Reviewer 3
We thank the reviewer for their time and feedback on the manuscript. The valuable comments and questions are carefully considered, and we discuss below how we intent to incorporate their suggestions in the next version of the manuscript.
In short (also considering comments from the other reviewers), we will focus on (a) extending our temporal analysis to increase our sample size of observed lake drainages, (b) extend analysis of our activeness parameter by comparing to the vulnerability map of Lai et al. (2020) and to the strain rates (c) provide more details on methods, specifically the NeRD algorithm and the definition of thresholds and (d) will implement major textual changes to better clarify and align our conclusions to our observations.
For some (minor) edits we have already started implementing changes, and for some comments we can already provide some provisional additional figures in this document. We thank you for understanding this was not feasible yet for all comments.

General

The goal of the work is very straightforward and well defined. Basically, there are three main goals: 1) map the overlap of damage and ponds; 2) find the draining ponds and identify in which conditions they occur; and 3) identify triggering events.
It is an interesting study, very important in the study of Antarctica. However, it is hard to define how impactful the findings are, since the methodology lacks more description of the metrics. Furthermore, the analysis should be deeper (see Major Comments).
I think the manuscript has 4 main issues:
1. observations do not support the claims in the manuscript
2. lack of a more robust statistical analysis
3. small number of observations
4. more descriptive methodology and how are the thresholds defined

We understand and acknowledge these concerns. Generally, concern 1-3 stem from the small number of observations and how these were used and analysed. Our study is a conceptual study and we concede that some claims have been too strong. In the next version of the manuscript we'll provide more analysis, clarification and nuance to better align these aspects.
In the next version of the manuscript we will (a) extend the time series to increase the number of observations to support our analysis and claims, to 2015-2024 (adding 6 more years). However, even then the study will still remain relatively small (covering one ice shelf) and the main effort of our revisions will be (b) aimed to rephrase our conclusions to align the observations to the claims. We think, even with the small sample size, this study can add to the scientific discussion of when/where hydrofracturing is actually observed to occur, adding more nuance to a general state of vulnerability of all fractures to hydrofracturing (e.g. Lai et al (2020)). However, we do agree that this study does not provide enough observations to make strong (Antarctic) generalisations, and we will adjust the text accordingly, focusing more on qualitative arguments rather than quantitative.
Considering issue (4): we will put more effort in describing and clarifying the methods. The NeRD algorithm and the thresholds specifically, both for lake detection and damage detection, will be better explained, and a supplementary section will be included to address this. Our code will be made available:

For the next review process, Google Earth Engine and Python code available through the links below. Please beware those are prelimary versions. For the final submission the repositories will be revisted, cleaned and prepared with documentation.
GEE: https://code.earthengine.google.com/?accept_repo=users/juliussommer/HydrofractureShackleton
Python Github: https://github.com/js-chemE/HydrofractureShackleton_2023

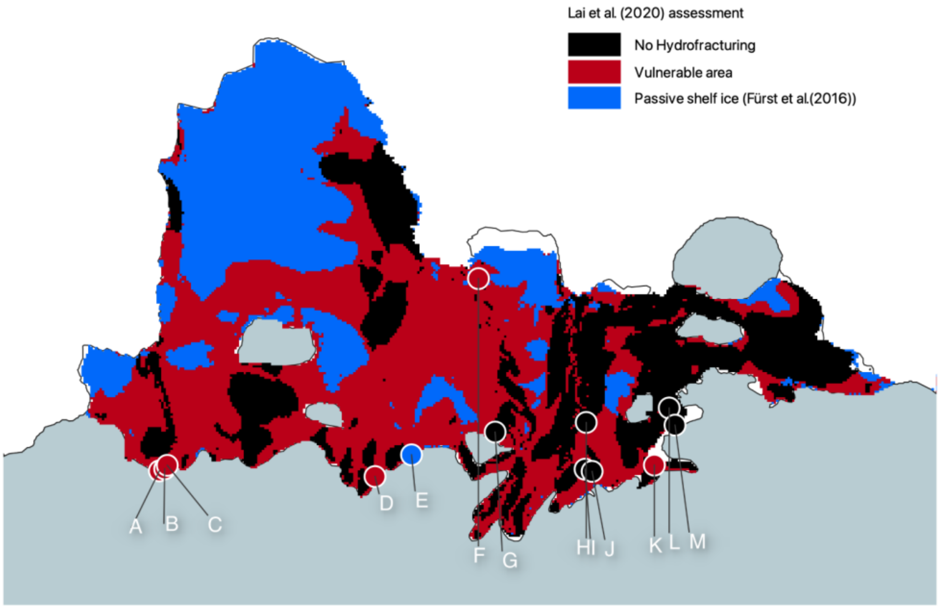
Major Comments

2.1 Sample size

The authors used a time span of 3 years for the analysis performed. They found 13 events, which is a small sample size. Why do you not cover a larger time span to have more data? In addition, covering other regions would be beneficial. Furthermore, the definition of threshold values seems quite arbitrary, indicating that they fit only to these present specific conditions.

Also, the authors highlight the need for more sophisticated statistical approaches. This is a major concern for me, because not only the dataset is small, but it also lacks meaningful statistical analysis.

We agree that the sample size is small. For the next version of the manuscript, we'll extent the time span of the study to cover 2015-2024 (governed by the availability of Sentinel-1 for consistent damage mapping). While this adds more datapoints, we don't expect an order of magnitude change in available observations of rapid lake drainages, and therefore we will also focus the text on a more qualitative analysis. The analysis will be extended by using a more comprehensive comparison to the vulnerability assessment of Lai et al. (2020), for which some preliminary data is included in this document below.
Considering the thresholds: we presume the reviewer is talking about the lake drainage threshold (80% of volume). This threshold has been chosen consistent with previous literature (see reviewer #1 for detailed discussion and references), and is therefore not chosen specifically for our conditions. Nonetheless, we will expand on this and include an uncertainty estimate to the number of drainages detected when varying to 75-90% of volume (similar as Miles et al., 2017).



2.2 Not supportive data

I think the data presented do not support the conclusions drawn. In fact, the authors repeatedly make an affirmation and draw it back after a few sentences, for example:
L139-140 = "However, all of the detected lake drainage events occur in areas of the ice shelf classified as medium to highly active"
And L151-L153 = "For example, drainages H, K, and M took place in areas with relatively low damage but high activeness, while drainages A, F, and G occurred in the least active regions, yet showed high levels of damage."
Furthermore, analysis of Figure B1 reveals that 1 of the events is in a low activeness area, undermining the claim in L139-140.

We concede that the wording of these claims are too strong, and inconsistent. To clarify, indeed the first sentence (L139-140) should not be 'all drainage events', but '10 out of 13'. Our sincere apologies for this error. We will adjust accordingly, and by extending the timeseries to 2015-2024, these descriptions of observations and the conclusions drawn will undergo major revisions.

2.2.1 Activeness parameter

The authors conclude that "all of the detected lake drainage events occur in areas of the ice shelf classified as medium to highly active", but they sum up to 90% of the studied region. Taking into account that they have only 13 drainage events, their conclusion is not supported by the data since further analysis should be made.

Even if drainage events are correlated to activeness, it could be the case that activeness is actually related to damage (meaning that damage is higher when it is caused by the flow, which is very plausible) and that drainage events are mainly driven by damage (also very plausible).

We completely understand the critical standpoint against the low amount of samples, and, again, will alleviate these concerns by extending the timeseries with 6 more years. The conclusions will be adjusted with those results, and more care will be taken to align the weight of the conclusions to the observations.

A note about the damage signal strength: this is uncorrelated to the ice flow, as it is purely based on how clear the damage feature is visible in the image, as a stark white line in a dark field on radar images (in optical images it is a dark line on white), which is mainly dependent on the look angle of the Sentinel-1 sensor and the amount of noise/speckle, where backscatter of the radar is generally highest for steep vertical walls. The strongest damage signals are actually found near and at the ice front, where large full ice-penetrating rifts yield the highest contrast between the ice and the ocean.
Apart from this, the activeness parameter is indeed correlated to damage. It is based on the orientation of the detected damage, not on the strength of the detected signal, but still, there has to be a detected damage feature to get an activeness assessment. The intent of the activeness parameter is to shift the focus of the detected damage signal strength: you can find small yet opening crevasses ('active') as well as large but stationary rifts ('passive').

2.2.2 Section "Lake Drainage Events in Periods of Increasing Tidal Heights"

As reviewer #1 said "6/11 drainage events in 2019 (more than half) started during the lowest (or even descending) phase of the tidal cycle (drainages M, L, F, H, J, E)". This simply invalidates the sentence: "Our findings unveil a compelling narrative of ice shelf dynamics, revealing an intricate interplay between tidal forces and supraglacial lake drainage events".
Furthermore, in a hypothetical case where all the drainage events occur in an ascending amplitude phase of the tidal, it will not imply what is said: if you have thousands of lakes in a ice shelf, and each time only a small fraction of lakes drain, the likelihood of a drainage event triggered by tidal flexure would be higher in the crest of the amplitude phase.

We refer to our response to reviewer #1, which we will repeat here as well: we would like to clarify that indeed drainage events M, E and K 'start' in the descending phase (3/13) and 4/13 in the lowest phase. However, with this, it is important to realise that the drainage events shown in figure 3 occur *between* the detected dates t1 and t2, since we only detect 'lake is present' at t1 and 'lake has drained' at t2 – it is not a draining that starts at t1 and ends at t3. Therefore, Figure 3 does (in our opinion) clearly suggest that drainage does not occur in the descending phase. We do acknowledge that we cannot differentiate if the drainage occurs in the lowest part of the cycle or the ascending part, so will adjust the text accordingly.
As for the higher likelihood of a drainage event triggered in the crest of the amplitude phase: this is not reflected in any of our observed drainages.

Minor comments

3.1 Methodology clarification

Some clarifications are needed on the two main metrics used in the study. First, I agree with reviewer #1 about the lack of information on how damage is calculated. Usually it is inverted from
$\mu =(1 - D)B2\dot{\varepsilon}n-1ne$ (1)
where μ is the ice viscosity, D is damage (which you want to invert), B is the ice rigidity, εe is the effective strain rate, and n is the flow law exponent. I think it is needed to specify how damage is calculated from remote sensing. Also, if possible, it would be interesting to relate the damage calculated in the present work (which reviewer #1 suggests changing the name to "satellite-derived damage" and I agree), and damage calculated from Equation 1.

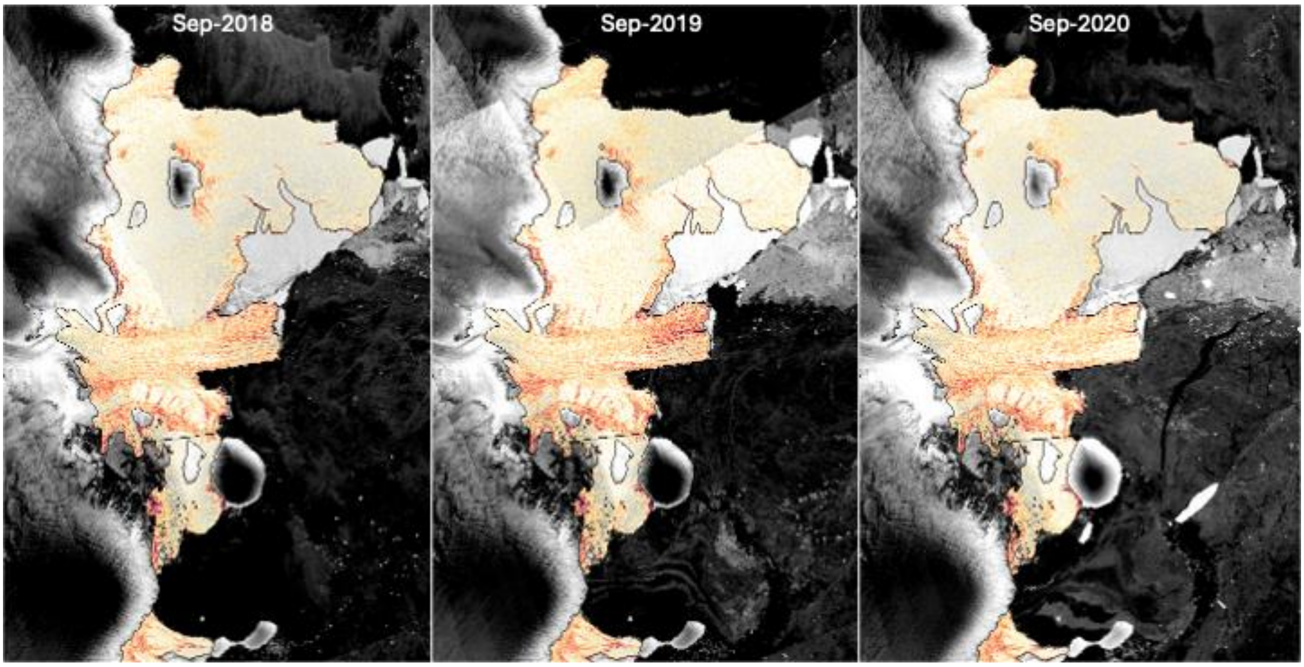
We see that in lieu of brevity the methods have been too concise, and will add more details on how exactly damage is calculated. In short, the NeRD method consists of the following steps: (i) create cut-out windows from the image (for which we use 10x10 pixels), (ii) apply the Normalised Radon transform to these windows, (iii) extract dominant feature signal strength and orientation for every window, (iv) quantify the damage signal value by removing noise from the signal and (v) postprocessing. In the post-processing step we clipped the product to the ice shelf bounds.

As the NeRD method is a published algorithm (Izeboud and Lhermitte, 2023), the description will remain short and to the point, as for extended sensitivity studies and evaluation of the method we refer to that publication. Nevertheless, we will include the produced annual damage maps (also included below, before downsampling to 3 km) and an assessment of the changes in detected damaged area for every year in the supplementary material.

We do find calling it satellite-derived damage a good suggestion and will implement this term as well as referring more strictly to 'damage signal' or 'detected damage' in the manuscript.

Lastly, we agree that it would definitely be interesting to relate the calculated damage maps to damage calculated from damage mechanics models, and to our knowledge this is an active field of research – refer to e.g. Gerli et al. (2024) and De Rydt et al. (2021) – but we consider it out of scope for this study.

- *Izeboud, M. and Lhermitte, S.: Damage Detection on Antarctic Ice Shelves Using the Normalised Radon Transform, Remote Sensing of Environment, 284, 113 359, https://doi.org/10.1016/j.rse.2022.113359, 2023.*
- *Gerli, C., S. Rosier, G. H. Gudmundsson, and S. Sun. 2024. 'Weak Relationship between Remotely Detected Crevasses and Inferred Ice Rheological Parameters on Antarctic Ice Shelves'. The Cryosphere 18 (6): 2677–89. https://doi.org/10.5194/tc-18-2677-2024*
- *De Rydt, J., R. Reese, F.S. Paolo, and G. H. Gudmundsson. 2021. 'Drivers of Pine Island Glacier Speed-up between 1996 and 2016'. Cryosphere 15 (1): 113–32. https://doi.org/10.5194/tc-15-113-2021*



Detected damage by the NeRD method at 300 m resolution, before downsampling to 3 km

3.2 Activeness vs. damage

I think a further analysis of the relationship between damage and activeness is required. It can be the case that the relationship between them is high, so any relationship between activeness and drainage occurrences is only due to damage.

The activeness parameter is correlated to damage, but only in the sense that it is determined in areas where damage is detected (binary): it is based on the orientation of the detected damage, not on the strength of the detected signal. It shifts focus to damage features that are likely undergoing change (opening, widening) from features that are stationary/passive. It is an interesting point though, that there could be areas of 'activeness' without visible damage features to show for it. To address this we will add a comparison of the strain rates, damage, and activeness parameter, as we expect this will shed more light on where and when damage with high activeness is found.

3.3 Motivation

The first sentence of the article ("Surface lake drainage can destabilize ice shelves, occurring either slowly via supraglacial channels or rapidly through crevasses") makes me wonder if you are not studying the opposite. Instead of analyzing how damage influences lake drainage, should not you analyze how lake drainage influences damage? Otherwise, if you really want to analyze how damage influences lake drainage, you should motivate that in the introduction. I think you are putting the cart before the horse.

Thank you for this insight, it's important to us that the introduction is very clear. We are mainly analysing the place and timing of lake drainages, and in that sense we are analysing how damage influences lake drainages: we hypothesised that just 'having' damage features does not necessarily lead to hydrofracturing – since damage is so abundent on many ice shelves in antarctica, and hydrofracturing is less widespread.

We will edit the introduction to (subject to small editorial changes):
Surface lake drainage can destabilize ice shelves, occurring either slowly via supraglacial channels or rapidly through crevasses driven by the weight of the water – a process known as hydrofracture (Nye and Perutz, 1957). While Greenland's lake drainages are relatively well-studied (Williamson et al., 2018a; McMillan et al., 2007), much less is known about similar processes in Antarctica. Understanding when and how surface lakes drain is crucial for assessing ice shelf stability. Previous studies, such as Trusel et al. (2022) have shown that lake drainages on the Amery Ice Shelf are linked to high-amplitude tidal cycles. Moreover, Lai et al. (2020) have shown large areas of Antarctica's ice shelves that are vulnerable to hydrofracturing if (existing) crevasses are inundated with meltwater. However, given the widespread presence of crevasses and other damage features on Antarctic ice shelves, it remains unclear to what extent pre-existing damage influences the likelihood and timing of lake drainage events. Here, we use observations of lake drainage events from remote sensing data to study their place and timing, examining whether damage alone is sufficient to indicate a potential of hydrofracturing, or if additional conditions, such as tidal forcing, are necessary to initiate lake drainage

Specific comments

L16: Is it the first time that "activeness" is used? If so, say that the manuscript introduce this concept. Otherwise, make a reference.

Thank you for your comment. The concept of activeness is newly introduced by us. And we will adjust the sentence accordingly:

L15: We therefore hypothesize that, apart from using the presence of damage features, another metric is needed to indicate a likelihood for occuring lake drainages. Specifically, we propose that in addition to the presence of damage (open crevasses, fractures, and rifts), a measure of the activeness of the damage feature (i.e. crevasse opening or propagation) can be used to identify where lake drainages are likely to occur on an ice shelf.

L44: Do you advect the features when you merge the images in the mosaic?

Thank you for your question. We did not account for advection when creating the mosaics. Since each mosaic covers a maximum of 10 days, the only area where advection would be significant is at the fast flowing Denman Glacier (max speeds of ~1500 m/yr (Miles et al. 2020) in its center, ~4.6 m/day, just enough to have 1 pixel of advection in the 30 m Sentinel-2 images in the selected period) -- ice flow speed quickly drops to <= 500 m/year to the sides of the ice tongue, where the majority of the melt lakes are observed. Moreover, in this short period of time the amount of repeat satellite overpasses is very limited, and we get just the minimum amount of image to stitch them together for a domain-covering mosaic, thus having almost no impact of advection on the mosaic calculation. We'll include a supplementary figure in the next manuscript to show the overlap of the used satellite images to construct each mosaic, and clarify in the text that advection is not taken into account.

L50: This sampling frequency (once a year) is very different from the optical image that you are going to contrast later on. How do you deal with that?

The damage maps are derived outside the melt season to ensure reliable detection. They show minimal year-to-year variability (as it can be seen above), and we will include them in the supplementary material.

L52: It is not clear where you use the velocity field. For sure for the "activeness" calculation, but do you also use to transport the features? Make it clear near the description of the velocity field calculation.

Thank you. Agreed and we will adjust the sentence as follows:

L52: Ice flow velocity data for 2019 are sourced from the ITS_LIVE campaign (Gardner et al., 2020) and used for the calculation of the activeness metric.

L61: Regarding the threshold 1800 m2 you make a citation for this value, but what is the reasoning of using this threshold?

Thank you for the comment. All subsequent steps where executed as indicated by the reference. But in order to clarify the purpose of reach step we have split the sentences:

L61: Outliers are removed in a subsequent step if they are not located on ice mass or have a misinterpreted depth of less than 0 m (Williamson et al., 2018). To minimize further noise, lakes with a surface area of less than 1800 m2 (2 or 18 pixels of L8 and S2 imagery, respectively) are removed as suggested by Williamson et al. (2018).

L65: Why 80%? I can also imagine 50% as a massive drainage event. this looks like a random choice, that you need to pick one, but if you lower the threshold, you would have many more drainage events, increasing the data you can use to infer, since 13 drainage events are not many.

This is in line with other literature, following Doyle et al.(2013), Fitzpatrick et al. (2014), Miles et al. (2017), Williamson et al. (2017). These references will be specified in the text.

- *Doyle, S. H., Hubbard, A. L., Dow, C. F., Jones, G. A., Fitzpatrick, A., Gusmeroli, A., Kulessa, B., Lindback, K., Pettersson, R., and Box, J. E.: Ice tectonic deformation during the rapid in situ drainage of a supraglacial lake on the Greenland Ice Sheet, The Cryosphere, 7, 129–140, https://doi.org/10.5194/tc-7-129-2013, 2013*
- *Fitzpatrick, A. A. W., Hubbard, A. L., Box, J. E., Quincey, D. J., van As, D., Mikkelsen, A. P. B., Doyle, S. H., Dow, C. F., Hasholt, B., and Jones, G. A.: A decade (2002–2012) of supraglacial lake volume estimates across Russell Glacier, West Greenland, The Cryosphere, 8, 107–121, https://doi.org/10.5194/tc-8-107-2014, 2014*
- *Miles, K. E., Willis, I. C., Benedek, C. L., Williamson, A. G., and Tedesco, M.: Toward monitoring surface and subsurface lakes on the Greenland Ice Sheet using Sentinel-1 SAR and Landsat 8 OLI imagery, Front. Earth Sci., 5, 1–17, https://doi.org/10.3389/feart.2017.00058, 2017*
- *Williamson, A. G., Arnold, N. S., Banwell, A. F., and Willis, I. C. (2017). A Fully Automated Supraglacial lake area and volume Tracking ("FAST") algorithm: development and application using MODIS imagery of West Greenland. Remote Sens. Environ. 196, 113–133. doi: 10.1016/j.rse.2017.04.032*

L67: Was not the are threshold 1800? Furthermore, why do you use these threshold? Give a reason and use the citation. Only the citation is not enough.

Thank you for your comment. We use two sets of thresholds in our analysis. The 1800 m² threshold is applied to minimize noise by excluding very small features that are likely spurious. In contrast, the 54,000 m² threshold is used to focus on lakes that are large enough to potentially drain and impact the ice shelf. This 54,000 m² value corresponds to about 60 pixels in Landsat 8 imagery, a size considered significant for water volume and hydrological impact (Williamson et al. (2018a)).

*Williamson, A. G.; Banwell, A. F.; Willis, I. C.; Arnold, N. S. Dual-Satellite (Sentinel-2 and Landsat 8) Remote Sensing of Supraglacial Lakes in Greenland. The Cryosphere **2018**, 12 (9), 3045–3065. https://doi.org/10.5194/tc-12-3045-2018 a.*

L72: you excluded 20 out of 25, so you have 5 drainage events. How then you have 13 drainage events in your results?

Apologies, but we are unsure where the count of 20 excluded events comes from. The manuscript states in **L72:** *"Twelve out of twenty-five events are removed, as they are judged to be refreezing lakes rather than draining lakes."*

This results in 13 events.

L78: Add "resolution" after "300 m".

Agreed and implemented.

L83: Missing citation. Makes sense to compare the angle of the fracture to its orientation, but quantifying it can be tricky. Is there any supporting studies for the use of that values?

We apologise for any confusion: the angle of the damage feature is a result of the NeRD method (specified just below this sentence). We will clarify in this sentence ("The obtained damage orientation <u>from the NeRD algorithm (Izeboud and Lhermitte, 2023)</u> is used to identify areas with a likelihood of active damage development, ..."). Moreover, as we will expand the explanation of the NeRD algorithm to provide more clarity on how damage (and its orientation) is detected in the method section beforehand, this will further aid clarity.

L88: If activeness is binary, how you produce an image like Figure 2 b)? It do not seems like a product of downsampling.

Thank you for your comment. Our activeness metric is indeed binary (0 or 1) at the original 300 m resolution. However, when we downsample by a factor of 10 using an average resampling method, each 3000 m pixel then represents the average (or proportion) of active pixels in that block, which naturally yields continuous values between 0 and 1. This will also be clarified in the text.

L88: I don't see the reason of donwsampling it.

It can be very tricky to properly assign which damage features should be linke to which lake drainage events based from these remote sensing observations, and what the appropriate lengthscale of such influence is. Furthermore, we hypothesize that an area of active damage might be indicative of a general structurally weakened ice zone, which might facilitate lake drainages through previously undetected (small) fractures. To capture this, we translated the detected damage and activeness parameter into a less localized representation, reflecting the overall integrity of the ice over a larger area.

L89: Why normalizing? You can say directly that damage varies between 0 and 1 and everything is already normalized.

Thank you for the question. We are using the NeRD algorithm (Izeboud and Lhermitte, 2023) which outputs damage values between 0 and 0.5. We then normalize these values using the maximum derived from the Shackleton Ice Shelf, which standardizes the data into a 0 to 1 range for easier comparison across the study area.

L89: Add "resolution" after "3000 m".

Agreed and implemented.

L101: Add "total" in "with total maxima".
Agreed and implemented.

L116: I think the definition of the thresholds should go to methods with a further explanation of the threshold used.
Agreed, we concur that for the sake of brevity we left out too much, and we will implement this.

L119-124: This is a very sounding result, supporting the hypothesis of a strong relationship between drainage events and damage. However, I would expect the same analysis regarding the activeness. The distribution of 10%, 71%, and 19% does not allow this analysis. Contrasting the area distribution and Figure B1, and do not see a strong relationship between activeness and drainage events. I would say that most of the signal of drainage events are due to damage.
This is insightful, and we will take this into consideration when we revise the analyses with the extended time series, which may or may not change these results.

L129: Same comment as for damage.
Understood, the thresholds will be more extensively described.

L139: This conclusion is not surpirse. It sums 90% of the studied area.
The reviewer is correct. We'll revise accordingly with the extended timeseries.

***L151:** What do you mean by "parallel trend". Be more specific. **L151-153:** This goes against the phrase: "However, all of the detected lake drainage events occur in areas of the ice shelf classified as medium to highly active". **L156:** You previously said the opposite. **L158:** As far as I understood, it is impossible to drainage event to occur without damage. Be more precise with this statement.*
Thank you for your comment. We agree that the wording needs to be more precise and consistent regarding the derived categorizations of damage and activeness. We have revised the text as follows:

L151 – L159: Intriguingly, activeness and damage do not always follow the same behaviour, suggesting that these two factors influence lake drainage in a more complex manner. For instance, drainages H, K, and M occurred in areas with medium damage but high activeness, whereas drainages A, F, and G were found in regions with medium activeness yet exhibited high levels of damage. By integrating the dominant orientation of fractures relative to ice flow into our activeness metric, we capture an independent characteristic of ice shelf damage that can vary from the overall damage intensity.

L161: I agree with reviewer #1 regarding the drainage events with respect to the ascending phase of tidal cycles.
We understand the concern. We would like to clarify that indeed drainage events M, E and K 'start' in the descending phase (3/13) and 4/13 in the lowest phase. However, with this, it is important to realise that the drainage events shown in figure 3 occur between the detected dates t1 and t2, since we only detect 'lake is present' at t1 and 'lake has drained' at t2 – it is not a draining that starts at t1 and ends at t3. Therefore, Figure 3 does (in our opinion) clearly suggest that drainage does not occur in the descending phase (even though, indeed, crevasse might be more prone to opening in that phase!). We do acknowledge that we cannot differentiate if the drainage occurs in the lowest part of the cycle or the ascending part, so will adjust the text accordingly.

L187: I think you do not have drainage events that last hours. If this is the case, remove the "few hours".
You are right and we'll remove this.

L190: Here you say that it is difficult to assign the drainage events to hydrofracturing, but in the discussion you did this.

Thank you for your observation. We acknowledge that our previous wording may have been unclear. While we can identify the occurence of drainage events (i.e. detecting a drained lake) using satellite imagery, we don't know the exact timing and speed of the drainage that occured between the satellite overpassess. This is the distinction we were trying to convey, and we'll edit the text to clarify this.

L197: If NeRD can not identify individual fractures, how then you measure the orientation of the fractures to calculate the activeness?

Apologies, this wording is misleading. NeRD returns one value for damage signal strength and one for damage orientation for every processing window of 10x10 pixels (30 m per pixel). It does not, however, return the exact location of the detected feature within the window, and neither its width or length. It is also possible there are multiple crevasses within the window, for which case the algorithm favors the feature with the strongest contrast (see Figure from Izeboud and Lhermitte (2023) below). So, what we mean is actually that NeRD does not detect the exact outlines of individual features.

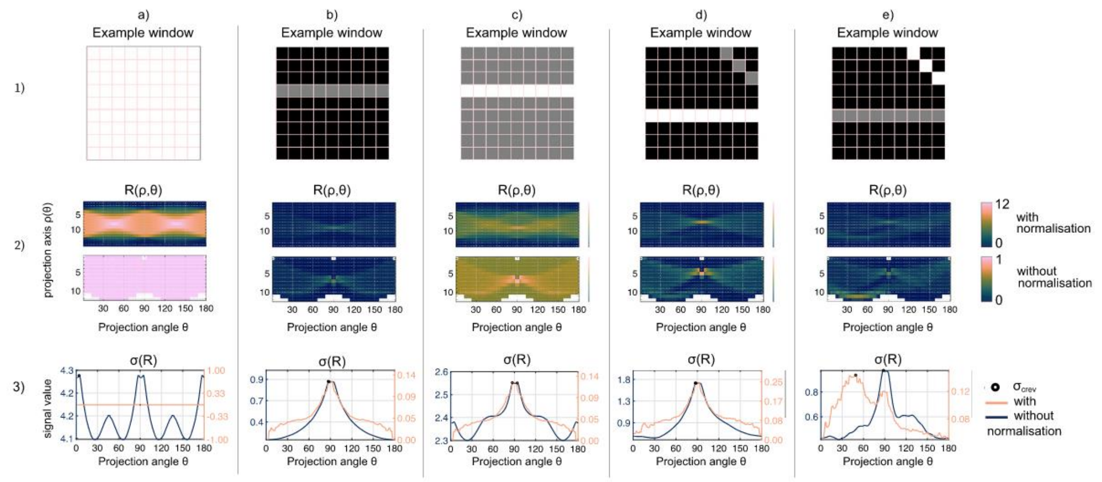


Fig. 2. Idealised scenario's to illustrate the differences between the Radon transform with and without normalisation. Panel a–e represent different scenario's to which the Radon transform is applied: a1–e1 show an idealised window with a hypothetical crevasse, a2–e2 show the corresponding 2-D feature space $R(\rho, \theta)$ without and with normalisation (respectively top and bottom), and a3–e3 the signal response $\sigma(\theta)$ with and without normalisation — from which σ_{crev} is extracted (black dot).

214-216: This sentence is another major concern for this study.
We understand the concern. Again, this comes back to the low sample size (this sentence refers to the 2 of 13 drainages that occur in low activeness areas), and we expect to improve upon these assessments with additional observations by extending the timeseries.

I would appreciate a hexabin graph with activeness and damage in the axes. This would allow us to see the correlation between both metrics and the occurrence of drainage events.

Thank you for the suggestion, we brought two plots from our previous discussions.
We could add something similar to Figure 2 and address the distribution of detected drainage events vs the distribution damage and activeness.

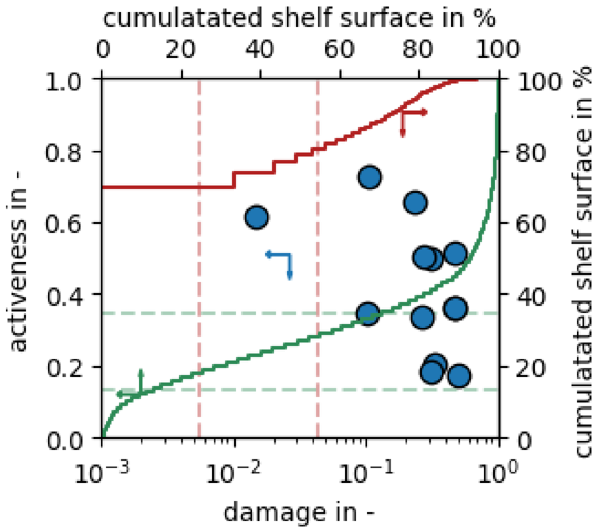


Figure is showing detected activeness versus logarithmic damage. The dots are the detected drainage events, dashed lines are the respective limits for medium and high characteristics. The secondary axis show the cumulative ice shelf coverage of damage and activeness of the melt season 2019/2020.

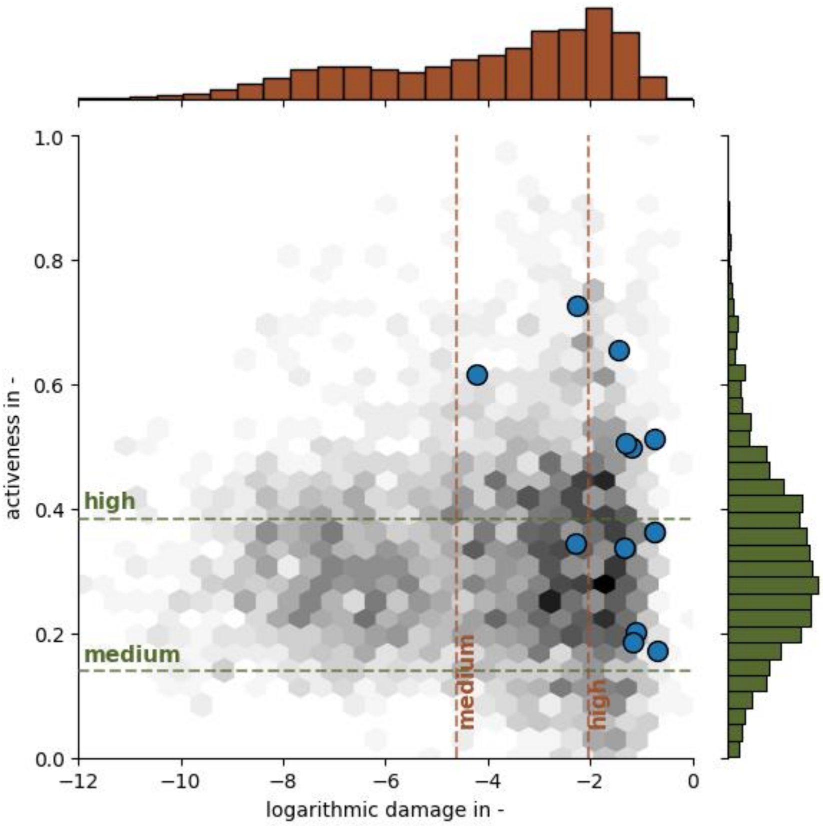


Figure is showing detected activeness versus logarithmic damage (please note that for swift plotting we applied the logarithm instead of making the scale logarithmic). The dots are the detected drainage events. The dashed lines are the respective limits for medium and high characteristics. The histograms show the distribution over the ice shelf damage and activeness of the melt season 2019/2020.

Table A1: Bring it to the main body of the text, it is too important. I suggest including two more columns: Classification of damage and activeness (low, medium, high).

Good idea, we'll add the extra columns. We agree that it would be good to have this in the main body. However, the Brief Communications format only allows for three display items (tables/figures) so we are very limited in our flexibility here, unfortunately.

Figure 1: Define LIMA as an optical imagery mosaic from Landsat.

Agreed and implemented.

Figure 2: Add ", respectively" at the end of the first sentence.

Agreed and implemented.