Mengzhen Qi et al.'s

A 14-yr Circum-Antarctic Iceberg Calving Dataset Derived from Continuous Satellite Observations

Reviewed by Chad A. Greene of NASA Jet Propulsion Laboratory
February 15, 2021

In this paper, Mengzhen Qi, Yan Liu, and their collaborators present a new dataset that pinpoints the year, location, and magnitude of nearly two thousand Antarctic iceberg calving events detected since 2005. This dataset is a first of its kind, and it is very impressive. I have now spent a couple of days exploring the dataset, and I can attest to the quality of the work and the highly useful nature of the data.

The utility of this data dataset is exemplified by the fact that I am only writing this review because I stumbled upon the present paper in my search for shapefiles that could offer some insights about where calving has occurred each year in recent history. This dataset was what I needed, and it has already helped me in my research. The methodology of the data generation is well conceived, and it is communicated in this paper with clarity and grace. This is exactly the kind of data that will benefit the community, and I will be happy to see the manuscript published in ESSD.

I have made a number of suggestions below, and at times I go into some depth when describing my line of thought. Please don't let the lengthy nature of my review be confused with any sign of irreparable shortcomings. Nor should the authors feel obligated to respond to my every fleeting thought. Rather, my intent is to spark ideas and find ways in which this very good paper could meet my own personal vision of a great paper. I reiterate, this dataset has already proved invaluable in my own research, and that fact alone is strong indication that this paper is well worth publishing.

General Concerns

Underreported values of calving flux: The 771.1 Gt/yr value of calving flux presented here is significantly lower than the 1265 Gt/yr value reported by Rignot et al., 2013, the 1321 Gt/yr reported by Depoorter et al. 2013, or the 1026 Gt/yr "steady state" calving flux reported by Liu et al., 2015. The abstract states that the *total* pan-Antarctic calving flux is 771.1 Gt/yr, when in fact this number merely represents the total amount of calving that was *measured*.

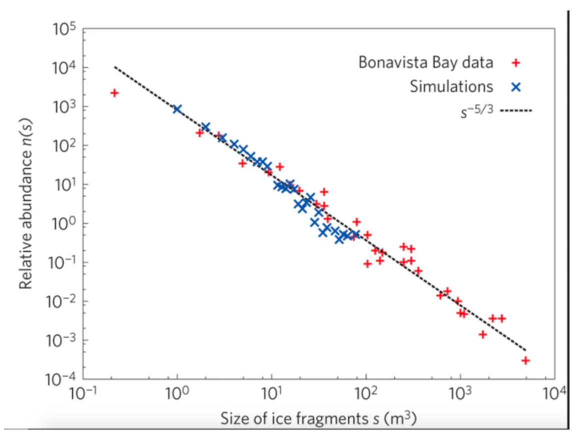
I suspect that some portion of the under-reporting is due to the 1 km^2 area threshold that must be met for an iceberg to be counted. The value of 771.1 Gt/yr is not far off from Liu's 755 Gt/yr, which is not surprising, as the present dataset appears to be a continuation of the same work, or it at least follows a similar method. Liu et al. also did not include icebergs smaller than 1 km^2

in their count, although they did not speculate about how much ice they might be missing from their count as a result.

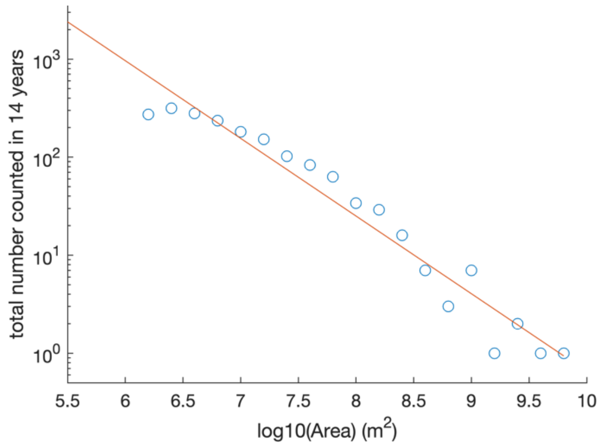
The wording of the present manuscript suggests that the entire Antarctic coastline is captured annually in this study, but I could not find any explicit statements about whether small ice shelves or marine-terminating glaciers without ice shelves were included. From the wording of the manuscript alone, I would assume that they are included, except that the final tally closely matches Liu's value of 755 Gt/yr, and the description of the method appears very similar to the method employed by Liu, who did not include marine-terminating glaciers or any ice shelves smaller than 10 km^2. This is an important point because calving estimates with a static flux gate suggest at least 100 or maybe 200 Gt/yr come from marine-terminating glaciers and small ice shelves.

Regarding the 1 km^2 threshold that must be met for icebergs to be included in estimates of calving flux, the authors point out that although small icebergs are high in number, their contribution to the total calving flux can be neglected because the icebergs are so small. This argument needs evidence.

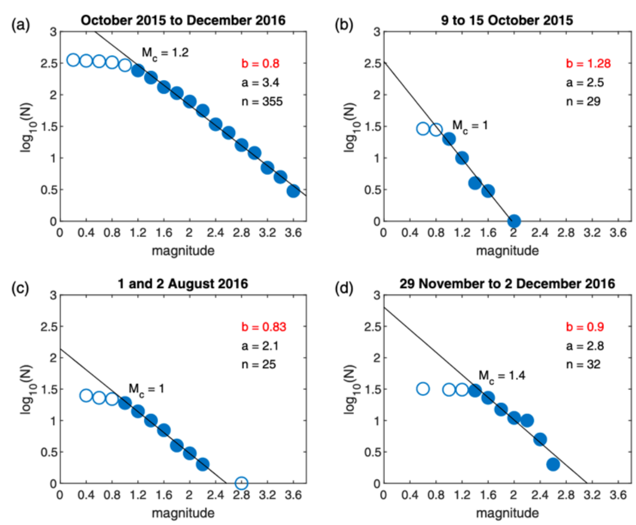
What we know about the relationship between iceberg size and relative abundance is that they generally follow a Pareto distribution. This is also described as the Gutenberg-Richter relationship. Here, for example, is one figure from the Åström et al. 2014 paper that's cited in the present manuscript:



In the figure above, we see that smaller icebergs are more common larger icebergs. It's a very simple relationship, but this log-log form is convenient because it makes it possible to make predictions about icebergs that have not been directly observed. (By the same principle, for example, hydrologists may use just a few decades of data to identify a "hundred-year" flood event.) Here's the Pareto distribution of iceberg abundance as a function of area among the 1786 calving events reported in the present dataset:

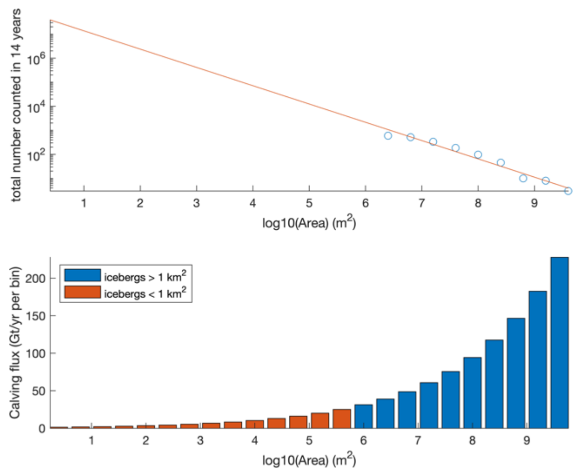


In the plot above, you'll notice that the linear relationship falls off on the left hand side. From this, one might conclude that the Gutenbrg-Richter relationship has failed to describe nature, and that very small icebergs are somehow less abundant than the slightly larger ones. I suspect, however, that nature is actually behaving as expected, but that the downturn on the left hand side of data indicates a slight under-counting of the smaller icebergs. We see a similar behavior in this earthquake data by [Leva et al., 2021](), where the smallest earthquakes are naturally high in number, but are difficult to detect due to their low amplitudes:



My main point here is that an absence of small icebergs in the observations should not suggest that they do not exist or that they can be wholly neglected from continent-wide tallies of calving.

In addition to adding some essential caveats to the 771.1 Gt/yr value presented in the abstract, I suggest including a few sentences in the Discussion section about what *isn't* captured in this dataset. Such a discussion might even exploit the Gutenberg-Richter relationship. For example, below I've assumed a Pareto distribution for the 1786 calving events in this dataset, and by fitting a line in log-log space I'm able to extrapolate this relationship down to a hypothetical iceberg size of 1 m^2. This gives us some insight about how many 1 m^2 icebergs are likely to calve each year, and how many 10 m^2 icebergs, and how many 100 m^2, and so on. By knowing the number of icebergs that likely exist, but are too small to be included in this dataset, we can start to build some intuition for how much calving flux might be missing from the dataset.



The predicted calving flux in each bin of the histogram above was calculated as

$$A*rho*H*N_{pred}*10^{-12}/14,$$

where A is the mean area of the bin, rho is the density of ice, H is the mean ice thickness of 223 m in the dataset, N_{pred} is the number of predicted icebergs in each bin of area, 10^{-12} converts from volume to Gt, and dividing by 14 gives the average value per year, as the dataset contains 14 years of data.

By summing up all of the calving flux of the hypothetical icebergs smaller than 1 km^2, I find that their total contribution is likely on the order of 100 Gt/yr. I will warn that the linear fit and the resulting estimate of missing calving flux are sensitive to the chosen bin size, but it's at least a first-order attempt to try and understand what is and isn't included in this dataset. This is the type of consideration I would like to see in the Discussion section.

Is Antarctica growing in extent? One possible way to reconcile the 771.1 Gt/yr value reported here with the >1000 Gt/yr calving flux reported in previous studies—and that is if the ice sheet is growing in extent. The static flux gates used in previous studies may indeed measure a flow of >1000 Gt/yr passing by, but if the ice does not calve after passing through the flux gate, then the ice sheet grows, and the methods presented here would capture a lower and more accurate measurement of true calving behavior. Such was suggested by Liu et al., 2015 to reconcile their measurement of 755 Gt/yr with their own estimate of 1026 Gt/yr of "steady state" calving.

I will note that if you add ~150 Gt/yr of calving from marine-terminating glaciers, plus ~100 Gt/yr of icebergs smaller than 1 km^2, plus the 271 Gt/yr rate of ice shelf growth due to area extent reported by Liu et al, to the 771 Gt/yr reported here, the result is about 1300 Gt/yr, which falls between the flux-gate calculations of Rignot et al. and Depoorter et al. Is it a coincidence that this perfectly closes the mass budget and reconciles the difference between the two measurement techniques? It would be insightful to read the authors' take on this in the Discussion section.

If the ice sheet is in fact changing in extent, then it seems worth reporting an annual time series of ice sheet area that can be directly compared with the annual time series of calving area. Given that this work involved the development of annual coastlines that purportedly cover the entire continent, it should be trivial to plot a time series of the area enclosed by the coastline each year, and this would provide valuable context for understanding what is or isn't included in the calving flux data.

Section 3.3.3 is relatively innocuous, but it's unclear what value there is in this type of subjective binning. The thresholds were not set by any sort of natural clustering in the size-vs-frequency distribution, and no evidence is given that there is any meaningful glaciological distinction between behavior of "low-frequency" and "high-frequency" calving events. Accordingly, the parameters in Table 3 seem somewhat arbitrary, and I'm not sure what can be gained by classifying a calving event as either "low-frequency" or "high-frequency". I recommend either removing this section from the paper or expanding a little bit on why the bins are meaningful and how they might be interpreted in a glaciological or oceanographic context.

ENSO analysis seems out of place: The attribution of calving events to ENSO anomalies and surface melting looks very interesting, but the analysis feels out of place in this data paper, particularly as the topic is only introduced in the final closing remarks of the paper. If these scientific results are compelling as they appear to be, then they should be described in detail in a separate paper, where they can be discussed at length while being given a chance for proper peer review.

The Discussion section does very little discussing. In its present form, this section repeats several numbers from earlier in the manuscript, but there is no discussion of what the numbers mean or how we should interpret them. There's also no mention why the annual calving fluxes in this paper differ so greatly from those reported by Rignot et al., 2013 or Depoorter et al., 2013.

Given that this is a data paper, I suggest including a genuine discussion about what insights can be obtained directly from the data or what benefits are gained by any new methods presented here. To be clear, a list of characteristics of the data is not very insightful on its own. Simply stating that Antarctica calves 127.6 icebergs per year does little for readers without any discussion of why that number is important. A sentence that lists the years of elevated calving flux is somewhat meaningless without talking about why calving rates are high some years or whether interannual variability is dominated by a few large calving events or by broadband increases in calving. Likewise, if there is some sort of insight to be gained by knowing that the formal estimate of calved area uncertainty is 17.1 km^2, then by all means, *discuss* that here. Does the area uncertainty represent some limitation of the dataset? *Discuss* that here. Can the measurement uncertainty be used to offer readers any words of caution about how to interpret the data? *Discuss* that here. Can we use these 14 years of observations to understand calving processes that occur on multi-decadal timescales? *Discuss* that here. If there's anything else that feels important to understanding this data, then please discuss it here.

Data Issues

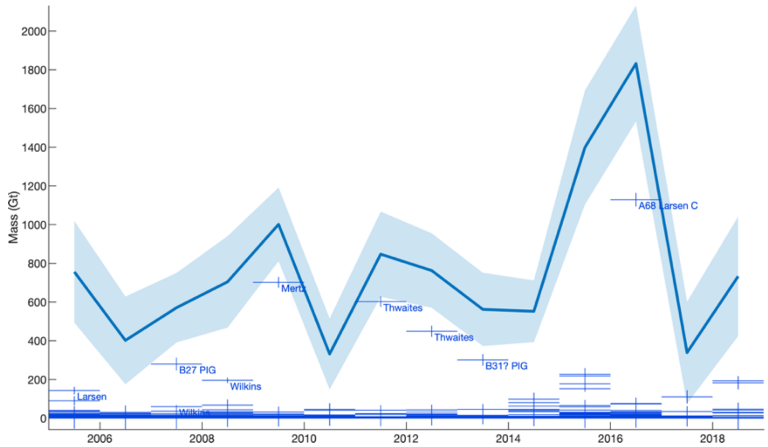
Missing coastline shapefiles: A handful of papers have been published recently, each claiming to have mapped the Antarctic coastline at annual resolution, but to my knowledge, no such data has been made publicly available by any group. Given the aggressive open-data policy of ESSD and my own personal interest in obtaining such a dataset, I feel compelled to ask about the whereabouts of the annual coastlines that were developed to generate this calving flux dataset. Will the annual coastlines be made available?

File format: The dataset is currently made available as a .rar compressed file. That's a somewhat uncommon format, at least in the United States, and it required me to download special software to decompress the file. Users may experience less of a barrier if instead the data are zipped up in an ordinary, open-format .zip file.

Polygon type: I experienced a very minor issue that I could not read the shapefile data in Matlab, because the polygons were saved as PolygonM or PolygonZ format rather than simple Polygon format. To get around this, I had to open each shapefile in QGIS and re-save as Polygon format before I could open them in Matlab. I work with a lot of shapefiles in Matlab, but this is the first time I've encountered this particular issue. I'm not sure if the inability to read PolygonM or PolygonZ format is specific to Matlab, but it may help more people use the data if it's saved as a plain Polygon format.

Named icebergs and known collapse events: A few well known calving events occurred during the study period, but they can't be directly queried in the shapefile data. It would be helpful if the attributes of the shapefiles contained iceberg names and approximate dates of major events such cases as the Mertz Glacier tongue calving event of 2010, or iceberg A68 at Larsen C in July 2017, or the successive collapse events at Wilkins Ice Shelf.

I suggest adding iceberg names and calving dates because the first thing I did upon reading the shapefiles into Matlab was plot them in the form of this time series seen below. Immediately I got curious about the handful of outliers that appear to dominate interannual variability on the continental scale. After a bit of googling, I think I've correctly linked some of the polygons in the shapefiles with known events, but I'm not 100% sure I've gotten them right. I suspect I won't be the only user of this data who wishes to link these polygons to known calving events.



Minor Comments

Line 28 is problematic as it currently reads, “In total, 1786 annual calving events occurred on the Antarctic ice shelves from August 2005 to August 2019.”

The wording of this sentence could easily be misinterpreted, because it says without qualification that the *total* number of calving events during the study period was just 1786. No doubt this is an underestimate, likely by an order of magnitude or more. For example, the previous paper by Qi et al. reports that 2032 calving events were detected just within the final four years of the present study’s period of investigation.

I suggest a very simple fix, which is to say something like “we detect a total of 1786 calving events larger than 1 km^2...”

Line 187-192: Calving area uncertainty is also influenced by velocity uncertainty, and should probably be accounted for here.

Table 4 took me a while to understand, mainly because I was thrown off by the unitless use of the word *frequency*. The usage of the word frequency is not incorrect, per se, but it is slightly more difficult to parse than simply discussing the number of events that were counted.

- I think the title should be something like “Number of calving events detected in MODIS and SAR..”
- Likewise, the “Calving frequency” row should be renamed “Number of calving events”.
- It’s not immediately clear what is meant by the row labeled “Calving area”. Does it mean “Total calved area”?
- The Scale column contains only text descriptors of Area categories that were binned based on somewhat arbitrary thresholds, and to interpret them in this table the reader is tasked with going back to Table 3 to figure out what these categories mean. I

recommend removing the subjective labels like “Medium-scale” and “Extra-large-scale” throughout the paper and replacing with the area range.

Section 5.1 contains several uses of the phrase *calving frequency* to mean the total number of calving events that were detected. The distinction may just be a minor wording preference of mine, but it feels significant as the *calving frequency* implies an intrinsic property of the system, whereas the *number of detected calving events* is unambiguous and directly describes what has been measured. For clarity, I recommend replacing *calving frequency* throughout the paper. This would also make room for another phrase used in this paper, *calving recurrence interval*, which might be easily confused with *calving frequency* if both phrases are used in the same manuscript.

Figure 5 is interesting, but for I would

- change the phrase “calving frequency” to “number of calving events”,
- change the phrase “calving area” to “total calved area”,
- remove the entire right-hand column, as the colored bins of area shown in the left hand plots are a natural proxy for frequency.

Section 5.3 dives into the recurrence interval data, but the concept of the recurrence interval hasn’t been adequately described. A fair attempt was made to give sort of a dictionary definition of the term on Line 215, but then by Line 216 the meaning gets lost in vague terms about calving that occurs in the same neighborhood, which depends on some threshold of distinguishing between in-neighborhood and out-of-neighborhood events, and by now I’ve lost any sense of what the recurrence interval can tell us.

After Section 3.3.3 in which the quantification method is described, the topic of recurrence intervals is comes up again in Section 5.3 without any conceptual bridge to help readers understand physically what is meant by this sentence that begins on Line 307:

Calving events with a recurrence interval of 3 had the highest frequency,...

Already I’m confused, as I have no physical intuition for what the recurrence interval tells us about glaciers, or if this is just a characteristic of the detection limits of the method. Without setting the stage for understanding what the recurrence interval actually tells us, readers are likely to left wondering how once-every-three-year calving events can have a higher frequency than annual calving events. After hitting this confluence of conceptual roadblocks so early in the sentence, it’s difficult then to understand the remainder of the sentence, which continues,

...accounting for 18.8% of the total and occurring 335 times in 14 years, followed by those with recurrence intervals of 5, 2 and 4, accounting for 18.5%, 15.2% and 14.3%, respectively.

Again, what do any of these numbers mean? My interpretation is that the methods presented in this paper may not fully capture the types of calving events that occur more often than every three years, because the high-frequency events are more likely to be smaller and go undetected. If that's the case, that's absolutely fine, but explore the concept further, so the folks who use this data will understand what it tells us and what the limitations are. If, on the other hand, there's something glaciologically meaningful here, then please help readers understand it.

Figure 6 is difficult to interpret. I know it's supposed to be communicating something, but I keep gazing into the figure, hoping it will reveal its secrets to me. After some inspection I see that area and mass well correlated, but that is to be expected. And it appears that high numbers of the small icebergs combine to represent a fair portion of the total detected mass, but still

- I'm not sure what the take-home message of this figure is.
- I think the units of the horizontal axis are supposed to be years, but there are no data in the 1-year bin, so I must wonder,
- Are there truly *no* places in Antarctica where calving occurs every year, or is this just a limitation of the detection method? Or am I misunderstanding the meaning of recurrence interval?
- I also suspect there's some intention behind the shading of the left and right side of the figure, but I can't figure out what it's trying to communicate.
- The task of interpreting this figure is not made any easier by the caption, which contains only a verbless sentence fragment. From the main body of this manuscript, it is clear that the authors are excellent writers, so by all means, use your talents to fill this space with vivid descriptions of what's happening in the figure! Help me understand what I should be noticing in the graphic, and help me understand why it's important.

Line 344 says "the Totten Ice Shelf was collapsing every year." The word "collapsing" may be a bit too strong, so consider changing to something like "We detect calving events at Totten Ice Shelf every year."

Lines 350 to 363 rehash numbers from earlier in the manuscript without reframing or adding any new perspectives. I think this paragraph can be deleted without any detriment to the paper.

Lines 363 to 373 convey a level of enthusiasm that should absolutely be harnessed to develop this analysis further. The early results look very interesting indeed, but I don't think the Discussion section of a data paper is the appropriate place to present new scientific findings about correlations between ENSO and iceberg calving. I recommend removing these two paragraphs and Figure 9 from the paper.