Review of 'Reflective properties of melt ponds on sea ice' by A. Malinka et al.

The manuscript describes a new numerical model to calculate the spectral reflectance of melt ponds on Arctic sea ice, mostly determined by three independent variables. The authors find good agreement between simulated and observed spectra from in-situ measurements during three different field campaigns. This allows them to derive water depth, under pond ice thickness and transport coefficients for each of the ponds. Given the ongoing changes of the Arctic sea ice cover towards longer melt periods and increasing melt pond fractions, the manuscript describes a timely topic, which is well suited for publications in The Cryosphere.

Over all, I suggest publication after minor revisions, which mostly comprise some additional discussion and sharpening of the main conclusions.

General comments:

- It is not clear to me what the NEW elements of this model are, compared to existing models and theoretical approaches. It seems that most relations and assumptions are taken from existing studies. Since this is a mostly methodological manuscript, the following aspects need to become obvious:
    - What are the additional and new insights into radiative transfer of melt ponds?
    - How can or should this model be used in future (the outlook at the very end is rather unspecific and too general)?
    - What kind of scientific merit do the authors expect from this and following studies (applications of the model).
- The authors conclude that only three independent parameters are needed to characterize melt ponds and thus to retrieve an appropriate optical characterization from them. They do discuss and show results of pond depth and substrate thickness, but I am missing an analysis and more discussion and details on the transport coefficient. In that respect, the role of the three main parameters should be discussed in the discussion and be concluded at the end of the manuscript. How do they impact the model (not only in equations) and what sensitivity do we expect and receive?
- The comparison with in-situ observations show differences of under-pond ice thickness and water depth of 50% and some even significantly higher. I do not follow the argumentation that this is satisfactory, in particular since there is very little discussion about this (see comments below). I consider these differences as more significant than the discussion reveals. In particular with respect to the under-pond (substrate) thickness, which should be the most important parameter to determine pond albedo. Note: I am puzzled about the term "substrate". Why not under-pond ice thickness?

Specific comments:

Abstract: The abstract may be significantly improved by adding more results and a statement that explicitly names the additional benefit and further applications of the model:

- Page1/Line15 (P1/L15): … are examined: What is the result of the examination?
- P1/L16: several => three field campaigns
- P1/L17: "good performance" this is rather relative, good in what measure?
- Why are the three main parameters not mentioned in the abstract? How do they perform?
- What does this model stand out for and what is the (likely) future benefit of this study/model?

Introduction

- Recent studies by different groups show the increasing fraction and importance of melt ponds. Also shifts in melt onset and melt season duration are observed and discussed in various ways. I am missing this aspect in the introduction, while this would add to the motivation of this study and model development.
- In addition, there are various approaches to parameterize melt ponds in circulation models of various complexities. This should also be included and could even link to the role of light transmittance into and through sea ice (the remaining after reflection). This could also well link the introduction to the final part of the conclusions (see comments below)
- P2/L4: Include also "water" properties.

Model descriptions

- This section is most detailed. It could be improved by distinguishing better between existing models and theories and highlighting new ideas and findings.
- The role of the resulting three main parameters should be highlighted.
- It would add value to the manuscript if the model is made available for other users. How is the model implemented? How (numerically) costly are the simulations?

Model verification

- P13/L16-19: The realization of the validation and comparison should be described in more detail.
    - How did the authors derive that these are the three main parameters. What other parameters were analyzed?
    - What about the transport coefficients? How were they studied/discussed?
    - How are the thicknesses retrieved?
- It is a disadvantage that most ponds were not open ponds as it is assumed in the model development. I do see the constrains through the given data set, but this weakens the verification and needs more consideration. Why is there e.g. no thin surface ice in the model?
- P14/L14: Add the year (2008) into the main text.
- Section 4.4 should be the main discussion of the comparisons. This is too short and somewhat superficial.
    - Where do these rather large differences of 50% come from? I do see various reasons in e.g. pond depth distributions, non-planar interfaces, footprint of sensors compared to pond properties. But this needs to be discussed in more detail.

    - What precision may/can be expected in such models?
    - What determines the uncertainties? Which of the given assumptions might not be ideal, but what would it mean to adapt this? It is most likely not realistic within this study, but some additional discussion would be useful and interesting for further studies.
- With respect to those differences: As discussed, impurities are mostly low in the ponds, so the result is mostly based on scattering (not absorption). In this case, the retrieved spectral shape may be expected to be in good agreement, while amplitude is the main aspect of evaluation. But if then the simulated differences are still around 50% for the under-ice thickness this is somewhat surprising to me. I agree that the RMSE match is quite good if not excellent, but may be not because of the right thicknesses, but other parameters in the model. This should be discussed more.

Conclusions

- Given that ponds may be described by the three parameters: How would future applications look like? What is the main benefit from this conclusion? (P16/L16)
- P16/L27: This raises the question: How much of the model has been used before and what is new (see above)?
- P16/L30: “can be useful”: This is somewhat vague. How can it realistically be used?
- The last lines of the manuscript are not convincing to me. How would these improvements be implemented? What are the next applications or which part of these results is most promising. This needs a more thoroughly discussion and a more specific outlook.
- The conclusions section misses a conclusion on the uncertainties and deviations from the field measurements (Section 4.4). At the same time, I suggest to highlight that the validation was done against quite a suite of field measurements and variable pond conditions. This is a valuable aspect and could be stressed more. Many studies limit their validation to a single data set (e.g. one field experiment).

Table 1

- I think that this is not needed.

Table 2

- The pond code names seem to be an internal coding with almost no use for other studies. Using station names and dates as identifiers that link to field reports, Polarstern station numbers, and Pangaea data sets is suggested.
- I suggest to re-arrange the columns and group retrieved/measured/difference (absolute, and %) for each: ice thickness and water depth. This eases evaluation of the performance.
- RMSD values could be given in units of e.g. 10^-3 to save space and ease reading