

Peer review of “Data-driven emulation of melt ponds on Arctic sea ice” by Driscoll et al.

The manuscript presents a Machine-Learning (ML) model intended for estimating melt-pond fraction on sea ice based on meteorological variables and developed and tested against spectrometer satellite data of Melt-Pond Fraction (MPF). The work appears solid but the manuscript is somewhat short on methodology and review of earlier work. In addition, the manuscript seems to claim that melt ponds have not been taken into account in climate models before, which is not the case. Here follows some mostly minor comments and suggestions that the authors can take into account before acceptance for publication is considered.

Comments and suggestions

1. In particular in the abstract line 5, the manuscript seems to claim that melt ponds are not included in GCMs, and that melt ponds therefore play no role in IPCC future projections of sea ice. This appears not to be the case, indeed many models have included parameterisations of melt ponds, see e.g. Sterlin et al, *Ocean Modelling*, 2021, and references therein, as well as Roeckner et al., *Journal of Advances in Modeling Earth System*, 2012, and Diamond et al, *Journal of Climate*, 2023.

In addition to correct this aspects of historical uses of melt-pond parameterisations in ECMs, the manuscript should also provide a much more extensive review of historical efforts to provide melt-pond parameterisations, including for instance reviews of Lühtje et al., *J. Geophys. Res – Oceans*, 2006, and Flocco and Feltham, *J. Geophys Res. - Oceans*, 2007.

2. Test against SHEBA. In order to test performance of the developed ML model with other types of data and to results from other parameterisations (e.g. Holland et al, *J. Clim.*, 2012), the SHEBA data may be applied.

3. Section 2.1 first paragraph: Expand the description of the MPF data, e.g. regarding how MPF is obtained from these data. Are there missing data and due to what?

4. L116-119: The conclusion regarding change in pattern between the two observational data set being due to instrumental differences or climate changes would be better explored if training data were chosen for only MERIS, 2002-2011.

5. First two paragraphs of section 3.2: Please more extensively describe the procedure of constructing the model, including finding the hyperparameters, in a way that none ML experts can follow. Shortly describe all “ML words”. “Plateaux” → “converge”.

6. L69-72: Please first describe the earlier studies before indicating differences to yours.

7. Fig. 1 caption: Explain how the standard deviation is calculated, what are the input data?

8. L124: Write out the abbreviation “DMIOI-L4”.

9. Fig. 2 caption: In the first sentence before “note that” indicate the period averages are taken over. The “note that” part is quite awkwardly formulated, perhaps “Note that if at least one observation exists for a given grid point this point is included”.

10. L135-137: The sentence is awkwardly formulated.

11. L155: What is “level ice”?

12. L189: “unseen” seems to be exaggerated given that The ML model build on observational input. Perhaps it suffice to conclude that the model can reproduce MPF associated with stochastic weather variability.

13. Fig. 5: Why not show root-mean square error (RMSE) of model results? Perhaps also show anomaly relative to climatology of observations, in order to compare (and hereby regard climatology as a persistence model).

14. L197: What is “R2 score”?

15. Fig. 6: In (b) the scale of the shading is not indicated.

16. L255-259: This paragraph is not so clear.

Typos:

L92: “data is” → “data are”, data are plural. Same in Fig. 2 caption. And Line 235.

L96: Remove “one”.

L130: “of”->”as”.

L245: “it is” → “that it is”.