PML | Plymouth Marine Laboratory

Research excellence supporting a sustainable ocean

AI for Earth Observation: The NEODAAS perspective

Katie Awty-Carroll & Dan Clewley

UKRI Natural Environment Research Council

NEODAAS

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

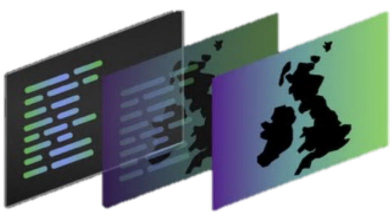
Summary

- The NERC Earth Observation Data Analysis and AI Service (NEODAAS) provides a range of services to EO researchers
- There is increasing interest in AI for Earth Observation and research in this area is growing exponentially
- Research into AI for EO can be hindered by barriers to entry
- To bridge this gap, NEODAAS has introduced a new AI service for NERC eligible researchers to get access to support for AI applications
- Through NEODAAS, researchers can access expertise in applying AI to EO data in addition to accessing our GPU cluster dedicated to EO
- NEODAAS is already working on a variety of internal and external projects, including mangrove mapping, tree crown segmentation, and ship track detection
- Future work will focus on expanding our training capability, developing robust pipelines, and researching solutions to common problems

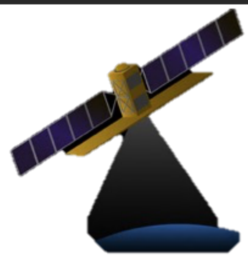
The NERC Earth Observation Data Analysis and AI service (NEODAAS)

- NEODAAS is hosted at Plymouth Marine Laboratory (PML), overseen by the National Centre for Earth Observation (NCEO) and funded by the Natural Environment Research Council (NERC)

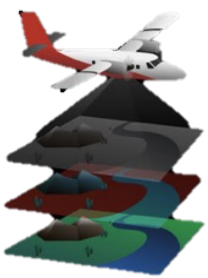
- NEODAAS provides a range of services to NERC eligible researchers:



Operational Satellite Data Processing for the Scientific Community



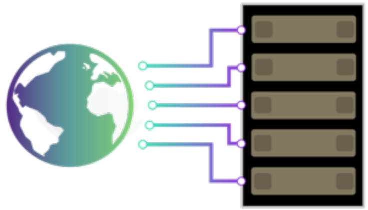
Near real-time support and rapid response using satellite data



Airborne data processing

Support and training

Development of new products

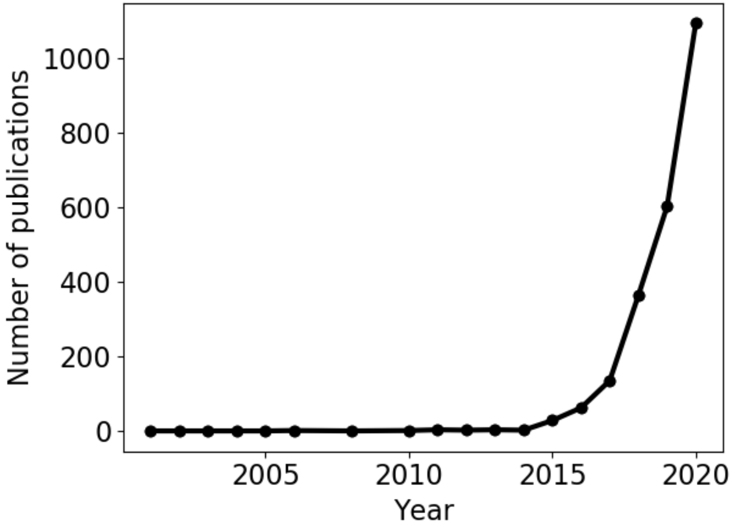


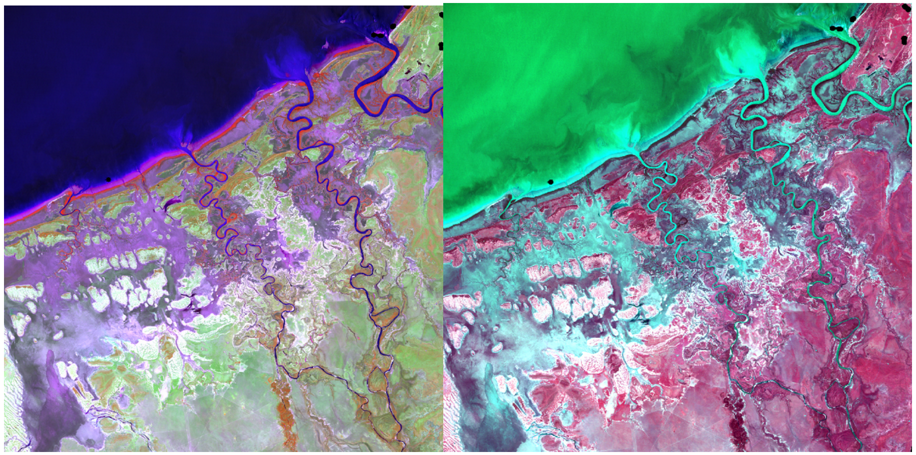
Massive GPU Cluster for Earth Observation (MAGEO) and AI service

Need for AI service

- AI applications in the EO field are growing exponentially, and this will likely continue into the near future
- EO data present unique challenges for machine learning, as well as advantages
- There is a need for training and up-skilling as current materials often only focus on “standard” images and datasets, and do not translate well to EO
- It is important for methodologies to be robust and consistent, with comparisons to existing baselines

Articles published per-year on the topic of "deep learning" and either "satellite" or "remote sensing", in the subject area of Earth sciences.

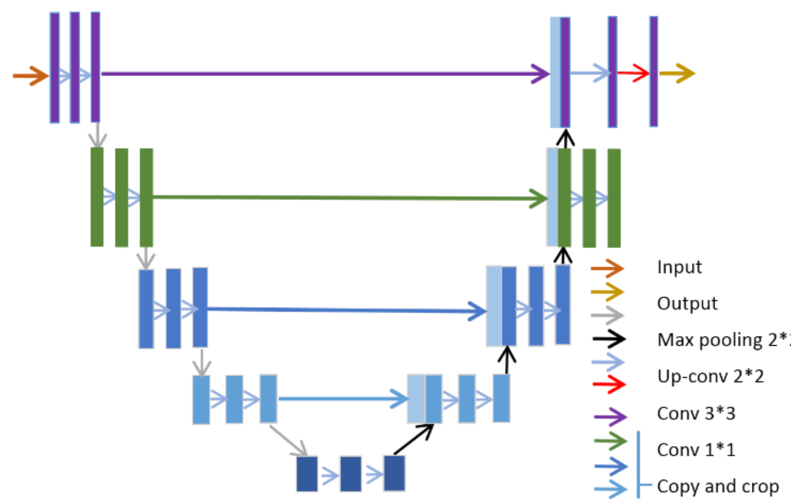




False colour Landsat 8 composites of the Gulf of Carpentaria, North Australia. Many standard approaches are not designed to operate on multispectral data.

Barriers to uptake

- Lack of specific training and examples based around AI for EO, including pre-processing, model selection, parameter tuning, and results interpretation
- Lack of access to appropriate software and hardware
- Lack of support for up-scaling AI workflows to large EO datasets
- Lack of transparency of machine learning models (the “black box” problem)
- Lack of models available for transfer learning and pre-training



Ronneberger et.al.

The NEODAAS AI service

- The AI service is an addition to the NEODAAS offering designed to help bridge the knowledge gap facing many EO researchers

- NEODAAS can now offer support throughout the full data pipeline, from acquisition and pre-processing of data through to application of machine and deep learning algorithms

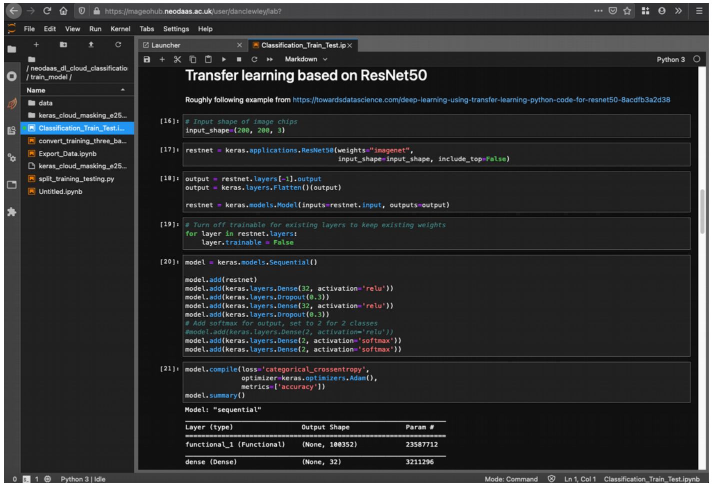
- NEODAAS users have access to a wide range of expertise in EO data processing and AI, in addition to optimised hardware and software for applying AI to EO data (MAGEO)

PML Plymouth Marine Laboratory

Research excellence supporting a sustainable ocean

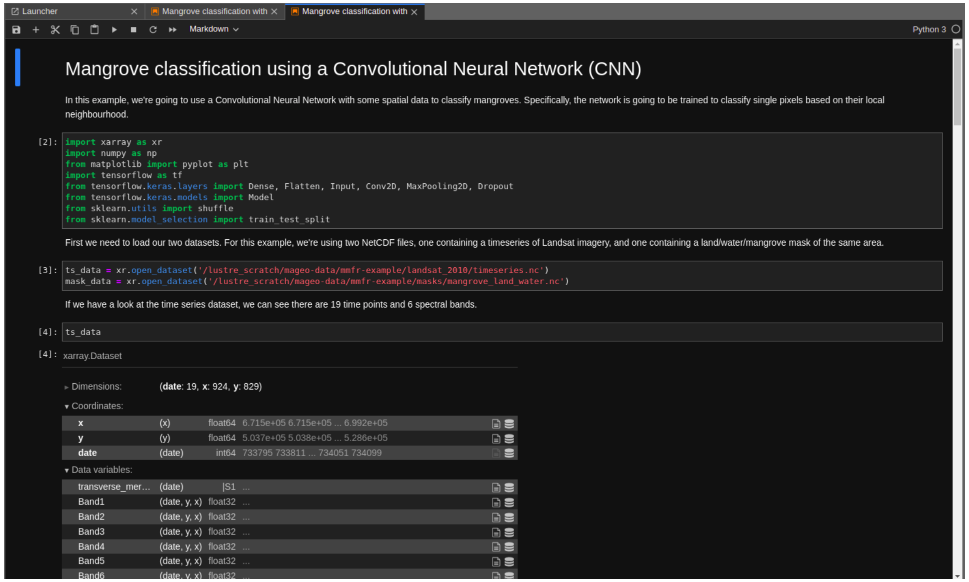
The MAssive GPU cluster for Earth Observation (MAGEO)

- MAGEO is a GPU cluster specifically designed for EO applications and became operational in early 2020
- MAGEO consists of 40 Tesla V100 GPUs totalling more than 204,000 GPU cores
- 0.5 PB Fast Storage + 6 PB existing storage
- AWS bill for the equivalent compute power to MAGEO would be ~$100,000 per month
- There is a JupyterHub web-based frontend for easy development and transfer of code
- Range of pre-installed, optimised libraries for data processing and machine learning

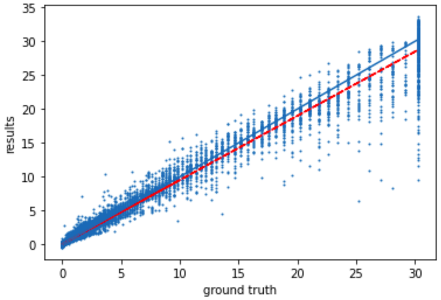


NEODAAS AI support - Training

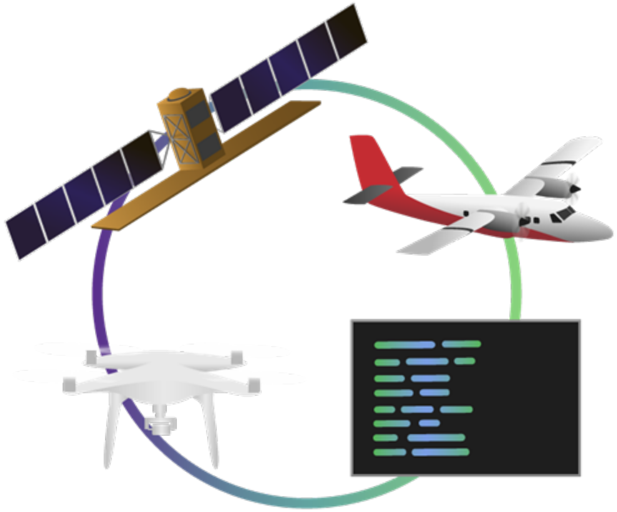
- Machine learning examples utilising Earth Observation data based on published and experimental work
- Ongoing user support during projects
- Training provision, e.g. for NERC CDT/DTP programmes



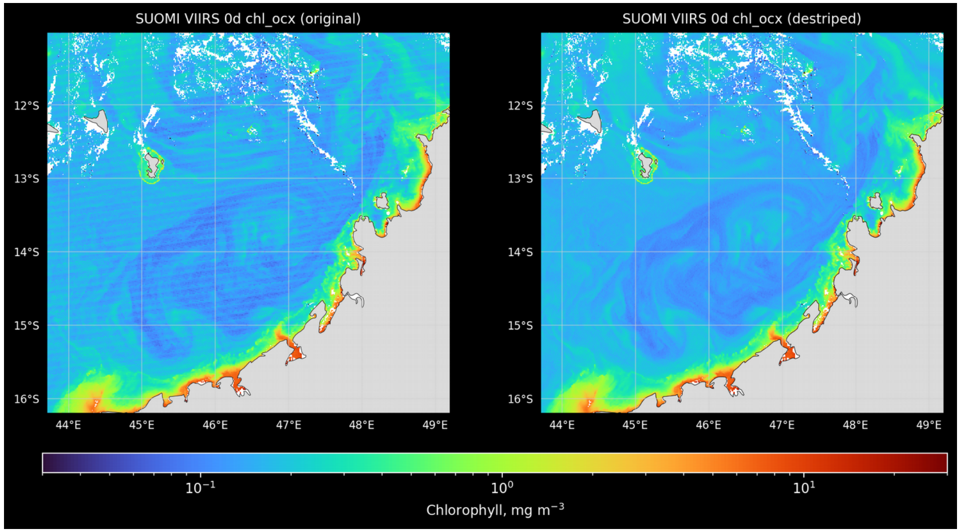

```
[ ]: import matplotlib.pyplot as plt
     import numpy
     plt.scatter(test_labels, results, s=1)
     plt.ylabel("results")
     plt.xlabel("ground truth")
     z = numpy.polyfit(test_labels, results, 1)
     print(z)
     p = numpy.poly1d(numpy.squeeze(z))
     plt.plot(test_labels,p(test_labels),"r--")
     plt.plot((0,test_labels.max()),(0,test_labels.max()))
     plt.show()

[[0.948295  ]
 [0.03109106]]
```

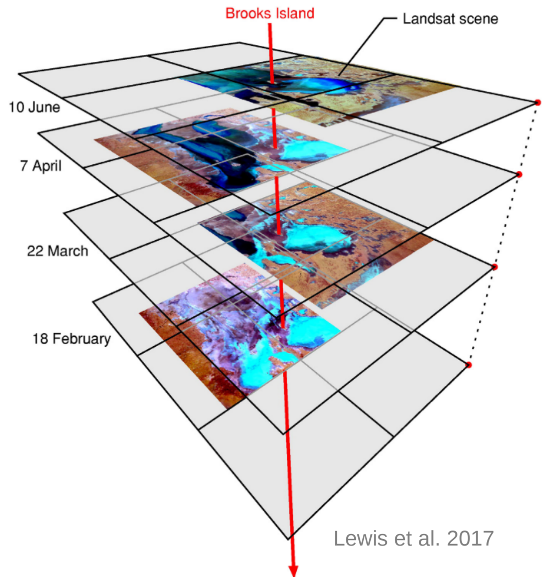


NEODAAS AI support – Data



- Custom data products for use with AI/ML
- Large scale processing (e.g. using the RAPIDS library) to transfer existing ML workflows to GPUs
- Data storage and management
- Data pre-processing, feature selection, and augmentation

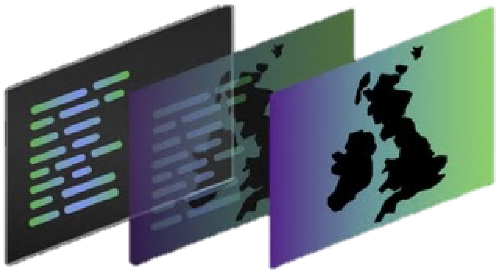


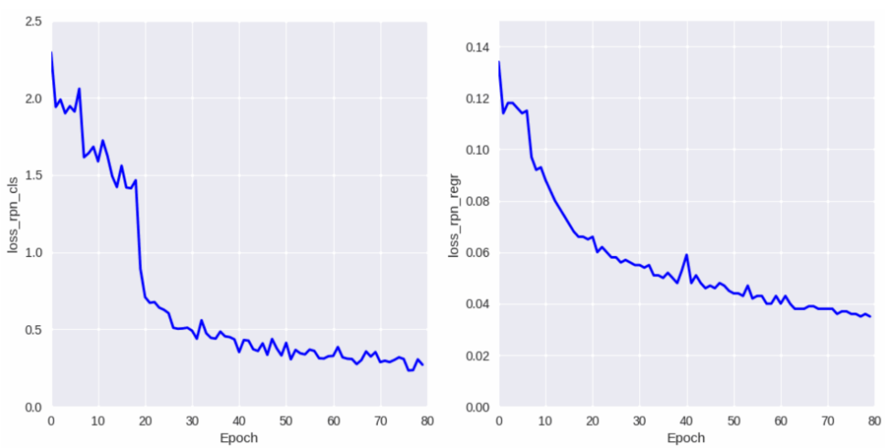


Lewis et al. 2017

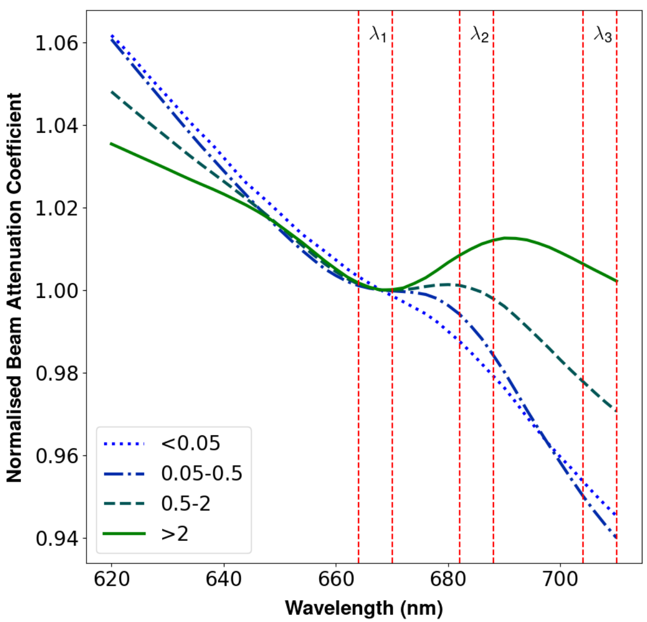
NEODAAS AI support - Research

- Support for customised software environments
- Expertise on model selection, training and optimisation
- Support for scaling up to multi-gpu and multi-Terabyte workflows, including large scale inference
- Support for inter-operability between platforms (e.g. JASMIN)





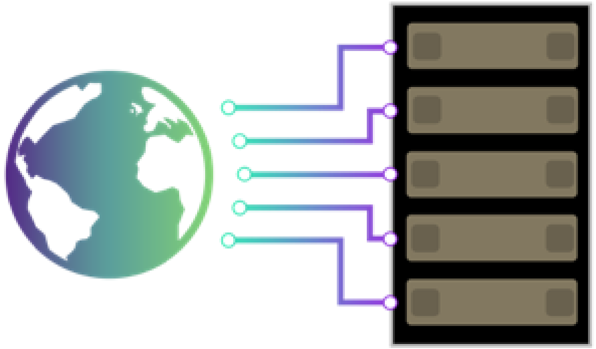
Classification and localisation loss (error) over time for a model being trained to detect vehicles in satellite imagery.



Relationship between particulate beam attenuation coefficient data and Chl-*a* concentrations at different wavelengths. The shape of Chl-*a* spectra at different concentrations can be exploited for prediction with ML.

How to apply for NEODAAS AI support

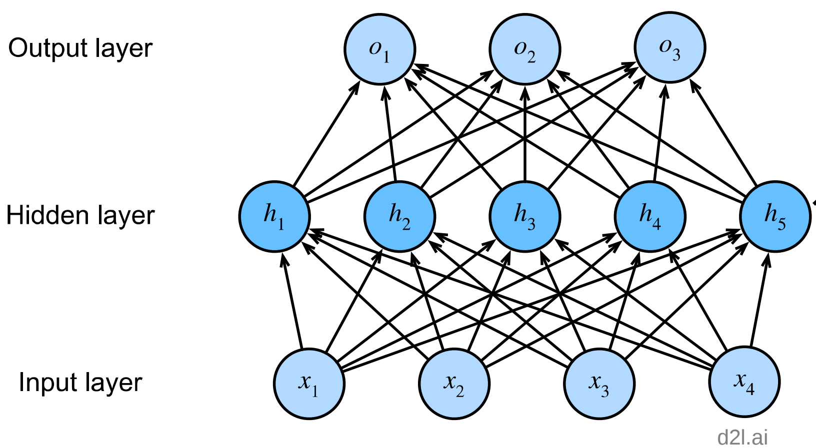
- NEODAAS services are available to UKRI eligible researchers
- AI support is accessed via extension to the standard NEODAAS request process
- Please contact us to discuss your requirements

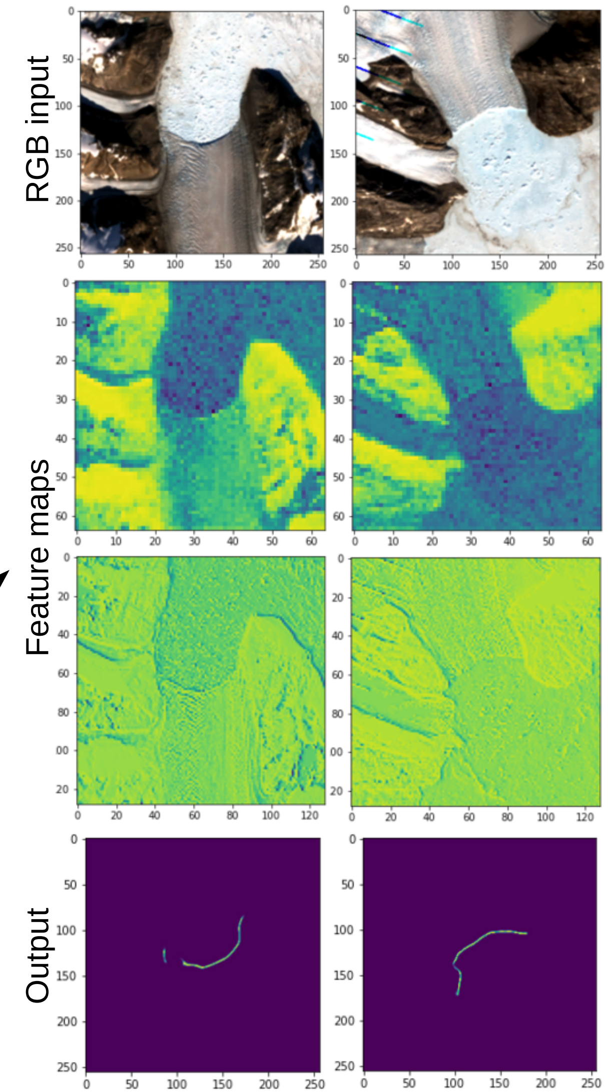


Examples

Glacial front detection

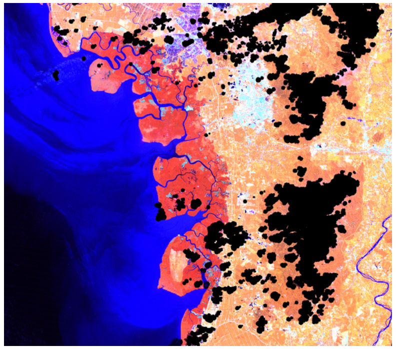
- Landsat 7 RGB imagery with hand-digitised fronts as training data
- Used EfficientNet with weights pre-trained on ImageNet dataset
- Feature maps can be extracted from the intermediate layers of the Convolutional Neural Network to aid in understanding how features are detected



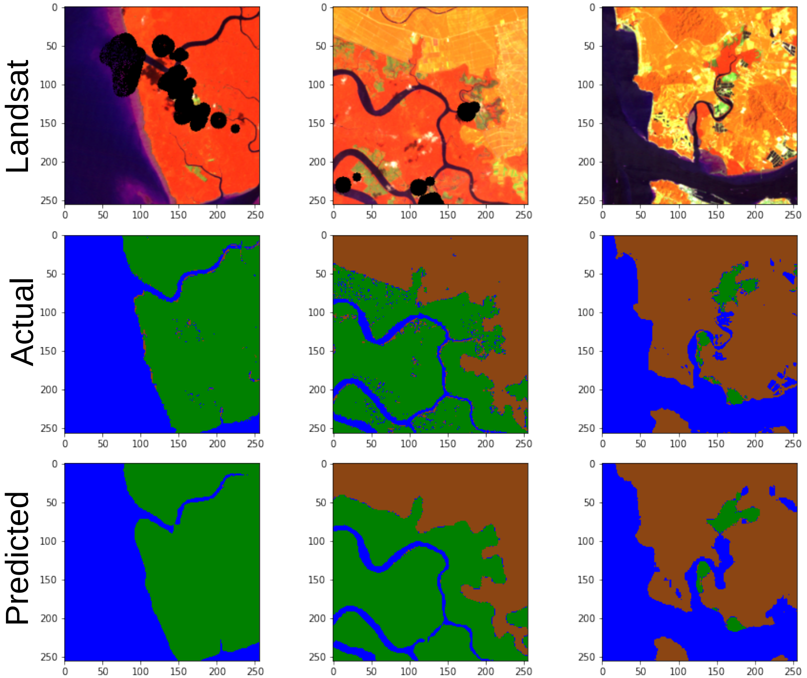


Mangrove mapping (Uni. of Aberystwyth)

- Data from full Landsat and Sentinel-2 archives across 10+ sites
- Open Data Cube used for data organisation and loading - https://www.opendatacube.org/
- Using U-Net to produce segmented maps of mangrove/water/other land cover based on 2010 Global Mangrove Watch baseline



False colour Landsat 5 image of the Matang Forest Reserve, Malaysia.



Mangrove mapping with deep learning.

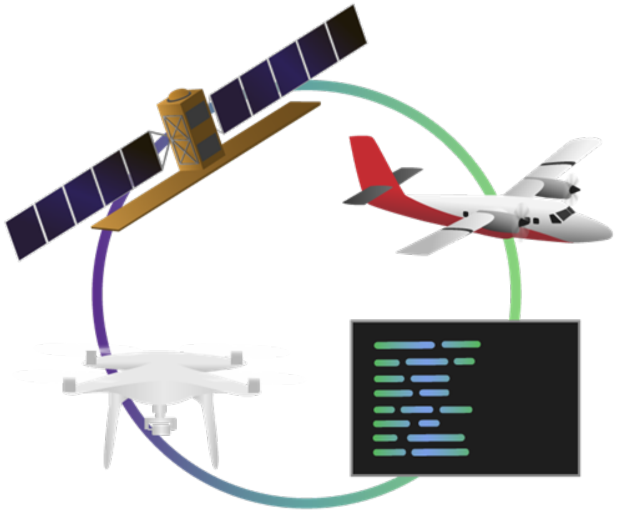
Tree crown segmentation (Uni. of Cambridge)

- MRes project on deep learning for segmentation of tree crowns in tropical forest
- LiDAR data used to create tree crown masks, then applied to RGB drone imagery
- Detectree algorithm developed based on Mask R-CNN - https://github.com/shmh40/detectree

Hickman and Ball, 2021

Conclusions and future work

- The NEODAAS AI service can provide support for researchers in applying AI to ground, airborne, and satellite data, meeting the increasing need for AI expertise in the EO field
- NEODAAS is continuing to expand its training offering to include practical courses on AI for EO
- We are developing robust, transferable pipelines for applying AI to EO data, in addition to looking at optimising pre-processing for GPU execution
- There is huge potential for research into model pre-training, for example, self-supervised learning as a pretext task
- Intermediate outputs such as feature maps can help reduce the black box aspect of AI



Thank you

- To access support from NEODAAS please contact helpdesk@neodaas.ac.uk
- More details about MAGEO and NEODAAS can be found at http://www.neodaas.ac.uk

- Dan Clewley (NEODAAS Manager)
  - dac@pml.ac.uk
- Katie Awty-Carroll
  - kawc@pml.ac.uk
  - @KawtyCarroll