REVIEW OF "Changes in Supraglacial Lakes on George VI Ice Shelf, Antarctic Peninsula: 1973-2020"

T. J. Barnes, A. A. Leeson, M. McMillan, V. Verjans, J. Carter, C. Kittel

General assessment:

The authors of this paper attempt to synthesize the satellite imagery record, local weather station data, firn model output, climate reanalysis model output, and CMIP future climate model output in order to better understand the future expansion of the George VI's surface meltwater drainage system. This is motivated by the recent findings of unprecedented melt in the 2019-2020 melt season. The motivations behind this study, to use the available data to give an assessment of present-day conditions of the surface drainage system to better predict its future evolution, is well-thought out, valuable, and promising for future study. Other studies have done similar multi-year assessments of Antarctic ice shelf surface hydrology systems (Langley et al., 2016, Stokes et al., 2020, Spergel et al., 2021). However, the authors present many different means of showing a disconnect between climate and surface conditions with lake coverage, but do not discuss that lake coverage, i.e where melt is observed ponding in satellite imagery, is mainly controlled by surface topography. The pre-existing surface depressions (discussed in Reynolds, 1981) must be filled to overspilling before water can drain over the surface into new areas/depressions. I am not familiar with the surface topography of GVIIS, but other ice shelves have more-or-less U-shaped depression cross-sections, so the addition of more meltwater does not change the surface area of the water body. It is only when all available space is filled with water that the surface area of water coverage expands via over-surface drainage. What the authors seem to describe with their analysis of similar meltwater lake coverage between 1989 and 2020 is not a dampened response to climatic forcing, but the fact that meltwater pond coverage increases as water flows and partially fills depressions nearby to meltwater production, but meltwater coverage plateaus as the partially-filled depressions fill, and only once overspilling occurs does lake coverage increase again significantly.

I would recommend refocusing this paper on the changes observed in melt pond distribution on GVIIS between the 1980s, as described in Reynolds, 1981, and where melt is observed today. There is a lot of value to giving a base-line and a detailed description of the inter-annual variability in the ice shelf's hydrology. I would recommend a thorough search of the literature to give a broader context to the authors' findings on GVIIS.

Grades are from 1 (excellent) to 4 (poor). Your MS is graded as follows::
Originality: 2
The purpose of the paper, to assess the decadal trends in a persistent surface drainage network has a lot of merit.

Scientific Quality: 4

The oversight of topography controlling where melt forms, and what that means for measuring meltwater lake coverage with satellite imagery, makes a lot of the analysis done in the paper unsuccessful in proving any climate-surface hydrology mechanisms.

I also have many questions about methods that are not addressed in the paper. The results are reliant on the threshold of NDWI, the comparison of imagery coverage across the years, and uncertainty in manual mapping. These three issues and the uncertainty they contribute should be discussed.

Significance: 4
As the paper is now, the points that are presented successfully (persistent, widespread melt on GVIIS; inter-annual variability in melt production leads to variability in meltwater lake coverage; meltwater being divided between ponds and firn pore spaces) are not novel enough to be significant. In its current form, the paper is unsuccessful in supporting a new mechanism for climate-surface hydrology interaction, the "dampening effect" of increased firn air content on lake coverage.

Presentation Quality: 3
Much of the paper is well written, but there are a few issues in the paper's presentation: 1) there remains a number of passages that could use scientific, quantitative terminology instead of conversational. 2) quantities such as averages, sums, etc, should be precise in what they are describing to avoid ambiguity. 3) The figures should be revised to be more readable, especially the time-series plots. 4) Much of the material presented in the supplementary materials is critical to assessing the paper, and should be brought into the main text.

Main Comments:

1. The paper needs to be reassessed after considering how surface topography affects where water pools. Several of the proposed mechanisms and causal relationships between climate, firn, and meltwater lake coverage need to be reconsidered, and revised if still true or removed if no longer true.
2. The methods by which lake pixels are selected need to be further explained. Moussavi et al. (2020) would be a good reference if the new method is to be kept, but I would recommend using Moussavi et al.'s available code for Landsat 8 imagery and discuss the process used to select the NDWI thresholds for Landsat 1-7. I also don't understand what the scaling of lake pixels derived from non-Landsat 7 imagery includes, but the uncertainty introduced by this needs to be discussed.
3. Many of the assertions about climate effects on meltwater lake coverage presented in the discussion/conclusion need to be supported by data or citation of the literature. The choice of MAR is discussed in the supplementary materials, but the authors also seem to use MAR output as a single point rather than a spatially-varying raster dataset.

4. Some sections need to be rewritten to clear up ambiguity in what was done, what is being extrapolated, etc.

Line by Line Edits:

L15. "build up strong" > "build up. Strong..."
L15. I think "precede" or "yield" is meant here instead of "entail."

L26. Adusumilli not Adusumili

L32. Consider "In considering SAM, we also analyze..." Instead of "Furthermore, thanks to..."

L33. What metric is this being referring to?

L38. "reduced buttressing due to ice shelf collapse is estimated to have led to.."

L40. "a theory supported by widespread lake drainage observed prior..."

L46. In George VI Ice Shelf, "Ice Shelf" should be capitalized

L47. The area number needs to be cited.

L55. Hottest day means highest daily average temperature recorded?

L57. What are the values of the two percentages being compared?

L59. Vulnerability - "vulnerability to meltwater-driven hydrofracture."

L72. How were the individual satellite images merged? In what software?
From Table S1, it looks like there are several different Landsat rows used: did the authors have a change in coverage of the area of interest depending on which row was used?

L75. In what software was this lake pixel identification conducted?

L77. I suppose I don't understand why Moussavi et al.'s code was not used, because it has published results and an extensive workflow to remove rocks, clouds, shadows, and other false positives.

L84. Where did the thresholds the study is using come from?

L89. Are shadows also removed during this pre-processing step?

L90. Is the removal all manual or automated?

L96. See Main Comment 2

L99. GVIIS

L110. What does it mean that the authors took into account points from Agosta et al. 2019?

L112. I think the authors mean that the limitations of MAR are acknowledged, but an analysis of these limitations is outside of the scope of this paper. This section should be moved to the discussion. It is not appropriate to include it in the data and methods section.

L113. What is the resolution, temporal and spatial of ArMAP? Have others tested it compared to ground truthed data nearby?

L115. I don't understand this sentence. What about ArMAP is connected to the study region? That it was created from Antarctic firn cores?

L122. Citation required for CMIP models.

L127. I think the authors should include any pertinent information from S6 here in the main text.

L132. 1. This should be in the discussion, not methods. 2. If this multi-variate method did not yield useful results, why is it included in the paper? Can the authors not just present the univariate correlations?

L136. "ratios with melt versus lake coverage." - What are the ratios?

L137. This is not a necessary statement: "Standard deviations were also calculated in order to define exceptional variance from the mean and thus to identify exceptional years of lake coverage versus controlling variables."

L143. "the previously reported peak of AP melt season"

L146. This does not belong in the initial introduction of the results. Also, try to adopt a more objective tone: "Of course" comes off as too conversational.

L150. It would be better to present the actual uncertainty in areal lake coverage in 2020 and 1989.

L151. lake coverage. There's not a total, unless I am missing a summation?

L156. Is there a figure associated with this results section? If not, one should be added.

L156. There is no significant linear trend, right?

L161. I believe "defined" is meant, not "quantified"

L161. "mean percent coverage"

L161. I would split this into a separate sentence. "We find the following years to have high lake coverage: ..."

L162. This statement should be in the introduction

L164. Reynolds, 1981 is missing from the bibliography.

L165. I don't understand what morphology means in this context. Shape of lake margins? Location?

L166. I don't know what this means: could the authors say from eastern to western grounding lines or from the northern to southern regions of our region of interest?

L167. western should not be capitalized

L167. "near" instead of "in the proximity of"

L168. Considering there was a decently in-depth study of why George 6's lakes form in particular areas (Reynolds, 1981), there should be more spatial information available to qualify and give more information to this finding.

L169. "form interconnected networks across surface depressions"

L169. Instead of "discrete" - "do not show the same connectivity between surface depressions".

L170. "such as" not "namely"

L170. Two points do not show evolution. One can describe that two different spatial distributions were observed in these two composite maps, but unless the authors see a graduated shift between these two different distributions, the authors can't extrapolate a time-varying process.

L171. reverse the order that (d) and (b) are presented.

L172. I don't know how the 9.1% was calculated.

L174. I would like to also see a description of lake area histograms during each period.

L179. There should be a discussion of spatial variability in temperature considering the temperature data is a raster not a point.

L181. "Melt" is MAR modelled melt production?

L181. surface temperatures as reported at the AWS or those from MAR?

L181. Accumulation always refers to snow accumulation?

L182. Is this the summer (DJF) of 2019-2020?

L182. This is the annual mean temperature? I find that number very difficult to believe.

L183. Again, is this daily mean snow accumulation or annual snow accumulation?

L183. total accumulation over 2019-2020? daily output?

L188. Are you describing a single monthly average 2m air temperature? From the AWS? If this is a key point, it should be clear where this temperature is from.

L192. is this the number of lakes or lake coverage? In either case, this is controlled by both pre-existing surface topography and by melt input. See Main Point 1.

L207. With this and other r values in the text, a p-value is required to show that the correlation is statistically significant.

L212. I wonder if correlating the specifics of SAM to melting on GVIIS is beyond the scope of this paper.

L225. high cloud cover could also increase the trapping of long-wave heat.

L231. Why is 2020 being excluded? It's an outlier?

L236. I'm having trouble following the logic here around the dampening mechanism. Reword possibly?

L236. There's a lot of supposition behind the dampening effect the authors are describing. I worry that this dampening effect is the result of pre-existing surface topography needing to be filled to overspilling in order to increase lake coverage. In any case, the build-up of firn air content is not described previously.

L250. I believe that firn air content did moderate over-surface drainage extension, but rain will also deepen the depressions on the ice, which will also 'dampen'/slow the expansion of water into new depressions.

L265. I don't think the authors presented lake densities in 1990 or 1991

L275. "We find that the build-up of firn from successive low lake seasons since 2013 contributed to the dampened response to climate forcing we see in 2020" - This point needs the presentation of more data previously in the paper.

L282. This presupposes that the available accommodation space in the existing depressions will be filled to over-spilling.

L283. Where is this statement on ice shelf viability from?

<u>Figures.</u>

Figure 1.
- labelled points and insets should be on the same panel, (a) in this case.
- My reading order of this figure is left to right, top to bottom. I would rearrange the panels to fit this reading order.
- I'd rather see one satellite image with labelled locations, and then maps of identified lakes in the remaining panels. I can't get much information from these satellite images presented like this.
- I would make the area of interest's outline thicker to differentiate it from the lat/lon gridlines
- what is e if it isn't an RGB true color composite?

Figure 2.
- What images/dates are these two maps composed of?
- I would put the legend in a box to the right of the figure rather than at the bottom of (a).

Figure 3.
- This figure is very difficult to parse. Have you considered making a scatterplot with each point labelled by year, like figure S7?
- Why is there a gap in the November FAC and accumulation time series?

Figure 4.
- I can't read any sort of trend in plot A.
- These maps should be focused on only the area being studied.
- What is ∂T here? Difference in annual mean? In summer mean?

Supplemental Material:

Table S3's minimum area column is unnecessary.

S53. I don't understand, the authors decreased the total area in images with more coverage?

S100. This shouldn't be in supplemental methods descriptions, but rather in the main text