Response:

Response to Reviewer #1

- Figure 2: In the image segmentation part, should "SNIC" be replaced with "SLIC"?

We thank the reviewer for this correction. We confirm that we utilized the SLIC (Simple Linear Iterative Clustering) algorithm, and the typo in Figure 2 has been corrected in the revised manuscript.

- L169: It would be good to refer to Figure S2 here.

We thank the reviewer for the suggestion.We have refer to Figure S2 in the corresponding sentence (L162).

- L174-179: Can the authors add any supplementary figures in determining the decision thresholds? Figure S2 was super helpful for understanding how the random forest settings are determined based on the out-of-bag error analysis, so it would also be good to show examples of P-R and ROC curves in the supplement to help readers' understanding.

We thank the reviewer for this constructive suggestion. We have added a new supplementary figure, Figure S3, which presents examples of the ROC and P-R curves, and also added a reference to Figure S3 in the main text (L168).

- L427: "a lower minimum detection threshold" -> It would be good to specify "a lower threshold for minimum iceberg size."

We have revised the phrase to "a lower threshold for minimum iceberg size" in the manuscript as suggested (L370).

Response to Reviewer #2

- Figure 1: Please also mention in the caption of Figure 1 that the time series refer to 1 pixel each.

We appreciate the reviewer's suggestion to clarify the spatial scale of Figure 1. We have revised the caption to explicitly state: Each time series corresponds to a single pixel.

- Figure 2: What is meant by Iceberg Ocean model? Is this what you call iceberg classes model in the abstract? Please use a consistent name. I would actually just call it a scaling.

We apologize for the inconsistency and confirm that 'Iceberg Ocean model' and 'Iceberg Classes Model' refer to the same concept. We agree that 'Iceberg Size Scaling' is a more accurate term and have standardized its usage throughout the manuscript, including Figure 2 and the main text (L10 and L207).

-L182-183: If you optimise the parameters for each year, the method is not tested on its generalisation capabilities! However, L318 says you leave one year out for testing in a cross-validation? Please clarify how exactly you do it and how much the parameters are tuned to each year.

We thank the reviewer for the constructive comments regarding our validation strategy. We acknowledge that, because our original approach optimized model parameters separately for each year, the previous description of the "leave-one-out cross-validation" procedure was ambiguous and potentially illogical.

To clarify the exact procedure: each year we generate approximately 2,000 manually labeled superpixel samples, of which 1,200 are used for training, 400 for parameter tuning (validation set), and another 400 as an test set. Taking 2018 as an example, after completing parameter optimization for that year, we previously attempted to use the test samples from other years as training data and then evaluated the model using the 2018 test set. Although this result suggested that the model parameters and accuracy did not vary substantially among years—indicating a certain degree of stability—we recognize that because the parameters were optimized

specifically for one year, applying them directly to other years' data cannot rigorously demonstrate the model's generalization ability.

Therefore, in the revised manuscript (L280-290), we have replaced the previous approach with a strict year-by-year accuracy assessment. Under this revised strategy, the model optimized for a specific year (e.g., 2018) is evaluated exclusively using the independent test set from that same year, ensuring full independence and methodological consistency.

In addition, to illustrate the magnitude of parameter adjustments across years, we provide the optimal parameters for each year (including classifier weights and decision thresholds) in the newly added Table S1 in the Supplement. As shown in Table S1, these parameters exhibit only minor interannual variations, indirectly demonstrating the stability of the model. In future work, we plan to integrate multi-year datasets to develop a unified, year-independent model capable of processing new data without retraining.

-L 221: Nong et al (2025) is not peer reviewed yet. If you base your methods on this paper, please describe and verify it yourself.

We thank the reviewer for this valuable comment. In the revised manuscript, we have removed the citation to Nong et al. (2025) and have provided a clear description of how we derive the area-volume relationship for small icebergs directly from the 10-class iceberg classification scheme proposed by Stern et al. (2016) (L210–216).