typically involves band thresholding to differentiate ice and water (Sulak et al., 2017; Moyer et al., 2019).
However, these approaches often use medium-resolution data (10 – 30 m pixel data, e.g. Landsat and Sentinel 2) that have
insufficient spatial resolution to identify small icebergs or distinguish between larger adjacent icebergs without more complex
processing. For example, convolutional neural networks (CNN) have been developed to downsample images, allowing the
delineation of smaller iceberg edges at sub-pixel scale (e.g. Rezvanbehbahani et al., 2020).

45

In optical imagery, the presence of ice mélange (mixture of icebergs and sea ice) in images also proves problematic for
automated band thresholding techniques. This arises due to the similar reflectance signal of mélange to that of icebergs,
potentially leading to the generation of erroneously large outlines. Additionally, prolonged cloud cover in some parts of the
polar regions and polar night can result in large gaps between observations using optical imagery.

50

SAR data has the potential for more continuous coverage as the active nature of the sensor can penetrate cloud cover, and does
not rely on solar illumination to acquire imagery (e.g. Soldal et al., 2019). However, a notable shortfall of both optical and
SAR data are that they are only capable of expressing a surface area of an iceberg, with volumes typically estimated using
empirical area-volume relationships derived from DEMs (Sulak et al., 2017; Schild et al., 2021).

55

Time-stamped ArcticDEM version 3 (Porter et al., 2018) tiles provide an under-exploited resource that allows the derivation
of both iceberg areas and their volumes. These data are obtained from optical stereo-image pairs acquired between 2009 and
2017, providing high spatial resolution DEMs (2 m posting), though have variable temporal coverage due to cloud
contamination and satellite image acquisition tasking. While this archive currently has poor return frequency compared to

60 optical and SAR satellite platforms, its spatial resolution and ability to determine iceberg volumes offers the potential for gaining insights that can also be applied to more frequently acquired optical and SAR derived data.

Due to the significant numbers of icebergs existing at any one time in the polar regions, time intensive manual delineation is not a practical approach to apply to ice sheet-wide analysis or even at a single glacier site. However, manually digitising

65 icebergs are viable options for: 1) creating training sets for supervised classification of semi-automated approaches for a selection of image scenes (Sulak et al., 2017); and 2) to generate highly targeted datasets of icebergs, e.g. the Canadian ice island drift, deterioration and detection (CI2D3) database (Crawford et al., 2018).

Iceberg area distributions have previously been used to constrain glacier calving dynamics (Scheick et al., 2019) and determine

70 iceberg disintegration processes (Kirkham et al., 2017). Area-size distributions of icebergs have previously been described using power laws in particle modelling studies (Åström et al., 2021) and from imagery in areas adjacent to glacier termini, to gain insight into calving dynamics (Enderlin et al., 2016; Sulak et al., 2017; Scheick et al., 2019; Rezvanbehbahani et al., 2020). These relationships describe probability distributions of iceberg size, with Equation 1 describing the general form of these relationships,

75
$$p(x) = Cx^{-\alpha}, \ where \ x \geq x_{min} \tag{1}$$

where p(x) is the distribution, C is a constant and α is the exponent of the power law (or slope value). The value of α (reported hereafter including the negative sign in Equation 1) provides an indication of iceberg size distributions at the time of data acquisition with lower values suggesting a higher prevalence of smaller icebergs, whereas more positive values suggest

80 relatively larger icebergs dominate. Typical α values for Greenlandic and Antarctic environments have been reported between -1.2 and -3.0. As icebergs drift from Greenland's termini to the open ocean, their distributions have been observed to transition from being best described as power law distributions (suggested to be controlled by calving) to lognormal distributions as melting becomes the primary control on their disintegration (Kirkham et al., 2017).

85 When fitting icebergs to power law distributions and calculating α, it is important to determine a threshold which removes icebergs below a certain area size (x_{min}). Where smaller icebergs are included in the distribution, these can result in less robust fits with power laws because they follow different size distributions compared to larger icebergs. Including smaller icebergs in this analysis can therefore disproportionately skew the α value and potentially misrepresent the data (Scheick et al., 2019). Given the larger surface area to volume ratios of smaller icebergs, it is also more likely that their different size distribution

90 arises from more extensive modification by submarine and atmospherically driven melting. Defining the appropriate x_{min} value is therefore critical for investigations that seek to interrogate how iceberg size is impacted by glacier calving processes.

A further complexity of the x_{min} value is if the value is defined too high there will be significant data loss that will limit the explanatory value of the distribution. This is especially the case for glaciers where there is a high proportion of small icebergs.

95   For example, at Sermeq Kujalleq (Jakobshavn Isbræ), Scheick et al. (2019) defined an x_{min} value of 1,800 m^2 as it improved the fit compared to other x_{min} values tested, appropriately justifying this in terms of both their research question and the glacier being characterised by relatively large icebergs. In other studies, the resolution of imagery available has impacted the range of x_{min} values that can be defined. For example, CNN performed on Planet imagery (3 m optical imagery) resulted in x_{min} values of 288 m^2 and 387 m^2, while Sentinel 2 (10 m optical imagery) required values of 12,000 m^2 and 3,200 m^2 for Sermilik and

100 Kangerlussuaq Fjords, respectively (Rezvanbehbahani et al., 2020). This demonstrates how the availability of finer spatial resolution data can also allow the definition of smaller x_{min} values and the retention of more data.

Few studies have been able to directly estimate iceberg volume (e.g. Sulak et al., 2017), as optical and/or SAR imagery are (without significant further processing) limited to the extraction of iceberg areas only. The three-dimensional shape of an

105 iceberg above the waterline allows its volume to be inferred, though it does not always scale exactly with its planform area. For example, rafts of icebergs frozen together by mélange that occur at some glaciers will be relatively flatter and have a lower volume compared to single icebergs of the same area that have calved from a glacier. Applying iceberg area to volume conversions to these raft outlines would therefore lead to an over-estimation of their volumes.

110 One of the current difficulties faced by those studying the impact of icebergs on fjords is the lack of available iceberg outline and volume data that are suitable for use in numerical models of fjord circulation, stratification and iceberg melting (e.g. Moon et al., 2018; Davison et al., 2020a). This issue is compounded by the time and computational expense involved in the detection of icebergs (e.g. data collection, storage, memory and processing). A potential solution to this is offered by the Google Earth Engine (GEE) cloud computing platform (Gorelick et al., 2017) that provides the ability to rapidly access and process data

115 from multiple different satellites, offering the potential for ice-sheet-wide and global analysis (e.g. Shugar et al., 2020).

This study provides a GEE workflow and easy to use graphical user interface (GUI), using 2-m strip ArcticDEM version 3 data (Porter et al., 2018) to automatically detect icebergs at three marine-terminating glaciers on the west coast of Greenland. The aim of this study is to demonstrate the ability of the workflow to automatically generate a large and reliable dataset of

120 icebergs from glaciers of varying size and fjord conditions. In doing so it aims to allow users to gain detailed insight into iceberg area-volume relationships, and identify how these vary between glaciers.

2.0 Study sites

Three different marine-terminating glaciers were selected to conduct analysis, identified on the basis of their different fjord environments, iceberg sizes and data availability: (1) dense large iceberg coverage: Sermeq Kujalleq (Jakobshavn Isbræ)

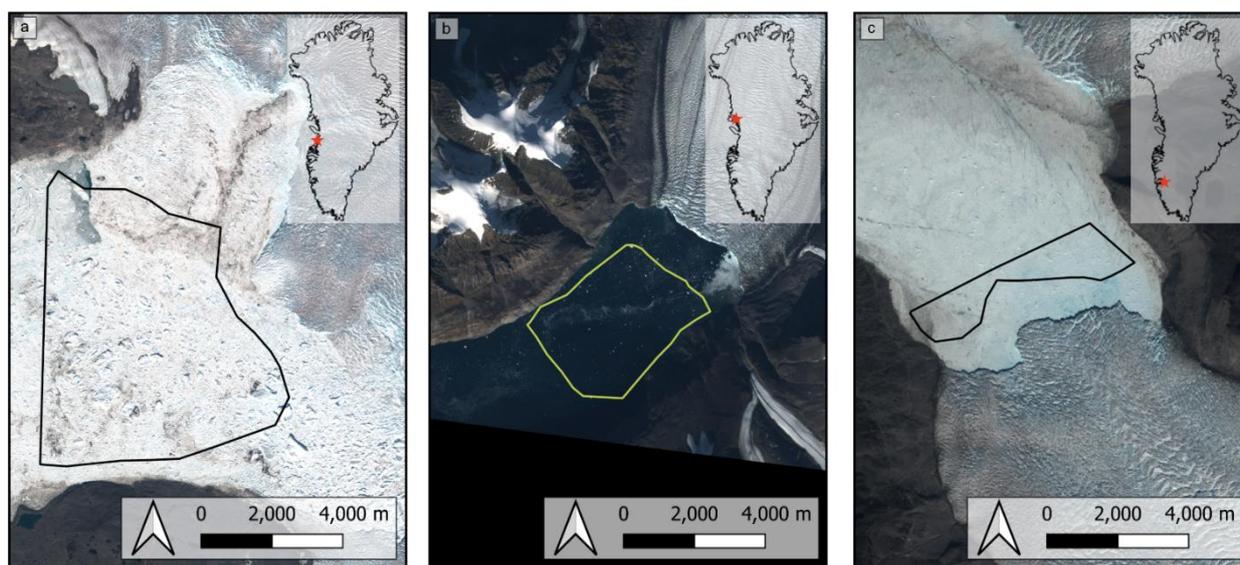
125 (herein SKJI); (2) mix of dense iceberg coverage and frequent open water: Umiammakku Isbræ (herein UI); and (3) dense small iceberg coverage with occasional open water: Kangiata Nunaata Sermia (herein KNS) (Figure 1). Regions of interest (ROI) at each glacier were identified to maximise ArcticDEM data availability and reduce the impact of winter/spring seasonal advance of the caving margin during the study period of 2009-2017.

130

Figure 1. Sentinel 2 imagery of the near terminus region for a) Sermeq Kujalleq (4th August 2021): 69.16o N, 49.91o W, b) Umiammakku Isbræ (22nd August 2021): 71.42o N, 52.26o W and c) Kangiata Nunaata Sermia (9th July 2021): 64.25o N, 49.50o W ROIs are mapped by black (a and c) and yellow (b) bounding boxes.

135 SKJI accounts for 45% of the total drainage of Greenland's central west sector, with a mean ice discharge (2010-2018) of 43.64 Gt yr-1 (Mankoff et al., 2019; Mouginot et al., 2019) (Figure 1a). Ice mélange buttressing of its terminus can inhibit calving, influence flow and allow advance (e.g. Amundson et al., 2010; Cassotto et al., 2021). Between 2011 and 2017, SKJI experienced a range of grounding line depths varying from 828 m and 980 m (Morlighem et al., 2017; Khazendar et al., 2019), producing icebergs as large as 700-1,000 m across, forcing ice mélange down-fjord because of full-thickness calving events

140 (Amundson et al., 2010; Walter et al., 2012). The retreat of SKJI from an annually floating terminus which calved larger icebergs has led to a seasonally grounded terminus, causing much smaller icebergs to be calved during the summer months (Scheick et al., 2019).

UI has a mean ice discharge (2010-2018) of 1.36 Gt yr-1 (Mankoff et al., 2019) (see Figure 1b). The terminus was grounded

145 in 2013 and 2015 between 230 m and 500 m (Carroll et al., 2016; Morlighem et al., 2017; Fried et al., 2018). Between 2003
and 2008, prior to the study period, UI experienced a substantial (4 km) rapid retreat of its terminus (Bartholomaus et al., 2016; Fahrner et al., 2021).

KNS is the largest marine-terminating glacier south of SKJI on the west coast of Greenland with a mean ice discharge (2010-
150   2018) of 4.92 Gt yr-1 (Mankoff et al., 2019) (see Figure 1c). It has retreated over 23 km from its Little Ice Age maximum position (Lea et al., 2014a; Lea et al., 2014b), but has remained relatively stable in the last decade (Davison et al., 2020b; Fahrner et al., 2021). The glacier's fjord is typically filled with mélange of small icebergs and brash ice and currently has a relatively shallow grounding line depth of approximately 250 m (Morlighem et al., 2017). While the development of a channelised, subglacial hydrological system at KNS increases localised calving activity due to greater submarine melt and
155   plume surfacing, it decreases terminus-wide calving and suggests high levels of runoff could decrease the number of calving events (Bunce et al., 2021).

3.0 Data and Methods

3.1 ArcticDEM data

The availability of ArcticDEM within GEE and its high 2-m spatial resolution (10 cm vertical accuracy) is used to create an
160   automated workflow to delineate icebergs and validate these against manually digitised outlines. The workflow is also packaged within a GUI (see: https://github.com/ConnorShiggins/Google-Earth-Engine-and-icebergs). To ensure a consistent level of high-quality data, analysis is automatically limited to only include DEMs generated from stereopair images acquired on the same day. In doing so, this limits the effect of iceberg drift, ensuring that only the highest quality DEMs are analysed. DEMs acquired between the months of July and October are analysed to avoid the presence of seasonal floating ice tongues
165   that form and persist through winter and spring that could lead to erroneous results. The data availability for each glacier is variable, with KNS having 16 available images from 2013-07-04 to 2017-08-26, SKJI 20 images, ranging from 2011-07-08 to 2017-08 -09 and UI 3 images between 2012-07-04 and 2017-07-03.

3.2 Workflow description

The only user defined input required for the code to execute is a ROI (Figure 2). The workflow dynamically filters the
170   ArcticDEM image collection to retain DEMs with >80% coverage of the ROI, before scenes with low image quality (e.g. cloud affected) are removed by calculating the 90th percentile of a scene's elevation, and ensuring that it is within ± 10 m of the WGS84 geoid.

To allow for potentially poor spatial registration in the Z dimension of the DEM and different tidal states at the time of data
175   acquisition, sea level is automatically calculated for each individual DEM. This is achieved by assuming that when DEM elevation values over the fjord are plotted as a histogram with 0.25 m bin widths, its peak (i.e. the most common elevation in

the DEM) represents sea level at the time the image was acquired (Appendix A, Figure A1). This allows each DEM to be registered to a common base level (i.e. 0 m above sea level) for consistent iceberg identification, and calculation of iceberg freeboard height and volume.

180

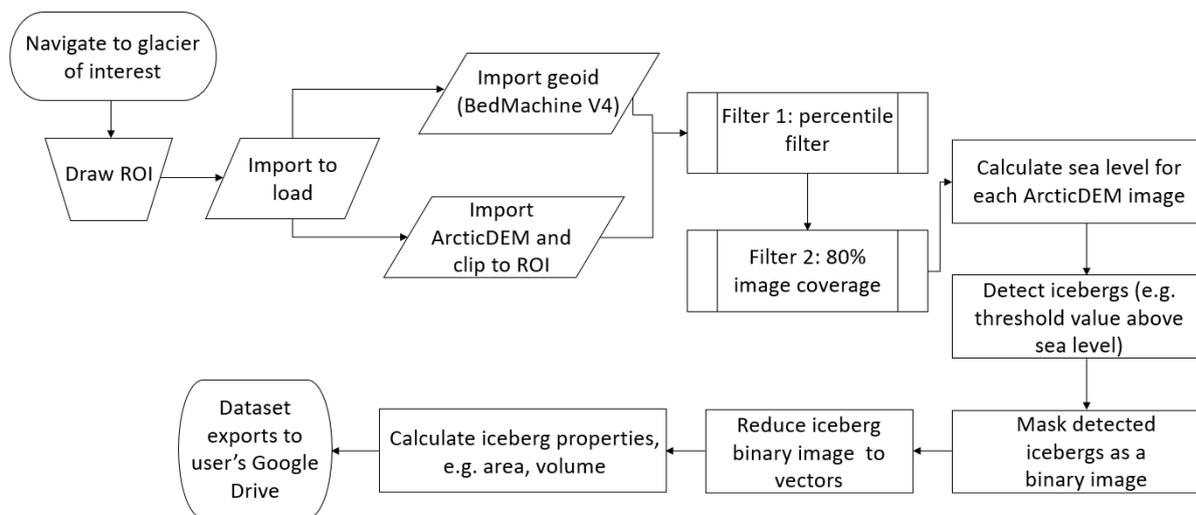



Figure 2. Workflow model of the automated iceberg detection in GEE.

To delineate iceberg outlines, it is necessary to separately define a threshold value above sea level where icebergs can be

185 confidently delineated without multiple icebergs being erroneously merged. Consequently, derived iceberg areas and volumes from the workflow represent minimum estimates. Potential threshold values for each glacier were explored, using increments of 0.1 m between 0.1 m and 1.5 m for KNS and UI (glaciers where small icebergs dominate), whereas this was increased to 0.5 m increments between 1.0 m and 5.0 m for SKJI where dense concentrations of large icebergs exist. From these results, the most appropriate iceberg detection threshold was evaluated through visual comparison to manually digitised iceberg

190 outlines. Through visual comparison, the most appropriate threshold was determined to be 1.5 m above sea level for KNS and UI, and 3.0 m for SKJI. The workflow uses these values to identify any area that exceeds these thresholds above sea level as an iceberg. Depending on the type of fjord environment (e.g. densely packed, open water) and the research question being addressed, the user can potentially alter the default iceberg detection threshold of 1.5 m above sea level.

195 Within the workflow, areas of the DEM that exceed the threshold are converted to a binary image (1 = iceberg, 0 = no iceberg) which are then vectorised into iceberg outlines. Iceberg specific metadata (e.g. area, volume) are appended to each outline automatically, using DEM input data where needed. The final part of the workflow removes any large object (> 100,000 m^2) in case of false iceberg detection by erroneously delineating fjord edges and/or the glacier termini before the user can choose to export results in their preferred file format (e.g. CSV, Shapefile or GeoJSON).

3.3 Iceberg distributions

Iceberg areas and volumes from each glacier are fitted to power law distributions as described in Equation 1 using the 'powerlaw' package in Python (Alstott et al., 2014). To allow consistent comparison of how power law distributions evolve through time x_{min} values are kept the same for every image, defined as 500 m^2 for KNS and UI, and 1,000 m^2 for SKJI. The lower x_{min} value of 500 m^2 for KNS and UI was chosen as they produce smaller icebergs compared to SKJI, meaning that 1,000 m^2 value would have resulted in significant data loss. Both values assigned for the three glaciers allowed reduced skewing of the α exponent and more robust fits to power law distributions. The x_{min} values set are also within the range used by previous studies (e.g. Sulak et al., 2017; Scheick et al., 2019; Rezvanbehbahani et al., 2020).

4.0 Results

4.1 Workflow evaluation

The ROI at each glacier varied from 5.3 m^2 to 41.0 m^2 with the number of detected icebergs across all available images ranging from 6,973 at UI to 147,714 at SKJI (Table 1). For each individual glacier, iceberg distributions obtained from automated and manual delineation methods were found to be qualitatively and quantitatively comparable (Figures 3 and 4; Table 1).

Table 1. Data from the three glaciers, including the ROI size, the date of the ArcticDEM image which was manually validated, number of images in the entire collection, number of icebergs detected, both automated and manual power law slope values (with one sigma) for area with corresponding x_{min} (total iceberg volume below and above the respective value: SKJI =1,000 m^2, UI and KNS = 500 m^2) and the execution time.

	ROI (km^2)	Date	Images	Total number of icebergs detected (icebergs per scene per km^2)	Automated power law slope	Manual power law slope	$X_{min\,(m^2)}$	Total iceberg volume below x_{min} value (km^3)	Total iceberg volume above x_{min} value (km^3)	Execution time (minutes)
SKJI	41	2011-09-10	20	147,714 (180)	-1.88 ± 0.06	-1.91 ± 0.06	1,000	0.984 (10%)	8.629- (90%)	~120 225
UI	9.6	2012-07-04	3	6,973 (242)	-2.16 ± 0.12	-2.17 ± 0.13	500	0.016 (24%)	0.051 (76%)	6
KNS	5.3	2013-08-21	16	9,051 (107)	-2.38 ± 0.16	-2.77 ± 0.16	500	0.021 (22%)	0.074 (78%)	8

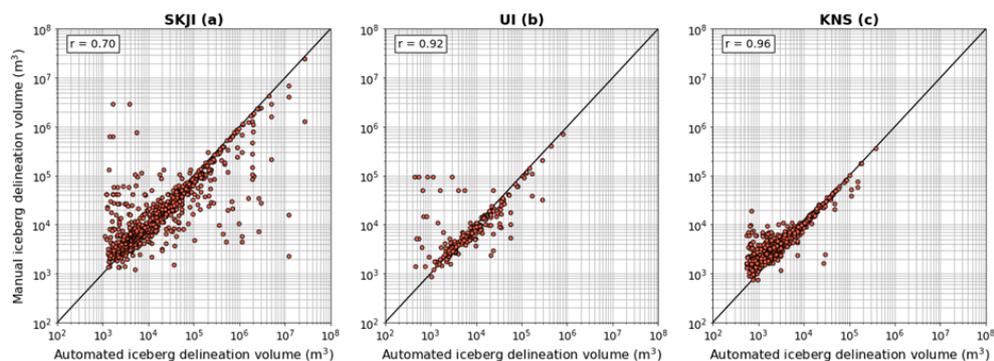
230

	SKJI (a)	UI (b)	KNS (c)
n	805.0	198.0	507.0
Automated volume sum (km³)	0.2127	0.0049	0.0045
Manual volume sum (km³)	0.1065	0.0044	0.0046
Difference (%)	49.0	11.0	2.0
RMSE (km³)	1.767e-05	6.682e-05	2.437e-05

Figure 3. The relationship between the iceberg volume for both the manual and automated delineation methods for each glacier and respective summary statistics. The Pearson's correlation coefficient is also highlighted (SKJI = 0.70, UI = 0.92, KNS = 0.96).

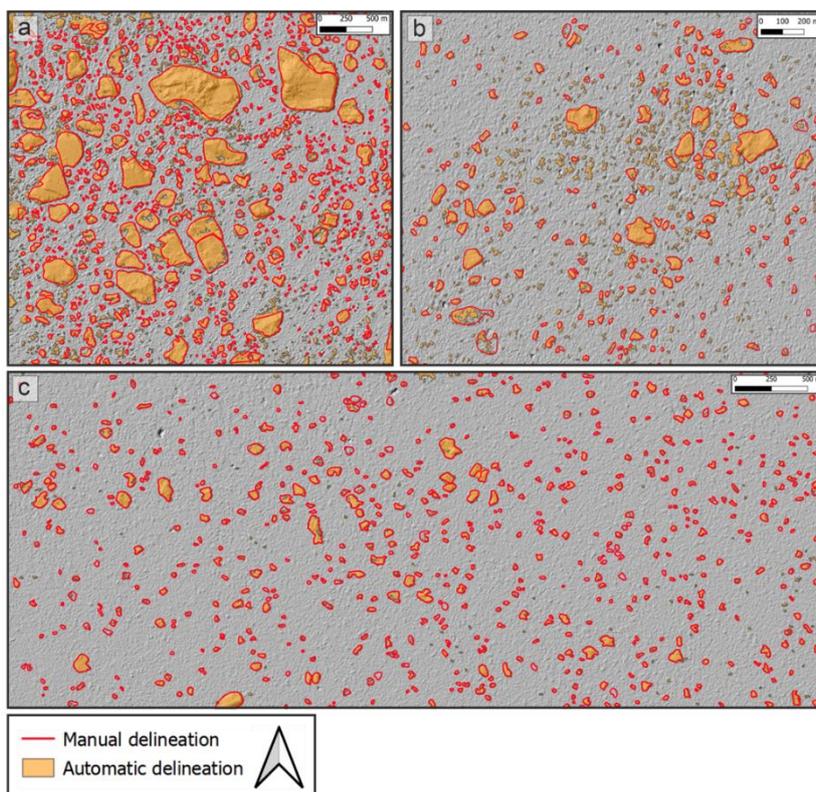



Figure 4. Manual (red lines) and automated (orange) delineation of the iceberg subset for a) SKJI (2011-09-10), b) UI (2012-07-04) and c) KNS (2013-08-21).

265

For each ArcticDEM scene throughout the study period, sea level ranged from 23 m to 32 m at SKJI, KNS' range was 31 m

to 39 m and UI ranged from 26 m to 31 m. Visual comparison between manual and automatically delineated data for each

threshold showed that threshold values of 1.5 m above sea level for KNS and UI, and 3.0 m above sea level for SKJI (Figure

5) provided the best visual correspondence and provided more concordant power law fits with manually digitised outlines

270   (Figure 6; see Methods).

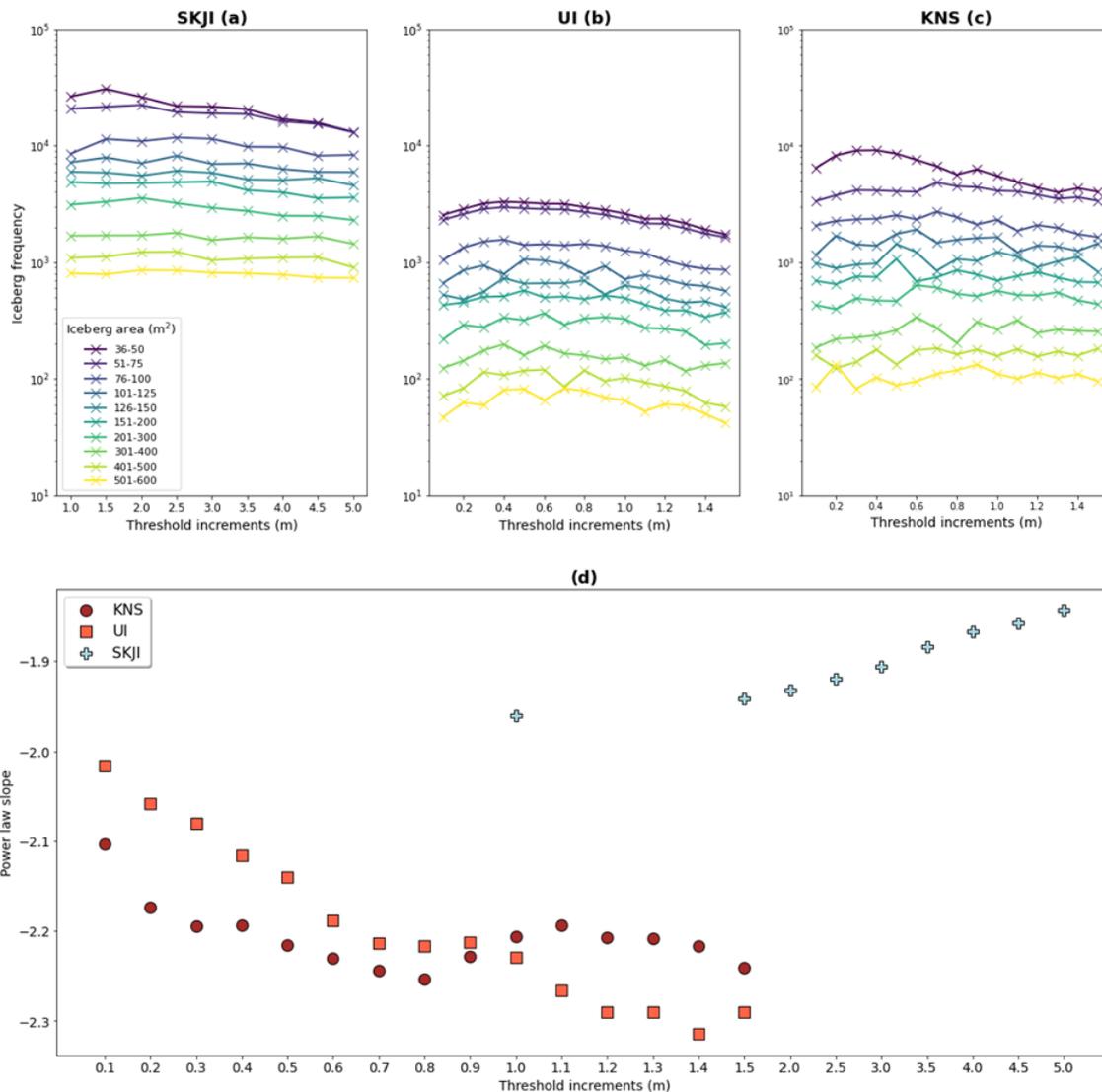



Figure 5. Iceberg frequency for each threshold increment tested. a) SKJI's increments were every 0.5 m above sea level between 0 and 5 m above sea level. Increments of 0.1 m between 0 and 1.5 m above sea level were ran for both b) UI, c) KNS. d) shows how the α value for each glacier changes, depending on which threshold increment is chosen to detect their respective icebergs. Note the log-scale and normalised y-axis and the x-axis for SKJI (a) is on a different scale to that of UI (b) and KNS (c).

Iceberg area size distributions of both the manually and automated methods are found to follow power law distributions for the x_{min} values applied (Figure 6). Results reveal that SKJI has the least negative power law slope (α = -1.88) of the three glaciers followed by UI (α = -2.16) and KNS has the most negative values (α = -2.38), correctly highlighting icebergs at SKJI are generally larger than those at UI or KNS (Figure 4). Good correspondence between automatically and manually delineated iceberg area α values were observed for SKJI and UI where they differed by 0.03 and 0.01 for SKJI and UI respectively. However, this increased to 0.39 for KNS suggesting either manual digitisation underestimates the areas of small icebergs, or that the automated workflow underestimates the number of small icebergs for the defined x_{min} value (Figure 6). Power law relationships applied to iceberg volume distributions for each of the glaciers showed similar results, however the difference in the α value is minimised to 0.02 and 0.01 at SKJI and UI respectively, and reduces to 0.20 at KNS.

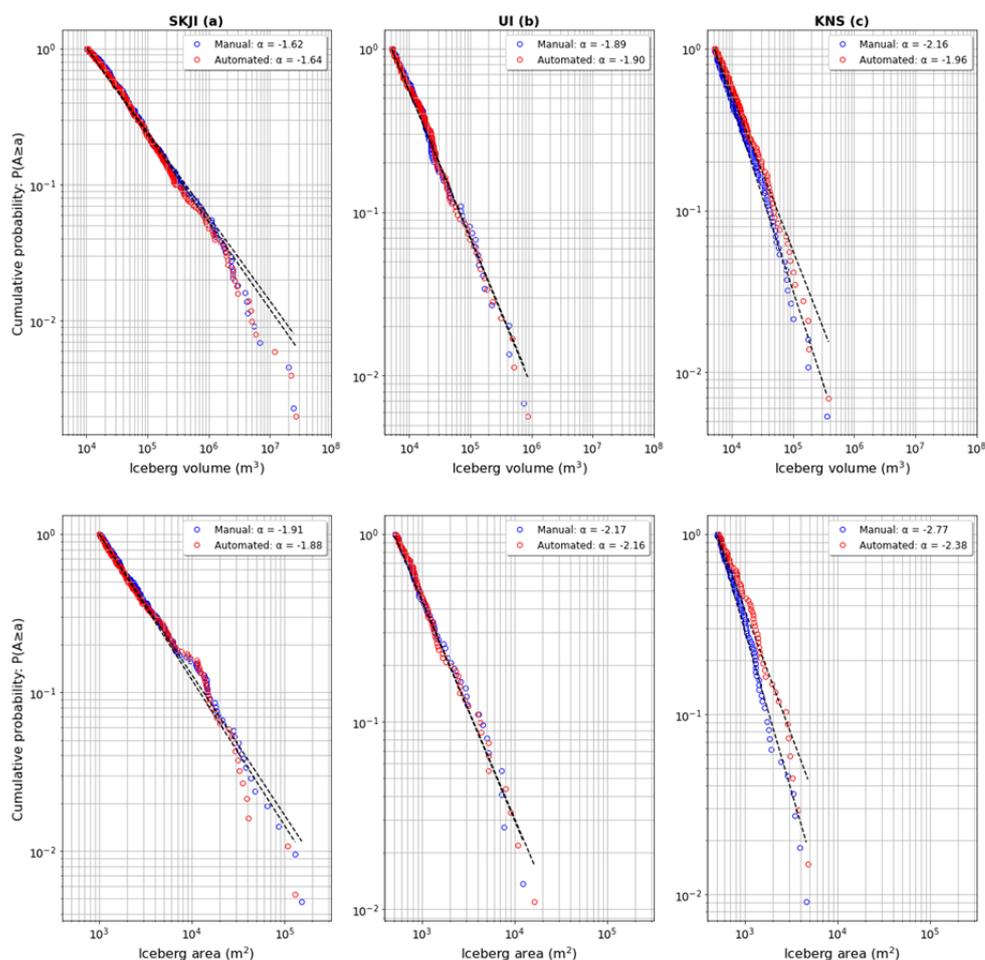



Figure 6. Power law plots for the manual (blue open circles) and automatically (red open circles) delineated icebergs. For iceberg area: a) SKJI has an x_{min} value defined at 1,000 m^2, compared to 500 m^2 for both b) UI and c) KNS. For defining an x_{min} for iceberg volume distributions, the respective area x_{min} value was converted using the equation of Sulak et al. (2017; their equation 5) yielding a value of 10,270 m^3 for SJKI and 5,135 m^3 for UI and KNS.

330 **4.2 Iceberg area and volume distributions**

The three-dimensional nature of DEMs allow the volume of each iceberg to be calculated assuming neutral buoyancy, allowing the derivation of the relationship between planform iceberg area (A) and volume (V) (Figure 7). For the entire iceberg dataset can be expressed as

$$V = 14.90A^{1.16} \qquad (2).$$

335 The large nature of the dataset also allows equations describing the lower and upper confidence bounds to be derived, with the 5th percentile of the distribution described by

$$V = 7.55A^{1.18} \qquad (3)$$

And the 95th percentile of the distribution described by

$$V = 15.73A^{1.20} \qquad (4).$$

340 When compared to the previously published area to volume conversion equation of Sulak et al. (2017; their equation 5), their relationship would produce lower volumes for small area icebergs, and higher volumes for large area icebergs (Figure 7).

345

350

355

360

365

370

375

380

385

390

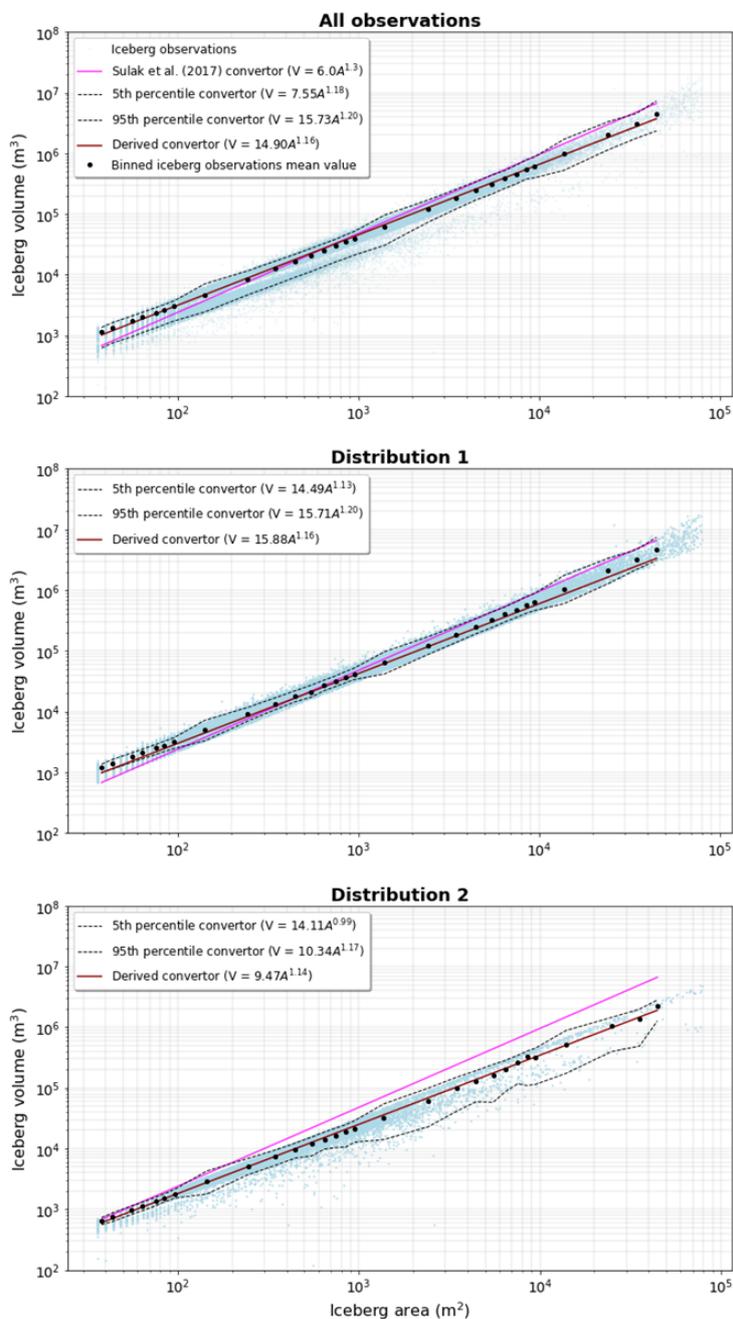



Figure 7. The mean iceberg area and volume for each size class (e.g. mean of 30 – 40 m², mean of 100 – 200 m², mean of 1,000 – 2,000 m², and so on) for the respective glacier, overlaid with the Sulak et al. (2017) conversion (pink) and the one derived here (brown). Bias in the calculation of each relationship towards more frequently occurring small icebergs through obtaining mean iceberg sizes
395 **and volumes in log₁₀(X+1) increments (black dots). Uncertainty in these distributions is also characterised by deriving similar relationships for the 5% and 95% limits.**

When observing how iceberg area scales with volume, two distinct distributions are identified at SKJI and KNS (Figure 8; Table 2). These two distributions are observed to persist between DEMs throughout the study period. While the lower distribution at both SKJI and KNS accounts for only 7.2% of the icebergs in the population, the divergence between the upper and lower distributions is found to proportionately increase with iceberg area (Figure 8 insets). These two distributions can be described in a similar manner to the overall distribution, with the equations for the upper and lower distributions shown in Equation 5 and Equation 6 respectively (Figure 7b, c).

$$V = 15.88A^{1.16} \tag{5}$$

$$V = 9.47A^{1.14} \tag{6}$$

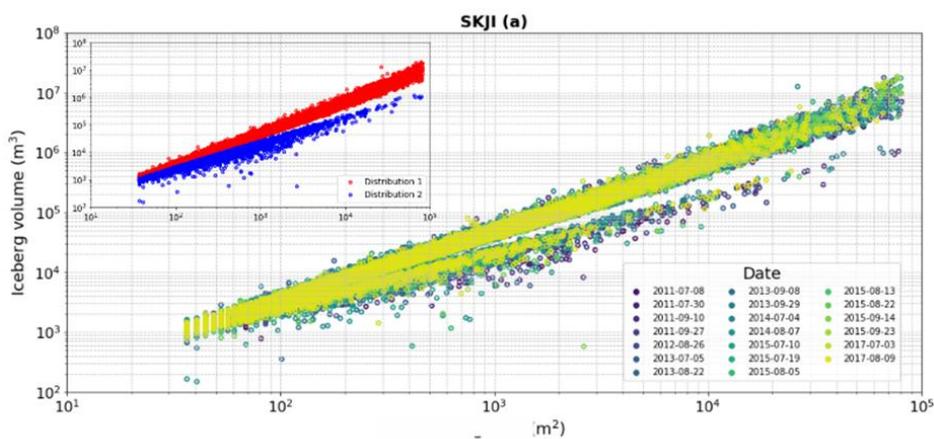
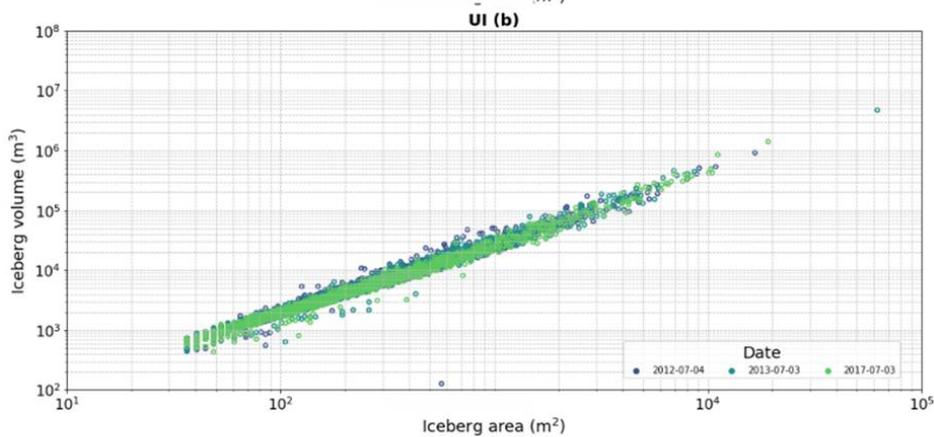
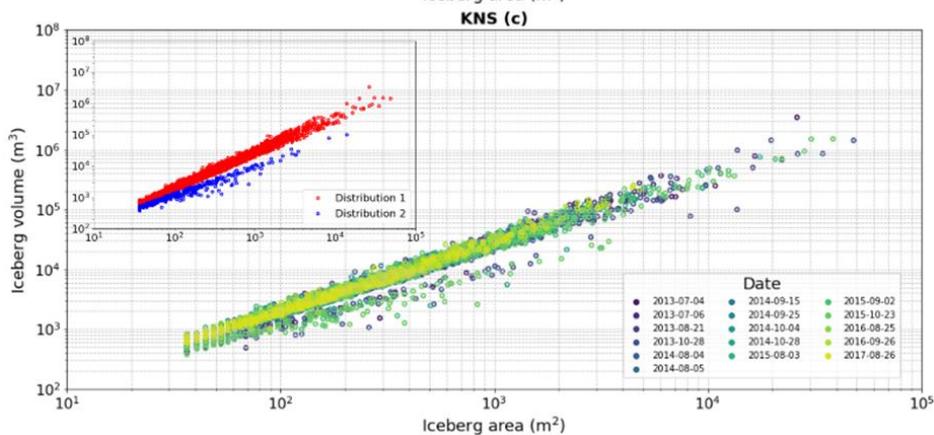



Figure 8. Iceberg area versus volume and freeboard height which is colour coded by each ArcticDEM scene date. The inset in both the SKJI (a) and KNS (c) panels show the two distributional branches identified in the data with the respective derived convertors for each distribution. Each distribution is separated by identifying local minima in probability distribution histograms of the entire dataset.

Table 2. Summary statistics of the two distributions outlined at SKI and KNS. (1) represents distribution 1 and (2) represents
455 **distribution 2. Standard deviation is abbreviated to SD.**

	SKJI (1)	SKJI (2)	KNS (1)	KNS (2)
Number of icebergs	136,673	11,041	8,697	354
Mean area (m^2)	1,036	434	383	307
Mean volume (km^3)	6.9×10^{-5}	8.0×10^{-5}	1.1×10^{-5}	2.7×10^{-6}
Maximum area (m^2)	79,820	77,192	47,520	13,536
Maximum volume (km^3)	0.017	0.001	0.003	9.5^{-5}
Area SD (m^2)	3,929	2,169	1,347	935
Volume SD (km^3)	4.0×10^{-4}	3.3×10^{-5}	5.8×10^{-5}	7.2×10^{-6}

In the entire dataset, small icebergs (area = <1,000 m^2; Rezvanbehbahani et al. [2020]) account for over 80% of the total
iceberg count for each glacier, however they only contribute to 10%, 37% and 35% of the total volume at SKJI, UI and KNS,
respectively (Figure 9). Consequently, while small icebergs dominate the distributions in the fjord of each glacier, compared
460 to larger icebergs they are found to account for a significantly smaller proportion of total iceberg volume.

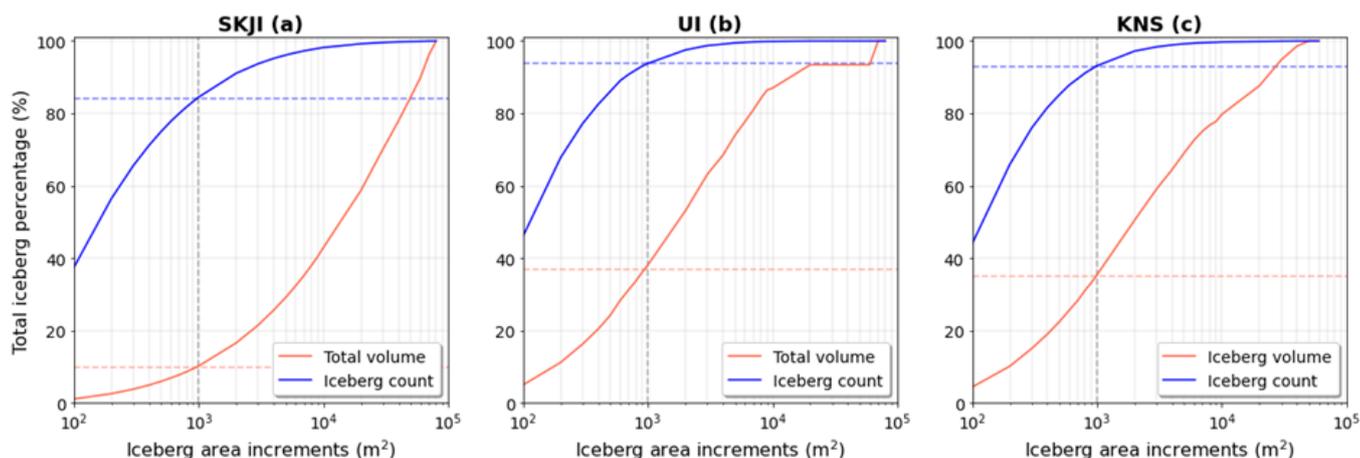



Figure 9. Cumulative iceberg volume (orange line) and count (blue line) plotted as percentage with their respective surface area.
The small iceberg threshold (1,000 m^2) is defined by a dashed grey line. The total volume made up by small icebergs is represented
by the orange dashed line and the total number is the dashed blue line. 84% of all the icebergs at SJKI are small, however they only
account for 10% of the total volume, whereas they account for 94% of the total count at UI and 37% total volume. 93% of the
465 **icebergs are small at KNS and account for 35% of the total volume.**

5.0 Discussion

5.1 Workflow

The automated workflow presented here is shown to successfully delineate icebergs and capture their area and volume size distributions and assign a range of metadata to each individual iceberg (Figure 6). The workflow therefore will allow users to

470   rapidly obtain iceberg data to interrogate glacier calving styles and iceberg freshwater fluxes. The application of the workflow to glacier fjords with a range of different iceberg densities and sizes demonstrates the utility of ArcticDEM data for iceberg detection and mapping across a range of different fjord environments typical of Greenland and elsewhere. As a result, this approach is suitable for pan-Arctic iceberg detection where availability of DEM data allow.

475   This new method executes quickly and is capable of successfully filtering ArcticDEM scenes by cloud contamination, ROI data coverage, and dynamically define sea level for each ArcticDEM scene to account for potentially poor image registration and local tidal state. While this results in the rejection of scenes with data gaps and partial cloud contamination that may have regions that are suitable for analysis, the automated image filtering steps implemented in the workflow removes the requirement for time consuming user-led data cleaning.

480

The detection thresholds defined (1.5 m for KNS and UI and 3.0 m for SKJI) are found to be suitable for correctly delineating iceberg outlines and subsequent size distributions (Figure 4). Though a mismatch in size distributions are found at KNS where small icebergs dominate, it is likely that this arises from operator bias in the manual delineation of these. This arises due to the operator being able to delineate across DEM pixels, whereas the workflow analyses each entire pixel individually. In this

485   instance, the workflow therefore provides a more complete footprint of small icebergs than a manual digitiser is able. Visual comparison of iceberg outlines produced by the workflow to multi-angle hillshaded DEMs (Figure 4) provide confidence that it is able to detect icebergs as small as 40 m^2 (10 pixels) is achievable.

Exploration of the workflow's sensitivity to increasing the detection threshold above sea level results shows that increasingly

490   higher thresholds detect only larger icebergs, though will result in smaller overall iceberg areas and volumes (Figure 5). The definition of this detection threshold value is dependent on whether smaller icebergs are important to include for the user's research question. Where only the largest icebergs are of interest, a higher detection threshold could be set with relatively little loss in the final iceberg areas and volumes. This is because volumetrically larger icebergs are more likely to have higher freeboard heights, and the iceberg margins omitted due to higher thresholds are likely to be small in terms of their relative area

495   and volume.

As a consequence, those wishing to explore power law size distribution relationships where small icebergs would be excluded from analysis could potentially set a higher detection threshold. Conversely, if a study is wanting to attain the maximum

number of icebergs, lower thresholds could be defined, though this risks that outlines of neighbouring icebergs being
500   erroneously identified as a single iceberg. This is highlighted by the fact that rafts of small individual icebergs frozen together
by mélange are correctly identified by the workflow as single floating bodies of ice, though the individual icebergs that they
are comprised of are not separated out.

While it should be emphasised that all results from the workflow are likely to represent minimum area and volume estimates,
505   it is suggested that for the majority of cases a threshold of 1.5 m should be sufficient.

5.2 Glaciological implications

The new area-to-volume conversions presented offer the potential for wide-scale application to iceberg area outline data that
have been derived from optical and/or SAR imagery (Equations 2-6; Figure 7). The large dataset generated also allows for the
quantification of uncertainties when scaling area to volume (Equations for upper and lower bounds; Figure 7). This will allow
510   iceberg volumes to be estimated from data sources that extend beyond the spatial and temporal availability of ArcticDEM and
that are more frequently acquired (e.g. Landsat satellites and Sentinel 1 and 2). Improved constraint of uncertainties in iceberg
volumes therefore provide new opportunities for temporally and spatially extending studies that seek to model fjord freshwater
fluxes (Davison et al., 2020a) and quantify iceberg volume distributions (e.g. Schild et al., 2021). While it should be
remembered that the conversion equations result in minimum volume estimates, inclusion of lower and upper limits will assist
515   in better quantification of ranges of potential iceberg volume from outlines alone.

We find evidence of two iceberg populations at SKJI and KNS across multiple ArcticDEM scenes between 2010 and 2017,
though only a single population at UI (Figure 8). The DEM surface expression of icebergs identified in the second distribution
tend to be flatter than those of distribution 1, resulting in lower overall volumes. Manual inspection of DEMs suggest the
520   majority of those in distribution 2 represent rafts of small icebergs that are frozen together by mélange. Though distribution 1
dominates the total dataset, studies using two-dimensional data (i.e. optical and/or SAR) should be aware that their methods
may identify these iceberg rafts as single icebergs. For glaciers where these two iceberg distributions exist using a single area-
to-volume conversion will therefore result in an overestimation of total iceberg volume.

525   Consequently, those using two-dimensional image data may be required to undertake further analysis to identify which
distribution individual icebergs are likely to fall within. Such techniques could include approaches that go beyond pixel level
analysis; for example, incorporating image texture as part of machine learning methods (e.g. Rezvanbehbahani et al. 2020).
Results show that small icebergs (area = <1,000 m^2) account for the majority of those identified (over 80% for each glacier)
yet contribute a smaller fraction of the total iceberg volume (10-37% of total volume; Figure 9). Consequently, small
530   differences in the number of large icebergs can have a disproportionate impact on overall fjord iceberg volume. At these

glaciers, large icebergs therefore represent comparatively larger freshwater reservoirs in their fjords, and account for a more significant proportion of overall ice mass loss from their source glaciers.

Expressing iceberg counts and volumes for each glacier as percentages (Figure 9) also offers the potential for empirically
535   estimating the evolution of iceberg populations for individual ice sheet outlets from frequently updated velocity derived glacier discharge data (e.g. Mankoff et al., 2019). Although this would assume a consistent calving style through time, such relationships could assist in estimating how the number and volumes of icebergs have evolved; may evolve in the future (through application to ice discharges from ice dynamic modelling [e.g. Choi et al., 2021]); and assessment of potential iceberg hazards.

540   **6.0 Conclusions**

This study presents a new workflow and GUI to automatically detect icebergs within Google Earth Engine using ArcticDEM. Results of this show excellent agreement with manually digitised iceberg outlines (r-values = 0.70, 0.92, 0.96), with mismatches occurring for the smallest of icebergs where the precision of manual digitisation is poorer compared to that of the workflow. Workflow results show evidence for two distinct iceberg populations at SKJI and KNS and one at UI representing:
545   (1) individual icebergs; and (2) small iceberg rafts frozen together by mélange. New area-to-volume conversion equations are presented for each distribution which also provide the means to determine upper and lower bound uncertainties. While smaller icebergs at each glacier are found to dominate the distributions (84-94% of the total count), their contribution to total volume and therefore freshwater flux are relatively small (10-37%).

550   Although ArcticDEM data are temporally and spatially limited relative to those obtained by optical and SAR satellite platforms, the results presented here offer the potential for extending studies into fjord iceberg cover and glacier calving that use iceberg outlines derived from these data. A new approach of expressing relationships between iceberg count and volume will also allow empirical estimation of iceberg size distributions from iceberg discharge observations. This would have benefits to those investigating iceberg freshwater fluxes within fjords, and who seek to model the evolution of mass loss from the GrIS.
555

The workflow and user-interface presented here allows users to generate their own large, reliable datasets for their glacier(s) of interest. Consequently, it opens the possibility of extending the results presented here to any location where suitable ArcticDEM data are available.

560

Appendix A: Automated sea level identification

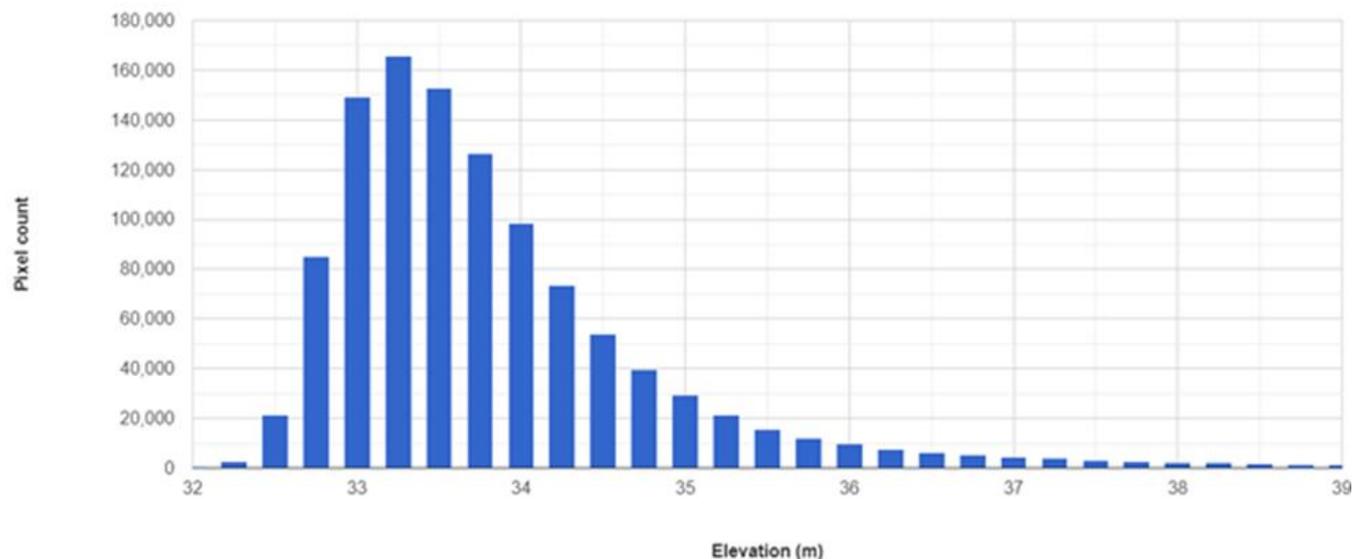



Figure A1. Example of an automated histogram calculated within GEE of elevation pixel count in an ArcticDEM image at
565   KNS (2013-07-04). The elevation with the highest pixel count is automatically selected as the sea level for that scene. In this
example sea level would be 33.25 m.

7.0 Code Availability

The user-interface guide to generate an iceberg dataset is available on the GitHub site
(https://github.com/ConnorShiggins/Google-Earth-Engine-and-icebergs) and a direct link to the tool can be accessed through
570   the Google Earth Engine Code Editor (https://code.earthengine.google.com/f38494aa9f268cf807cdefe8c107783e).

8.0 Acknowledgments

C.J.S. PhD studentship is funded by the University of Liverpool. J.M.L. is supported by a UKRI Future Leaders Fellowship
(Grant No. MR/S017232/1). We thank Dominik Fahrner for providing terminus positions for Umiammakku Isbræ. We are
grateful to the Polar Geospatial Center for the development and accessibility of ArcticDEM.

575   ## 9.0 Author Contributions

C.J.S. and J.M.L conceived the study. C.J.S. produced the dataset, conducted all the analysis and led the manuscript writing.
J.M.L and S.B. offered conceptual and coding advice and contributed to the writing of the manuscript.

10.0 Competing interests

The authors declare no conflict of interest.

580 ## 11.0 References

Alstott, J., Bullmore, E. and Plenz, D.: powerlaw: a Python package for analysis of heavy-tailed distributions, PloS one, 9, e85777, https://doi.org/10.1371/journal.pone.0095816, 2014.

Amundson, J.M., Fahnestock, M., Truffer, M., Brown, J., Lüthi, M.P. and Motyka, R.J.: Ice mélange dynamics and implications for terminus stability, Jakobshavn Isbræ, Greenland, J. Geophys. Res. Earth Surf, 115,
585 https://doi.org/10.1029/2009JF001405, 2010.

Åström, J., Cook, S., Enderlin, E.M., Sutherland, D.A., Mazur, A. and Glasser, N.: Fragmentation theory reveals processes controlling iceberg size distributions, J. Glaciol, 67, 1–10. https://doi.org/10.1017/jog.2021.14, 2021.

Bartholomaus, T.C., Stearns, L.A., Sutherland, D.A., Shroyer, E.L., Nash, J.D., Walker, R.T., Catania, G., Felikson, D., Carroll, D., Fried, M.J. and Noël, B.P.: Contrasts in the response of adjacent fjords and glaciers to ice-sheet surface melt in
590 West Greenland, Ann. Glaciol, 57, 25-38, https://doi.org/10.1017/aog.2016.19, 2016.

Bigg, G.R., Icebergs: their science and links to global change. Cambridge University Press. 2015.

Bigg, G.R., Wei, H.L., Wilton, D.J., Zhao, Y., Billings, S.A., Hanna, E. and Kadirkamanathan, V.: A century of variation in the dependence of Greenland iceberg calving on ice sheet surface mass balance and regional climate change, Proc. R. Soc. A: Math. Phys. Eng. Sci, 470, https://doi.org/10.1098/rspa.2013.0662, 2014.

595 Bunce, C., Nienow, P., Sole, A., Cowton, T. and Davison, B.: Influence of glacier runoff and near-terminus subglacial hydrology on frontal ablation at a large Greenlandic tidewater glacier, J. Glaciol, 67, 343-352, https://doi.org/10.1017/jog.2020.109, 2021.

Carroll, D., Sutherland, D.A., Hudson, B., Moon, T., Catania, G.A., Shroyer, E.L., Nash, J.D., Bartholomaus, T.C., Felikson, D., Stearns, L.A. and Noel, B.P.: The impact of glacier geometry on meltwater plume structure and submarine melt in
600 Greenland fjords, Geophys. Res. Lett, 43, 9739-9748, https://doi.org/10.1002/2016GL070170, 2016.

Cassotto, R.K., Burton, J.C., Amundson, J.M., Fahnestock, M.A. and Truffer, M.: Granular decoherence precedes ice mélange failure and glacier calving at Jakobshavn Isbræ, Nat. Geosci, 14, 417-422, https://doi.org/10.1038/s41561-021-00754-9, 2021.

Choi, Y., Morlighem, M., Rignot, E. and Wood, M.: Ice dynamics will remain a primary driver of Greenland ice sheet mass loss over the next century, Commun. Earth. Environ, 2, 1-9. https://doi.org/10.1038/s43247-021-00092-z, 2021.

605 Cook, S.J., Christoffersen, P., Truffer, M., Chudley, T.R. and Abellan, A.: Calving of a large Greenlandic tidewater glacier has complex links to meltwater plumes and mélange. Geophys. Res. Earth Surf, 126, https://doi.org/10.1029/2020JF006051, 2021.

Crawford, A., Crocker, G., Mueller, D., Desjardins, L., Saper, R. and Carrieres, T.: The Canadian ice island drift, deterioration and detection (CI2D3) database, J. Glaciol, 64, 517-521, https://doi.org/10.1017/jog.2018.36, 2018.

610   Davison, B.J., Cowton, T.R., Cottier, F.R. and Sole, A.J.: Iceberg melting substantially modifies oceanic heat flux towards a major Greenlandic tidewater glacier, Nat. Commun, 11, 1-13, https://doi.org/10.1038/s41467-020-19805-7, 2020a.

Davison, B.J., Sole, A.J., Cowton, T.R., Lea, J.M., Slater, D.A., Fahrner, D. and Nienow, P.W.: Subglacial drainage evolution modulates seasonal ice flow variability of three tidewater glaciers in southwest Greenland. J. Geophys. Res. Earth Surf, 125, https://doi.org/10.1029/2019JF005492, 2020b.

615   Eik, K. and Gudmestad, O.T.: Iceberg management and impact on design of offshore structures, Cold Reg Sci Technol, 63, 15-28, https://doi.org/10.1016/j.coldregions.2010.04.008, 2010.

Enderlin, E.M., Hamilton, G.S., Straneo, F. and Sutherland, D.A.: Iceberg meltwater fluxes dominate the freshwater budget in Greenland's iceberg-congested glacial fjords, Geophys. Res. Lett, 43, 11-287, https://doi.org/10.1002/2016GL070718, 2016

Fahrner, D., Lea, J.M., Brough, S., Mair, D.W. and Abermann, J.: Linear response of the Greenland ice sheet's tidewater glacier
620   terminus positions to climate, J. Glaciol, 67, 193-203, https://doi.org/10.1017/jog.2021.13, 2021.

Fried, M.J., Catania, G.A., Stearns, L.A., Sutherland, D.A., Bartholomaus, T.C., Shroyer, E. and Nash, J.: Reconciling drivers of seasonal terminus advance and retreat at 13 Central West Greenland tidewater glaciers, J. Geophys. Res. Earth Surf, 123, 1590-1607, https://doi.org/10.1029/2018JF004628, 2018.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R.: Google Earth Engine: Planetary-scale
625   geospatial analysis for everyone, Remote Sens. Environ, 202, 18-27, https://doi.org/10.1016/j.rse.2017.06.031, 2017.

Khazendar, A., Fenty, I.G., Carroll, D., Gardner, A., Lee, C.M., Fukumori, I., Wang, O., Zhang, H., Seroussi, H., Moller, D. and Noël, B.P:. Interruption of two decades of Jakobshavn Isbrae acceleration and thinning as regional ocean cools, Nat. Geosci, 12, 277-283, https://doi.org/10.1038/s41561-019-0329-3, 2019.

Kirkham, J.D., Rosser, N.J., Wainwright, J., Jones, E.C.V., Dunning, S.A., Lane, V.S., Hawthorn, D.E., Strzelecki, M.C. and
630   Szczuciński, W.: Drift-dependent changes in iceberg size-frequency distributions, Sci Rep, 7, 1-10, https://doi.org/10.1038/s41598-017-14863-2, 2017.

Laidre, K.L. and Stirling, I:. Grounded icebergs as maternity denning habitat for polar bears (Ursus maritimus) in North and Northeast Greenland, Polar Biol, 43, 937-943, https://doi.org/10.1007/s00300-020-02695-2, 2020.

Lea, J.M., Mair, D.W., Nick, F.M., Rea, B.R., Weidick, A., Kjaer, K.H., Morlighem, M., Van As, D. and Schofield, J.E.:
635   Terminus-driven retreat of a major southwest Greenland tidewater glacier during the early 19th century: insights from glacier reconstructions and numerical modelling, J. Glaciol, 60, 333-344. https://doi.org/10.3189/2014JoG13J163, 2014a.

Lea, J.M., Mair, D.W.F., Nick, F.M., Rea, B.R., Van As, D., Morlighem, M., Nienow, P.W. and Weidick, A:. Fluctuations of a Greenlandic tidewater glacier driven by changes in atmospheric forcing: observations and modelling of Kangiata Nunaata Sermia, 1859–present, The Cryosphere, 8, 2031-2045. https://doi.org/10.5194/tc-8-2031-2014, 2014b.

640   Mankoff, K.D., Colgan, W., Solgaard, A., Karlsson, N.B., Ahlstrøm, A.P., Van As, D., Box, J.E., Khan, S.A., Kjeldsen, K.K., Mouginot, J. and Fausto, R.S:. Greenland Ice Sheet solid ice discharge from 1986 through 2017, Earth Syst. Sci. Data, 11, 769-786, https://doi.org/10.5194/essd-12-1367-2020, 2019.

Moon, T., Sutherland, D.A., Carroll, D., Felikson, D., Kehrl, L. and Straneo, F.: Subsurface iceberg melt key to Greenland fjord freshwater budget, Nat. Geosci, 11, 49-54, https://doi.org/10.1038/s41561-017-0018-z, 2018.

645 Morlighem, M., C. Williams, E. Rignot, L. An, J. E. Arndt, J. Bamber, G. Catania, N. Chauché, J. A. Dowdeswell, B. Dorschel, I. Fenty, K. Hogan, I. Howat, A. Hubbard, M. Jakobsson, T. M. Jordan, K. K. Kjeldsen, R. Millan, L. Mayer, J. Mouginot, B. Noël, C. O'Cofaigh, S. J. Palmer, S. Rysgaard, H. Seroussi, M. J. Siegert, P. Slabon, F. Straneo, M. R. van den Broeke, W. Weinrebe, M. Wood, and K. Zinglersen.: BedMachine v3: Complete bed topography and ocean bathymetry mapping of Greenland from multi-beam echo sounding combined with mass conservation, Geophys. Res. Lett, 44, 11-051,
650 https://doi.org/10.1002/2017GL074954, 2017.

Mouginot, J., Rignot, E., Bjørk, A.A., Van den Broeke, M., Millan, R., Morlighem, M., Noël, B., Scheuchl, B. and Wood, M.: Forty-six years of Greenland Ice Sheet mass balance from 1972 to 2018, PNAS, 116, 9239-9244, https://doi.org/10.1073/pnas.1904242116, 2019.

Moyer, A.N., Sutherland, D.A., Nienow, P.W. and Sole, A.J.: Seasonal variations in iceberg freshwater flux in Sermilik Fjord,
655 southeast Greenland from Sentinel-2 imagery, Geophys. Res. Lett, 46, 8903-8912, https://doi.org/10.1029/2019GL082309, 2019.

Porter, Claire; Morin, Paul; Howat, Ian; Noh, Myoung-Jon; Bates, Brian; Peterman, Kenneth; Keesey, Scott; Schlenk, Matthew; Gardiner, Judith; Tomko, Karen; Willis, Michael; Kelleher, Cole; Cloutier, Michael; Husby, Eric; Foga, Steven; Nakamura, Hitomi; Platson, Melisa; Wethington, Michael, Jr.; Williamson, Cathleen; Bauer, Gregory; Enos, Jeremy; Arnold,
660 Galen; Kramer, William; Becker, Peter; Doshi, Abhijit; D'Souza, Cristelle; Cummens, Pat; Laurier, Fabien; Bojesen, Mikkel. https://doi.org/10.7910/DVN/OHHUKH, Harvard Dataverse, V1, 2018.

Rezvanbehbahani, S., Stearns, L.A., Keramati, R., Shankar, S. and van der Veen, C.J.: Significant contribution of small icebergs to the freshwater budget in Greenland fjords, Commun. Earth. Environ, 1, 1-7. https://doi.org/10.1038/s43247-020-00032-3, 2020.

665 Scheick, J., Enderlin, E.M. and Hamilton, G.: Semi-automated open water iceberg detection from Landsat applied to Disko Bay, West Greenland. J. Glaciol, 65, 468-480, https://doi.org/10.1017/jog.2019.23, 2019.

Schild, K.M., Sutherland, D.A., Elosegui, P. and Duncan, D.: Measurements of iceberg melt rates using high-resolution GPS and iceberg surface scans, Geophys. Res. Lett, 48, https://doi.org/10.1029/2020GL089765, 2021.

Shugar, D.H., Burr, A., Haritashya, U.K., Kargel, J.S., Watson, C.S., Kennedy, M.C., Bevington, A.R., Betts, R.A., Harrison,
670 S. and Strattman, K.: Rapid worldwide growth of glacial lakes since 1990, Nat. Clim. Change, 10, 939-945, https://doi.org/10.1038/s41558-020-0855-4, 2020.

Soldal, I.H., Dierking, W., Korosov, A. and Marino, A.: Automatic detection of small icebergs in fast ice using satellite wide-swath SAR images, Remote Sens, 11, https://doi.org/10.3390/rs11070806, 2019.

Sulak, D.J., Sutherland, D.A., Enderlin, E.M., Stearns, L.A. and Hamilton, G.S.: Iceberg properties and distributions in three
675 Greenlandic fjords using satellite imagery, Ann. Glaciol, 58, 92-106. https://doi.org/10.1017/aog.2017.5, 2017.

Walter, F., Amundson, J.M., O'Neel, S., Truffer, M., Fahnestock, M. and Fricker, H.A.: Analysis of low-frequency seismic signals generated during a multiple-iceberg calving event at Jakobshavn Isbræ, Greenland, J. Geophys. Res. Earth Surf, 117, https://doi.org/10.1029/2011JF002132, 2012.

680

685

690