General comments:

This manuscript presents a time series of surface meltwater area over the Amery Ice Shelf, East Antarctica, which has been automatically generated using an existing optical band threshold meltwater detection algorithm alongside a new method for calculating image visibility in Google Earth Engine. The authors use this to estimate cumulative total lake areas between 2005 and 2020 and compare this to modelled melt predictions from a regional climate model. Finally, they discuss controls on the temporal variations in surface meltwater, including the summer Southern Annular Mode index.

Although the method for surface meltwater delineation in this study is not novel, it goes beyond previous approaches (Moussavi et al., 2020) by quantifying image visibility (i.e. area of surface water missed due to cloud cover and image data coverage) with the aim of developing a consistent time series. This is a useful step forward given that most surface meltwater assessments do not explicitly quantify this. In addition, while the relationship between surface melting and large-scale climatic variability has been investigated on the Antarctic Peninsula (e.g. Bevan et al., 2018), it has so far rarely been explored as a control on meltwater on East Antarctic ice shelves. Therefore, it is my view that the findings from this manuscript are of broad interest to the cryospheric community.

In general, I would like to complement the authors on their well-written and clearly-structured manuscript. The study aims and methods are generally well explained and justified, and there is good consideration of wider controls on lake formation and variability. The results build upon previous work that has reported surface meltwater on this ice shelf by specifically accounting for image visibility in the interpretation of lake area time series. It is encouraging to see that this method can be applied to Landsat 7 imagery as well as Landsat 8.

I would like to see the authors expand their discussion around how their approach of scaling up mapped lake areas could be overpredicting lake coverage. Surface meltwater lakes tend to be highly clustered, so some pixels (regions) are much more likely to be covered by surface meltwater than others. This could produce potentially large lake area overestimates, and if applied over other regions of Antarctica could affect the accuracy of surface meltwater time series.

Secondly, although the authors discuss misclassification errors in Section 4.1, I think this could expanded, including how these were verified. I question whether this one study area alone is sufficient to be able to determine the applicability of this approach on a pan-ice sheet scale. Misclassifications could be a lot higher across other ice shelves (especially in dirty ice/high debris cover/topographically-complex regions). I think this should be acknowledged more in the text when discussing transferability across other ice shelves.

Lastly, the abstract is quite a bit lengthier than this journal's suggested limit of 150-250 words and could provide a more concise summary.

Once the authors address this point and my specific comments below, I can therefore recommend that this manuscript is suitable for publication in The Cryosphere.

Specific comments:

Line 37: widespread ~~across~~ → 'around' (as meltwater is restricted to the ice-sheet margins)

Line 42: I find 'concentrated at the break of slope' an odd phrase to describe katabatic winds, as they descend the escarpment region from the interior – suggest rewording.

Line 57: ~~affect~~ the flow of grounded ice → accelerate?

Line 109: I question whether this one study area alone is sufficient to be able to determine the applicability of this approach on a pan-ice sheet scale. Although it is a large ice shelf with large numbers of lakes, it is a relatively 'uncomplicated' area – the performance could be different in regions of dirty ice/with high debris cover or wind-blown sediment, for example the McMurdo Ice Shelf. I don't think you can be completely confident about its continental applications until at least multiple test sites have been trialled.

Line 115: (Figure 1) – Could you use a more up-to-date basemap and coastline in this figure, for example from the Antarctic Digital Database which has a high resolution coastline updated since the 2019 'loose tooth' calving event (https://data.bas.ac.uk/collections/e74543c0-4c4e-4b41-aa33-5bb2f67df389/)?

Line 158: specify this is the NDWI adapted for ice (red and blue bands), typically written $NDWI_{ice}$.

Line 163: It isn't immediately clear what period a time window covers - is this variable, and are there always two time windows per month?

Line 215: I found this Figure quite difficult to get my head around, especially VoI vs TotV – could it be clarified somehow?

Line 259: I suggest maybe clarifying that an 'island' refers to an ice lid within the lake centre.

Line 266: Why was this Depoorter et al. (2013) grounding line dataset used rather than, for example, the MEaSUREs dataset (Rignot et al., 2016) derived from DinSAR measurements from 1992-2014?

Line 274: Please specify ERA-interim resolution.

Line 275: Is the box covering the study region that is referred to here the same red box shown in Figure 1, or a different extent? How were mean monthly summed melt values extracted from within this box – did you include values from any RACMO grid cells that intersected this box?

Line 285: temperatures, and therefore surface melting?

Line 293: which part of the ice shelf/grounded ice were false positives generally located?

Line 295: how did you estimate lake depth?

Line 300: (Figure 5) Consider adding insets to show locations on ice shelf.

Line 355: (Figure 7) Colours are difficult to distinguish (e.g. between December and January lakes on Panel a). Could you also indicate the maximum elevation of the study region?

Line 363: 'The highest <cumulative> number of lakes were observed' because there are higher lake numbers in individual time windows in 2014, 2018 and 2019.

Line 365: I would refer the reader to the relevant sub-panels here, e.g. Fig 8b. Similarly, Fig 8a-b on line 373.

Line 392: I find it quite difficult to see the seasonality discussed in the text given the bars are quite close together – maybe you could choose one melt season to do an additional seasonal plot? Also a question regarding numbers of lakes in Panel (a) – how did you account for the Scan Line Corrector striping 'dissecting' larger lakes into multiple polygons, and could this potentially have affected the lake totals presented pre-Jan 13?

Line 406: 2013/2014 modelled melt total is higher than in 2016/17, so this doesn't appear to be correct?

Line 409: Again, I find it hard to pick out these minor interannual variations in modelled melt rates using the axis scale you have chosen.. for example, distinguishing whether melt peaks in December or January in different years. Also, wondering why you have chosen to present modelled melt in units kg m^2 rather than the typically-used mm w.e. day^{-1}?

Line 453: 'whereby the temporal resolution can be varied' – please clarify.

Line 555: I agree that Figure 11 shows surface scouring and blue ice exposure, likely generated by katabatic winds, but can you really say that this was enhanced katabatic wind scouring compared to other melt years? Is there evidence of this in the satellite imagery for other melt seasons?

Line 571: '[…] does not specifically include a representation of blue ice', and because it is difficult for regional climate models like RACMO to resolve these melt feedbacks at this scale (27 km).

Line 589: Good point.

Line 593: Your analysis focuses on lake area, so I suggest you remove the reference to volume here.

Technical corrections:

Line 51: italicize 'in situ'

Line 53: 'ice shelf flexure'

Line 156: hyphenate 'mask based'

Line 265: hyphenate 'meltwater filled'

Line 281: delete 'the' large-scale

Line 296: hyphenate 'sediment laden'

Line 481: hasn't → has not

Line 504: Roi Bedouin → Roi Baudouin

Line 573: 'under predicts' → one word

Line 594: aren't → are not

References

Bevan, S. L., Luckman, A.J., Kuipers Munneke, P., Hubbard, B., Kulessa, B., & Ashmore, D. W. (2018). Decline in surface melt duration on Larsen C Ice Shelf revealed by the advanced scatterometer (ASCAT). Earth and Space Science, 5. https://doi.org/10.1029/2018EA000421.

Rignot, E., J. Mouginot, and B. Scheuchl. (2016) MEaSUREs Antarctic Grounding Line from Differential Satellite Radar Interferometry, Version 2. [Indicate subset used]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: https://doi.org/10.5067/IKBWW4RYHF1Q.