Review of Buckley et al. Observing the Evolution of Summer Melt on Multiyear Sea Ice with ICESat-2 and Sentinel-2

The authors provided a mainly methodological paper that integrates various remote sensing sources to track the evolution of melt pond characteristics (melt pond fraction and depth) on multiyear sea ice in the perennial ice zone north of the Canadian Arctic Archipelago in the 2020 melt season. This region contains some of the oldest and thickest ice in the Arctic and is distinct from the broader and largely first-year ice dominated Arctic pack. Multispectral optical data from Sentinel-2 and Worldview-2/3 satellites are used to estimate melt pond fraction (MPF) by image classification, and the resolution limitations of Sentinel-2 (10m) are evaluated in the context of the classifier by using the high-resolution Worldview-2/3 images (~1m) as verification. This is useful and insightful, as it provides new information about the limitations of the widely available but resolution constrained Sentinel-2 to estimate and study MPF broadly. As perhaps expected, the Sentinel-2 classifier underestimates MPF due to its inability to resolve smaller scale features such as small ponds and interconnecting meltwater channels due to pixel mixing. Despite being expected, it is good to see this problem being addressed rigorously and the limitations quantified. The authors found Sentinel-2 MPF is biased low "by up to 20.7% and averaging 7.2%, when small ponds are widespread across the surface". The authors also provide a method for adjusting the Sentinel-2 MPF based on input Worldview imagery when available. This approach is limited by the availability of high-resolution optical imagery, which should be less of a problem in the future as more open access data of that type becomes available.

The authors further compare two novel melt pond depth retrieval algorithms that use ICESat-2 altimetry data as input (ATL03 geolocated photon height) and discriminate pond surfaces from bottoms to resolve pond depth. They are unique, such that a comparison like this provides detail essential to their effective application to studies of atmosphere-ice-ocean processes during advanced stages of sea ice melt, when melt ponds play a key role in energy and meltwater budgets, among other things. The limitations of the algorithms are explicitly and thoroughly addressed in the context of the detection limits of ICESat-2, e.g. the pulse width imposes a 0.2m limit on pond depth, which means depth would not likely be resolved on seasonal sea ice due to their shallow nature. As the authors clearly indicate, one of the algorithms (DDA) is superior to the other (UMD-MPA) for detecting pond depth in a larger proportion of melt ponds in the studied area because it can resolve smaller ponds.

Sea ice and melt pond fractions are combined with melt pond depth estimates to examine the evolution of the advanced melt season in 2020 and the evolving parameters are discussed in the context of known key periods (early melt, maximum melt, etc.), previous observations from field campaigns SHEBA and MOSAiC, and model parameterizations (CESM and CICE). In general, the authors draw similarity between their results and those of others except when comparison to the modelling approaches where there is no observed relationship between pond depth and fraction.

The paper is original and presents new insights on the combined evolution of melt pond fraction and depth, thus providing an important third dimension to the study of these important features using remote sensing data. The methods are valid and suitable and connections to parallel and ongoing work on the melt pond depth algorithm (DDA) are clearly and helpfully stated. Enough detail is provided to understand both application potential and limitations of the described methods. The topic is definitely appropriate for *The Cryosphere*, though the authors should first address the below major and minor comments.

Major comments

The paper is imbalanced in that it presents as an analysis of the 2020 melt season due to its relevance in a climate context and in terms of the sea ice record – anomalously warm spring, early melt, and the second lowest September sea ice extent on record. It is more focused on the methods associated with melt pond fraction and depth retrieval and limitations. By comparison, the melt season analysis is rather cursory and mostly limited to a time series analysis of the derived parameters. The authors could strengthen the paper by either placing more emphasis on the methodological components as the core theme of the paper (e.g. section 5.2.4 belongs in results), or by expanding the melt season analysis so that the 2020 melt season evolution, as it is characterized by the retrieved melt pond properties, is better understood in the context of being an anomalous season. The latter would benefit from a comparison of 2020 conditions to other years which understandably would be limited to 2021 and 2022 by Sentinel-2 and ICESat-2 availability.

The authors should better describe the combined and relative importance of melt pond fraction and depth for the study of sea ice melt season evolution and atmosphere-ice-ocean interactions. This could be more explicitly addressed through the stated research goal(s) of the paper. Outlining the potential benefits of merging depth estimates with pond fraction, i.e. for volume estimates, and for improved understanding of melt evolution, is needed. How tracking the depth in only the larger ponds would influence the results is also of relevance. This is needed partly due to the imbalanced nature of the paper, as mentioned above. Detail regarding the importance of these observations mainly comes at the end of the conclusions, whereas the emphasis is on the methods early on and only general statements are made (e.g. line 83, understanding the evolution of melt).

Minor Comments (L=line)

L90: Clarify what area is the multiyear ice region.

L120: Give the spectral range as done for Worldview below.

L146: “Worldview”

L168: Delete Maxar

L173: Pixel misclassification is discussed in section 3.3 and 3.4 and both discussions refer to the same figure. Should be in one section.

L178: Add “image” after Worldview.

213: Use more precise terminology than chunks.

L218: Delete second comma

L229: Use MPF instead of melt pond fraction.

L233: area “of” …

L238: Use SIC after Sentinel-2.

L253: MPF

L259: Use ATLAS only, it was defined earlier.

L371: “false”

L379: Does the dead time effect also happen for leads as it does for melt ponds?

L402: The Perovich et al. 2002b paper is cited, but a little more context about the study would still be appropriate.

L406: How do we know that this is due to seasonal evolution versus differences in the area covered by the Worldview scenes and the ice topography etc.?

L429: “imagery”

L436: Provide detail on how the presence of ice lids is determined. Is it the dark grey color mentioned, in which case can this be confused with a drained melt pond?

L555: Can delete “, and inability to track smaller ponds”

L614: “…being biased high…”

Figure 6 caption: word missing “Examples (c) and (d) the subsurface”. Show?