PML | Applications Ltd

www.pml-applications.co.uk

Building The Welsh Data Cube

Dan Clewley, Carole Planque, Richard Lucas, Colin Sauze, Sebastien Chognard, Suvarna Punalekar, Clive Hurford, Chris Owers, Abigail Sanders, Pete Bunting, Emmanuel Nwokocha, Ben Calton

PRIFYSGOL ABERYSTWYTH UNIVERSITY

LIVING WALES

Sêr Cymru II

Ariennir gan yr Undeb Ewropeaidd
Funded by the European Union

Cronfa Datblygu Rhanbarthol Ewrop
European Regional Development Fund

Llywodraeth Cymru
Welsh Government

Tyfu Ymchwil Rhyngwladol Ragorol yng Nghymru
Growing International Research Excellence in Wales

LIVING WALES

Cronfa Datblygu Rhanbarthol Ewrop
European Regional Development Fund

Cymru Fyw

“ Sicrhau bod lloerennau’n gweithio dros ein hamgylchedd ”

Living Wales

“ Making satellites work for our environment ”

Centre for Alternative Technology
Canolfan y Dechnoleg Amgen

1872 PRIFYSGOL ABERYSTWYTH UNIVERSITY

Llywodraeth Cymru
Welsh Government

Why build a data cube for Wales?

- Framework for products being generated as part of Living Wales
- Support use of Earth Observation data in government
- Link with other international projects (e.g., Australia, Switzerland)

Promoting and building national capability in earth observation and economic success whilst ensuring long-term care and maintenance of the environment and resources

Prosperity for Wales

Providing new opportunities for economic development in all sectors by providing open access and usable earth observation and derived products to the population.

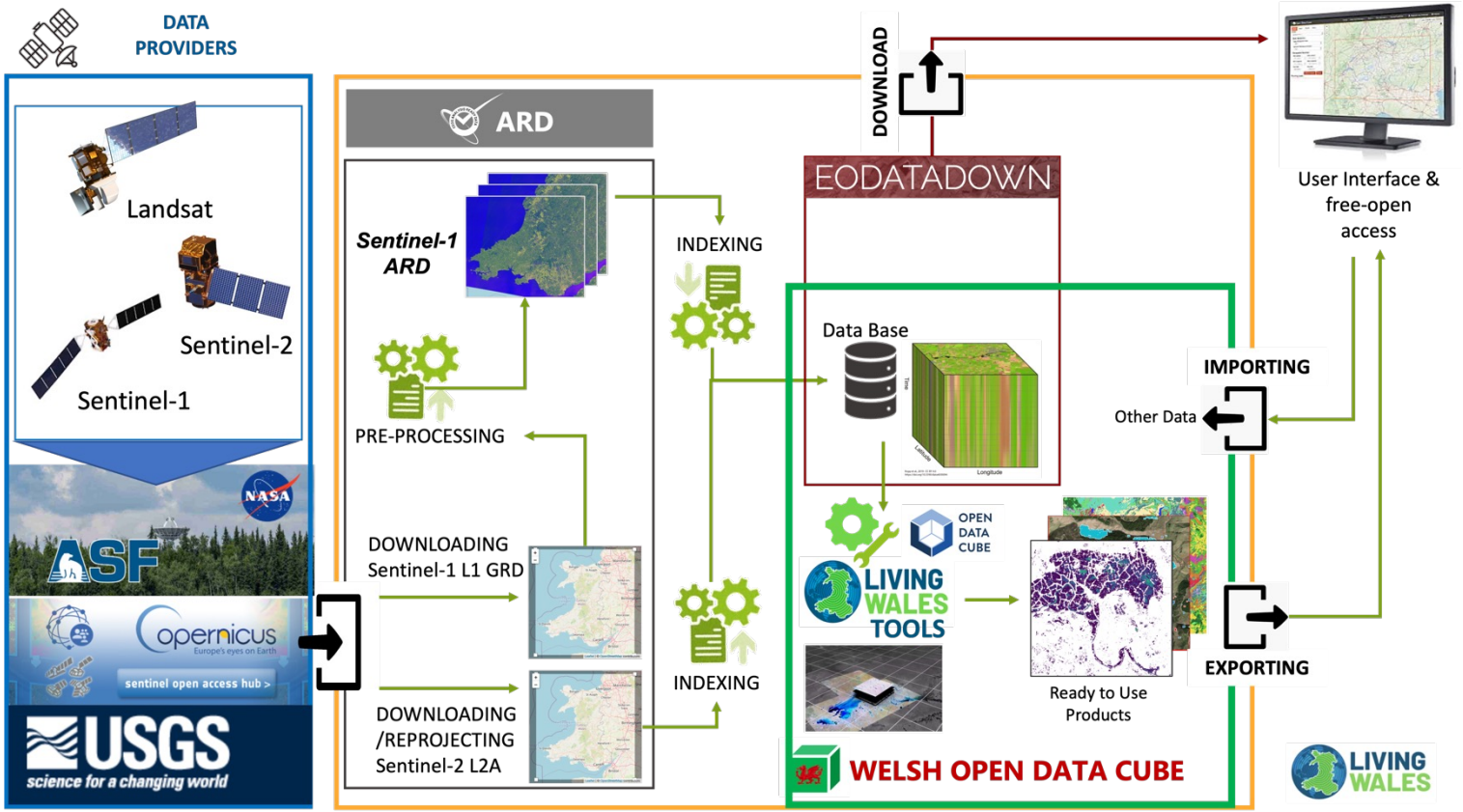
Sustainability for Wales

Providing a long-term system for understanding, monitoring and planning landscape change that is applicable at a national level and based on historical and near real time earth observations.

Resilience for Wales

Ensuring maintenance and promoting enhancement of the state and function of Welsh landscapes and their ability to respond to adverse environmental change through integration of earth observation data.

System overview

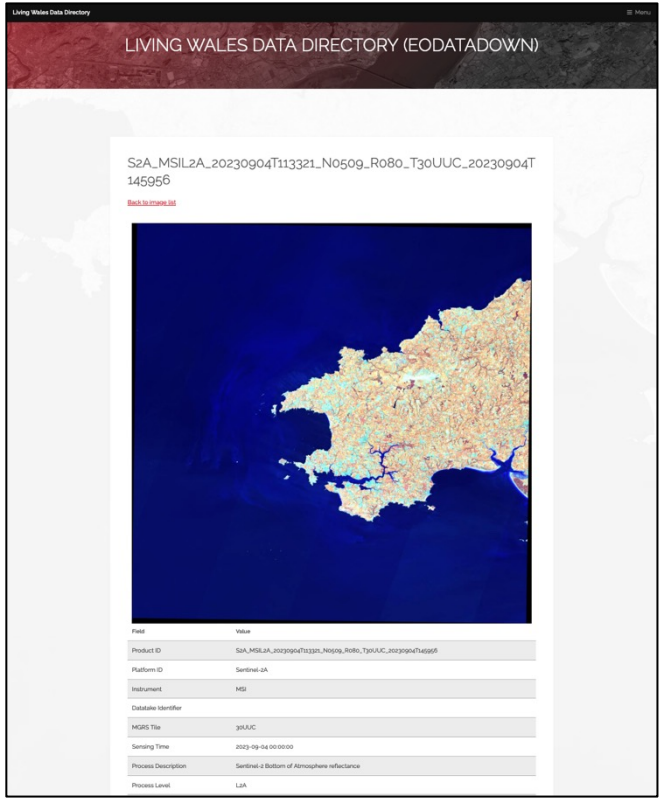


Core technologies: EODataDown

- Python library with postgres database
  - Classes for different providers / sensors
  - https://github.com/remotesensinginfo/eodatadown

- Performs the following
  1. Data search
  2. Data download
  3. Conversion to ARD
  4. Quicklook generation
  5. Map tile generation

- Flask web app for data search and download



https://livingwales.aber.ac.uk/

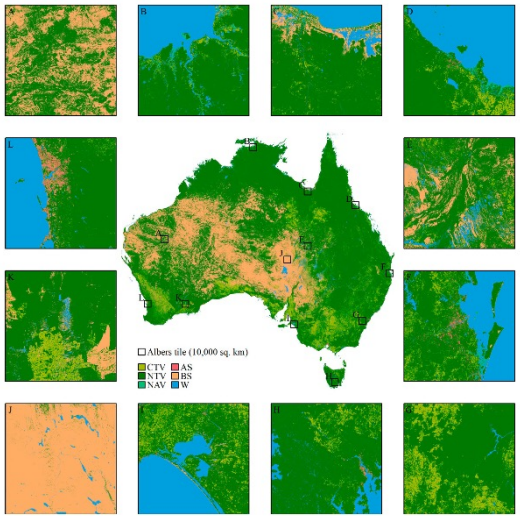
Core technologies: OpenDataCube

- Python library(s) with postgres database
  - https://github.com/opendatacube

```
import datacube
dc = datacube.Datacube()
dataset = dc.load(product="sen2_l2a_gcp",
                  x=(-4.095, -4.076),
                  y=(52.407, 52.422),
                  time=("2018-01-01", "2018-12-31"),
                  output_crs= "epsg:27700",
                  resolution= (-10,10))
```

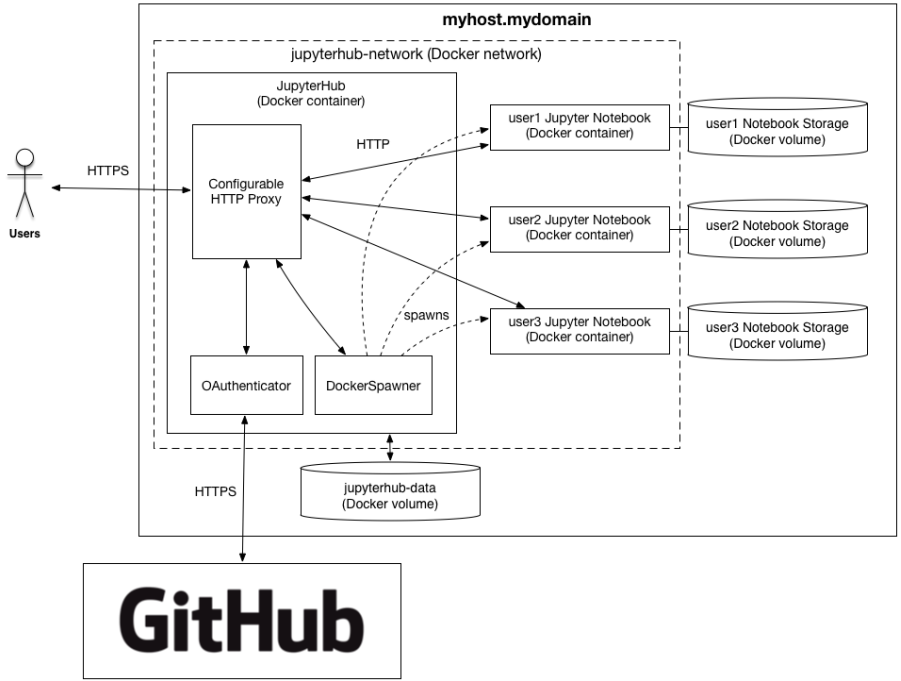
- Index data in database, query with Python library
  - Returns as xarray dataset
  - Can use dask

- Can use ‘virtual products’ to apply algorithms on the fly



Core technologies: JupyterHub

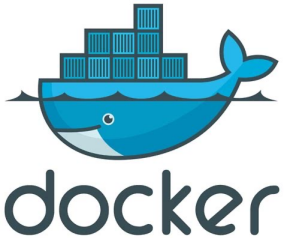
- Provides flexible toolkit for analysis
- Using Zero to JupyterHub with Kubernetes
  - https://z2jh.jupyter.org/en/stable/
- Investigated other technologies (e.g., the littlest jupyterhub) z2jh determined to be most futureproof

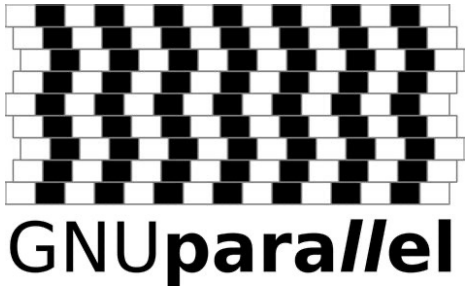


Current Deployment

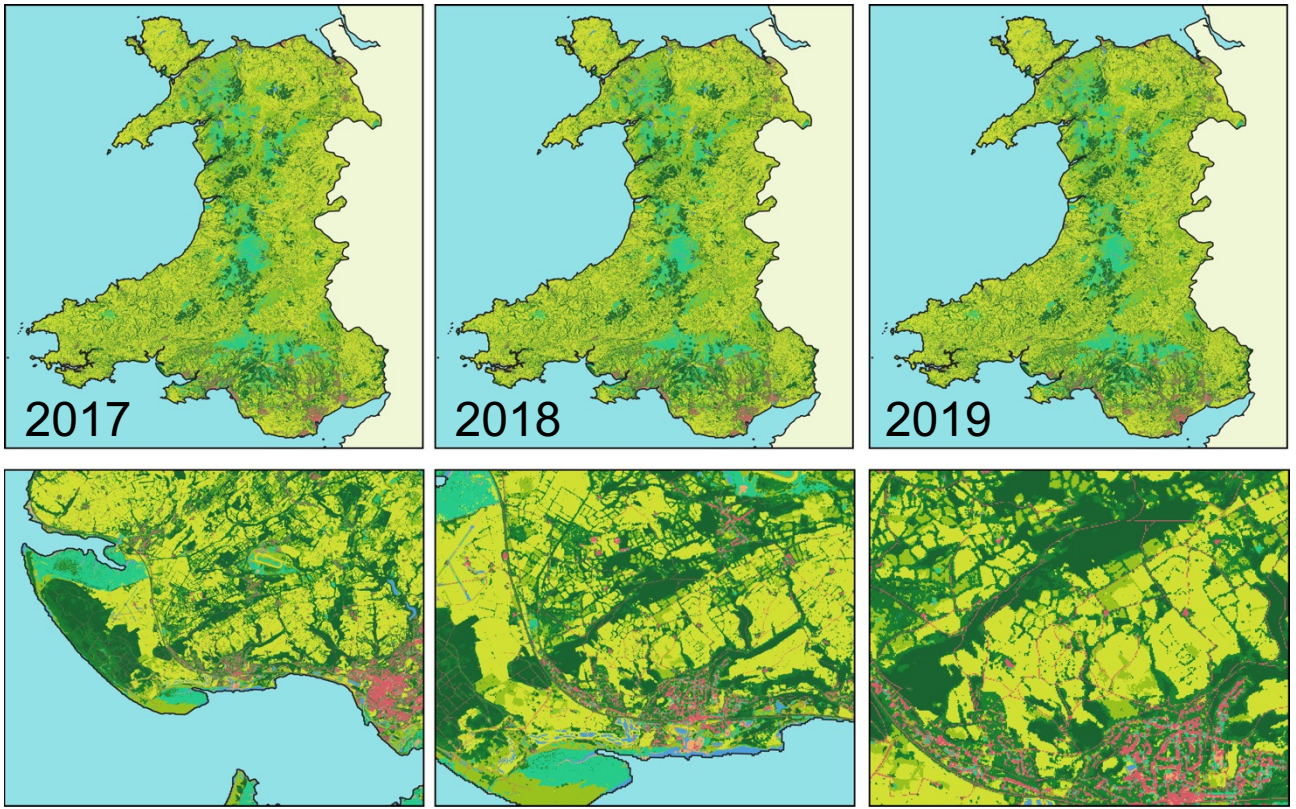
- Single server within Supercomputing Wales

- Technologies:
  - Singularity for EODataDown (avoids elevated permissions)
  - Docker for EODataDown front end
  - microk8s for JupyterHub
  - GNU parallel for batch processing as needed

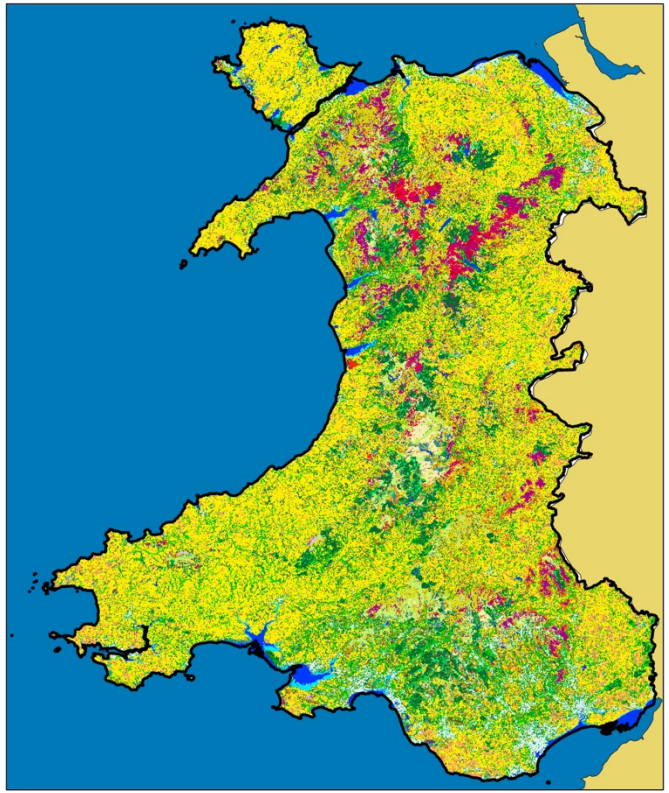


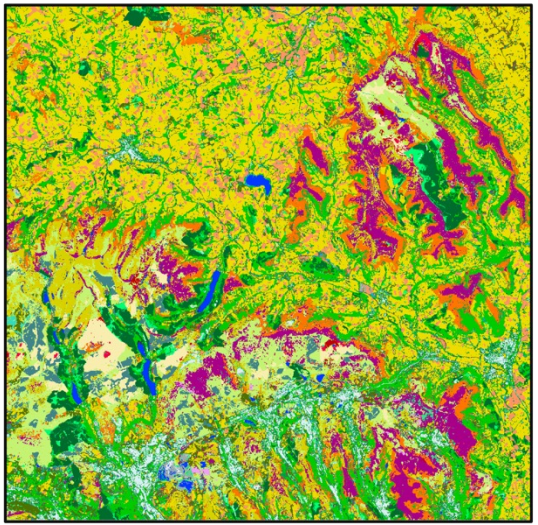


Outputs: National Land Cover Maps



Outputs: National Habitat Maps





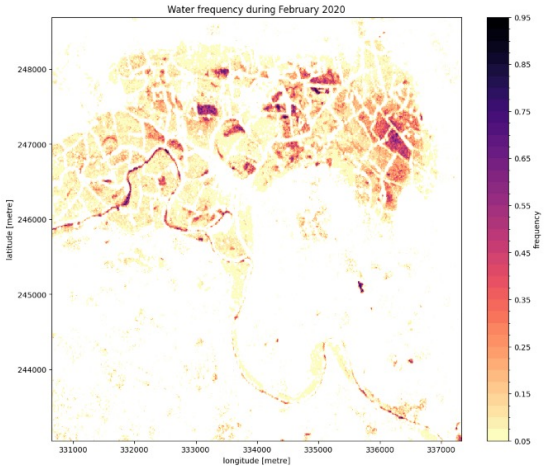
Living Wales habitat map (2020) translated from the FAO LCCS land cover map

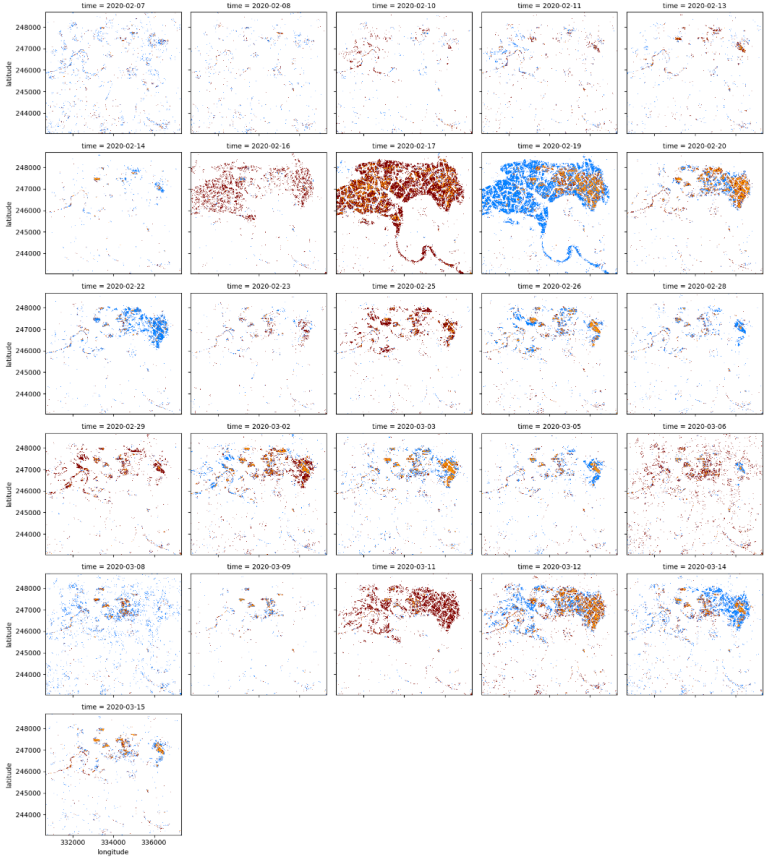
Habitat classes

- Semi-natural grassland
- Juncus rushes
- Molinia grassland
- Broadleaved woodland
- Coniferous woodland
- Acid grassland
- Improved grassland
- Marshy grassland
- Bracken
- Dry dwarf shrub heath
- Wet dwarf shrub heath
- Blanket bog
- Raised bog
- Modified bog
- Fen
- Swamp
- Open water
- Saltmarshes
- Sand dune
- Dune grassland
- Cultivated (arable)
- Natural bare surfaces
- Artificial bare surfaces

Outputs: Flood mapping

- Demonstrated can track floods
- Using Sentinel-1 SAR data
- As new data is downloaded, processed and indexed in OpenDataCube can track in real time

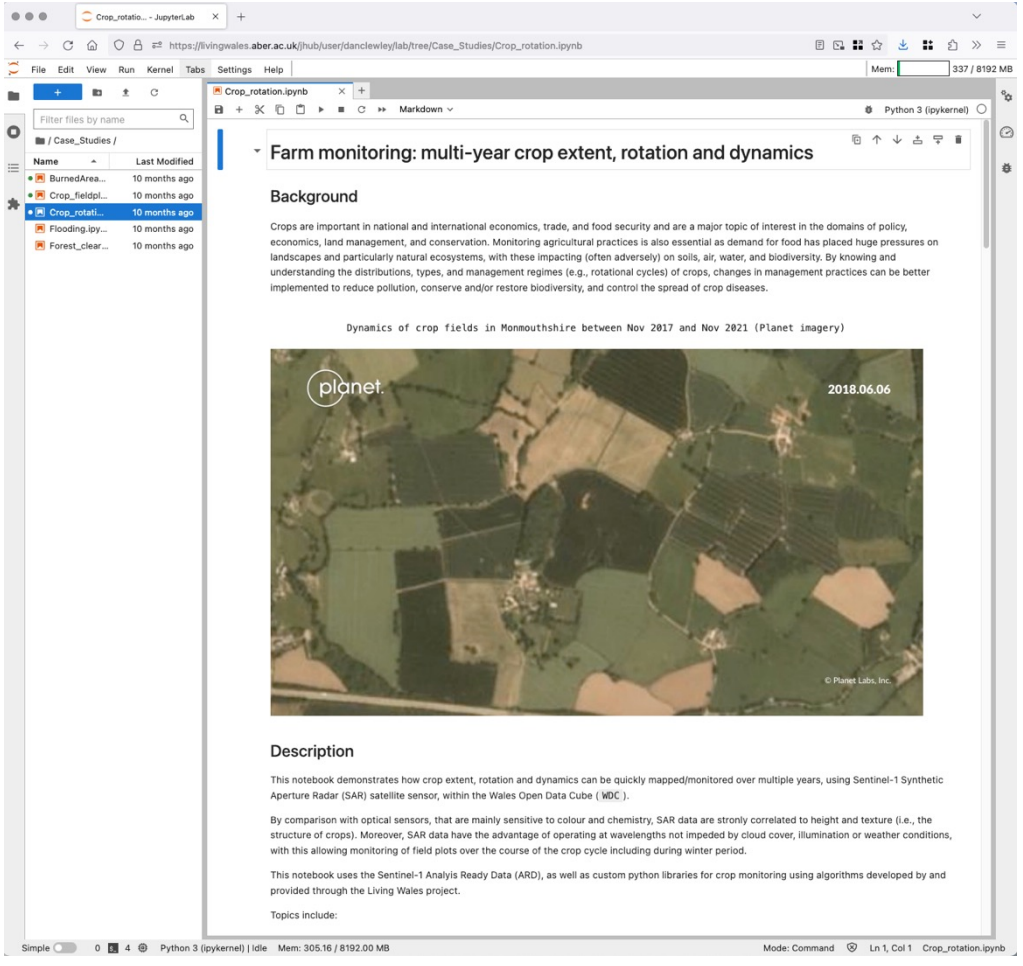




Outputs: Training

- Jupyter Notebooks guiding through common tasks

- Demonstrate going from large dataset to something which can be:
  - Added directly to reports
  - Used in Excel
  - Used in GIS package



Future plans

- Continue to support more users
- Continue developing products
- Continue international links
- Infrastructure moving to Microsoft Azure
  - Allows better separation
  - Can scale better to meet demand (more users, more resources)
- Better integration with DataMapWales

DataMapWales

Data and maps from the Welsh public sector

Thank you