Response to the comments of Commenter 1

CC1 General Comment: This is an impressive piece of work and I am glad that it is already online as a citable discussion paper. I went through the manuscript - not necessarily with a reviewer's eye but with the intention to learn something new and to provide ideas how readability and usability of this publication could eventually be improved further. Perhaps one or more comments could assist here.

We would like to thank the Commenter for his very helpful insights, which helped us shape further the structure of the manuscript while taking additional care on different aspects that strengthen the manuscript. We are very grateful for the contribution, which we carefully addressed and took into consideration. Our responses to each point suggested by the Commenter appear below in **blue text**, positioned directly beneath each respective comment.

- L1-2: This is a very high-level title which, at least from my side, raises very high expectations. While you manage to come up with an impressive compilation of various data sets and also attempt to describe sensing and processing techniques I suggest to come up with an alternative structure to guide the reader a bit better. I was wondering whether it would make sense to begin - as you did - with the relevance of melt ponds for studying the Arctic climate and selected (!) studies where observations of melt ponds changed and/or influences our knowledge about Arctic summer sea ice conditions. I think in that section if would not matter whether you are referring to satellite, airborne or ground-based studies - simply because all have their different application areas and examples. Then, I would come up with the section where you describe the 3 main different observational tools: ground-based, airborne, spaceborne in 3 sub-sections. In each of these you would refer to the measurement technique, provide a list of the sensors and their characteristics, experiments / expeditions / satellites and at the end of each of these 3 subsections come up with the limitations of use / knowledge gaps / and room for improvement.
- Following this structure suggestion together with inputs. I would try to provide tables and/or appendices that are clearly linked to these 3 sub-sections, i.e. in-situ, air-borne and satellite.

We appreciate this thoughtful structural suggestion to organize by observational platform, and we have carefully considered this alternative organization. Section 1 (Introduction) was revised to better consolidate the motivations for melt ponds, while outlining the structure of the manuscript, and Section 2 (Melt Pond Properties and Seasonal Evolution) was updated to integrate their relevance and key findings. These two initial sections underwent significant changes, following the Commenter's suggestion (while also considering Reviewers 1 and 2, by incorporating an overview/synthesis from melt pond studies - which were previously examined in a dedicated Section 4.2, and are now part of Appendix C).

Following suggestions from Reviewers 1 and 2, we have improved the logical flow of Section 3 within its existing structure by clearly distinguishing between spaceborne observations (3.1), in situ/field campaigns (3.2), and processing techniques (3.3), while strengthening the links between them. Specifically, we have added a synthesis paragraph at the beginning of Section 3 that previews overarching challenges with forward references to where each is

addressed in detail for each sensor type. We have improved consistency in terminology (i.e., using EO as a broader term and 'remote sensing' when referring specifically to satellite/airborne observations, with explicit definitions in the introduction). We also enhanced cross-referencing between sections to make connections more explicit, particularly linking Section 3.1 (sensor capabilities) to Section 3.3 (processing techniques) and to Section 4 (datasets). This revision also ensures that the focus of this section stays on existing datasets. These changes directly address the reviewer's comment regarding the need for clearer linkages to the main text. We hope that the substantial structural revisions, including the platform-level organization emphasized by the reviewers, and the prioritization of physical measurement principles, ensure consistency with the reviewers' feedback while improving the overall flow and readability of the manuscript.

- I don't think it makes a lot of sense to try to distinguish between pan-Arctic and regional spaceborne applications because in the long run, applications such as from Sentinel-2 MSI might become pan-Arctic as well once there is enough coverage. I guess, if described properly, readers will understand that 10 km x 10 km large super-high resolution satellite images are not pan-ARctic and are not suitable for climate studies but - like the air-borne data sets - are perfect for algorithm development and evaluation.

We have considered this feedback alongside suggestions from Reviewers 1 and 2. While maintaining the pan-Arctic versus regional categorization for its current practical utility in distinguishing dataset characteristics and intended applications, we acknowledge the Commenter's important point about the evolving nature of these distinctions. We have added a clarifying statement noting that certain platforms currently categorized as regional (such as Sentinel-2 MSI) may achieve pan-Arctic coverage in the long term as data accumulation continues, and that this distinction should be understood as reflecting current data availability rather than fundamental sensor limitations. The revised text ensures that all spaceborne applications (such as S2 MSI-based products, which were also pointed out by the Commenter) are appropriately categorized, with clear descriptions of their spatial coverage, temporal extent, and suitability for different research applications (e.g., algorithm development and validation versus long-term climate studies).

- Figure 4: Signature row needs more information; mixes onset and mature melt pond existence; overly simplistic and partly misleading

We recognize that Figure 4 attempts to synthesize complex, stage-dependent signatures into a single schematic, inevitably simplifying the temporal evolution. To address this limitation, the text preceding Figure 4 has been revised to explicitly acknowledge (i) temporal variability between melt onset and mature pond stages, (ii) signature sensitivity to environmental conditions, and (iii) dependence on ice type (FYI versus MYI). Readers are also explicitly referred to Section 3.1, where these aspects are discussed in greater detail.

In response to the reviewer's concern, Figure 4 has been further revised to improve clarity and completeness: (i) clarification of onset detection applications was made within the image; (ii) radar altimetry has been added, (iii) the figure layout has been restructured to more clearly distinguish between *retrieved parameters* (i.e., directly measured physical

quantities) and *main applications* (i.e., higher-level derived uses) and (iv) the rows containing mission examples, retrieved parameters, signatures, and main applications, derived from our systematic literature review and supported by Appendix C (formerly Section 4.2), have been refined to ensure better separation and consistency across sensor types.

Finally, the figure caption has been revised to better contextualize the schematic nature of the figure and to clarify its intended scope and limitations.

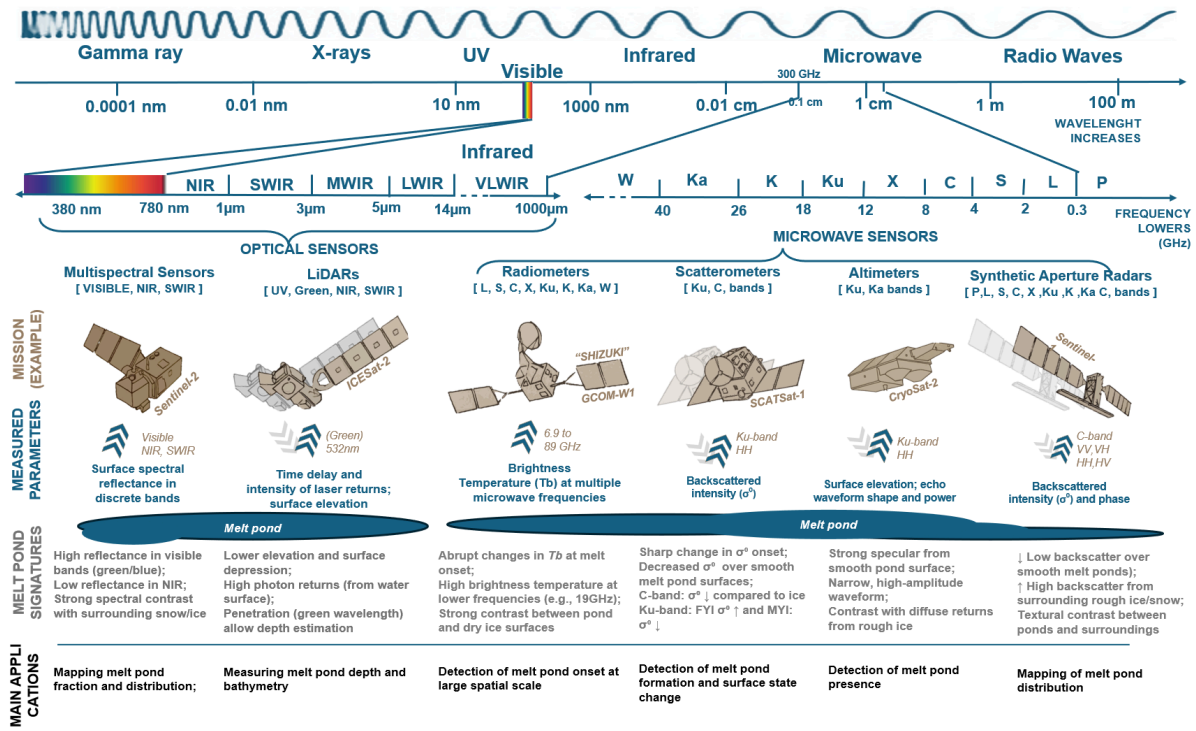

Figure 4: Overview of Earth observation methods for melt pond detection across the electromagnetic spectrum. The figure shows six sensor types (multispectral sensors, LiDARs, radiometers, scatterometers, radar altimeters, and synthetic aperture radars) organized by wavelength range (optical and microwave). For each sensor type, the figure presents: (top row) example satellite missions and operating wavelengths; (middle row) measured parameters and physical principles; (second last row) characteristic melt pond signatures and (bottom row) main applications.

- L196: Add Lee et al., 2020

Added.

- L249-252: Clouds are an issue for ICESat-2; limited daily coverage may miss seasonal development steps

We thank the Commenter for this relevant note, which we added to the revised text.

- L254-256: Statement about microwave atmospheric independence not entirely correct; water vapor affects ~22 GHz and ~90 GHz; wind and salinity effects

The text was updated to correct microwave atmospheric independence, and addresses water vapor, wind-induced surface roughening, and water salinity effects.

- L295++: Be more specific about penetration depth (dry ice/snow only, not water; melt pond depth has no influence on radar backscatter); clarify "seasonality"; provide rule of thumb about wind-roughened vs. smooth pond detection

We expanded on this by clarifying the actual influences of penetration. The term "seasonality" was further clarified following the Commenter's suggestion. Additionally, we added guidance on the detection thresholds for wind-roughened versus smooth ponds.

- L360-369: Don't mix land surfaces and sea ice; concentrate on sea ice; mention sensors first, then applications

L360-369 was revised to concentrate on sea ice and maintained consistency with the other missions' descriptions.

- L378-380: Check The Cryosphere 2016 and 2020 for more recent literature on melt pond uncertainties in SIC retrieval; 2nd sentence needs reference

More recent literature from The Cryosphere (Kern et al., 2016, 2020) was included addressing melt pond uncertainties in SIC retrieval, and references were added to the previously uncited sentence.

- L489/490: Link does not work

Link updated.

- L500: Link back to earlier sensor descriptions; structural improvement needed

We added explicit links to clarify that this section presents the main technical approaches for processing melt pond observations (detailed in Appendix A), building on the sensor characteristics and signatures described in Section 3.1 (spaceborne observations) and Section 3.2 (in situ and field campaign observations). The section was restructured to establish clear connections between processing families (Figure 8) and sensor types and measured parameters introduced in Figure 4.

- L501: "Early approaches" doesn't make sense when citing 2015+

Changed.

- 505: Kern et al. 2016 did not develop the approach; should cite Tschudi et al., 2008 and Rösel et al., 2012

Corrected.

- L515-527: Combine methods and data sources (satellite sensors) in one go; clarify organizational logic

We added transitional text clarifying the organizational logic of the section and improving readability. The revised structure now presents methods and data sources within a sensor-specific framework progressing to multi-sensor integration and sensor-independent approaches, reflecting the methodological evolution of the field.

- L563/564: No systematic independent intercomparison exists; be careful about accuracy claims; dataset producers have biased views

We appreciate this note, and the text was revised addressing the dataset common challenges; the lack of a systematic independent intercomparison study; and mentioning that datasets with reported high accuracy were based primarily on validation by the data producers.

- L602: Refer to datasets by author team name unless there was an actual expedition

Updated.

- L602-615: Don't mix satellite, airborne, and ground-based observations in this section

L602-615 were restructured to group them by platform following the suggestion of the Commenter, improving readability.

- Section 4.2: Goes beyond review scope; should focus on datasets and processing approaches; much repetition of earlier content; place 1-2 key studies per dataset upfront

We have carefully considered the Commenter's suggestions and implemented a substantial revision. Specifically, Section 4.2 has been relocated to Appendix C, in alignment with Reviewer 1's feedback. Although this appendix extends beyond the immediate scope of the review, it compiles over 40 studies on melt ponds, providing valuable context that reinforces both the motivation and relevance of melt pond research. Furthermore, it offers insights into the methodological evolution within the field and aids in identifying future research directions, thereby establishing clear links across the manuscript. Throughout the text, readers are referred to this appendix for further consultation.

- L703-707: Doesn't this downgrade SHEBA and MOSAiC importance? Their role needs emphasis

The text was revised to ensure it does not unintentionally downplay these critical campaigns, while emphasizing their irreplaceable role in providing detailed measurements that cannot be obtained remotely, and their importance for melt pond studies. Moreover, following additional suggestions from Reviewers 1 and 2, this section was revised to better balance the level of description across campaigns and to explicitly highlight how each advanced our understanding of melt ponds by including their key findings.

- L757-762: Little hope to disentangle SAR signatures between melt-pond-free ice, melt ponds, and leads

Indeed this is a fundamental challenge, which has been now emphasized at L757-762

- L765: (1) SIA is physically correct quantity, not SIE; (2) Melt ponds don't influence SIE computation (>15% threshold); they influence SIA; this lack of SIC accuracy drove community to SIE over SIA

We appreciate this correction, which has been made in the revised version of the manuscript.

- L863-865: (1) Are commercial satellite images openly accessible? Coverage? (2) Emphasize the impressive amount of such data available; (3) These are NOT in-situ measurements—they require retrieval; communicate uncertainties; OIB validation is inter-comparison, not true validation

We addressed each point: satellite openness, coverage, licensing needs, the volume of available data were included. We explicitly noted that both airborne and satellite MPF products are retrieved values (not in situ measurements) requiring algorithms with associated uncertainties. We clarified that comparisons between satellite-based and airborne-derived MPF constitute inter-comparisons between two retrieval products at different spatial resolutions, not true validation against ground truth measurements.

- L885-890: Effort required not worth the outcome; better to have 2-3 datasets for intercomparison to assess uncertainty

The text has been revised to clarify that standardization is intended to *enable* efficient intercomparison of a small number of representative datasets (e.g., 2–3), to assess classification uncertainty, rather than focusing on a single dataset or exhaustive analyses.

- Appendix B: (1) Check names (TransArc may be in IceWatch); verify IceWatch availability to present; (2) Rename "PANGEA" dataset since multiple datasets are in PANGAEA database; (3) Check ship-based observation link (https://www.cen.uni-hamburg.de/en/icdc/data/cryosphere/seaiceparameter-shipobs.html); (4) Is list exhaustive? Why two appendices (B and C)? Consider combining

(1) Appendix B has been revised to correct and clarify dataset names. We have verified IceWatch availability, and added the recommended ship-based observation link. We thank the reviewer for this suggestion.
(2) PANGEA dataset name was revised.
(3) The list is not intended to be exhaustive but representative.
(4) Appendices B and C are retained because they serve different purposes: Appendix B summarizes observational datasets, while Appendix C lists research studies and applications, some of which use or led to the datasets in Appendix B.