We are grateful to the editor and reviewers for the time they devoted to reviewing this manuscript and for the constructive comments, all of which were taken into account in the revised manuscript. Below, we address all comments point by point. The comments are reproduced in blue, and the authors' responses are provided in black. The underlined text indicates our corresponding changes in the revised manuscript and in the updated supporting information, both of which are provided with tracked changes. The locations of the underlined text are highlighted in **bold black text**.

Key revisions to the manuscript include:

- The manuscript title was revised to explicitly indicate the 23-year study period and clarify the temporal resolution (8 days) of the dataset.
- The Introduction was substantially updated with Antarctic-specific studies on supraglacial lake monitoring, meltwater impacts on ice dynamics and ice-shelf stability, and melt–albedo feedback. c
- The Methods section was substantially reorganized into seven subsections corresponding to the processing stages in Fig. 2.
- The Sentinel-2 reference data extraction procedure was described in greater detail in Sect. 3.1, including manual sample labeling, three-class Random Forest (RF) classification, training and validation sample selection, and class-specific accuracy assessment.
- More details were added to the model development, optimization, and postprocessing procedures (Sects. 3.4 and 3.7), including hyperparameter grid search, five-fold cross-validation, independent testing, and the application of surface melt and historical lake extent constraints with spatial and temporal consistency criteria.
- The accuracy assessment in the Results and Discussion section was expanded to focus more explicitly on SGL fraction retrieval. Additional metrics, including bias and mean absolute error (MAE), were incorporated for a more comprehensive evaluation of retrieval accuracy.
- The Results and Discussion section was expanded to better connect the dataset findings with previous glaciological studies. Additional discussion was provided on SGL persistence, spatial distribution, seasonal evolution, and interannual variability. The potential influence of cloud contamination on long-term estimates was also discussed, and corresponding cloud masks were made available as auxiliary data.
- Figures 1–8 were redrawn, and Table 2 was revised.
- New figures and tables were added to the Supporting Information.

Author Response to Reviewer #1

Review

Supraglacial lakes are widely distributed across the Antarctic Ice Sheet during summer and have significant impacts on ice sheet hydrology and ice shelf stability. Existing snapshot mapping of supraglacial lakes based on high-resolution satellite imagery is insufficient for capturing their rapid changes. This paper utilizes MODIS and Sentinel-2 satellite remote sensing imagery and employs machine learning methods to construct an 8-day Antarctic supraglacial lake dataset. This work is very timely and has potentially important implications for better understanding the Antarctic Ice Sheet. The paper is well-structured. I recommend publication after minor revisions and provide the following comments for the authors' consideration.

AR1: We thank the reviewer for the positive assessment of our work and for recognizing the scientific value of the dataset. We also appreciate the constructive comments and suggestions, which have helped us improve the clarity and quality of the manuscript. We have carefully considered each comment and revised the manuscript accordingly. Our point-by-point responses are provided below.

General comments

1. The supraglacial lakes extracted from Sentinel-2 satellite remote sensing imagery are used as training data, which greatly influences the accuracy of the machine learning algorithm. Currently, the description of the Sentinel-2 supraglacial lake extraction workflow is relatively brief, only stating that a supervised classification algorithm was used and results with classification accuracy greater than 95% were selected. I suggest further explaining the production of the training data for this supervised classification algorithm, as well as addressing the potential misleading of overall accuracy metrics caused by the small targets (supraglacial lakes) versus the large background area of snow and ice. This would help demonstrate the reliability of the Sentinel-2 supraglacial lake training data.

AR2: We appreciate you highlighting this issue. We acknowledge that the original manuscript provided only a brief description of the Sentinel-2 lake extraction procedure. We have therefore expanded Sect. 3.1 to clarify the reference sample selection, classification method, and accuracy assessment procedure.

We also recognize that overall accuracy (OA) alone can be misleading in scenes dominated by snow and ice. In fact, we retained only classification results with lake-water precision at least 95%, but this criterion was not clearly stated in the original manuscript. We have now clarified this criterion and provided additional

accuracy metrics for all 19 Sentinel-2 reference images in Table S2, along with lake classification maps in Fig. S1. These additions further demonstrate the reliability of the Sentinel-2 reference data.

The corresponding description has been added to Sect. 3.1, Lake extraction from Sentinel-2 imagery.

> High-resolution Sentinel-2 imagery is used to derive reference SGL data, corresponding to Step 1 ("Lake extraction") in Fig. 2. Sentinel-2 image selection is guided by previously reported SGL distributions (Corr et al., 2022; Kingslake et al., 2017; Stokes et al., 2019; Tuckett et al., 2025) and supplemented by manual visual interpretation. Images from multiple regions and melt seasons are selected to capture diverse SGL morphologies and surface conditions (Fig. 1 and Table S1).
>
> For each selected Sentinel-2 image, an independent RF classifier is developed in GEE using nine spectral bands (B2–B8, B11, and B12) as input features (Breiman, 2001). Reference samples representing three surface features (lake, slush, and ice) are manually labeled through visual interpretation of the imagery. These samples are randomly divided into training (70%) and validation (30%) subsets. Each RF classifier is trained on the training subset, while the validation subset is used to assess classification performance based on overall accuracy (OA), precision, recall, and F1-score (F1), with particular emphasis on the lake class. Only classification results meeting the predefined precision criterion (lake-class precision $\geqslant$ 95%) are retained. The classification performance metrics for individual Sentinel-2 images are summarized in Table S2.
>
> The resulting three-class maps are converted into binary SGL maps, in which lake is labeled as 1 and slush and ice are labeled as 0. Separately classifying slush and lake helps reduce the misclassification of slush as lake. Sentinel-2 classification results are presented in Fig. S1. **(L130–143)**

The additional accuracy metrics and classification results are presented in Table S2 and Fig. S1 of the Supplementary Information.

Table S2. Lake classification accuracy for 19 Sentinel-2 reference samples.

Sample	Lake Precision	Lake Recall	Lake F1 Score
George_VI_2018-02-04_1	1.00	1.00	1.00
George_VI_2018-02-04_2	1.00	1.00	1.00
George_VI_2020-01-19_1	0.97	0.95	0.95
George_VI_2020-01-19_2	1.00	1.00	1.00
George_VI_2020-01-19_3	0.99	0.93	0.96
George_VI_2020-01-19_4	1.00	1.00	1.00
George_VI_2020-01-19_5	1.00	1.00	1.00
George_VI_2021-01-23	0.96	0.90	0.93
Bach_2021-01-23	1.00	1.00	1.00
Amery_2020-02-17	1.00	1.00	1.00
Amery_2019-01-13_1	0.95	0.94	0.94
Amery_2019-01-13_2	0.96	0.95	0.96
Amery_2019-02-22	1.00	0.93	0.96
Riiser_Larsen_2019-02-17_1	1.00	1.00	1.00
Riiser_Larsen_2019-02-17_2	1.00	1.00	1.00
Roi Baudouin_2020-01-29_1	1.00	1.00	1.00
Roi Baudouin_2020-01-29_2	1.00	1.00	1.00
Roi Baudouin_2018-02-06	0.97	0.91	0.92
Roi Baudouin_2020-01-18	1.00	1.00	1.00

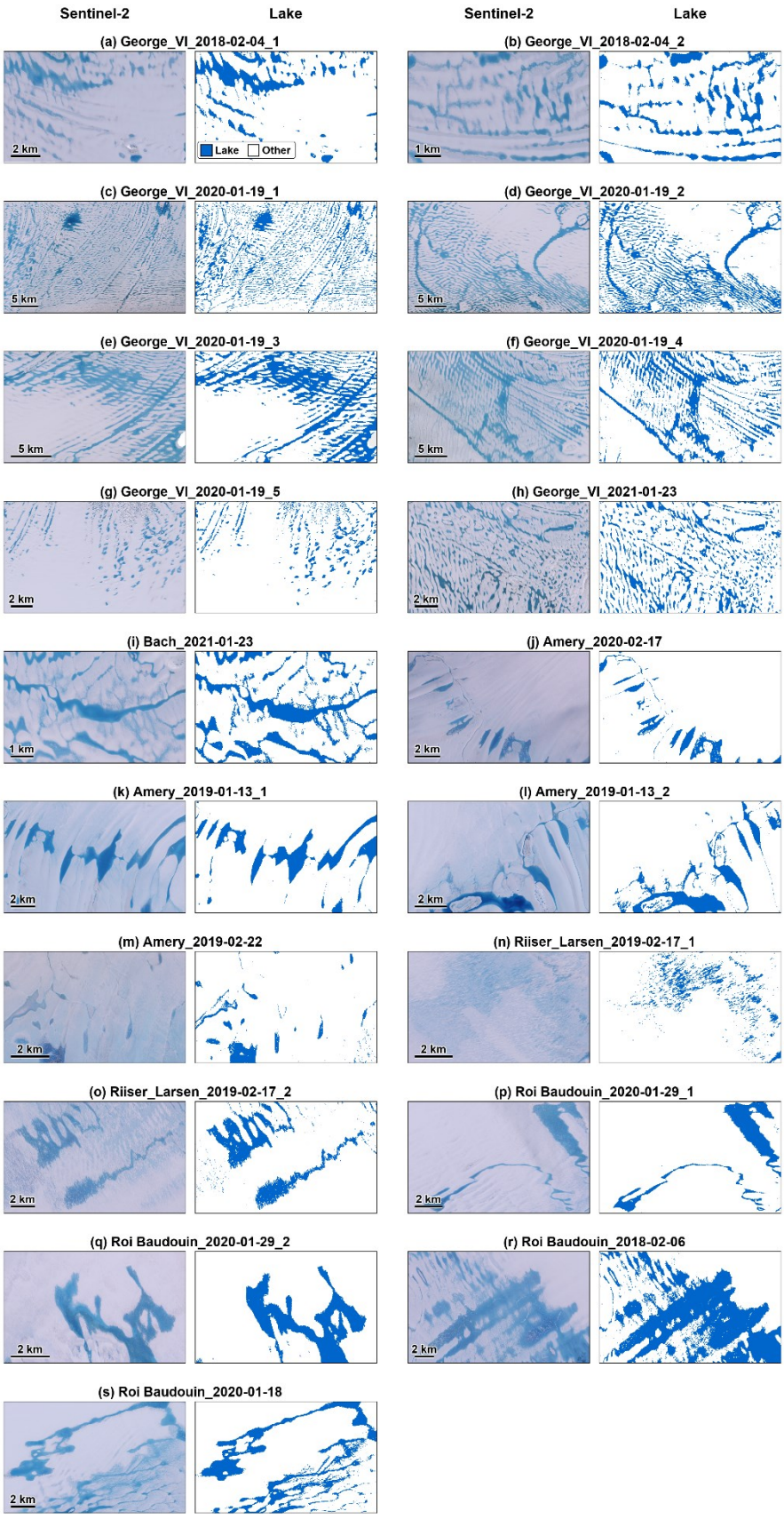


Figure S1: SGL extraction results from Sentinel-2 imagery. (a)–(s) show the true-color Sentinel-2 images and corresponding binary classification maps for the

19 Sentinel-2 image samples used in model cross-validation.

2. In the accuracy assessment section, the paper mentions both classification and regression accuracies, which is somewhat confusing. Classification yields potential water pixels, while regression calculates the SGL fraction only within those water pixels. Therefore, the final accuracy assessment should rely on the SGL fraction results rather than the water versus non-water classification results. If so, I suggest using "performance of SGL fraction retrieval" instead of "performance of SGL classification and fraction retrieval" in the results assessment (Section 4.1), as the latter may cause confusion. Additionally, in Table 2 of the accuracy assessment, metrics such as bias, MAE, and RRMSE could be added. In Figure 5, maps of Sentinel-2-derived SGL and SGL fraction, and difference maps between Sentinel-2-derived SGL and SGL fraction and MODIS-derived SGL and SGL fraction, could be included to better reflect the product accuracy.

AR3: Thank you for this helpful comment. We agree that the accuracy assessment should focus on SGL fraction retrieval, as this is the final output of the dataset. To clarify their different roles, we have revised Sect. 4.1 to place greater emphasis on the accuracy of SGL fraction retrieval.

In addition, following your suggestion, we have expanded the accuracy assessment in Table 2 by including bias and MAE. We have also revised Fig. 5 to include the Sentinel-2 reference SGL maps and spatial difference maps comparing the Sentinel-2 reference data with MODIS-derived SGL fractions. These maps provide a more intuitive visualization of spatial retrieval errors.

The corresponding description has been revised in Sect. 4.1, Accuracy assessment.

The performance of SGL fraction retrieval is evaluated using the correlation coefficient (r), root mean square error (RMSE), bias, and mean absolute error (MAE) (Legates and McCabe, 1999; Willmott, 1981), while SGL pixel identification is additionally assessed using OA, precision, recall, and F1 (Foody, 2002).

The performance of the RFC and RFR models is initially evaluated through five-fold cross-validation using datasets from 19 regions. The samples are divided into five approximately equal subsets. In each iteration, four subsets are used for training, while the remaining subset is used for validation. The RFC model achieves OA values ranging from 0.89 to 0.91 and F1-scores of 0.93–0.94 across the five folds (Table 1), indicating good and consistent classification

performance. For SGL fraction retrieval, the RFR model achieves correlation coefficients ranging from 0.84 to 0.86, RMSE values of 0.12–0.13, and MAE values of 0.080–0.084. Bias remains close to zero across all folds (−0.002 to 0.002), indicating limited systematic overestimation or underestimation of SGL fractions. The overall results yield an r of 0.85 and an RMSE of 0.12 (Table 1 and Fig. 4), indicating good agreement between the estimated and reference SGL fractions. The small variations in these metrics across the five folds further indicate stable model performance. **(L237–248)**

150 Additional accuracy metrics have been incorporated into Table 2 in Sect. 4.1.

Table 2. Classification and regression accuracy of independent test samples.

Region	data	OA	r	RMSE	Bias	MAE
Amery	01/27/2017	0.90	0.79	0.13	−0.01	0.08
Nivlisen	01/26/2017	0.88	0.74	0.14	−0.07	0.10
Amery	01/02/2019	0.85	0.74	0.13	0.04	0.11
Amery	01/03/2019	0.86	0.76	0.14	0.11	0.12
Tracy Tremenchus	01/29/2019	0.82	0.75	0.14	0.02	0.09
Voyeykov	12/28/2019	0.82	0.73	0.15	0.10	0.11
Nivlisen	01/26/2020	0.85	0.75	0.14	−0.10	0.12

Figure 5 has also been updated to include spatial difference maps comparing Sentinel-2 reference fractions with MODIS-derived SGL fractions.

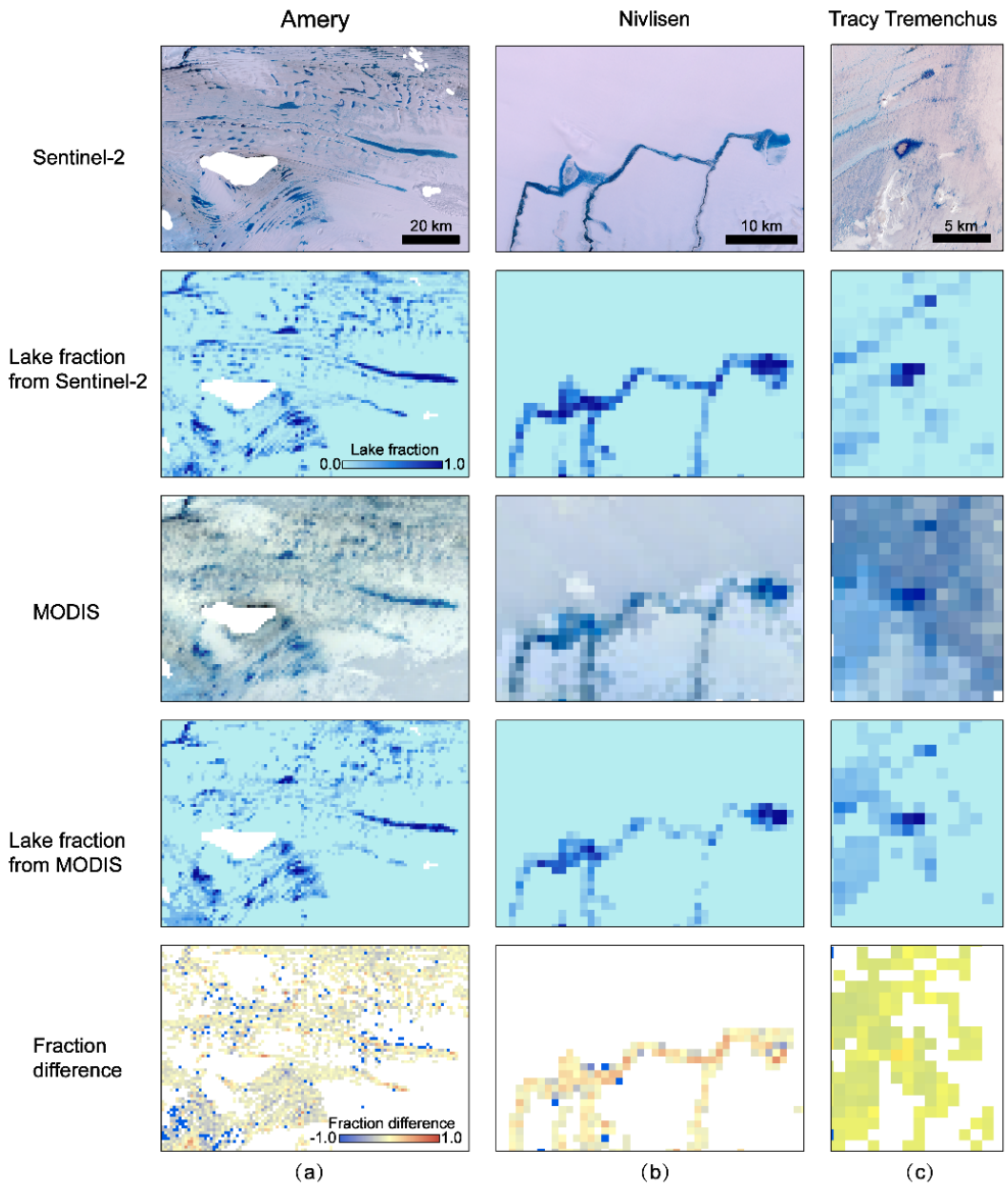


Figure 5: Comparison of Sentinel-2-derived reference SGL fractions and MODIS-derived SGL fractions for representative independent test samples. (a) Amery Ice Shelf, 27 January 2017; (b) Nivlisen Ice Shelf, 26 January 2020; (c) Tracy Tremenchus Ice Shelf, 29 January 2019.

Minor comments

160 1.line 27, "impermeable snow and ice surfaces", should it be "impermeable ice surfaces"?

AR4: We appreciate you bringing this to our attention. We agree with the suggested correction and have replaced "impermeable snow and ice surfaces" with "impermeable ice surfaces".

The corresponding sentence has been revised in the Introduction.

> As the climate warms, rising air temperatures can enhance surface melting in Antarctica, promoting the accumulation of meltwater within topographic lows on impermeable ice surfaces and the formation of supraglacial lakes (SGLs) (Dirscherl et al., 2021; Stokes et al., 2019; Trusel et al., 2015). **(L27–30)**

170 2.line 31, add a few recently published papers on the impact of meltwater on ice sheet dynamics and stability.

AR5: We appreciate your suggestion. We have updated the Introduction by incorporating recent studies on the impacts of surface meltwater on Antarctic glacier dynamics and stability.

The corresponding description has been revised in the Introduction.

> Surface meltwater reaching the glacier bed can increase subglacial water pressure and accelerate ice flow, as observed in Greenland (Das et al., 2008; Gjerde et al., 2025; Tedesco et al., 2013; Wei et al., 2026a) and more recently in East Antarctica (Sugiyama et al., 2026). Satellite observations have also revealed
180 seasonal glacier speed-ups on the Antarctic Peninsula, potentially associated with surface meltwater input and other environmental factors (Boxall et al., 2022; Wallis et al., 2023). Furthermore, supraglacial lake filling and drainage can induce ice-shelf flexure and promote hydrofracture, potentially contributing to ice-shelf disintegration (Banwell et al., 2013, 2024; Scambos et al., 2000; Trusel et al., 2022). **(L33–39)**

The newly added references are as follows:

> Banwell, A. F., Willis, I. C., Stevens, L. A., Dell, R. L., and MacAyeal, D. R.: Observed meltwater-induced flexure and fracture at a doline on George VI Ice Shelf, Antarctica, Journal of Glaciology, 70, e47, 1–14,
190 https://doi.org/10.1017/jog.2024.31, 2024.

> Boxall, K., Christie, F. D. W., Willis, I. C., Wuite, J., and Nagler, T.: Seasonal land-ice-flow variability in the Antarctic Peninsula, The Cryosphere, 16, 3907–3932, https://doi.org/10.5194/tc-16-3907-2022, 2022.

> Gjerde, G., Behn, M. D., Stevens, L. A., Das, S. B., and Joughin, I.: Seasonal drainage-system evolution beneath the Greenland Ice Sheet inferred from transient speed-up events, The Cryosphere, 19, 6149–6169,

https://doi.org/10.5194/tc-19-6149-2025, 2025.

Sugiyama, S., Kondo, K., Minowa, M., and Watanabe, A.: Acceleration of an Antarctic outlet glacier driven by surface meltwater input to the base, Nature Communications, 17, 6102, https://doi.org/10.1038/s41467-026-72724-x, 2026.

Trusel, L. D., Pan, Z., and Moussavi, M.: Repeated tidally induced hydrofracture of a supraglacial lake at the Amery Ice Shelf grounding zone, Geophysical Research Letters, 49, e2021GL095661, https://doi.org/10.1029/2021GL095661, 2022.

Wallis, B. J., Hogg, A. E., van Wessem, J. M., Davison, B. J., and van den Broeke, M. R.: Widespread seasonal speed-up of west Antarctic Peninsula glaciers from 2014 to 2021, Nature Geoscience, 16, 231–237, https://doi.org/10.1038/s41561-023-01131-4, 2023.

Wei, J., Yang, K., Zhu, Y., Wang, Y., and Guo, X.: Satellite monitoring of Greenland wintertime buried lake drainage and potential ice flow response, Remote Sensing of Environment, 333, 115157, https://doi.org/10.1016/j.rse.2025.115157, 2026a.

3.lines 38 and 43, add spatial resolution numbers here, e.g., XX-XX m.

AR6: Thank you for this suggestion. We have added the spatial resolution ranges of Landsat, Sentinel-2, WorldView, and MODIS to clarify the spatial capabilities of these satellite sensors.

The corresponding sentence has been revised in the Introduction, as follows:

High-spatial-resolution optical satellite data, such as Landsat (15–100 m), Sentinel-2 (10–60 m), and WorldView (0.3–30 m), are widely used for SGL monitoring (Dell et al., 2022; Halberstadt et al., 2020; Kingslake et al., 2017; Koehler et al., 2026; Moussavi et al., 2020). **(L43–46)**

To improve temporal coverage, medium- to low-resolution satellite imagery, such as data from the Moderate Resolution Imaging Spectroradiometer (MODIS; 250–1000 m), has been used for SGL monitoring due to its high revisit frequency (Arthur et al., 2020; Hubbard et al., 2016; MacDonald et al., 2019; Spergel et al., 2021). **(L49–51)**

4.line 66, specify the resampling method.

AR7: Thank you for this suggestion. We have clarified the resampling method used for MODIS imagery. Both MOD09GA and MOD09A1 are resampled from their native 500 m resolution to 1 km using mean aggregation.

The corresponding sentence has been revised in Sect. 2.1, MODIS surface reflectance data.

> To improve retrieval robustness and reduce sensitivity to noise and residual cloud contamination, both MOD09GA and MOD09A1 are resampled to a common spatial resolution of 1 km using mean aggregation (hereafter, all MODIS pixels refer to 1 km pixels unless otherwise specified). **(L72–74)**

5.Figure 1, add acquisition dates to the Sentinel-2 and MODIS imagery.

AR8: We appreciate your helpful suggestion. We have revised the caption of Fig. 1 to clarify the acquisition dates of the Sentinel-2 and MODIS images used for the sample regions.

The corresponding caption has been revised as follows:

> Figure 1: Representative SGL samples derived from MODIS and Sentinel-2 imagery across the Antarctic Ice Sheet. (a)–(l) True-color composite images from Sentinel-2 (a, c, e, g, i, k) and MODIS (b, d, f, h, j, l), acquired on (a, b) 17 February 2019, (c, d) 6 February 2018, (e, f) 13 January 2019, (g, h) 22 February 2019, and (i–l) 19 January 2020. (m) Geographic distribution of sample locations across Antarctica. **(L83–86)**

6.line 78, why use Sentinel-2 Level-1C products instead of Level-2A products?

AR9: We selected Sentinel-2 Level-1C (L1C) products for two main reasons. First, L1C top-of-atmosphere (TOA) reflectance data have been successfully used in previous studies of Antarctic supraglacial lakes (Moussavi et al., 2020). Second, L1C products provide more complete coverage for the early years of our reference dataset, particularly 2017–2018, when Level-2A (L2A) surface reflectance products had limited coverage over Antarctica.

The corresponding description has been revised in Sect. 2.2, Sentinel-2 top-of-atmosphere reflectance data.

> This study uses Sentinel-2 Level-1C (L1C) top-of-atmosphere (TOA) reflectance products, which have been used in previous Antarctic SGL mapping studies (Moussavi et al., 2020). L1C products are selected to ensure consistent data

availability across the reference dataset, particularly for 2017–2018, when Level-2A (L2A) surface reflectance products had limited coverage over Antarctica. Details of the selected imagery are provided in Table S1. **(L91–94)**

7.line 92, why use the CryoSat-2 DEM instead of the Reference Elevation Model of Antarctica (REMA)?

AR10: We selected the CryoSat-2 DEM because elevation data are used only for terrain slope screening rather than detailed topographic analysis. The 1 km resolution of the CryoSat-2 DEM matches that of our MODIS-based SGL dataset and is sufficient for applying the slope threshold.

8.lines 100-104, these lines overlap in content with lines 106-110 and should be removed.

AR11: Thank you for pointing out this redundancy. We have removed these lines to avoid repetition and improve the clarity and conciseness of the manuscript.

9.line 123, "known lake locations", is this referenced from another paper?

AR12: Thank you for this comment. The previously mentioned “known lake locations” refer to SGL distributions documented in earlier studies. To clarify the source of this information, we have revised the description in Sect. 3.1 and added explicit references to the relevant studies.

The corresponding sentence has been revised in Sect. 3.1, Lake extraction from Sentinel-2 imagery.

Sentinel-2 image selection is guided by previously reported SGL distributions (Corr et al., 2022; Kingslake et al., 2017; Stokes et al., 2019; Tuckett et al., 2025) and supplemented by manual visual interpretation. **(L131–132)**

10.line 128, why say "max pooling" rather than "max aggregation"?

AR13: Thank you for this suggestion. We agree that “max aggregation” is more appropriate in the context of spatial resampling. We have replaced “max pooling” with “max aggregation” to ensure consistent terminology throughout the manuscript.

The corresponding sentence has been revised in Sect. 3.2, Spatial aggregation of Sentinel-2 SGL maps.

First, max aggregation is used to generate binary SGL labels at 1 km resolution. **(L147)**

11.line 143, the design of $NDWI_{ice}$ is actually intended to better distinguish open water from easily confusable features such as low-albedo slush. This is because the $NDWI_{ice}$ index shifts the two bands of NDWI (green and NIR) toward shorter wavelengths, replacing them with blue and red.

AR14: Your comment has helped us better understand this issue. We have revised the description to clarify that $NDWI_{ice}$ shifts both spectral bands of NDWI toward shorter wavelengths, replacing green and near-infrared with blue and red, respectively. We also explain that this modification helps distinguish open water from spectrally similar features, particularly low-albedo slush.

The corresponding description has been revised in Sect. 3.5, Preliminary screening of potential SGL pixels in MODIS imagery.

> NDWI uses the green and near-infrared bands to enhance open-water signals (McFeeters, 1996), whereas $NDWI_{ice}$ replaces green and near-infrared with blue and red, respectively, to better distinguish open water from surrounding snow, ice, and low-albedo slush (Yang and Smith, 2013). **(L183–186)**

12.lines 148-149, the band combinations for these two water indices are written incorrectly; NDWI uses green and NIR, while $NDWI_{ice}$ uses blue and red. Please verify.

AR15: Thank you for identifying this error. We agree that the band combinations for NDWI and $NDWI_{ice}$ were incorrectly assigned in the original manuscript. We have corrected the equations to ensure that NDWI uses the green and near-infrared bands (MODIS B4 and B2), while $NDWI_{ice}$ uses the blue and red bands (MODIS B3 and B1).

The corresponding equations have been corrected in Sect. 3.5, Preliminary screening of potential SGL pixels in MODIS imagery.

$$NDWI = \frac{B4 - B2}{B4 + B2} \quad (1)$$

$$NDWI_{ice} = \frac{B3 - B1}{B3 + B1} \quad (2)$$

13.line 151, these thresholds are somewhat lower than those in previous supraglacial lake extraction studies, what might be the reason?

AR16: We appreciate you raising this point. At the 1 km pixel scale, lake water may occupy only a small fraction of a pixel, resulting in relatively low water-index values

due to the dominant spectral response of surrounding snow and ice.

To avoid missing these potential SGL pixels, we adopt relatively permissive thresholds for initial screening, rather than directly applying thresholds developed for higher-resolution imagery. These thresholds are calibrated using Sentinel-2 reference data (Text S1 and Fig. S2). The initial screening only identifies candidate SGL pixels, which are subsequently evaluated by the machine-learning models.

The corresponding description has been revised in Sect. 3.5, Preliminary screening of potential SGL pixels in MODIS imagery.

330 Due to the coarse spatial resolution of MODIS, SGLs may occupy only a small fraction of individual pixels, with surrounding snow and ice weakening the water-related spectral signals. To minimize missed detections of potential SGL pixels with low lake fractions, the NDWI and $NDWI_{ice}$ thresholds are empirically calibrated using reference lake samples derived from high-resolution Sentinel-2 imagery instead of directly adopting thresholds from previous studies (Text S1). The resulting thresholds are 0.08 for NDWI and 0.04 for $NDWI_{ice}$ (Fig. S2). **(L191–194)**

14.line 171, five-fold cross-validation, two references are cited here, did the authors follow their methods?

340 **AR17:** Thank you for this comment. We have reviewed and updated the references to better reflect the methods used in our study. Specifically, we have replaced Kuhn and Johnson (2013) and Wong (2015) with more relevant references. Bischl et al. (2023) provides methodological guidance on grid search, Probst et al. (2019) discusses random forest hyperparameter optimization, and Arlot and Celisse (2010) provides a comprehensive overview of cross-validation methods. These references support our methodological choices rather than indicate that we followed their specific experimental procedures.

We have also expanded the description to specify the hyperparameters, their search ranges and step sizes, and the five-fold cross-validation procedure.

350 The corresponding description has been revised in Sect. 3.4, Random forest model development and optimization.

For both models, two key hyperparameters, the number of trees and the maximum number of leaf nodes per tree, are optimized using a grid search

strategy (Bischl et al., 2023; Probst et al., 2019). The number of trees ranges from 5 to 200, while the maximum number of leaf nodes ranges from 5 to 100, both in increments of 5. Model performance is evaluated using five-fold cross-validation (Arlot and Celisse, 2010) on samples from 19 regions, while samples from the remaining seven regions are reserved for independent testing. **(L168–173)**

The newly added references are as follows:

Arlot, S. and Celisse, A.: A survey of cross-validation procedures for model selection, Statistics Surveys, 4, 40–79, https://doi.org/10.1214/09-SS054, 2010.

Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A.-L., Deng, D., and Lindauer, M.: Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges, WIREs Data Mining and Knowledge Discovery, 13, e1484, https://doi.org/10.1002/widm.1484, 2023.

Probst, P., Wright, M. N., and Boulesteix, A.-L.: Hyperparameters and tuning strategies for random forest, WIREs Data Mining and Knowledge Discovery, 9, e1301, https://doi.org/10.1002/widm.1301, 2019.

15.line 179, "supraglacial meltwater accumulation" can be changed to "supraglacial lake formation".

AR18: Thank you for this suggestion. We agree that “supraglacial lake formation” is more precise in this context. We have revised the wording accordingly.

The corresponding sentence has been revised in Sect. 3.7, Postprocessing of SGL fraction maps.

Surface melt is widely recognized as a prerequisite for supraglacial lake formation (Leeson et al., 2012; Stokes et al., 2019). **(L218–219)**

16.Figure 4, it is suggested to add metrics such as n, bias, MAE, and RRMSE.

AR19: We are grateful for your constructive suggestion. We have revised Fig. 4 to include the sample size (n), bias, and MAE, providing a more comprehensive assessment of model performance. We have not included RRMSE because SGL fractions can be close to zero, making relative error metrics sensitive to the choice of normalization. Instead, we retain RMSE on the original 0–1 fraction scale.

The additional accuracy metrics have been incorporated into Fig. 4 of the revised

manuscript.

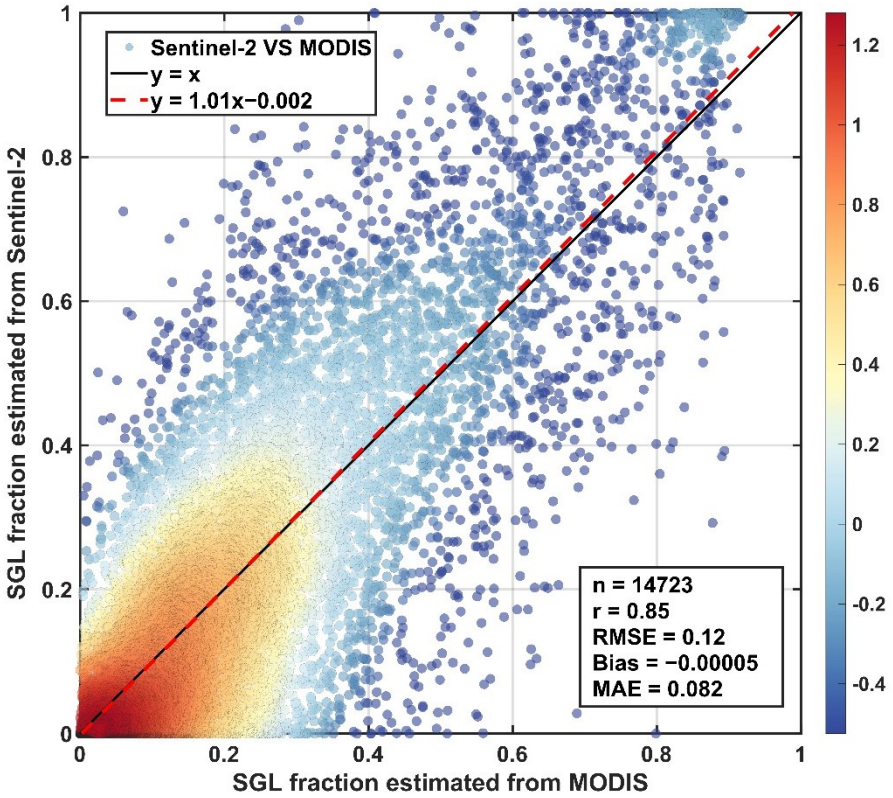


Figure 4: Validation of MODIS-derived SGL fractions against Sentinel-2 reference fractions. The black line indicates the 1:1 relationship, whereas the fitted regression is indicated by the red dashed line.

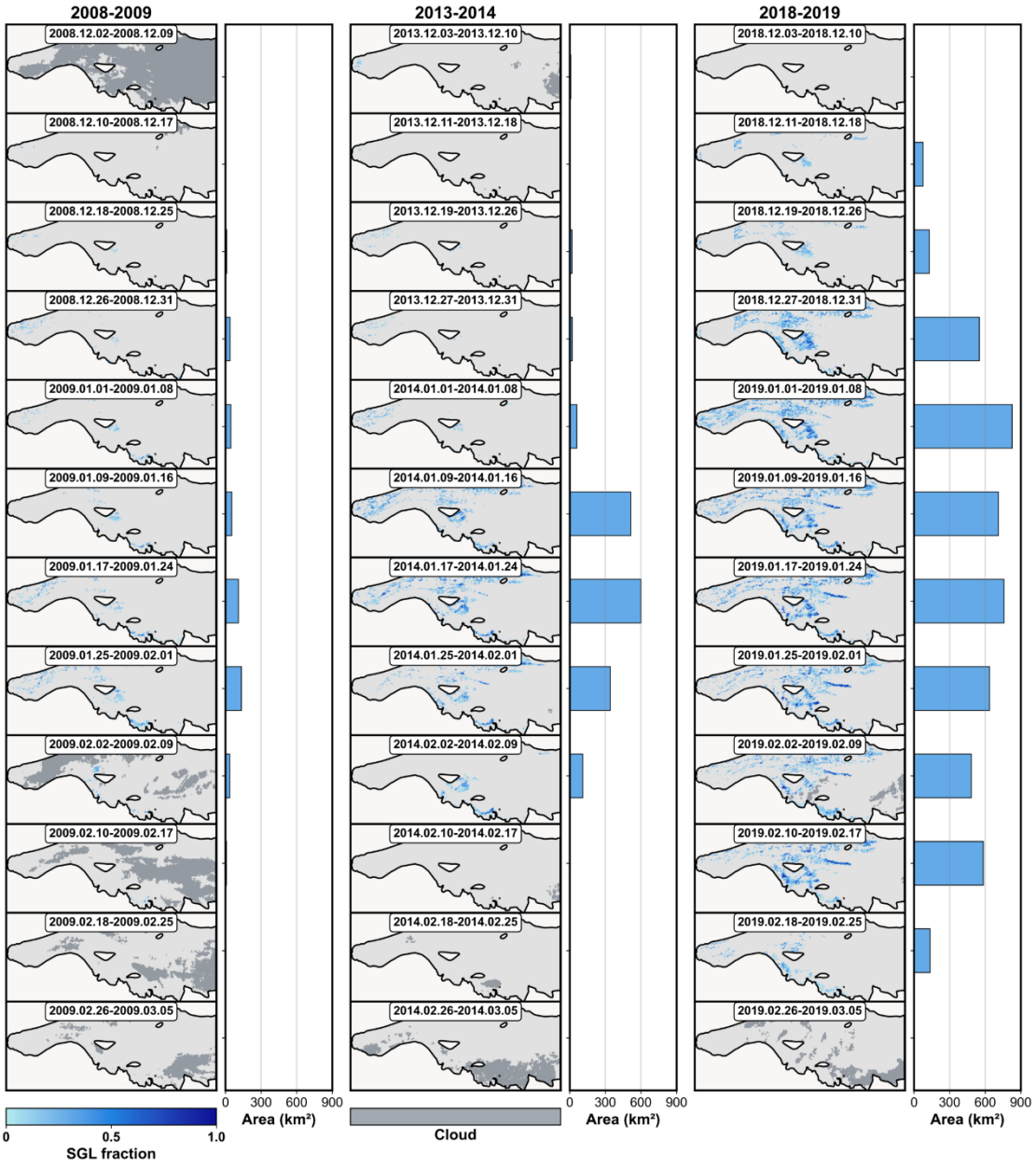
17.Figure 6, the area value ranges in the three columns are inconsistent, making direct comparison difficult. It is suggested to unify the range to 0–600 km². Does this figure's pronounced interannual variability in SGL area potentially suffer from cloud cover effects (even though the 8-day composite may largely mitigate this)? The SGL fraction in the figure is difficult to discern; consider focusing the map on regions with SGL distribution.

AR20: Thank you for this helpful comment. We have revised Fig. 6 in three ways. First, we unified the area axis range across the three columns to improve interannual comparison. Second, we adjusted the spatial extent to better highlight the main SGL distribution zones and improve the visibility of the SGL fraction patterns. Third, we added cloud information to each 8 d panel so that the potential influence of cloud cover can be assessed more directly.

The revised figure shows that cloud cover is present mainly during the early and late

parts of the melt season, when SGL area is generally low. In contrast, the main period of lake development and peak SGL extent occurs during mid-season, when cloud contamination is relatively limited in the examples shown. This suggests that, although cloud cover may still affect individual 8 d composites, it is unlikely to be the main cause of the pronounced interannual differences in SGL area displayed in Fig. 6.

Figure 6 has been revised accordingly in the manuscript.



410 Figure 6: Temporal evolution of SGLs in representative areas with high SGL concentrations on the Amery Ice Shelf from early December to early March during three melt seasons: (a) 2008–2009; (b) 2013–2014; and (c) 2018–2019.

18.Figures 7 & 8, nice figures! From these two figures, some extracted features appear as very wide channels and are classified as SGL. If such wide channels are limited in distribution, including them in the SGL discussion is fine; however, if they are widespread, is it useful to discuss them separately in future?

AR21: Thank you for your positive assessment of Figs. 7 and 8 and for this helpful comment. We agree that some of the mapped water bodies exhibit elongated, channel-like shapes. However, measurements from the imagery indicate that these features range from several tens of metres to more than 1 km in width. Their substantial widths are also clearly visible in the high-resolution Sentinel-2 imagery presented in Fig. 5.

Given their considerable widths and the 1 km spatial resolution of our MODIS-based dataset, we consider it reasonable to include these features in the SGL category. Nevertheless, we agree that distinguishing lakes from wide supraglacial channels could provide further insights into surface hydrological processes. We have therefore expanded the discussion of data limitations to address this issue and highlight the value of future studies.

The corresponding discussion has been revised in Sect. 4.5, Uncertainty discussion.

A further limitation concerns the distinction between SGLs and wide supraglacial channels. Some mapped water bodies exhibit elongated, channel-like shapes, but their widths range from several tens of metres to more than 1 km, as illustrated by the Sentinel-2 imagery in Fig. 5. Given their substantial widths and the 1 km spatial resolution of our dataset, these features are included in the SGL category. Future studies combining high spatial resolution with frequent temporal observations could better distinguish supraglacial lakes from river channels and quantify their distinct yet interconnected roles in supraglacial meltwater storage and transport (Liang et al., 2022; Yang et al., 2021). **(L434–440)**

The newly added references are as follows:

Liang, X., Zhang, W., Li, Y., Lu, Y., and Yang, K.: Evolution of surface meltwater on the Bach Ice Shelf of the Antarctic Peninsula, Chinese Journal of Polar Research, 34, 149–158, https://doi.org/10.13679/j.jdyj.20210029, 2022.

Yang, K., Smith, L. C., Cooper, M. G., Pitcher, L. H., van As, D., Lu, Y., Lu, X., and Li, M.: Seasonal evolution of supraglacial lakes and rivers on the southwest Greenland Ice Sheet, Journal of Glaciology, 67, 592–602,

https://doi.org/10.1017/jog.2021.10, 2021.

19.Figure 9, again, nice figure! My question is similar to that for Figure 6: do the annual results here suffer from cloud cover and data quality differences? Even though the 8-day composite may largely mitigate this, the 8-day composite MODIS data product still has certain issues. Should this be considered when analyzing multi-year SGL area changes?

AR22: Thank you for your positive assessment of Fig. 9 and for raising this important concern. We agree that cloud cover and differences in data quality may affect annual SGL area estimates and the interpretation of interannual variability.

To address this concern, we have expanded the uncertainty discussion to explicitly acknowledge the potential influence of these factors on interannual SGL area changes. In addition, we have made the cloud masks corresponding to the SGL fraction products available as auxiliary data. These masks allow users to examine cloud-affected areas and assess differences in valid observation coverage when interpreting the dataset.

The corresponding description has been revised in Sect. 4.5, Uncertainty analysis, and Sect. 5, Data and code availability.

> First, the inherent limitations of remote sensing data introduce observational uncertainties. Clouds and cloud shadows frequently obscure the ice-sheet surface, potentially leading to underestimation of SGL fractions or temporal gaps in retrievals (Arthur et al., 2020; Husman et al., 2023; Moussavi et al., 2020). To mitigate this issue, the 8-day composite MOD09A1 product is employed, as it preferentially selects observations with lower cloud contamination and favorable viewing conditions (Sun et al., 2016; Xiang, 2013). However, residual clouds and cloud shadows may still persist, and the availability and quality of valid observations can vary between years. These differences may introduce biases in annual SGL area estimates and affect the interpretation of interannual variability. To facilitate further assessment, we have made the cloud masks corresponding to the SGL fraction products available as auxiliary data. These masks allow users to examine cloud-affected areas and evaluate variations in valid observation coverage when interpreting seasonal and interannual SGL changes. **(L426–434)**
>
> Cloud masks corresponding to the SGL fraction products are also provided as auxiliary data. **(L468–469)**

The newly added reference is as follows:

480 Sun, L., Wei, J., Wang, J., Mi, X., Guo, Y., Lv, Y., Yang, Y., Gan, P., Zhou, X., Jia, C., and Tian, X.: A Universal Dynamic Threshold Cloud Detection Algorithm (UDTCDA) supported by a prior surface reflectance database, Journal of Geophysical Research: Atmospheres, 121, 7172–7196, https://doi.org/10.1002/2015JD024722, 2016.

20.Figure 10, this shows the 23-year long-term mean. I suggest adding standard deviation in Figure 10b to illustrate the degree of interannual variability.

AR23: Thank you for your positive assessment of Fig. 10 and for this helpful suggestion. We agree that including standard deviations would provide valuable information on the interannual variability of SGL area.

490 However, Fig. 10b contains multiple regional time series, and adding standard deviation shading to each curve would result in substantial overlap and reduce readability. Therefore, we have retained the original figure to preserve a clear comparison of the regional mean seasonal patterns. To address your suggestion, we have added Fig. S4 to the Supplementary Information. This new figure presents the 23-year mean seasonal evolution of SGL area and the corresponding standard deviations for Antarctica and nine subregions in separate panels.

The additional results are presented in Fig. S4 of the Supplementary Information.

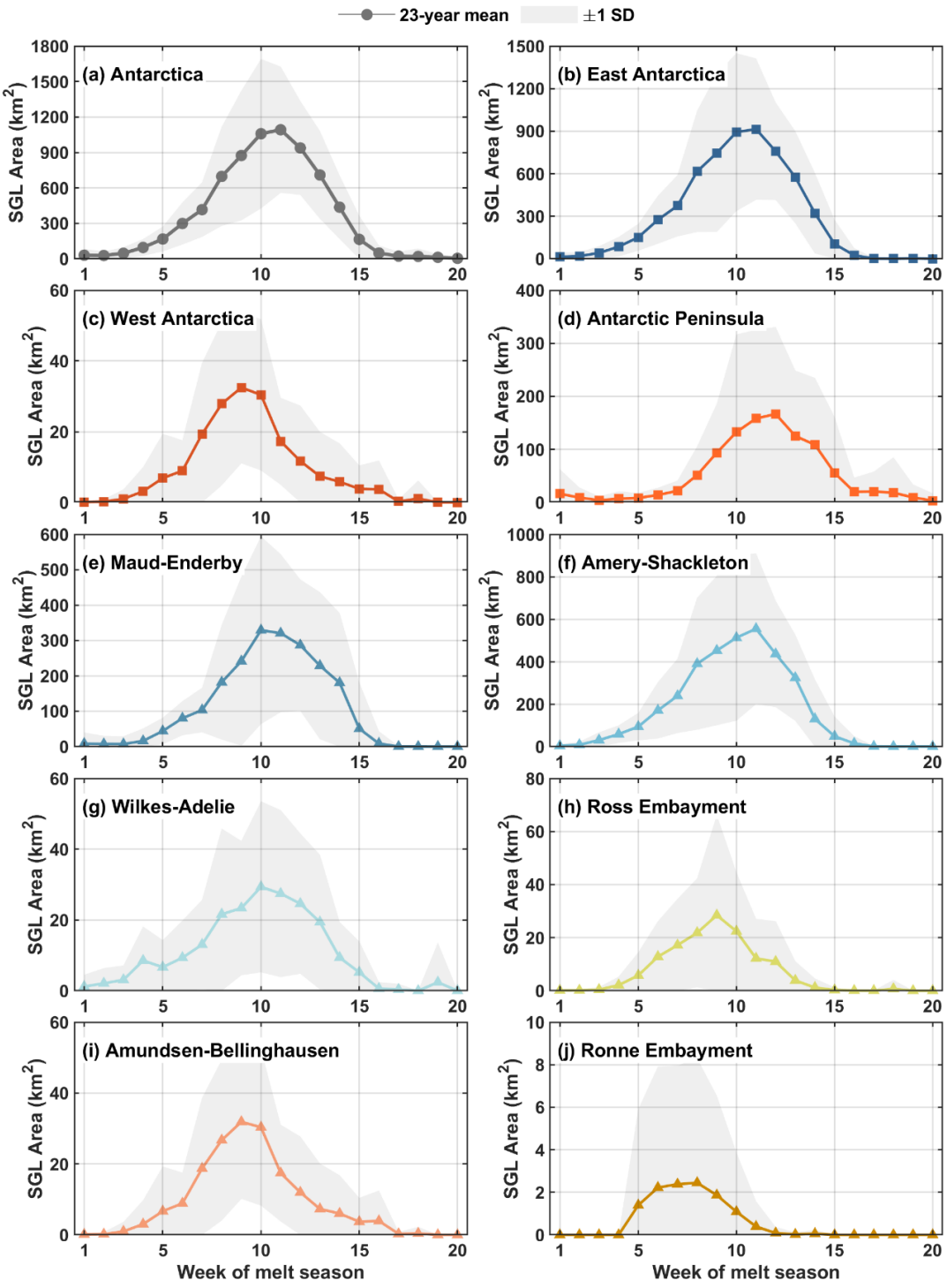


Figure S4: Seasonal evolution of the 23-year mean SGL area and interannual variability across Antarctica and its subregions.