Response to Referee #3

In this document, we outline our responses to comments from the referee, including modifications that we intend to make to the manuscript where necessary.

Referee comments are shown in black and our responses in blue.

Review report on "Year-Round High-Resolution Sea Ice Freeboard Retrieval Using ICESat-2 ATL03 Photon Data" by Liu et al.,

Sea ice freeboard is a critical parameter not only a proxy for snow depth estimation, but also critical for sea ice thickness extraction. This manuscript introduces the High-Resolution Freeboard Method (HRFM), a novel framework that retrieves sea ice freeboard directly from ICESat-2 ATL03 photon data at 5 m along-track resolution. The approach combines a two-stage denoising strategy (Kernel Density Estimation with adaptive histogram-based filtering) and a machine-learning surface-type classifier trained on 25 coincident Sentinel-2 scenes covering winter and summer conditions across diverse ice types. The authors validate surface heights against ATM (Airborne Topographic Mapper) data, compare their freeboard estimates with the operational ATL20 product, and demonstrate improved performance in preserving ridge heights, mitigating after-pulse effects, handling melt-season complexities, and enabling consistent weak-beam utilisation.

The study is highly relevant to the scope of TC Journal and timely to the cryosphere society, addressing a critical need for year-round, high-resolution sea ice thickness monitoring. The manuscript is well-structured, the methodology is clearly described, and the results represent a substantial advance over existing ICESat2 sea ice products (ATL07/ATL20), which convinced me that this study warrants publication in TC Journal. However, before final acceptance. There are numerical issues that either need further clarification or significant improvement. Please see my major and minor comments below that I hope the authors can consider during the revision.

We sincerely thank the reviewer for the positive evaluation of our manuscript. We're grateful to the reviewer for taking the time to check our manuscript, providing valuable comments that improve the clarity of the work.

Major Comments / Required Revisions

1. The authors should specify explicitly what annual cycle was investigated either in the title or in the abstract. Why was this annual cycle selected in this study?

The primary objective of this manuscript is to develop and evaluate the HRFM method, rather than to investigate the interannual variability of sea ice freeboard in a specific year. We are currently extending HRFM to all available ICESat-2 years to generate a comprehensive multi-year high-resolution sea ice freeboard dataset.

2. The key algorithm parameters are mentioned throughout the text (e.g., horizontal scaling a=25/100, histogram bin size 0.1 m, after-pulse offsets 0.45/0.9 m, 10-km window for sea-level reference, 3σ outlier removal, 10th percentile density threshold for refined denoising), a single table summarising all critical parameters with their values and justifications would greatly enhance reproducibility. Please add such a table (e.g., in Section 3 or as an appendix) in the revised manuscript.

We thank the reviewer for this valuable suggestion. In the revised manuscript, we will add a table in Section 3 to summarize the critical parameters.

3. In the abstract and results (Sect. 4.1), the authors state that weak-beam retrievals achieve "comparable accuracy" to strong beams. How "comparable" it was? Please provide a concrete assessment. Please consider adding a statistical test to show whether the remaining difference is meaningful or not.

In general, weak beams are expected to perform less favorably than strong beams because of their lower transmitted energy and consequently lower photon return rates. In the original manuscript, this statement that weak-beam retrievals achieve comparable accuracy was intended to indicate that the weak-beam results show similar performance to the strong-beam results when evaluated against independent reference data, including both the surface-height comparison with ATM measurements and the cross-validation of surface-type classification. We will report the validation statistics separately for strong and weak beams in abstract directly to improve clarity.

4. Table 2 shows thin-ice recall is only 0.50 (strong beam) and 0.43 (weak beam), with precision ~0.65. The authors merge thin ice into sea ice for two-class validation (Table A1, overall accuracy >0.98). However, thin ice is a distinct surface type with different radiative, thermodynamic, and mechanical properties. Please discuss:

a: How often does misclassified thin ice affect freeboard retrieval (e.g., when thin ice is wrongly labelled as lead and thus included in the reference sea level)?

b: Should users treat the thin-ice class as "experimental" or apply the two-class version for freeboard estimation? A clear recommendation would be helpful.

In this study, the three-class experiment was mainly designed to examine whether ICESat-2 ATL03 photon-based features have the potential to distinguish thin ice from leads and snow-covered sea ice. The results in Table 2 indicate that such potential exists, but they also show that the thin-ice class is still difficult to identify robustly. Thin ice can have photon-rate, density, height-variability, and background characteristics that overlap with both leads and snow-covered sea ice, which explains the relatively low recall and moderate precision reported in Table 2.

Because of this limitation, the operational HRFM freeboard retrieval in this manuscript does not rely on the three-class classifier. Instead, thin ice is merged into the sea-ice category, and a two-class lead/sea-ice classifier is used for freeboard estimation. This choice is made specifically to reduce the risk that thin-ice surfaces are incorrectly treated as leads and then used in the reconstruction of the local reference sea level. Therefore, the low thin-ice recall in the three-class experiment does not directly propagate into the freeboard estimates presented in this study.

Regarding point (a), if the three-class classifier were used directly for Arctic-wide freeboard retrieval, misclassified thin ice could affect the reference sea level when thin ice is incorrectly labelled as lead. Since thin ice generally has a positive freeboard relative to the local sea surface, including such segments in the lead samples would tend to raise the estimated reference sea level and could consequently lead to an underestimation of surrounding sea-ice freeboard. However, for large-scale retrievals it is difficult to quantify this effect rigorously because the true lead/thin-ice identity is not known outside the labelled validation scenes.

Regarding point (b), we will provide a clear recommendation in the revised manuscript. For sea ice freeboard retrieval, we recommend using the two-class lead/sea-ice classifier.

5. The reference sea level is computed by aggregating lead segments within a 10-km along-track window. The choice of 10 km is plausible but not justified. Could the optimal window size vary with ice regime (e.g., compact vs. marginal ice zone, winter vs. summer)? A brief sensitivity test (e.g., 5 km, 10 km, 20 km) on a representative track would strengthen the method. If such a test already exists, please cite or summarizeit.

We adopted a 10 km window primarily to be consistent with the ICESat-2 ATL07/ATL10 sea ice freeboard algorithm (Kwok et al., 2026) and with the physical rationale discussed by Kwok et al. (2019). The 10-km length is selected to minimize the contribution of residual sea surface slopes to uncertainties in freeboard calculations; the general expectation is that the local sea surface height is relatively constant over a length scale corresponding to the first mode baroclinic Rossby radius of deformation, which is on the order of 10 km for polar latitudes above 60° latitude.

6. Figure 10 shows seasonal mean differences between HRFM and ATL20 up to ±0.04 m, with sign reversals between winter and summer. The authors attributed these to differences in surface-height retrieval, lead detection, and reference sea-level construction. A rough quantitative decomposition (e.g., how much of the winter difference is due to after-pulse removal vs. lead sampling density) would greatly

strengthen the discussion. Even an approximate breakdown based on the examples in Figs. 8 and 12 would be valuable.

We agree that a quantitative decomposition of the seasonal differences between HRFM and ATL20 would strengthen the interpretation of Fig. 10. However, a rigorous component-by-component decomposition is challenging because sea ice freeboard retrieval is a sequential and highly coupled procedure. Differences introduced at the initial signal-photon identification stage can propagate into the segment height estimates and also affect the derived features used for surface-type classification. The estimated surface heights further influence lead identification and the reconstruction of the local sea-surface reference. Therefore, the effects of after-pulse removal, surface height retrieval, lead detection, and sea-level reference construction are not fully independent and cannot be simply separated in an additive manner.

In addition, HRFM and ATL07/ATL20 use substantially different processing strategies, and ATL20 does not provide all intermediate variables required for a strict attribution analysis, such as the photon-level filtering results, intermediate lead-selection decisions, and the exact local sea-surface reference construction at each processing step. This further limits the possibility of performing a fully or partly quantitative decomposition of the HRFM–ATL20 differences.

Nevertheless, we agree that even an approximate analysis would be useful. In the revised manuscript, we will attempt to provide a rough quantitative assessment based on representative cases. If the resulting estimates are physically meaningful and robust, we will include them in the revised discussion. Where a strict decomposition is not possible, we will explicitly state this limitation.

7. The classifier was trained on scenes south of 82°N because of Sentinel-2's orbital limit. How confident are the authors that the classifier works north of 82°N (e.g., the central Arctic) where no coincident optical imagery is available for validation? Please add a discussion of potential extrapolation risks and whether the physical feature pace (photon rate, density, height STD, background rate) is expected to remain valid poleward.

Although direct Sentinel-2-based validation is not possible north of 82°N, the monthly mean freeboard fields retrieved by HRFM show spatial patterns that are broadly consistent with the ATL20 products over the entire Arctic, including regions poleward of 82°N. This agreement provides indirect evidence that the classifier and the subsequent freeboard retrieval framework remain effective in the central Arctic. Based on the current experiments, background rate is the feature most likely to become less reliable under low solar elevation conditions. We have described the limitation in Section 5.3.

Minor comments

Except for the points that are addressed separately below, we will revise the manuscript following the reviewer's suggestions.

-Line 70: "detector after-pulses" should be introduced earlier (it is defined in Sect. 3.1.2 but referenced here). Consider adding a brief definition at first mention.

-Line 125: Use consistent version notation (e.g., "Version 6" rather than "V06", or define "V06" at first use).

-Figure 8 caption: "The green dashed lines represent the ATL07 signal selection envelope" :please describe this explicitly in the caption.

-Section 5.1, line 585: "lower transmitted energy level (~80% of outer beams)" – please verify and provide a citation for the energy difference between middle and outer beams.

This is reported by Kwok et al. (2019): The photon rate distribution for the weak beams (GT1L, GT2L, GT3L—dashed lines) are consistently at ~1.5 photons/shot. The strong beam photon rate distributions (in GT1R, GT2R, GT3R—solid lines), however, indicate that GT2R (Beam 3) has consistently weaker surface returns (~0.81 of GT1R and GT3R): The mean rates from GT1R and GT3R are ~6.8 while those from GT2R are ~5.5. This is attributed to the expected variations in the custom construction of the optical component used to split the laser energy into the six beams. These values are consistent with pre-launch expectations.

-Figure 5 caption: "It's a 10-km profile" should read "It is a 10-km profile"

-Language: The manuscript is generally well written, yet there are a few grammatical issues that remain (e.g., "It's" in figure captions, occasional missing articles). I see you have English co-authors, and please ask them to proofread the language.

We will carefully proofread the revised manuscript to correct the remaining grammatical issues.

- The text fonts in almost all figures are quite small. Please consider enlarging them to improve the readability of the figures

- The authors provided this data availability statement: The Sentinel-2 imagery was derived from Google Earth Engine at https://developers.google.com/earth-engine/datasets/catalog/sentinel-2. I am not sure if this is adequately enough by TC, or should authors provide the Sentinel-2 imagery data somewhere for readers to explore. I leave this for the TC handling editor to decide.

We will improve the data availability by providing a detailed list of the Sentinel-2 scenes and the corresponding ICESat-2 tracks used for classifier training and validation, as shown in Table 1 in our response to Referee #2.

References

Kwok, R., Markus, T., Kurtz, N. T., Petty, A. A., Neumann, T. A., Farrell, S. L., Cunningham, G. F., Hancock, D. W., Ivanoff, A., and Wimert, J. T.: Surface Height and Sea Ice Freeboard of the Arctic Ocean From ICESat-2: Characteristics and Early Results, JGR Oceans, 124, 6942–6959, https://doi.org/10.1029/2019JC015486, 2019.

Kwok, R., Kurtz, N., Wimert, J., Petty, A., Cunningham, G., Markus, T., Hancock, D., Ivanoff, A., Bagnardi, M., Herzfeld, U., Trantow, T., and ICESat-2 Science Team: Ice, Cloud, and Land Elevation Satellite (ICESat-2) Project Algorithm Theoretical Basis Document (ATBD) for Sea Ice Products, version 7, https://doi.org/10.5067/KPMXUOH7TNIY, 2026.