Zhongyang Hu et al. 2021: Improving surface melt estimates over Antarctica Using Deep Learning.

In this article, authors develop a deep machine learning model with the aim of improving surface melt estimates from a regional climate model, RACMO2. The ML model is applied to three locations on the Larsen C (and remnants of Larsen B) ice shelf, which adds complexity to the scientific question of whether the ML model can improve surface melt estimates. The technique is novel, and likely to lead to further investigation of ML techniques for improving model simulations of various glaciological mechanisms. The study is well thought out and evaluated, and the results are presented clearly. Although I list numerous specific concerns below, they are quite minor, and I would be happy to see the publication of this manuscript once they are implemented.

Specific comments

Ln 22: I would include additional references for hydrofracture where they explain this process more thoroughly than the IMBIE paper. Perhaps Kuipers Munneke et al. 2014 (https://doi.org/10.3189/2014JoG13J183) or Gilbert and Kittel 2021 (https://doi.org/10.1029/2020GL091733).
Ln 26: here you talk about melt volume importance, but you haven't mentioned melt volumes yet. You have only talked about melting in terms of mass loss and ice shelf stability. I would perhaps open this paragraph with an estimate of sea level rise associated with AIS melt volumes (either present estimates or future projections).
Ln 34: change to 'face difficulties in accurately estimating surface melt…'
Ln 41+44: change to 'physically-based models'
Ln 58: In number 2 you contradict why Larsen C is ideal. Rather than say that high-quality observations are scarce in Antarctica, instead say that high-quality observations are available in this region, which is rare for Antarctica.
Ln 64: remove first 'in'
Ln 67: some contradiction between 'no ice shelves' on line 65 and 'almost only' on line 67. The wording also needs changing as 'almost only' isn't right. Perhaps 'Ice shelves on the Antarctic Peninsula are mostly located on the eastern coast.
Ln 70-73: You need some citations here for these values.
Ln 82: indicator **of** surface melt
Ln 94: wording needs consideration. How can a model be adapted for its impact on SMB and SEB? The model doesn't have an impact on the SMB. Perhaps you mean adapted for more accurate representation of SMB?
Ln 95: Was RACMO2 forced by ERA-Interim or ERA5? Do you have a figure of the domain that you used for Larsen C?
Ln 100-103: Any citation for the albedo scheme so that readers can investigate further?
Ln 106: change to 'and the difference between observed and simulated albedo values ($\Delta\alpha$)'
Ln 209: How often has this interpolation had to occur due to persistent cloud cover? How many days of missing values are there? How frequent is the MODIS overpath at this location and at what approximate time of day? How does the correction to MODIS vary during the winter with a much lower solar zenith angle for a persistent time?

Ln 212: I would specify that you mean the MLP model here- as RACMO2 is also a model, so the current sentence 'it is vital to assess the model performance' could be misunderstood as referring to RACMO performance.
Ln 217: Can additional surface melt only be positive? As in your introduction you mentioned that it was also important to correct RACMO where it overestimates melt. Is this why you set negative corrected melt to 0, so that melt as a whole cannot be negative, or only the change in melt cannot be negative. Perhaps this needs rephrasing.
Ln 218: Did you attempt to apply the model outside of the austral summer? Did you turn on the model specifically at December 1 and off again at February 28/29? What about in years where the melt season starts early or ends late, which can be the case (e.g 2010), especially in the presence of föhn winds: see King et al. 2017. (https://doi.org/10.1002/2017JD026809). What difference could the model have outside of the summer?
Ln 264: Earlier you say that the MLP is not applied outside of austral summer, yet here you discuss May and August. So is the MLP applied year round? Or are the winter values from RACMO without being corrected?
Ln 308: Include the R2 for RACMO too, so that the reader than read that correlations are higher.
Ln 320-331: Point to some figures or tables here, or include some results of statistics to back up your analysis, as it is currently quite qualitative.
Ln 357: Are you able to say why it was erroneously corrected?
Ln 361: Have others also found timing offsets between RACMO and observations previously? A citation would strengthen this section. Perhaps this is covered more in the discussion though.
Ln 371: In which year?
Ln 380-390: AWS18 is located in a region with blue ice, where albedo is generally low and the valleys are relatively narrow. It could be that RACMO2 fails to capture the blue ice zone and could also have land use discrepancies with the topography poorly resolved in 27km resolution. It could be useful here to mention the blue ice zone and/or include some references to this, as it would be unlikely if RACMO2 was able to represent these surface conditions.
Ln 435: 'actual albedo' is a little misleading as the reader is unsure whether this comes from AWS observations or MODIS. You have to read the Figure caption to understand. I would write (from MODIS) or something similar in the text.

Figure 7: What are the white and grey bars in the background? Include this info in the caption