Machine Learning for Cloud Masking: A Res-Unet model for Sentinel-3 SLSTR masking

Johan Sjöberg[1], Sam Jackson[2], Karel Adamek[1], Wesley Armour[1], and Jeyarajan Thiyagalingam[1,2]

Background

The SLSTR-instrument

- 11-channel temperature radiometer mounted on the ESA Sentinel-3 satellites
- Applications include including sea surface temperature monitoring, fire detection and climate studies.

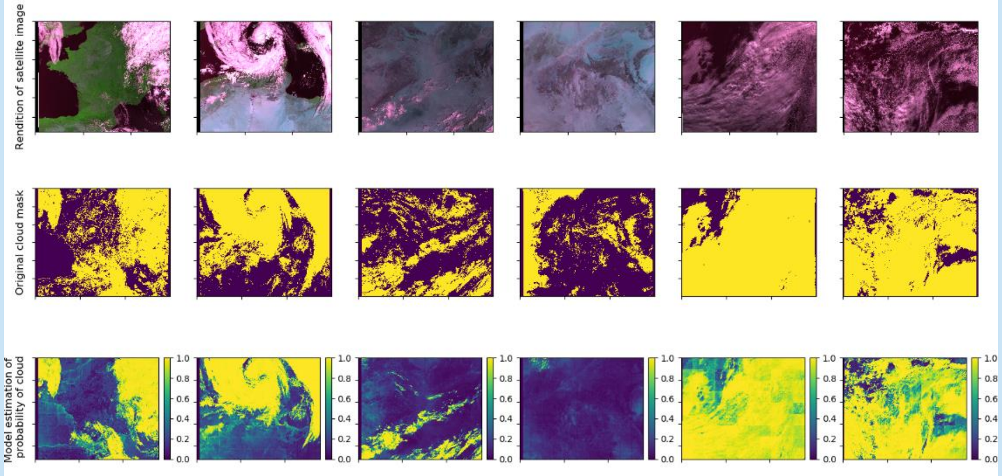
SLSTR Bayesian Cloud Mask

- SLSTR-product includes a Bayesian cloud mask
- There are several issues with this cloud mask:
    1. A high false alarm rate and a sensitivity to ocean fronts
    2. Long processing times (up to 13 min)
    3. It requires external data input including numerical weather predictions and is thus not autonomous

Project Goal

1. Develop a machine learning model that performs cloud masking on SLSTR-data
2. Ensuring that this model is accurate both compared to the SLSTR Bayesian cloud mask and outperforms the Bayesian mask when testing on actual 'ground truth' data
3. Ensuring that the model does this in real time without external data input

Example masks from the Bayesian and model masking algorithms



Model Architecture, Data, and Training Methodology

- Model based on Ronneberger's Unet but including short skip connections and operating on 128 by 128 pixel patches
- This architecture has been shown to perform well on a wide set of remote sensing applications
- Training was performed on a set of 2448 SLSTR-files using the Bayesian mask as a target
- Testing was partly done on 112 SLSTR-files and their Bayesian mask
- Due to the issues with the Bayesian mask, the model mask was additionally tested against CALIOP LIDAR data and hand labelled pixels[3]
- In order to test against CALIOP data, an SLSTR-CALIOP pixel collocation algorithm was developed. A pixel perturbation analysis was performed to understand what time difference was acceptable

Res-Unet model mask results and Bayesian mask comparison

	Bayesian mask			CALIOP LIDAR			Hand Labelled pixels		
	Accuracy	Recall	Specificity	Accuracy	Recall	Specificity	Accuracy	Recall	Specificity
Res-Unet	84.3 %	89.4 %	74.1 %	69.5 %	84.1 %	75.4 %	57.8 %	82.4 %	35.3 %
Bayesian	NA			79.3 %	78.7 %	75.3 %	58.7 %	92.9 %	7.7%

The model is highly accurate when tested against the Bayesian mask and outperforms the Bayesian mask on a specificity basis when tested against 'ground truth' datasets

Machine Learning based models can achieve high levels of accuracy in cloud masking both compared to the Bayesian mask and compared to the 'ground truth' without the need for external data

1. Oxford e-Research Centre, Engineering Science, University of Oxford, 7 Keble Rd, Oxford, UK, OX1 3QG , 2. Rutherford Appleton Laboratory, Science and Technology Facilities Council, Harwell Campus, Didcot, OX11 0QX ,3. The model was tested on this dataset by Sam Jackson. The dataset was kept internal to RAL Space and was provide courtesy of Brockman Consulting