Sea Ice Concentration & Machine Learning

An end-to-end workflow for estimating concentration maps from SAR images

Data Processing Pipeline — Cloud Development, Robust & Scalable

Machine Learning Design — Synthetic Aperture Radar, Custom Loss Function

Close Navigation Support — Python Search Interface, Results On-Demand

Existing Model Reliability — Retrospective Comparison, In Situ Observations

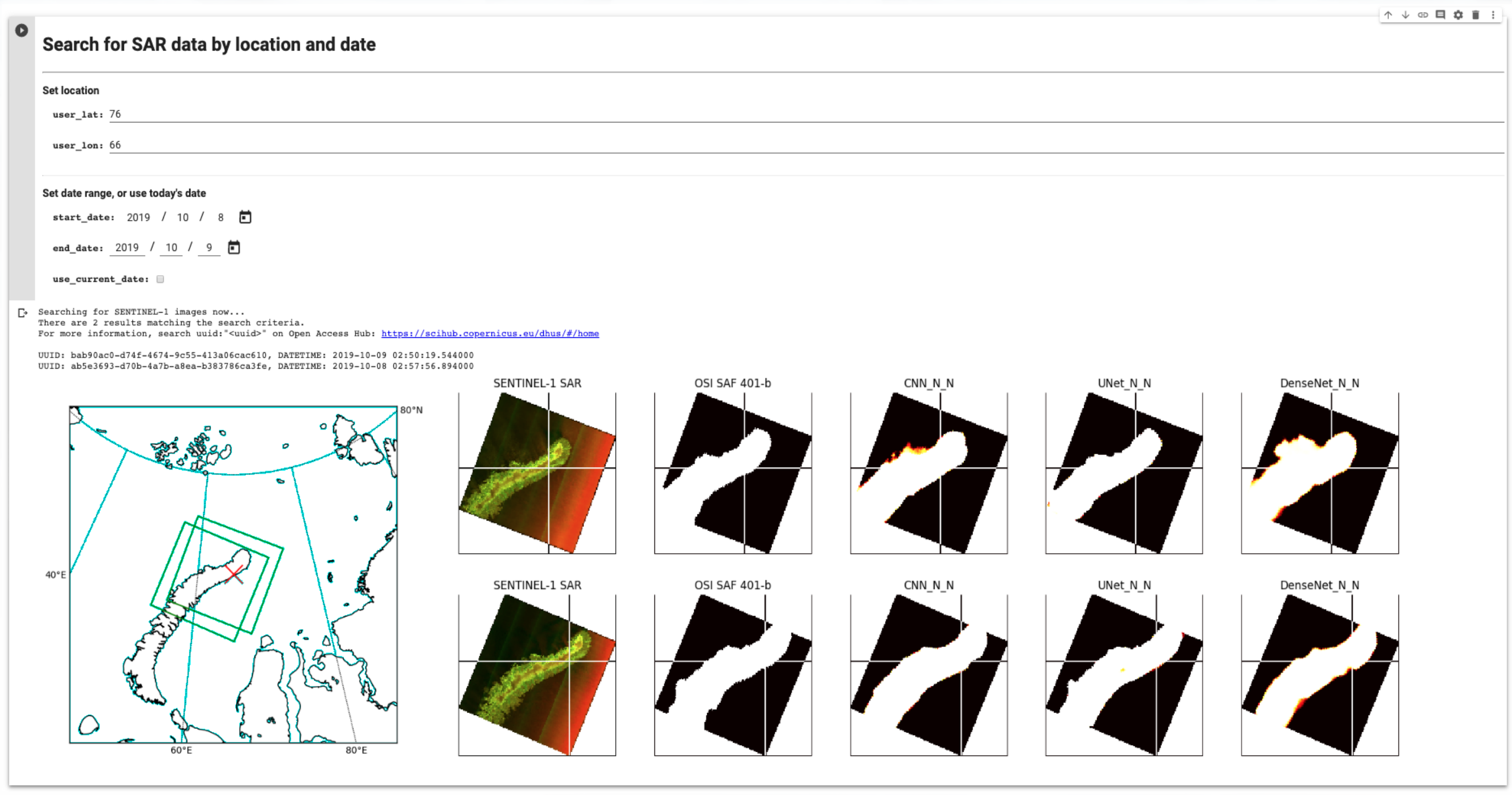

Figure 1: Search Interface developed to display estimation results from a selection of trained models. SAR images are found by entering a location and data of interest, then each image is rotated into alignment and displayed.

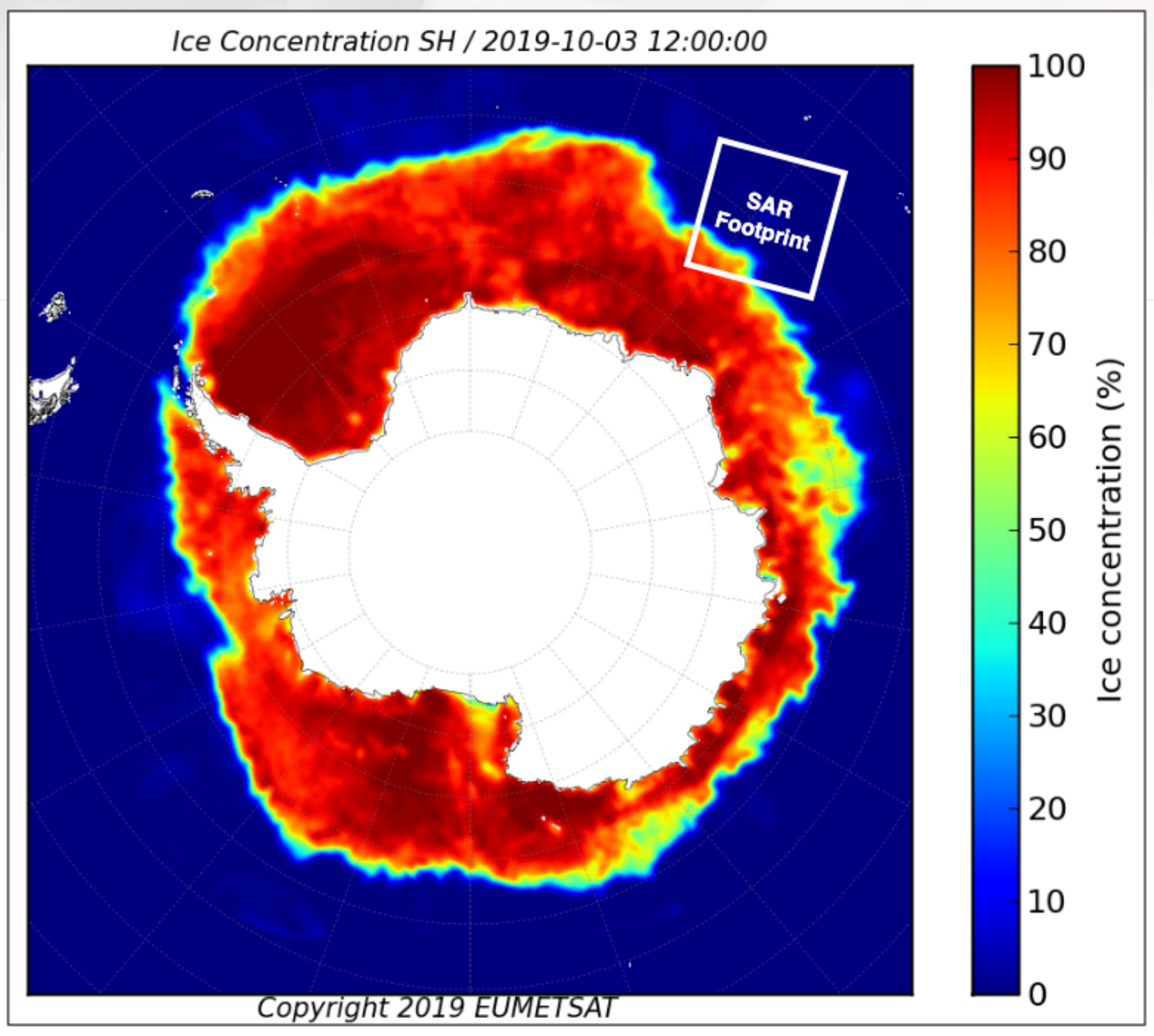

Figure 2: The large SIC map is a PMW-derived product, and a SAR footprint has been overlaid to illustrate the difference in coverage.

1. What's wrong with current models?

Sea ice concentration (SIC) is an important metric used to characterise polar sea ice behaviour. Understanding this behaviour and accurately representing it is of critical importance for climate science research, and also has important uses in the context of maritime navigation. Current models used to estimate concentration maps are derived from passive microwave (PMW) radar satellite data, but these instruments offer poor spatial resolutions, and are susceptible to interference due to atmospheric variations. Additionally, these models are usually calibrated for Arctic conditions, and evidence from recent in situ observations in the Southern Ocean suggests that their estimates are significantly biased and are not reliable in the Antarctic marginal ice zone (MIZ).

2. Why Machine Learning?

The reasons for pursuing a machine learning approach to generating SIC estimates were based on the hypothesis that a sufficiently well-designed machine learning structure would be able to overcome the inaccuracies inherent in the existing PMW-derived models, even if these inaccuracies are not fully-understood. This is a non-trivial claim, but the results have shown positive signs indicating that the underlying theory has promise.

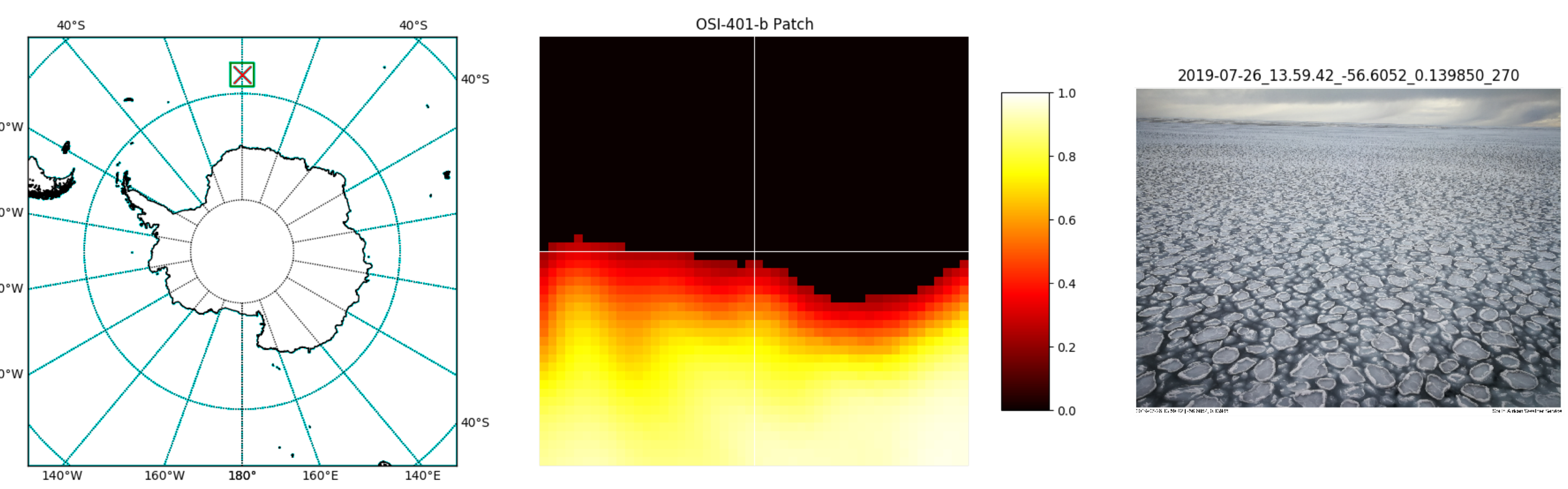

Figure 3: An example of the comparison tool output, showing that PMW estimations differ significantly from reality in the MIZ.

3. Results

To evaluate the performance of different model architectures and training strategies, 18 models were trained using different datasets and architectures. All training performance metrics and loss trajectories were recorded and analysed. All but three of these were able to achieve a weighted mean absolute error (MAE) less than 10%. The best performance results on unseen test data from each architecture are shown in the table below. One of the key additions to the training process was the custom weighted loss function, which used the uncertainty of each PMW estimation to prevent each new model from learning uncertain details an inaccuracies. Both the weighted MAE and the pure MAE are reported here for comparison.

	Best FCNN	Best U-Net	Best DenseNet
Weighted MAE	6.31%	3.24%	7.97%
Pure MAE	7.15%	3.69%	9.04%

4. Insights

Spatio-Temporal Resolution

Even after the original synthetic aperture radar (SAR) measurement has been decimated, the pixel spacing is still an improvement over PMW data. Since the neural network architectures were designed to replicate the input dimension in their estimates, all of the SIC estimates are at a resolution better than the original data with which they were trained. The temporal resolution of SAR acquisitions is also superior to PMW products. Usually two SAR images are available in the last 24 hours for high latitudes, while PMW SIC products are only published the following morning. This means that the most recent data available could be over 24 hours old, so the estimates generated from much more recent SAR data present a clear advantage.

Learning vs Numerical Models

Figure 4 was generated using the DenseNet architecture. The model prediction tile captures the small feature near the centre of the SAR image fairly well, which is not reflected at all in the PMW data (the 'target'), and so it would seem that the learned model has been able to differentiate in some way between features which are accurately represented in the training labels, and others which may be misleading.

Reliability of Existing Models

The comparison results shown in Figure 3 shows that there are significant errors in the concentration estimates within the Southern Ocean MIZ. The observation shown has been judged by experts to represent 100% ice concentration by surface area, and yet PMW estimate shows 0%. While this sample is one of the more drastic examples of this error, it illustrates beyond doubt that models designed and calibrated for the Northern hemisphere will not yield reliable results if they are naïvely applied in the Southern Ocean.

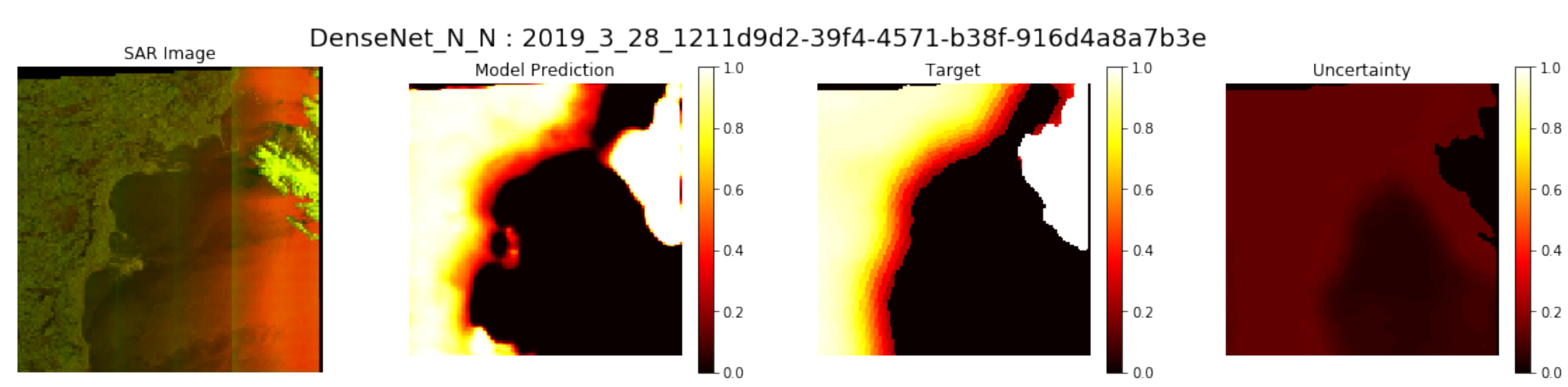

Figure 4: An example SAR image, DenseNet estimation, and PMW concentration and uncertainty labels. The DenseNet model was able to capture the small central feature, despite it not being represented in the training label.

5. Recommendations

- Define a more effective 'certainty' metric for use in the objective function.
- Investigate the effects of data augmentation in this context which have appeared to behave counterintuitively.
- Propose more effective ways to utilise the benefits of multi-stage training.
- Consider alternative satellite data sources which may have been overlooked.
- Consider scaling up the neural network input dimension to make better use of the high resolution SAR data which is so freely available.

Project GitHub Repository
Python Code, Datasets, Model Files, Project Report

Special Thanks

Christo Rautenbach
Chief Marine Scientist at South African Weather Service

Marc De Vos
Senior Marine Scientist at South African Weather Service

Stefan Dominicus
Final Year Student
BSc(Eng) Mechatronics