From Decision Trees to Deep Learning: Enhanced Supraglacial Lake Detection in Antarctica

Celia A. Baumhoer[1], Jonas Koehler[1] & Andreas J. Dietz[1]

[1]German Aerospace Center (DLR), Earth Observation Center (EOC), German Remote Sensing Data Center (DFD), Wessling, Germany

Motivation

Monitoring the seasonal evolution and dynamics of supraglacial lakes (SGL) is essential for understanding their possible destabilizing effects on the Antarctic ice sheet. Spaceborne remote sensing offers the best solution, providing continuous, large-scale, and long-term observations. However, extracting reliable information from optical data remains complex due to limitations in spatial transferability, cloud cover, polar night, and the spectral similarities of frozen lakes with surrounding ice.

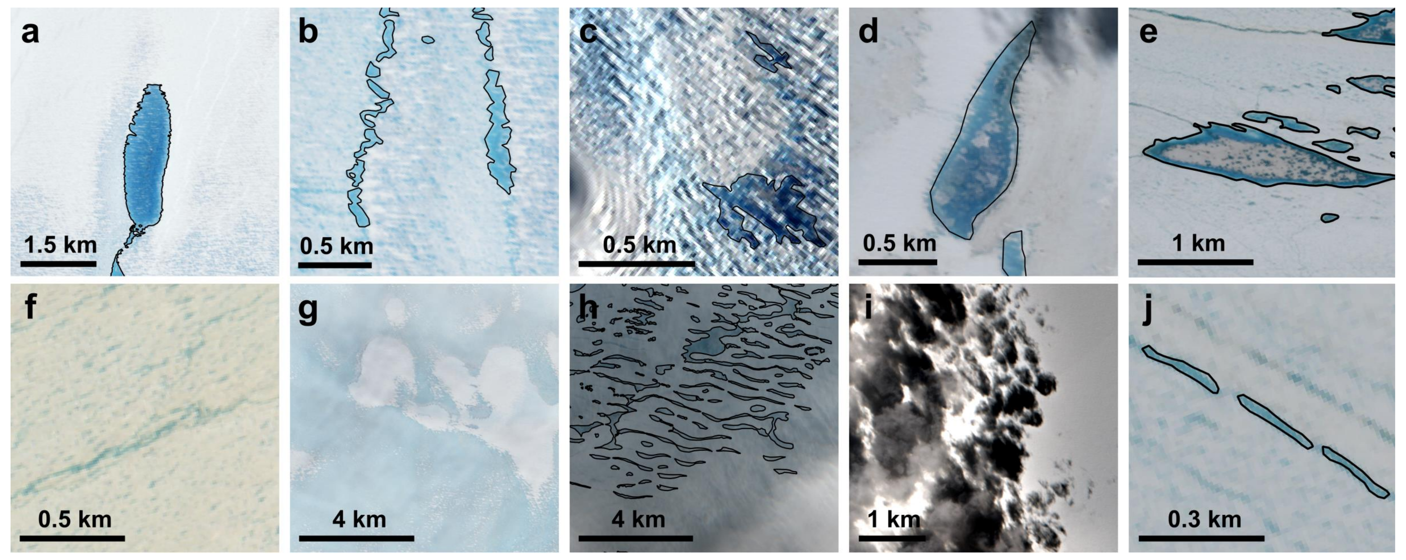

Fig. 1: Visual examples of various surface features included in the train and test datasets. (a) oval lakes at Roi Baudouin, (b) shallow lakes at Roi Baudouin, (c) lake on crevasses at Shirase Glacier, (d) thin floating ice at Amundsen, (e) thick floating ice at Amery, (f) slush at Riiser, (g) blue ice at Nivlisen, (h) lakes under clouds at George VI, (i) dark clouds over Amundsen, and (j) thin streams at Amery. Black boundaries indicate labeled lakes [3]. Background: Copernicus Sentinel-2 Data.

Performance Comparison

This study evaluates whether a deep learning (DL) based mapping approach outperforms a pixel-based Random Forest (RF) classification algorithm for SGL detection in Antarctica. As a benchmark, we utilized an RF model trained on 14 regions and 24 input channels, including Sentinel-2 spectral bands, spectral indices, and topographic variables [1]. The DL approach is based on a U-Net (see Fig. 4) with residual connections and an ASPP module (Atrous Spatial Pyramid Pooling) [2].

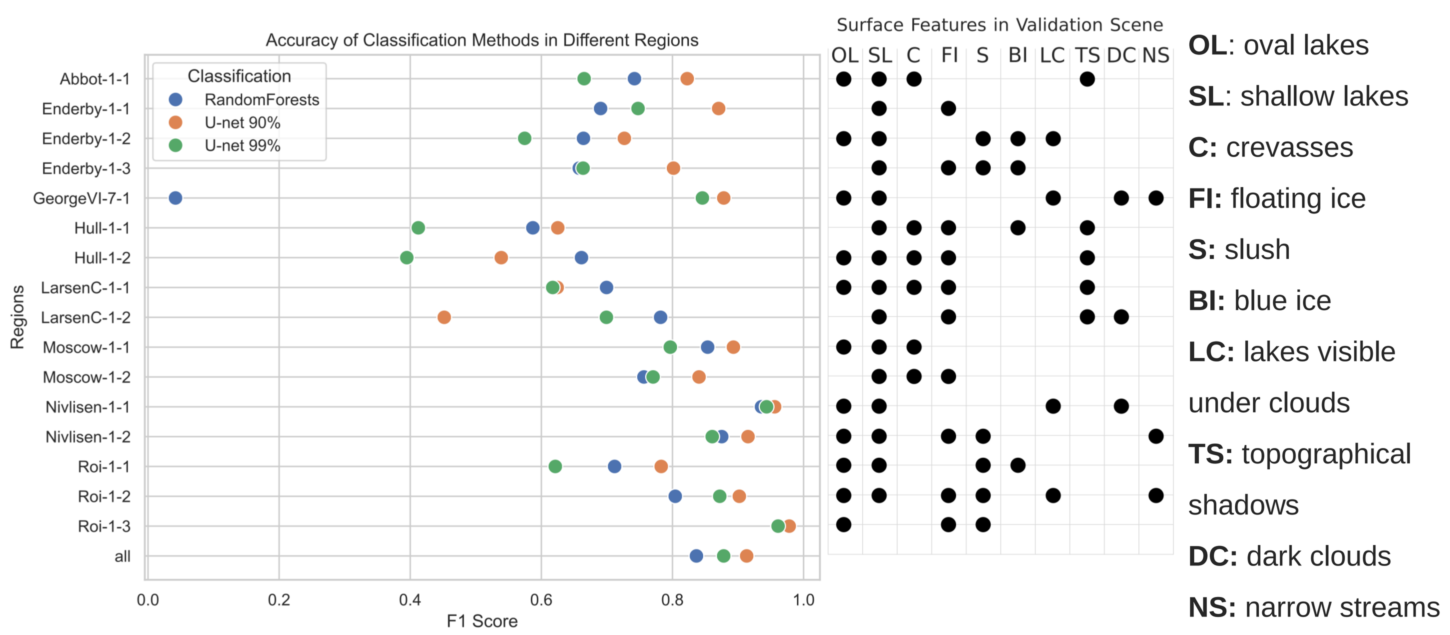

Fig. 2: Accuracy assessment of the Sentinel-2 U-net based supraglacial lake detection approach and comparison to the existing Random Forests-based approach for each validation scene. The table on the right shows which surface features are present in the scene [3]. See Fig. 1 for visual examples of these surface features.

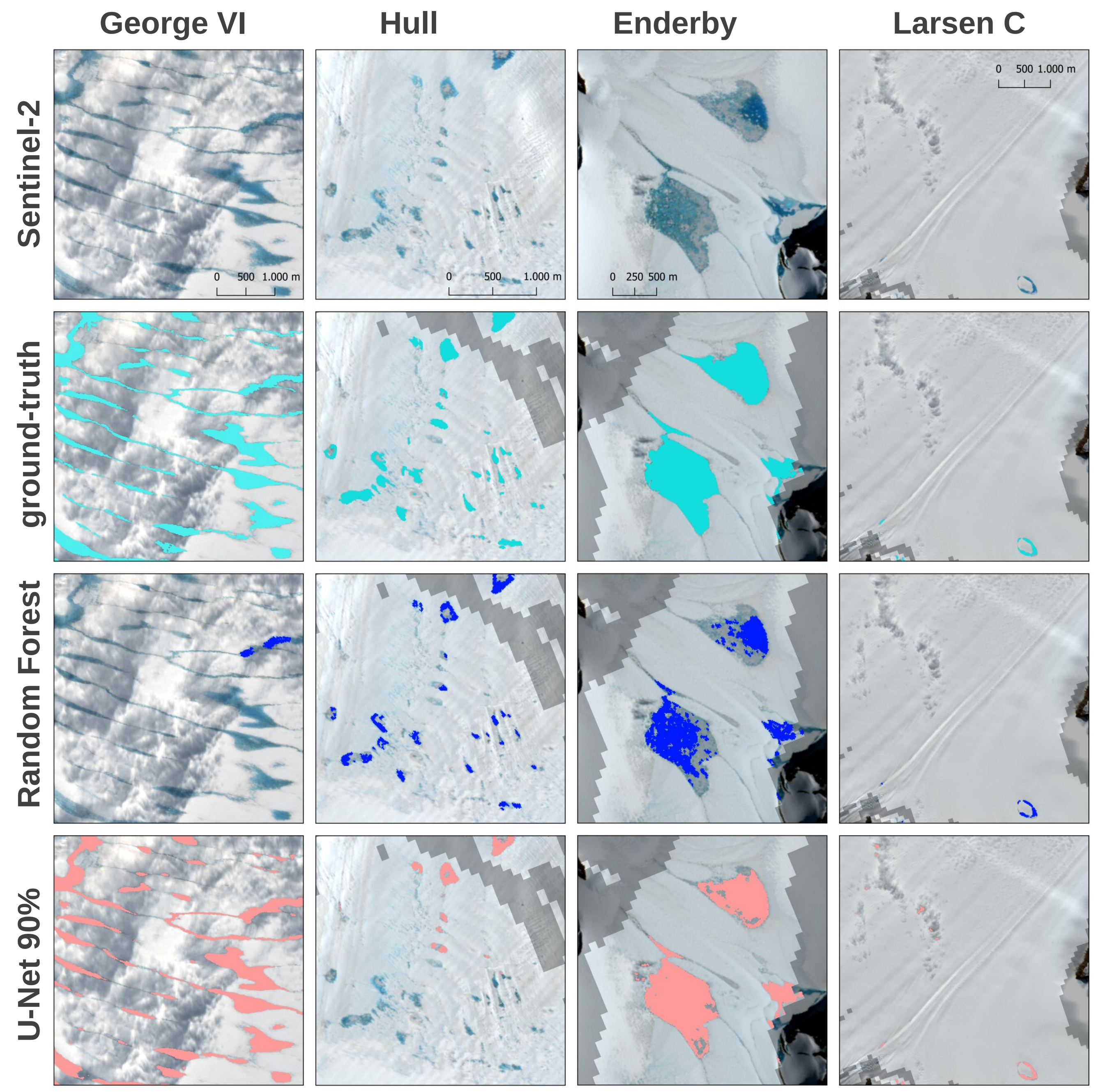

Fig. 3 : Examples of SGL predictions above different surface features and comparison of the Sentinel-2 U-net and RF-based classification approaches. Columns show subsets of four different validation sites. Rows show the original RGB Sentinel-2 image, the ground-truth SGL labels and the results of the evaluated classifiers.

Validation Results

The RF approach achieved a producer's accuracy, user's accuracy, and F1 score of 0.750, 0.945, and **0.837**, respectively, whereas the CNN-based workflow achieved scores of 0.915, 0.912, and **0.913**, respectively. In scene-specific comparisons, the CNN outperformed the RF approach in 13 of the 16 validation scenes. Key advantages of the CNN approach include its ability to detect lakes under thin clouds and floating ice, resulting in less fragmented lake area estimates and requiring fewer input features.

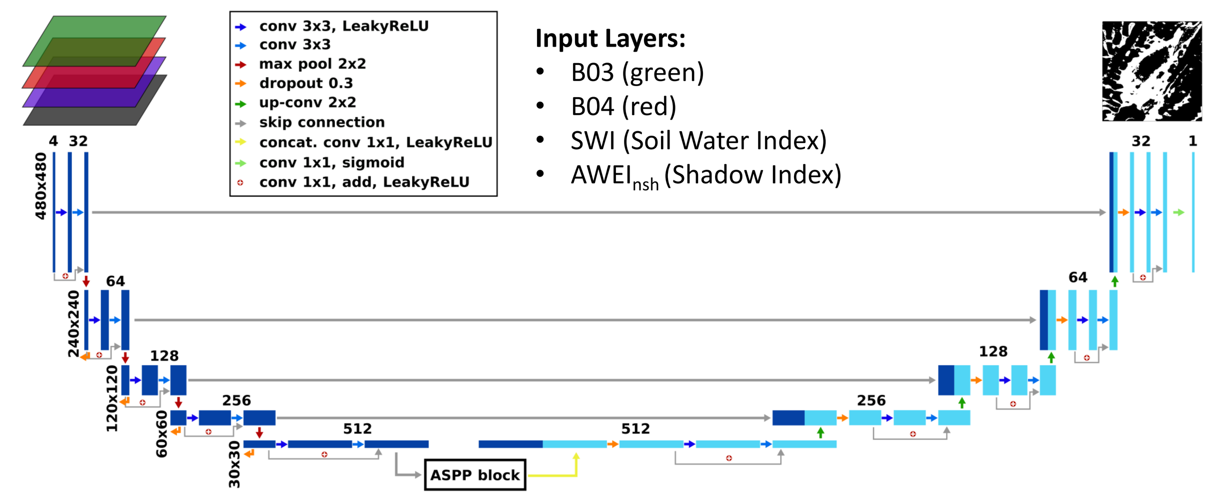

Fig. 4 : Architecture of the U-Net model used for Sentinel-2 DL-based supraglacial lake detection. The U-Net was only trained on the four input channels (red, green, SWI, AWEI$_{nsh}$). Channels were selected based on the highest feature importance from the RF-approach [1,3].

Resulting Dataset

By a combination of the DL approach for Sentinel-2 data [3] and a similar approach for SAR Sentinel-1 data [2] it is possible to create a combined dataset. The dataset provides bi-weekly SGL maximum extent information during austral summer months allowing the analysis of spatio-temporal lake dynamics at high spatial and temporal resolution over a time period of almost 10 years for most Antarctic ice shelves (see Fig. 5).

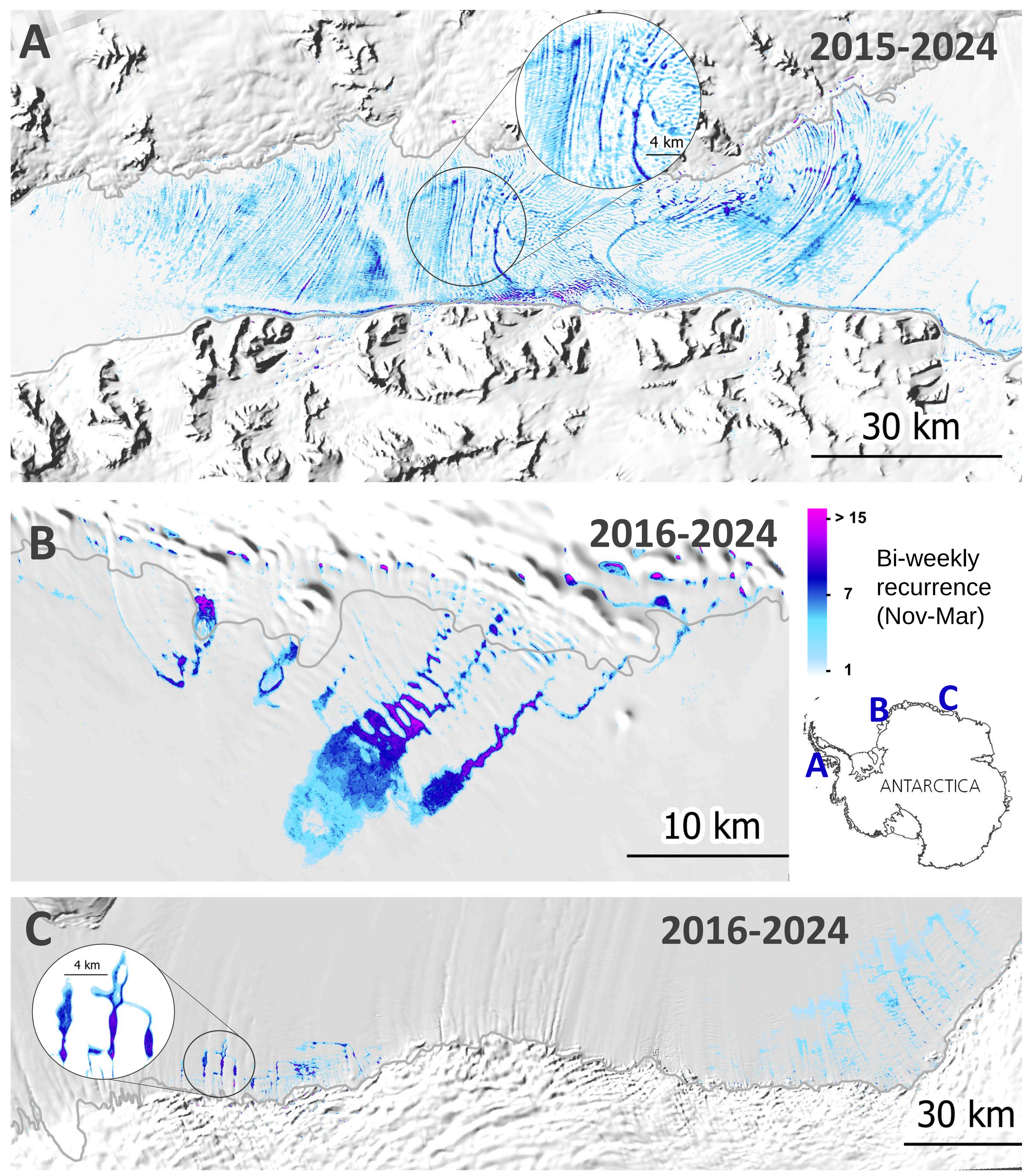

Fig. 5 : Exemplary supraglacial lake recurrence maps for selected ice shelves during the Sentinel-Era. (A) George VI Ice Shelf, (B) Riiser-Larsen Ice Shelf, and (C) Baudouin Ice Shelf. Background: REMA DEM.

Outlook

The presented approach will be used to generate a complete inventory of bi-weekly SGL occurrence maps derived from both Sentinel-1 and Sentinel-2 data for all Antarctic ice shelves. Early examples of such a product were already presented in this work. In future studies, this data can be used for validation of, and as an input into hydrological, firn and surface mass balance models to improve their ability to model surface melt.

References

[1] Dirscherl, M., Dietz, A.J., Kneisel, C., Kuenzer, C., 2020. Automated Mapping of Antarctic Supraglacial Lakes Using a Machine Learning Approach. Remote Sensing 12, 1203. https://doi.org/10.3390/rs12071203

[2] M. Dirscherl, A. J. Dietz, C. Kneisel, and C. Kuenzer, 2021. „A Novel Method for Automated Supraglacial Lake Mapping in Antarctica Using Sentinel-1 SAR Imagery and Deep Learning", Remote Sensing, Bd. 13, Nr. 2, Art. Nr. 2, 10.3390/rs13020197.

[3] Koehler, J., Baumhoer, C., Wouters, B., Lhermitte, S., Dietz, A., in review. Deep learning-based mapping of supraglacial lakes in Antarctica: A method using combined optical and SAR data. GIScience & Remote Sensing.

Dr. Celia Baumhoer
Team Polar and Mountain Regions
German Aerospace Center (DLR)
Celia.Baumhoer@dlr.de | https://celia.Baumhoer@github.io