

An 8-day Antarctic supraglacial lake dataset from MODIS
Shuo Wei, Lei Zheng, Qi Liang, Teng Li, Xiao Cheng
Review by Ian Willis

This is a novel paper which directly addresses the classic "mixed-pixel" dilemma of tracking small, highly dynamic features (supraglacial lakes, SGLs) using moderate-resolution satellite sensors. The workflow shown in Fig 2 and described in Section 3 seems sensible and robust. I liked the two-stage hierarchical modelling (RFC + RFR) where, instead of forcing a single regression model to calculate fractions across the entire continent blindly, the authors first use a Random Forest Classifier (RFC) to binarize whether a pixel contains water and then pass it to a Random Forest Regressor (RFR) to compute sub-pixel fractions. This two-stage constraint should reduce false-positive noise in dry snow zones.

A vulnerability in many random forest papers is spatial autocorrelation (evaluating a model on pixels directly adjacent to the training pixels). The authors avoided this trap by reserving 7 completely independent geographic regions spanning multiple melt seasons exclusively for validation. Achieving an Overall Accuracy (OA) > 0.82 and an $r > 0.73$ on these unseen areas proves the model generalizes well.

Training the models using the high-temporal-resolution daily product (MOD09GA) matched to exact high-spatial-resolution Sentinel-2 passes, but deploying the final operational dataset on the 8-day composite (MOD09A1), is sound, logical and enables the production of pan Antarctic maps of SGL area, persistence and fraction at 1 km resolution (Figs 6-9), as well as time series averaged across regions of seasonal changes in SGL area showing how these vary over the 23 year record (Fig. 9).

The creation of a "Level 1" product that applies an active melt flux constraint (Zheng et al., 2025) and a historical spatial mask (Tuckett et al., 2025) also increases the robustness of the final data sets as it presumably goes some way to cleaning out false positives, for example due to cloud-shadows that often slip past standard optical masks.

I have a few main points and lots of small, suggested edits / comments that I'd encourage the authors to consider.

Main points

1. The current title is ambiguous and implies the data set could be only 8 days long! If the authors want more people to read the paper, they should improve the title to convey 8 day refers to temporal frequency not total duration of data set. e.g. "A 23-Year, 8-Day Temporal Resolution Antarctic Supraglacial Lake Dataset from MODIS."
2. I think the methods section (Section 3) could be much improved. I'm not convinced someone could replicate the work from some of the descriptions given at present. I find it hard at times to reconcile what is written, with the

processing steps summarised in Fig 2. I think the authors need to first outline the overarching methodology and refer broadly to the boxes and arrows of Figure 2, for example clarifying that the RFC and RFR model are developed using the daily MODIS data, but that the final model works with the 8-day MODIS products. Then the authors need to go into more detail about each of the steps. It'd be very helpful if 7 subsections of Section 3 could map onto the 7 stages of the workflow outlined in the 7 boxes of Fig 2 using the same names.

3. Building on Point 2, an example of where more detail is needed is the description on how the Sentinel-2 lakes were detected which is currently exceedingly brief L124-127. Reference is made to Niu et al (2023) with no further detail. This is particularly concerning because the Sentinel-2 dataset is regarded as “ground-truth” and so is the absolute anchor of the model developed in this study. If the training labels are flawed, the entire 23-year MODIS output inherits those flaws.

Furthermore, the text says the authors used "a supervised classification algorithm... (Niu et al., 2023)". Looking at the references, Niu et al. (2023) specifically developed a highly complex deep learning network (a U-Net with Attention Blocks). The authors should explicitly name the type of classifier (e.g., "an attention-based U-Net convolutional neural network") instead of using the generic umbrella term "supervised classification algorithm" and tell us more about it and how it was used in their study.

The authors state that "only images meeting the classification accuracy threshold ($> 95\%$) are selected". But unless I missed it, they don't explain how this 95% accuracy was calculated. Did they manually digitize independent validation polygons for every single one of the 26 Sentinel-2 scenes to run a confusion matrix? Who verified it?

The authors don't explain how the Sentinel-2 binary 1/0 map accounts for the real-world gradients of shallow water, deep water, and wet slush before it got averaged up to the 1 km scale. If the Sentinel-2 model mistook wet slush for open water (a classic problem), the MODIS regression model will consistently overestimate actual liquid lake volumes.

4. The authors need to check all the references they've used to support all their statements in the Introduction. Many are inappropriate as they refer to regions other than Antarctica, or they refer to papers where completely different methods were used to the ones suggested. There are also many recent papers that are highly relevant, but which are not cited. I've given more details on these in the line-by-line comments below.
5. I'm not convinced the melt flux dataset of Zheng et al as used is providing a sensible constraint. It appears individual melt flux values from SSM/I on specific days are used to constrain the optical ponding data. But passive microwave sensors are incredibly sensitive to the instantaneous presence of liquid water within the near-surface snow/firn matrix. As soon as the surface temperatures

drop and freezing begins, that liquid water refreezes or stays trapped. SSM/I will instantly register zero active melt flux.

However, SGLs that filled during peak melt days do not simply vanish when active melting stops. Because of the sheer thermal mass of deep water and the ice's topography, surface lakes can persist for days or weeks after active melt flux drops to zero. An optical satellite image taken on a cold, non-melting day will still show massive, clear blue water ponding.

Another way to use Zheng et al. (2025) would be to use it as a cumulative lag constraint. You could:

- Calculate Cumulative Melt Energy: Integrate Zheng et al.'s daily melt flux over a rolling window (e.g., a 3-day, 5-day, or 10-day antecedent period) leading up to your optical image acquisition date.
- Establish a Threshold: If the cumulative melt flux over the previous week is zero, and your optical imagery claims there are lakes, you can confidently flag those optical pixels as false positives (such as cloud shadows or blue ice exposures mistranslated as water).
- Identify Lake Drainage Events: If Zheng et al. show massive active melt flux for 5 days, but your subsequent optical imagery shows zero water ponding, it provides a strong proxy indicator that a hydrofracture or a rapid subglacial lake drainage event occurred.

6. There are several places in the text and on Figures where I'd recommend writing out the names 'Roi Baudoin' and 'Nivlison' in full instead of using abbreviations. This seems especially relevant given that Tracy Tremenchus is written in full.
7. I appreciate this is a methods paper, but as a glaciologist, I was disappointed that the paper did not discuss any of the glaciological implications of the findings or compare the findings with results of similar previous work. For example, how do the persistency statistics compare to previous work? How does the statistic of 65% of water appearing within 10 km of the grounding line compare? Where is the rest of the water, i.e. what % > 10km above the GL vs > 10km below the GL? What have you been able to say with this new data set about water ponding on ice shelves that has not been said before?

Line by Line Comments

L15. Could you incorporate some of the information about how the persistence rate was calculated. Depending on how much space you have in the abstract, you could write "... approximately 83% of lakes exhibit a multi-year mean persistence rate below 5%, appearing in fewer than one out of twenty 8-day observations per melt season." Or shorter "... with approximately 83% of lakes exhibiting a mean seasonal observation persistence below 5% across the 23-year record."

L27. Echelmeyer et al., 1991 wrong reference as this is for GrIS not AIS

L29-30. Lüthje et al., 2006; Tedesco et al., 2012 are for GrIS not AIS. Could either make sentence refer more generically to ice sheets and / or keep the reference to Antarctica and substitute more relevant refs, e.g.

Dell, R. L., Willis, I. C., Arnold, N. S., Banwell, A. F., & de Roda Husman, S. (2024). Substantial contribution of slush to meltwater area across Antarctic ice shelves. *Nature Geoscience*, 17(7), 624–630.

which explicitly links surface water to albedo reduction.

L31-32. All these refs are for Greenland. I am not aware of any published study currently showing that lake drainage on Antarctic Peninsula glaciers causes glacier acceleration. There are a few studies measuring seasonal variations in velocity from satellites and suggesting it may be due to meltwater penetrating to the bed, e.g.,

Boxall, K., Christie, F. D. W., Willis, I. C., Wuite, J. & Nagler, T. Seasonal land-ice-flow variability in the Antarctic Peninsula. *Cryosphere* 16, 3907–3932 (2022).

Wallis, B. J., Hogg, A. E., van Wessem, J. M., Davison, B. J. & van den Broeke, M. R. Widespread seasonal speed-up of west Antarctic Peninsula glaciers from 2014 to 2021. *Nat. Geosci.* 16, 231–237 (2023).

and there is a recent study measuring in-situ acceleration in response to rainfall events:

Sugiyama, S., Kondo, K., Minowa, M. et al. Acceleration of an Antarctic outlet glacier driven by surface meltwater input to the base. *Nat Commun* 17, 6102 (2026).

L37-8. There are many other refs that could be cited here besides Niu et al, (2023) and some, at least, deserve to be acknowledged.

L39. None of these refs are for Antarctica. The ref for the Pamir seems particularly misplaced. Should refer to some of the many papers that have used Sentinel or Landsat data for Antarctica, e.g.,

Dirscherl, M. C., Dietz, A. J., and Kuenzer, C.: Seasonal evolution of Antarctic supraglacial lakes in 2015–2021 and links to environmental controls, *The Cryosphere*, 15, 5205–5226, <https://doi.org/10.5194/tc-15-5205-2021>, 2021

Moussavi, Mahsa, Allen Pope, Anna Ruth W. Halberstadt, Luke D. Trusel, Leanne Cioffi, and Waleed Abdalati. 2020. "Antarctic Supraglacial Lake Detection Using Landsat 8 and Sentinel-2 Imagery: Towards Continental Generation of Lake Volumes" *Remote Sensing* 12, no. 1: 134.

Koehler J, Baumhoer C, Wouters B, Lhermitte S, Dietz A. Advancing supraglacial lake mapping in Antarctica: A deep learning approach using Sentinel-2 imagery. *Environmental Data Science*. 2026;5:e38. doi:10.1017/eds.2026.10065

Tuckett, P. A., Ely, J. C., Sole, A. J., Lea, J. M., Livingstone, S. J., Jones, J. M., and van Wessem, J. M.: Automated mapping of the seasonal evolution of surface meltwater and its links to climate on the Amery Ice Shelf, Antarctica, *The Cryosphere*, 15, 5785–5804, <https://doi.org/10.5194/tc-15-5785-2021>, 2021.

Halberstadt, Anna Ruth W., Colin J. Gleason, Mahsa S. Moussavi, Allen Pope, Luke D. Trusel, and Robert M. DeConto. 2020. "Antarctic Supraglacial Lake Identification Using Landsat-8 Image Classification" *Remote Sensing* 12, no. 8: 1327. <https://doi.org/10.3390/rs12081327>

Dell RL, Banwell AF, Willis IC, Arnold NS, Halberstadt ARW, Chudley TR, Pritchard HD (2022). Supervised classification of slush and ponded water on Antarctic ice shelves using Landsat 8 imagery. *Journal of Glaciology* 68(268), 401–414. <https://doi.org/10.1017/jog.2021.114>

L44-45 and L 48-49. Some of these references are totally wrong here. Most are either wrong methodology or wrong continent.

L52. This is the first mention of ‘SGL fraction’. Suggest explicitly define what you mean here. Is this at a pixel scale?

L92-3. State why surface slope is required. Why and how is it used?

L93 Slater et al reference is not in the reference list.

L95-97. State briefly why and how these two data sets are used?

98-99. You need to make it clear you’re working at the 1km MODIS resolution here I assume, not at the S2 resolution.

L103. Put ‘ground truth’ in quotes as it’s not strictly ground based data.

L106-112. These repeat L98-104. Suggest delete 98-104.

L129. When you say ‘MODIS pixel scale’ I assume you’re talking about the resampled MODIS scale, i.e. 1km. Your ref to 1 km and ‘MODIS pixel scale’ here is ambiguous.

Fig 2. Box 1 on the left and Box 2 on the right have a spelling mistake ‘extreation’

L136. Again, are you working at 500 m or 1 km when you say ‘MODIS SGL pixels’.

L138. “...mixed surface signals within individual pixels” is unclear. Unless I’ve misunderstood, your 1 km MODIS pixels are made up of 100 x 10 m Sentinel-2 pixels. The whole point is that your 1 km pixels likely contain a mix of ‘water’ and non-water’ Sentinel-2 pixels.

L172. Clarify what is meant by ‘continuous’ here. Across space, time, both?

L188-9. I find the statement “...they are conditionally retained when both spatial and temporal consistency criteria are satisfied (Fig. 2)” unclear. Fig. 2 refers to “neighbourhood lakes ($3 \times 3 \geq 1$)” and “Lakes in either adjacent observation ≥ 1 ”. I think this procedure needs explaining more clearly. See my Main Point 2.

L178-189. Further to the above, I find it difficult to reconcile what is written in this section, with the ‘postprocessing’ workflow box depicted in Fig 2. See my Main Point 2.

L195-6. I assume you’re using ‘validation’ and ‘evaluation’ to mean the same thing here? I assume this cross-validation stage is only working with the 19 calibration images mentioned on L132 and not the 7 validation images mentioned on L133. Clarify this.

Fig 7. and 8. You mention earlier on the paper where the Rock dataset comes from, but I don’t think you mention where the Blue Ice data set comes from. It’d be useful to put the source of both in the Fig Headings for these figures.