Response to the comments of Reviewer 1

Please find our response to the comments from reviewers in **blue** font under the respective comments.

We thank the reviewers for the time and effort to thoroughly read and evaluate our manuscript, and we are grateful to the reviewers for the constructive comments and criticism about this manuscript which help us to improve our work. We have changed the content in the manuscript according to the comments and in the following we give detailed answers to the questions in the comments.

Anonymous Referee #1

Received and published: 30 September 2025

Abstract:

The statement of the research objective should be clearly articulated. The goal of investigating the optical and physical properties of refreezing ponds must be explicitly stated. What exactly is the spectral range? What are the surface characteristics of the ice-cover? This is not entirely clear in its current form. The abstract should clearly communicate that the primary objective is to generate a dataset focused on albedo-based classification of surface states, with particular emphasis on refrozen melt ponds (which remain under-documented compared to seasonal albedo variations). The exact goal should be distinctly stated—**what** and **why**. If there are multiple objectives, they could be presented in order of priority. For example, the development of the methodology could be considered one of the objectives, along with the spatial and temporal variation of albedo and the relationship between physical and optical properties.

Thanks for your constructive comments. We have rephrased the abstract to clarify the goals of this study and to add more specific information of the observation. The revised abstract is attached below:

"Sea ice plays an important role in the heat transfer into the Arctic Ocean whereas the presence of melt ponds on sea ice complicates the scenario. However, the refreezing pond is less focused and documented in comparison with the well-established seasonal variation. To better evaluate the effect of melt pond on the freezeup of sea ice, we conducted a series of observations with 81 melt ponds in the central Arctic during fall freezeup, 2012–2020. The melt ponds are categorized into five types based on the surface state to effectively investigate the various characteristics. The total albedo of each type is 0.14 (water pond), 0.20 (water-ice pond), 0.25 (ice pond), 0.39 (ice-snow pond), 0.74 (snow pond), respectively, showing the increase on albedo in

August and September (0.0036 d^{-1}) due to the changes of the surface state. The albedo dependence on the surface state, ice lid, pond depth and underlying ice is examined using both in-situ measurements and modified radiative transfer model, with result indicating the dominance of surface state followed by ice lid thickness. The total albedo of ice ponds decreases with increasing pond depth, and the raising of ice lid thickness reduces the albedo while rises that of ice-snow ponds. In addition, further analysis reveals the capacity of different ratios of spectral albedo on the distinction between snow-covered pond and unponded ice, potentially improving the melt pond retrieval algorithms."

General comments: The albedo categorization is unclear. What is the basis for the threshold? Since the distinction of 'classes' or types is vital, it may benefit from a brief explanation of the rationale behind this threshold. Why does refreezing specifically matter, and how does your classification fill a particular gap?

In this study, melt ponds are classified based on their surface states. We apologize for the misleading sentence in the abstract (Line 12, rewritten in the revised version), which was tend to emphasize the albedo difference between various types of melt pond, but not to describe the method of categorization. The surface state of each melt pond was classified based on both the description in the station documents and the recorded images during observation. To be specific, the state of melt ponds within a radius of 1–2 meters (depends on the field of view of sensors) centered on the projected position of the optical sensor is taken into consideration.

Previous studies have conducted extensive observations and analyses on the formation and evolution of melt ponds from June to July, and the temporal variation of melt pond albedo during this period is well understood. However, the surface states of melt ponds are almost always water during this stage. As a result, current understanding of ice lid and especially snow cover on melt ponds, which is less focused, is relatively limited, causing a lack of capacity to quantify the temporal variation of melt pond albedo during refreezing. However, melt pond is one of the main factors on the albedo variation of sea ice in this period, since unponded ice has similar albedo due to the snow cover later than mid-August. To gain more comprehensive understanding of sea ice, study on melt pond during refreezing is necessary.

Considering that albedo of individual melt pond is affected by weather such as wind and snowfall, thus it is hard to effectively predict its temporal variation during refreezing. But classifying the refreezing ponds allows us to acquire the range and the typical albedo of each type. By combining the proportions and albedo of different types, the temporal variation of the overall albedo of melt ponds can be better evaluated. Therefore, the study will enhance the understanding of sea ice albedo in the Arctic, and will contribute to the simulation or prediction on temporal variation of sea ice.

Based on the content above, we added more explanation about the categorization in Section 2.3, and we added sentences in Section 5 about the significance and potential use of this study.

Introduction:

- **Line 35:** "The surface albedo of sea ice in summer is largely determined by the melt ponds" — This is an overstatement, as bare ice areas still contribute significantly.
- Thanks for the correction. This sentence has been changed to "...the surface albedo of sea ice in summer is significantly affected by the melt ponds".

- **After Line 45:** I would suggest adding 1-2 sentences on why the refreezing period specifically matters (e.g., the transition from net melting to ice growth and model parameterization challenges during shoulder seasons). This addition would provide better context and emphasize the importance of documenting this stage. This also connects more smoothly to the section discussing existing knowledge gaps.
- We have added some sentences to explain the importance of refreezing period at the end of Line 45: "Thus, the thermal characteristics of summer sea ice strongly depends on fraction of melt ponds, and it remains remarkable during refreezing period when the transition from ice melting to ice growth occurs. The ice growth at the base of the underlying ice is limited due to latent heat stored in pond water which is trapped by ice lid until the pond freeze completely (Flocco et al., 2015). Furthermore, these various thermal processes also pose challenges to the accurate simulation of models.".

In-situ Data:

- The comprehensive spatial coverage, especially from multiple Arctic basins, is commendable.
- Thank you for your recognition. We will further expand the coverage of this dataset in subsequent observations and studies.

- **Line 66-67:** Could you briefly explain the reasoning behind the constraint imposed on data collection (overcast skies), defined by visual observation? Were clear-sky conditions excluded? If so, could you explain why, since this would exclude direct-beam scenarios?

- Yes, the data obtained under clear-sky are excluded and there are several reasons. Firstly, both some of the authors and the former reviewers suggest that the albedo of melt ponds under diffusive light and direct light differs from each other. Therefore, the data used in this study should be measured under the same light condition. Secondly, most ice stations were set during overcast weather since it's overcast at most time of the mentioned Arctic expeditions in this study. The melt ponds observed under clear-sky are too few (0~2 each cruise, 5 in total) to get convincing results. In order to well illustrate the characteristic of melt ponds, observations under diffusive light were used. Besides, the time with clear skies is only 6% in August and September in the Arctic (Grenfell and Perovich, 2008), so the albedo under diffusive light is more representative of melt ponds in this period.

- **Line 35:** The term "Surface scattering layer" — It might help to provide a brief definition for non-specialist readers.
- A brief definition has been added to help understanding this term: "a thin upper layer of sea ice with high scattering capacity for incoming solar radiation".

Methodology:

- **Lines 65-66:** Was the constraint on data collection (overcast skies) based on visual observation? If clear-sky conditions were excluded, could you explain why, since this would exclude direct-beam scenarios?
- The observations were conducted no matter the weather was sunny or overcast. But the only data under overcast conditions were selected (based on the in-situ document, both visual record in pictures and weather record in texts) in this study to avoid the distraction on albedo caused by the difference between diffusive light and direct light. Please also see the detailed reply to comment "Line 66-67" in part "In-situ Data"

- **Lines 84-87:** It should be clearly stated which physical properties were measured (and how), and which instruments were used to measure ice lid thickness and pond depth. This would improve reproducibility and help clarify the relevance and quality of the collected data.
- Thanks for your comment. According to the in-situ record, ice lid thickness was measured by a metal ruler, the pond depth was measured by a metal tape measure and the substrate ice thickness was measured by an ice thickness gauge. The detail of measurement on physical properties such as ice lid thick

or ponds depth has been added: “the ice lid thickness was measured by a metal ruler, the pond depth was measured by a metal tape measure, and the substrate ice thickness was measured by an ice thickness gauge”.

- **Line 63:** Multiyear ice is mentioned. Were all stations clustered in certain regions or ice types? Could this potentially impact the representativeness of the data?
- The locations of ice stations were randomly picked and distributed in different regions (see Figure 1) so there is no impact on the representativeness of data at this aspect. All stations were set on multi-year ice, since seasonal ice in August or September is too thin to allow ice field operations on it. And the refreezing process of melt pond, especially the variation of surface condition is slightly affected by the age of sea ice since it depends mainly on the weather condition and events, so observations on multi-year ice still provide data reliable enough to support the results in this study.

- **Line 86:** Were locations randomly selected? If there was a deliberate choice behind location selection, could this introduce bias or affect the representativeness of the data?
- All the locations of ice stations were randomly selected with no deliberate choice.

- **Table 1:** It would be helpful to include information on which year each instrument was used (e.g., CNr4 measurements applied in 2012, 2014, and 2016). In the text, this could be elaborated to clarify the distinction between different radiometers and their corresponding wavelength ranges.
- We added information of instrument and the corresponding wavelength in the caption of Table 1 to clarify.

- **Lines 164-65:** The calibration process is improved by using different coefficients for different surfaces. However, is there a potential circular dependency? That is, could there be a risk of misclassifying a pond initially, which would then lead to the application of the wrong calibration coefficient, reinforcing the misclassification? Also, how confident is the visual classification? Could this be slightly subjective? It might be worth exploring this further.
- Thank you for the comments. As mentioned before, the classification is based on the description in the station documents and the recorded images during observation, not the value of albedo. That is, the change on albedo does not

affect the classification, so there is no potential circular dependency. The description in the station documents recorded by field operators may be slightly subjective but we also double-checked and classified all ponds based on photos recorded during observation (which will be provided as part of the dataset) as objective evidence.

- **Lines 165:** Water ponds or water-ice ponds? It might be useful to clarify this distinction.
- The calibration in Section 2.3 is aimed to solve the problem of poor consistency between the two instruments in the latter three categories (i.e. ice pond, ice-snow pond and snow pond). It can be seen that the deviation of the two instruments increases along with the formation of ice and snow, so the deviation of water-ice pond (<0.01) is even smaller than the average deviation of the three calibrated types (ice, ice-snow and snow) after calibration (0.049, 0.017, 0.104). In this case, we did not perform calibration for water-ice pond. Besides, the water ponds were only observed by one instrument so no calibration is performed.

Results:

- **Lines 176-177:** It would be preferable to avoid speculating about the reduction or slowdown unless statistical analysis is used to account for differences.
- Thank you for pointing this out. We removed the speculation about the reduction and only retained the objective comparison of different observations.

- **Figure 6:** Consider increasing the font size at the top of the graph for better readability.
- We increased the font size of Figure 6a.

- **Lines 189-194:** Could you clarify whether these percentages are based on the number of observations or area coverage?
- The percentages referred here are based on the number of observations. We added a statement to make it clear for readers.

- **Section 3.1:** A clearer separation between observations and interpretations would be beneficial. Consider distinguishing between: 1) Measured albedo changes; 2) Observed surface state changes and 3) Mechanism hypotheses

- Thanks for your valuable suggestion. We reorganized part of the sentences accordingly in Section 3.1 to improve the readability.
- **Line 215-217:** Why emphasize the SHEBA comparison if ice type is a key variable? It seems that since 0.14 is equidistant from 0.12 and 0.2 but considerably different from 0.4, it might be better to state that the results align more closely with the dark pond literature, rather than general melt pond values.
- We are sorry for the confusion caused by a missed key description here. The value 0.4 refers to the albedo of light ponds observed in SHEBA, while the albedo of dark ponds observed at the same ice station is ~0.2, which is consistent with other observations mentioned later. We corrected the statement to emphasize the alignment with dark pond literature. Besides, the result in SHEBA also shows that sea ice type is not the dominant factor in the albedo of melt ponds. The mention of sea ice type here is intended to describe the observation detailly, but now we have realized that it is unnecessary and may cause ambiguity, so we removed description about ice types in this statement.

- **Line 237-238:** Please carefully identify the source of noise. It is important to accurately describe instrumental limitations rather than referring to solar physics.
- This sentence is intended to state that the irradiance in ultra-violet wavelength is close to the minimum resolution of the instruments, thus the instrumental limitation caused the noise. We rewrite the statement to better describe this: “Since less energy of solar radiation concentrates in ultra-violet wavelength, approaching the minimum resolution of instruments and therefore leading to noises in 320–350 nm....”

- **Lines 385-393:** You claim that >90% of ponds are frozen by late August/early September, yet Figure 6a shows only around 30% ice ponds and approximately 50% ice-snow/snow ponds by early September (Lines 190-191). Please reconcile these numbers or clarify what "frozen surface" refers to—are water-ice ponds counted as frozen?
- We apologize for the misunderstanding caused by the ambiguous meaning of the term "frozen" here. The use of "frozen" is intended to distinguish ponds with surface partly or completely covered by ice lid (both bare and snow-covered) from ponds with completely unfrozen surfaces, which have been extensively studied. Therefore, water-ice ponds are also counted as frozen. We changed

the description to “ponds with surface partially or completely frozen” to clarify.

- **Lines 418-424:** The critical finding that existing MPF algorithms (PCA, LinearPolar) fail to identify snow ponds and underestimate MPF during freeze-up is currently buried. This finding could be elevated as a main result rather than being relegated to the discussion section.
- We thank you very much for your positive evaluation. However, only two commonly-used melt pond retrieval algorithms are compared so the evaluation is not comprehensive for all current algorithms. Additionally, the number of snow pond observed is kind of small to provide a more robust result, which is also a limitation we regret. For the reasons mentioned above, the statements here is not sufficient enough to become part of the results as we evaluated, so the purpose of this section is to explore the misdetection issue of commonly used algorithms on snow ponds and provide potential directions of improvement. Thus, it is more appropriate to be included as part of the discussion section. We will further investigate the limitation and the potential improvement on MPF algorithms in our future work.
- Besides, to draw readers' attention to this content, we also mentioned and emphasized it in the conclusion section as follows: “The spectral distributions of snow pond is distinctive by its relative high values from other ponds and the increase at short wavelength (<550 nm) from unponded ice, causing misidentifications of MPF algorithms but also enabling the effectivity of indicators such as α360/α490 and α412/α667, which may help to the further improvement of melt pond algorithms”.

- **Lines 441-445:** Could the error from observations (which your model improves upon) be quantitatively shown?
- We added quantitative results about error from observations in Figure 11a. The statement about reduced error is also added in Section 3.4 and in the summary: “…the average RMSE (root mean square error) of all ponds reduces from 0.168 to 0.026 after the modification”.

- **Lines 429-440 (Summary):** The summary appears to introduce new information that was not present in the results section. For instance, the comparison of slower observations in the Amundsen Basin (Line 437) seems to be new. A summary should synthesize previously presented findings.
- Thanks for pointing out this issue. We added statement about comparison with

observations in Amundsen Basin in Section 3.1. Besides, we also checked the rest of the summary section to ensure every statement is introduced in result section before presented here.

- **Line 439:** Typo: "aα412/α667" should be corrected.
- Thank you for pointing this out. We have corrected accordingly.