The Cryosphere Discuss.,
https://doi.org/10.5194/tc-2020-159-RC1, 2020

Interactive comment on "Improved machine-learning based open-water/sea-ice/cloud discrimination over wintertime Antarctic sea ice using MODIS thermal-infrared imagery" *by* Stephan Paul and Marcus Huntemann

Anonymous Referee #1

Received and published: 3 August 2020

A review of tc-2020-159, by Stephen Paul and Marcus Huntemann

I have completed a review of this manuscript and recommend publication after minor revisions, most of which are just for clarification.

In this paper, the authors present a significant advance in the accurate detection of cloud in polar regions using advanced machine-learning techniques, and go on to illustrate the utility of this algorithm by retrieving much better estimates of thin ice thickness in the vicinity of the Brunt Ice Shelf, West Antarctica. The novel approach and subse-

quent advances in cloud detection are sorely needed for polar MODIS users, and as such I believe this paper is highly suitable for publication in TC, after minor modifications outlined below.

L22-29: It would be nice to point to other polar MODIS applications which would benefit from such a better cloud mask. Examples include composite image generation, landfast ice mapping, possibly sea ice motion retrieval using image cross-correlation.

L29: This problem has also been reported in coastal leads, e.g. Fig. 6 from Fraser et al., 2009. DOI: 10.1109/TGRS.2009.2019726

L40: Why was this study region chosen? Is this applicable for both flaw leads and nonlinear coastal polynyas (e.g., Terranova Bay)?

L54: At this stage, It strikes me that it might be better to describe input data before discussing the methods (i.e., move section 2.2 to here). This may have just been my personal preference though.

L112: This parapgraph needs more explanation. E.g., the 29-23... metric. Also, why are some numbers in bold type? What are the 35 epochs? 100 whats in a batch? Huber needs a capital H too.

Section 2.2: You need to describe the version of the MODIS products – particularly the MOD29 product. This determines which MOD35 version went into the product. There have been some significant improvements to MOD35 over the years, so it's important to document which version.

There is no description of the gridding or MODIS destriping/de-bowtie here. These must have huge influence on the performance of the algorithm, so a description of these processes is needed, in my opinion. Many of the channels used suffer strongly from detector striping in particular.

L140: What is the resolution of the IST product?

Section 2.2.1: The destriping description may fit better here.

L158: What about the increased atmospheric path encountered for high incidence angles – is that more important than the geometry distortion?

L164: Does MOD/MYD29 also apply atmospheric correction to more accurately determine IST?

L167: I guess you're developing this algorithm for coastal, latent heat polynyas. It might be good to make this clear here. I doubt it would work for offshore/sensible heat ones (which is fine)!

L178: This sentence "Generally, " is somewhat ambiguous.

Fig 3: One sub-figure would benefit from including a distance scale.

Fig 3: "Examplary" typo.

L219: The cal/val split was done on a point-wise basis? This seems a bit strange. Isn't the point of the cal/val split to ensure independence between the calibration and validation datasets by withholding at a more basic level, e.g., scene level? What I'm trying to say is, two neighbouring pixels are unlikely to be completely independent. So if there's a 75% chance of each pixel getting into the training dataset, then it's pretty much guaranteed that the validation dataset won't be particularly independent of the training dataset. Could you comment on this?

L278: Again, the Fraser et al., 2009 reference which shows this in a flaw lead would be good to reference here.

Fig 4, 5: There is a mismatch between the actual extent of the Brunt Ice Shelf and the masked version, based on the Rtopo product. This is due to ice shelf advection in the time between the creation of both products. In this case, there are both areas of ice shelf outside the mask, and areas of water/sea ice within the mask. Other high-resolution coastal datasets have mitigated this by including a manually-updated ice

shelf extent product on a regular basis (e.g., Fraser et al., 2020, ESSD Discussions, https://doi.org/10.5194/essd-2020-99), but this level of mitigation is probably unwarranted here. However, could you comment on the effect this might have on the training algorithm?

Fig 6: This is a great way of showing the seasonality in bias. However I'm still hanging out for a good old-fashioned scatterplot comparing these two datasets. This would show highly complementary information to your time series.

L286: I think the "average" metric you use here may not be the best way to highlight how much better your algorithm performs! Have you considered also using RMS difference?

L303: Unclear which product that this statement corresponds to.

L309: "Good agreement" between what and what?

L337: Again, the suggested RMS statistic would better highlight your improvement.

Interactive comment on The Cryosphere Discuss., https://doi.org/10.5194/tc-2020-159, 2020.