Committee on Earth Observation Satellites

Jupyter for CAPD – UK Overseas Development Aid, Agritech, etc.

Esther Conway (NCEO/CEDA), Heiko Balzter (NCEO),

Mat Pryor (CEDA), Phil Kershaw (NCEO/CEDA)

WGISS - 49

Virtual meeting

22 April 2020

Objectives & Members

1. Establishing effective coordination and partnerships among CEOS agencies offering EO education and training

- Facilitate education and trainings
- Exploit the cumulative capabilities of CEOS Agencies
- Partner with local & regional partners to increase effectiveness and decrease duplication
- Focus on user needs for data and capabilities to inform action

2. Working with CEOS entities to address data accessibility and use

- Work closely with CEOS entities to increase data accessibility
- Publicize resources, datasets, and software
- Promote the use of dissemination systems to effectively reach areas that lack consistent internet access
- Organize workshops and training activities to provide individual and institutional capacity to effectively use EO resources

WGCapD Activities

E-Learning

Organize free E-Learning courses and webinars for participants around the globe

- **Webinars:**
  - Future Data Architecture
  - Sustainable Development Goals
- **MOOCs:**
  - Echoes in Space: Introduction to Radar Remote Sensing

Workshops

Organize free training workshops for participants around the globe

- **EO Training Workshops**
  - SAR trainings
- **Thematic Training Workshops**
  - Forest monitoring
  - Flooding
  - Etc.
- **Regional Workshops**
  - GEO – AmeriGEO & AfriGEO
  - Southeast Asia

Jupyter Notebook Collaboration ?

What is a Jupyter Notebook?

A Jupyter Notebooks allows you to combine

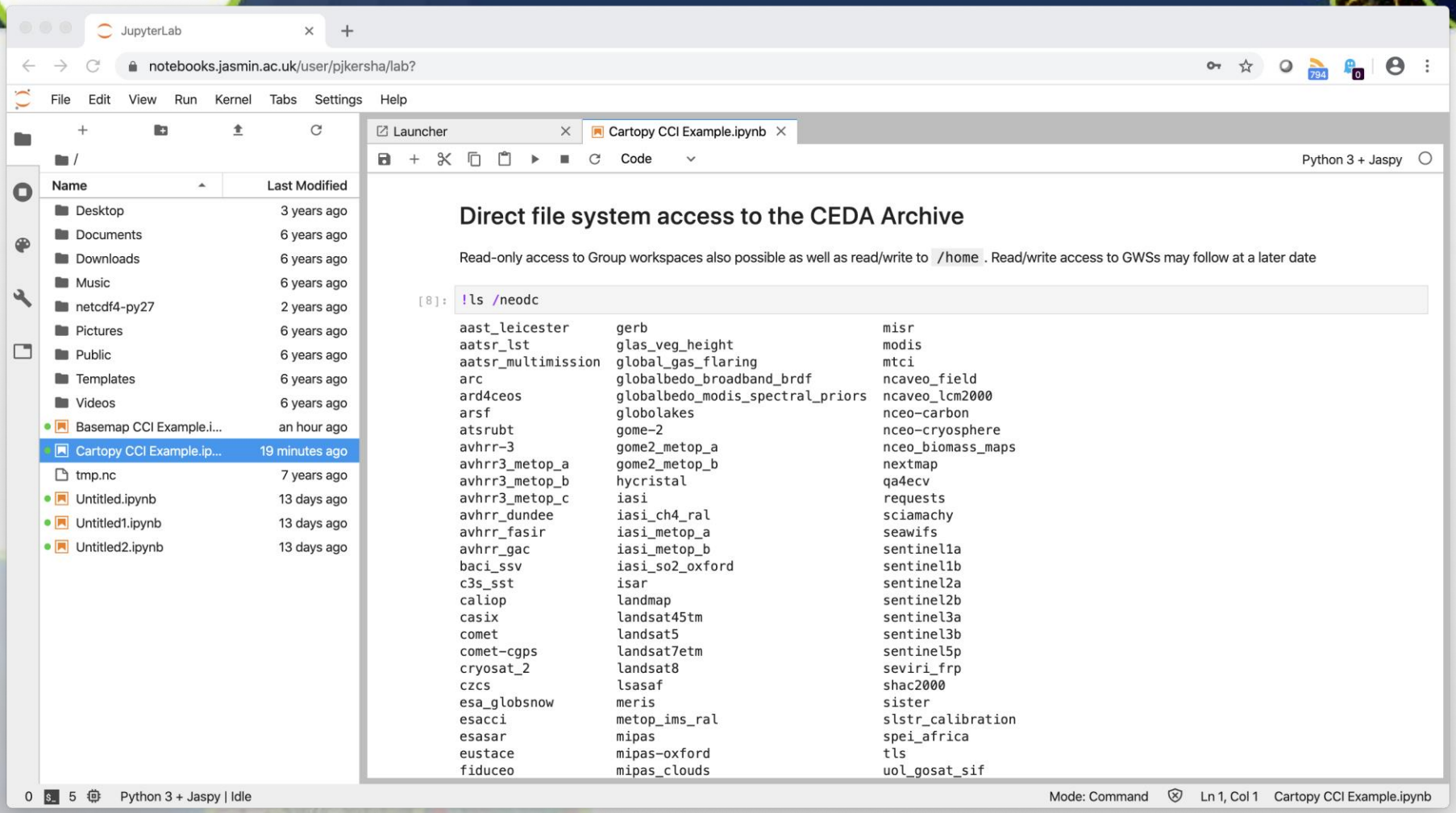
- **rich documentation**
- **live and adaptable code**
- **data visualizations**

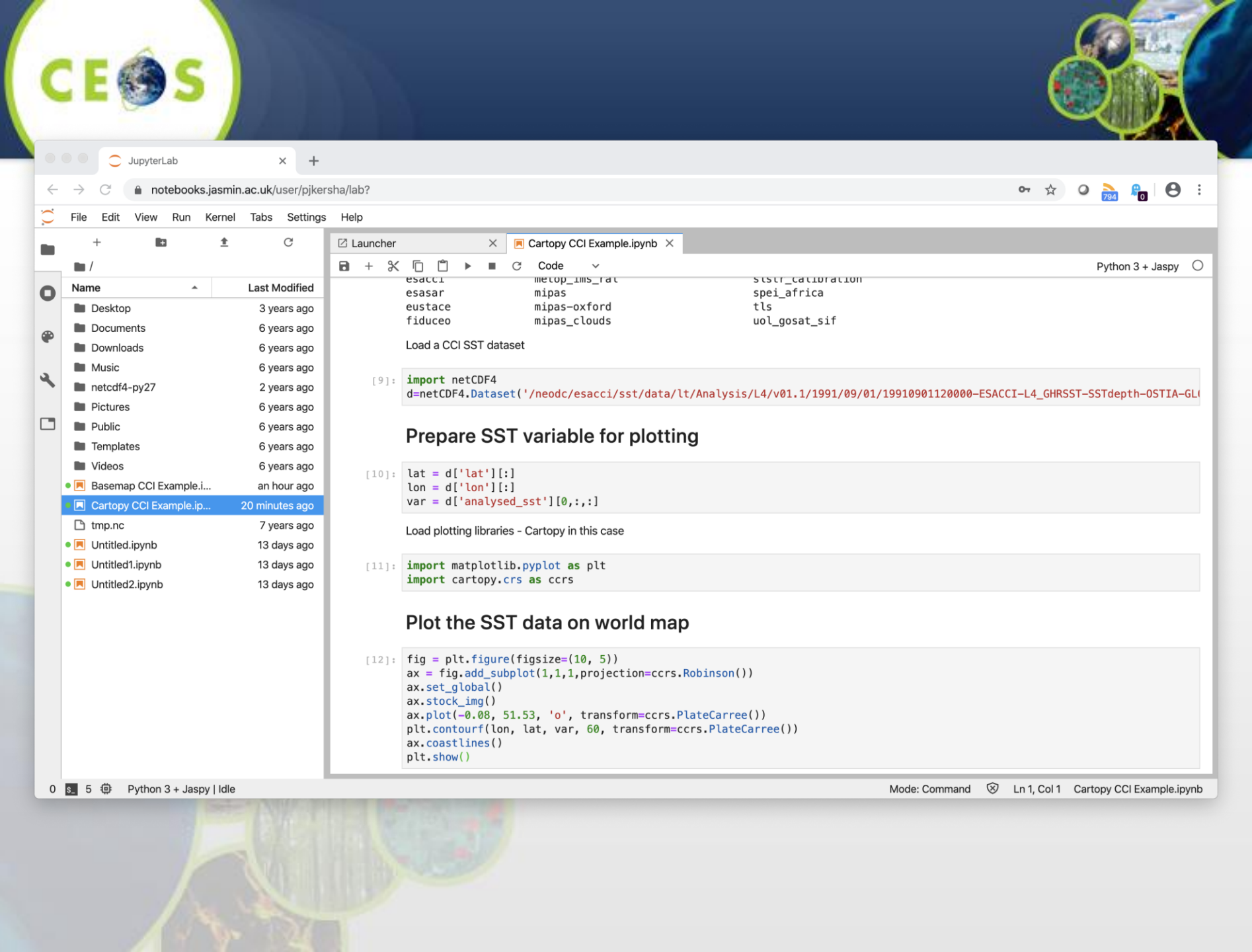
As a tool to share your data analysis with others, collaborate, teach, and promote reproducible science.

The Jupyter notebook began as an (IPython Notebook), this use of Python makes it ideal starting point for many people using EO data for the first time. In addition it currently supports around 40 programming languages, including Python, R and Julia (Ju-pyt-R)

Why might be this be useful for CEOS CapD

- Lowers barrier to use, Jupyter notebooks can be provided alongside data in an executable form (Click and Run)
- Good starting point for users wishing to improve Python
- User able to obtain meaningful result quickly
- A lot of freely available training materials to support user and help improve python skills
- Can be integrated with data cubes and CEOS Analysis Ready Data
- Jupyter Lab technologies means they be deployed as a webservice
- Jupyter Hub technologies mean they can train large classes (for more advanced countries)


JupyterLab
notebooks.jasmin.ac.uk/user/pjkersha/lab?
File Edit View Run Kernel Tabs Settings Help
/
Name
Last Modified
Desktop 3 years ago
Documents 6 years ago
Downloads 6 years ago
Music 6 years ago
netcdf4-py27 2 years ago
Pictures 6 years ago
Public 6 years ago
Templates 6 years ago
Videos 6 years ago
Basemap CCI Example.i... an hour ago
Cartopy CCI Example.ip... 19 minutes ago
tmp.nc 7 years ago
Untitled.ipynb 13 days ago
Untitled1.ipynb 13 days ago
Untitled2.ipynb 13 days ago
Launcher
Cartopy CCI Example.ipynb
Code
Python 3 + Jaspy
Direct file system access to the CEDA Archive
Read-only access to Group workspaces also possible as well as read/write to /home . Read/write access to GWSs may follow at a later date
[8]: !ls /neodc
aast_leicester aatsr_lst aatsr_multimission arc ard4ceos arsf atsrubt avhrr-3 avhrr3_metop_a avhrr3_metop_b avhrr3_metop_c avhrr_dundee avhrr_fasir avhrr_gac baci_ssv c3s_sst caliop casix comet comet-cgps cryosat_2 czcs esa_globsnow esacci esasar eustace fiduceo
gerb glas_veg_height global_gas_flaring globalbedo_broadband_brdf globalbedo_modis_spectral_priors globolakes gome-2 gome2_metop_a gome2_metop_b hycristal iasi iasi_ch4_ral iasi_metop_a iasi_metop_b iasi_so2_oxford isar landmap landsat45tm landsat5 landsat7etm landsat8 lsasaf meris metop_ims_ral mipas mipas-oxford mipas_clouds
misr modis mtci ncaveo_field ncaveo_lcm2000 nceo-carbon nceo-cryosphere nceo_biomass_maps nextmap qa4ecv requests sciamachy seawifs sentinel1a sentinel1b sentinel2a sentinel2b sentinel3a sentinel3b sentinel5p seviri_frp shac2000 sister slstr_calibration spei_africa tls uol_gosat_sif
0 5 Python 3 + Jaspy | Idle
Mode: Command Ln 1, Col 1 Cartopy CCI Example.ipynb


JupyterLab
notebooks.jasmin.ac.uk/user/pjkersha/lab?
File Edit View Run Kernel Tabs Settings Help
Launcher
Cartopy CCI Example.ipynb
Code
Python 3 + Jaspy
Name
Last Modified
Desktop 3 years ago
Documents 6 years ago
Downloads 6 years ago
Music 6 years ago
netcdf4-py27 2 years ago
Pictures 6 years ago
Public 6 years ago
Templates 6 years ago
Videos 6 years ago
Basemap CCI Example.i... an hour ago
Cartopy CCI Example.ip... 20 minutes ago
tmp.nc 7 years ago
Untitled.ipynb 13 days ago
Untitled1.ipynb 13 days ago
Untitled2.ipynb 13 days ago
esacci metop_ims_rat slstr_calibration
esasar mipas spei_africa
eustace mipas-oxford tls
fiduceo mipas_clouds uol_gosat_sif
Load a CCI SST dataset
[9]:
import netCDF4
d=netCDF4.Dataset('/neodc/esacci/sst/data/lt/Analysis/L4/v01.1/1991/09/01/19910901120000-ESACCI-L4_GHRSST-SSTdepth-OSTIA-GL
Prepare SST variable for plotting
[10]:
lat = d['lat'][:]
lon = d['lon'][:]
var = d['analysed_sst'][0,:,:]
Load plotting libraries - Cartopy in this case
[11]:
import matplotlib.pyplot as plt
import cartopy.crs as ccrs
Plot the SST data on world map
[12]:
fig = plt.figure(figsize=(10, 5))
ax = fig.add_subplot(1,1,1,projection=ccrs.Robinson())
ax.set_global()
ax.stock_img()
ax.plot(-0.08, 51.53, 'o', transform=ccrs.PlateCarree())
plt.contourf(lon, lat, var, 60, transform=ccrs.PlateCarree())
ax.coastlines()
plt.show()
0 5 Python 3 + Jaspy | Idle
Mode: Command
Ln 1, Col 1
Cartopy CCI Example.ipynb

JupyterLab

notebooks.jasmin.ac.uk/user/pjkersha/lab?

File Edit View Run Kernel Tabs Settings Help

/

Name	Last Modified
Desktop	3 years ago
Documents	6 years ago
Downloads	6 years ago
Music	6 years ago
netcdf4-py27	2 years ago
Pictures	6 years ago
Public	6 years ago
Templates	6 years ago
Videos	6 years ago
Basemap CCI Example.i...	an hour ago
Cartopy CCI Example.ip...	20 minutes ago
tmp.nc	7 years ago
Untitled.ipynb	13 days ago
Untitled1.ipynb	13 days ago
Untitled2.ipynb	13 days ago

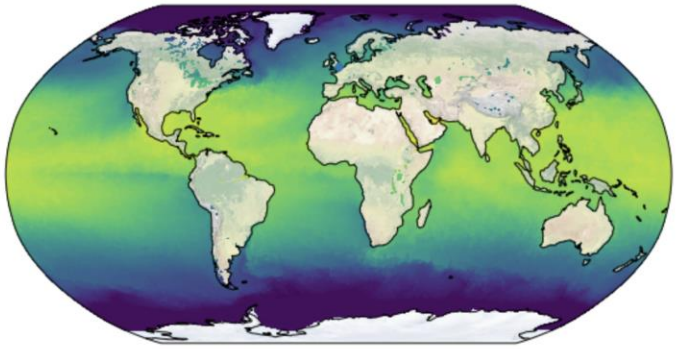
Launcher | Cartopy CCI Example.ipynb

Code | Python 3 + Jaspy

[11]:
```
import matplotlib.pyplot as plt
import cartopy.crs as ccrs
```

Plot the SST data on world map

[12]:
```
fig = plt.figure(figsize=(10, 5))
ax = fig.add_subplot(1,1,1,projection=ccrs.Robinson())
ax.set_global()
ax.stock_img()
ax.plot(-0.08, 51.53, 'o', transform=ccrs.PlateCarree())
plt.contourf(lon, lat, var, 60, transform=ccrs.PlateCarree())
ax.coastlines()
plt.show()
```

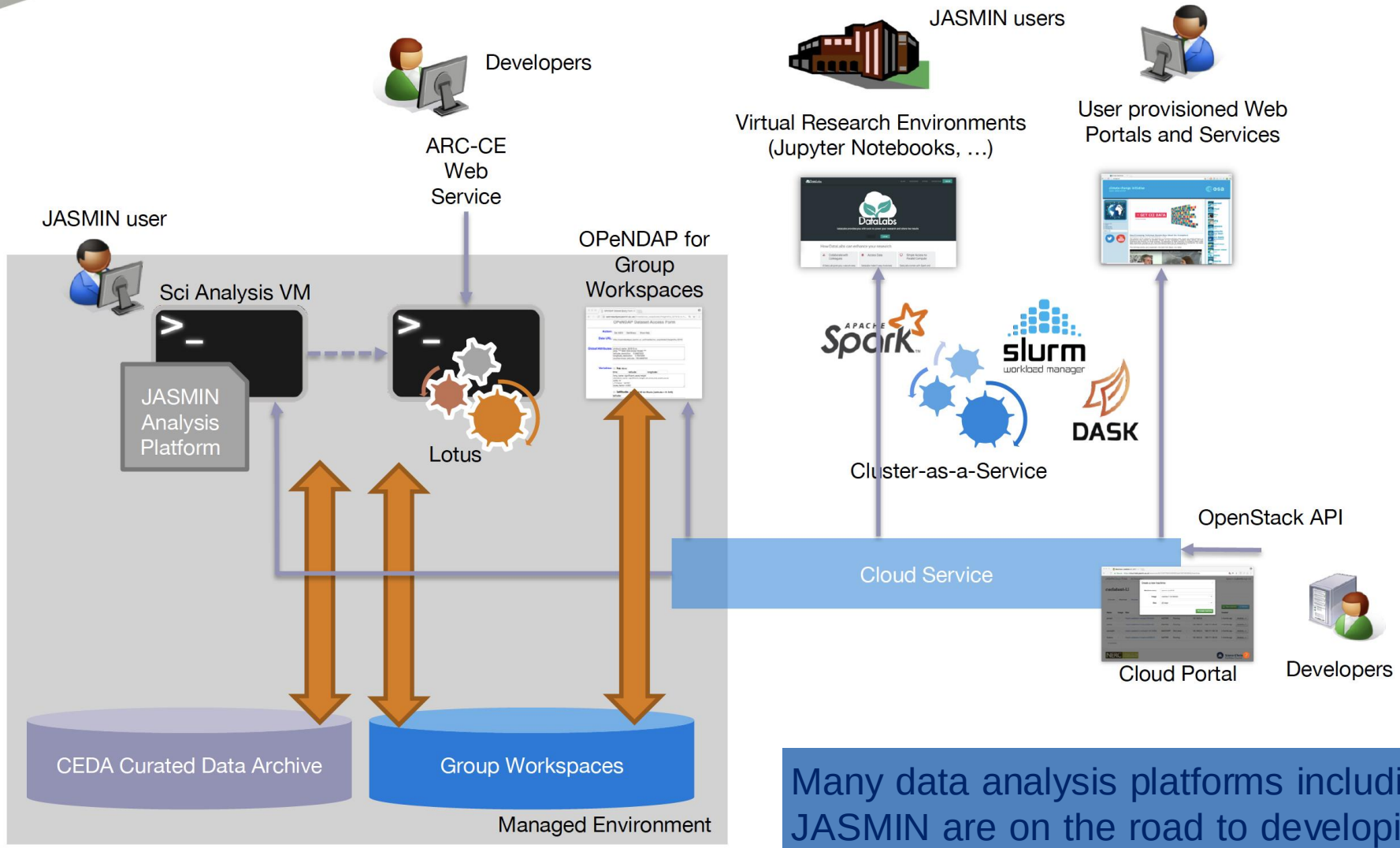


[ ]:

0 5 Python 3 + Jaspy | Idle

Mode: Command Ln 1, Col 1 Cartopy CCI Example.ipynb

JASMIN Service Evolution



Many data analysis platforms including JASMIN are on the road to developing Jupyter Notebook services which can read from central archives

Earth Analytics Interoperability Lab

OPEN DATA CUBE

Jupyter
WMS
CEOS ODC
Monitoring

Flux CD

HELM Tiller

Replication
Failover
Scheduling

Compute
Network
Storage

Where can they be deployed and what are the issues ? Why the need for a collaboration with WGISS?

- **Standalone computer notebooks can be supplied alongside the relevant data**
  - Issue around notebooks breaking due to dependencies on python version and libraries etc. Quality issues with code and documentation there are lot of notebooks and a lot rubbish amongst them). Need for CEOS standard and stamp of approval. Also an issue with large datasets and updates to datasets.
- **In classroom/institutional setting using Jupyter Hub**
  - Issue around many of the poorest countries not having this capability and has same dependency/quality issues as above.
- **As a centralised web services using a Data Cube (SEO, WGISS?)**
  - Issues with development overheads, compatibility with atmospheric pollution data and data cubes, challenges with diverse data from different domains, non standard ARD. Access to hosted processing and security access.
- **As different services on data analysis platforms**
  - Issue with development overheads, capacity and support. Dealing with diversity has its own overhead. May require catalogue of services and dealing with many administrations. Different agencies are at different point technically.

Example of ODA targeted from the NCEO ODA Programme

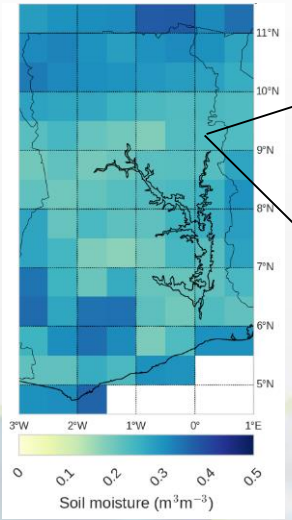
GCRF
Global Challenges Research Fund

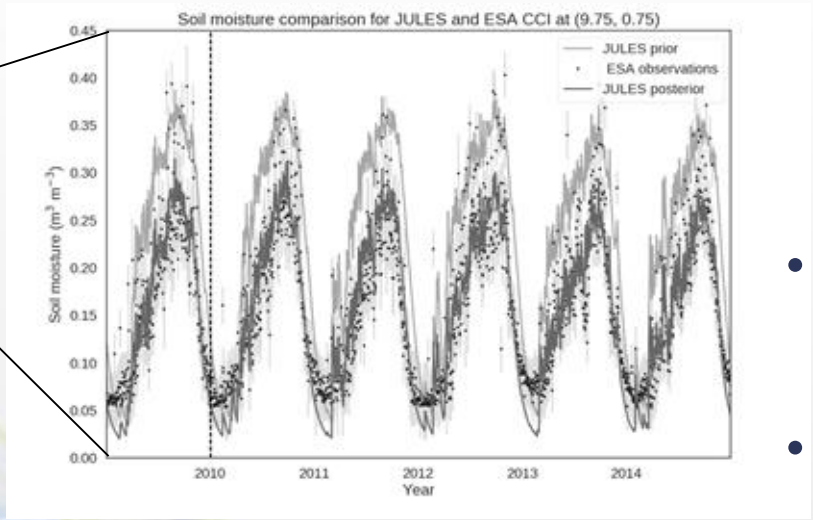
National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

NERC SCIENCE OF THE ENVIRONMENT

Heiko Balzter | hb91@leicester.ac.uk| University of Leicester

Soil moisture estimates for Ghana

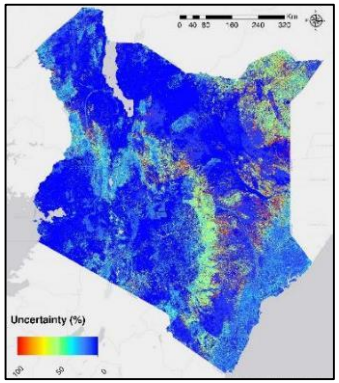


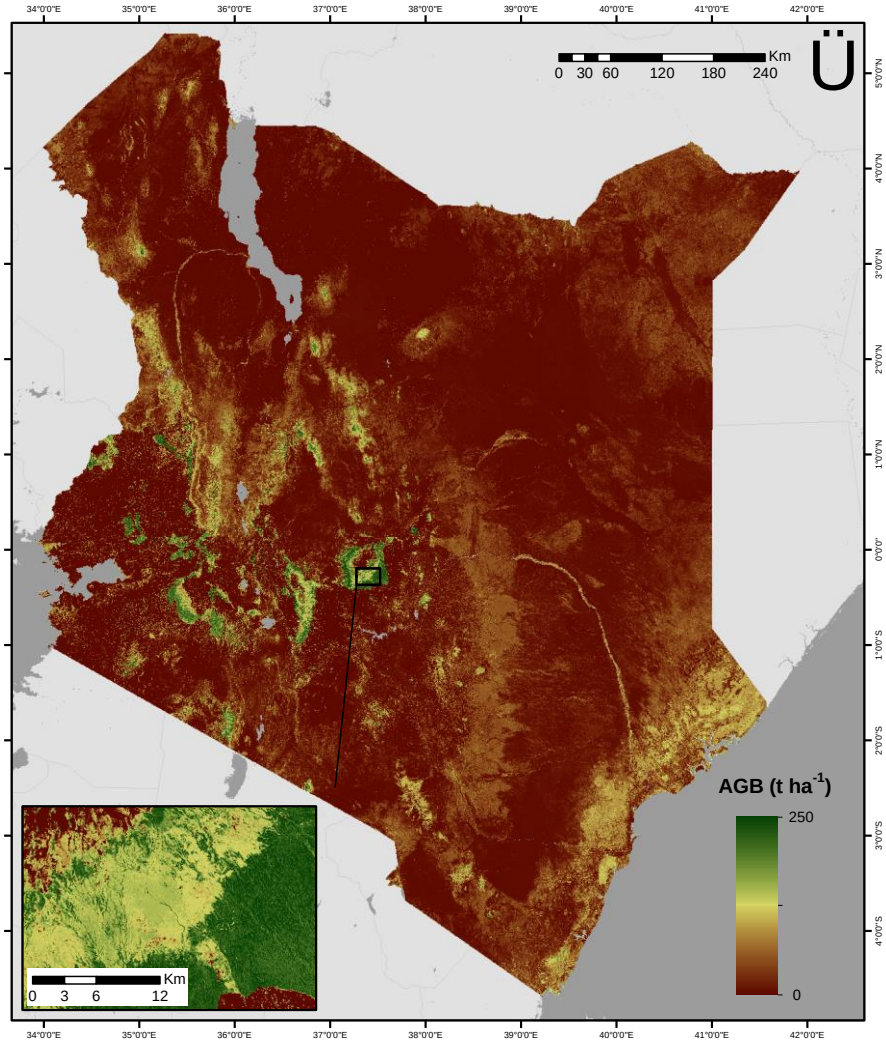


- Estimates of soil moisture over Ghana, found by combining model predictions with satellite observations.
- Prototype for larger scale Africa wide soil moisture dataset.
- Dataset description:

Pinnington, E., Quaife, T., and Black, E.: Impact of remotely sensed soil moisture and precipitation on soil moisture prediction in a data assimilation system with the JULES land surface model, Hydrol. Earth Syst. Sci., 22, 2575-2588, https://doi.org/10.5194/hess-22-2575-2018, 2018.

- The NCEO ODA AGB map is the most detailed and accurate biomass map for Kenya that has been produced to date. The accuracy assessment found coefficients of determination (R^2) between 0.64 and 0.90, and root mean square errors (RMSE) ranging from 35.9 t ha^{-1} to 65.2 t ha^{-1}
- The highest AGB densities are found in the dense tropical forests with values up to 530 t ha^{-1} (average 92 t ha^{-1}). Despite the low levels of biomass density (average 13 t ha^{-1}), the large area of wooded grassland and savannah in Kenya potentially stores up 58% of the total AGB in the country

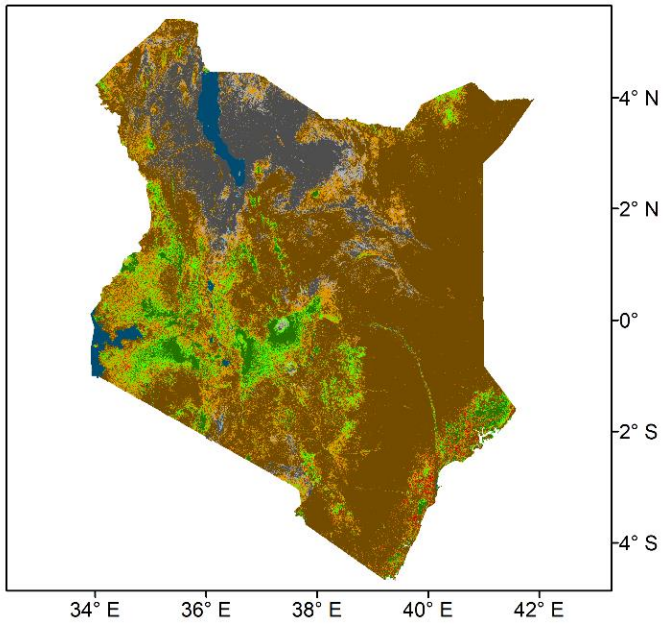




UNIVERSITY OF LEICESTER

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

Maps of stable and change classes in Kenya



Classification likelihood

- Likelihood scale defined according to IPCC (2010):
  - "unlikely": < 0.33
  - "about as likely as not" : 0.33-0.66
  - "likely": > 0.66

Country statistics 2014-2017 ("likely" classes)

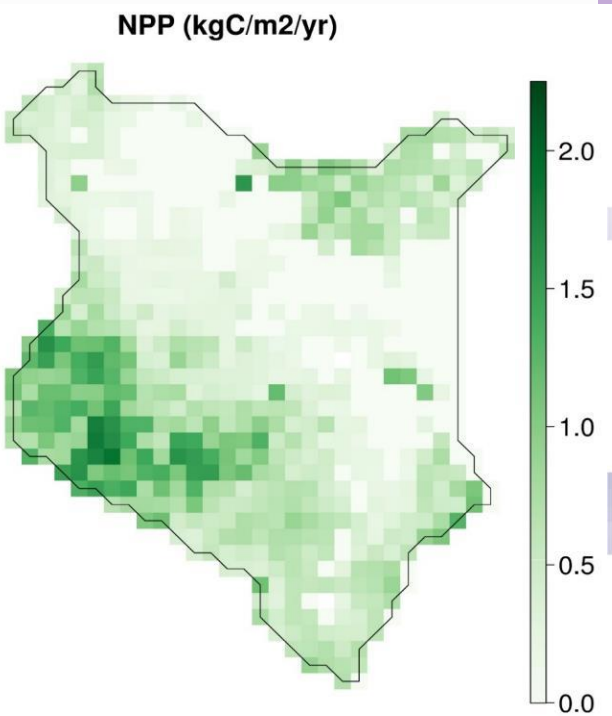
- Stable Forest (SF): 4.1%
- Stable Other Vegetation (SOV): 53.0%
- Stable Other Land (SOL): 17.0%
- Stable Water (SW): 2.0%
- Forest Loss (Floss): 0.1%

76.2% of the country classified as "likely", and 23.8% as "about as likely as not"

Kenyan forest C-cycle analysis

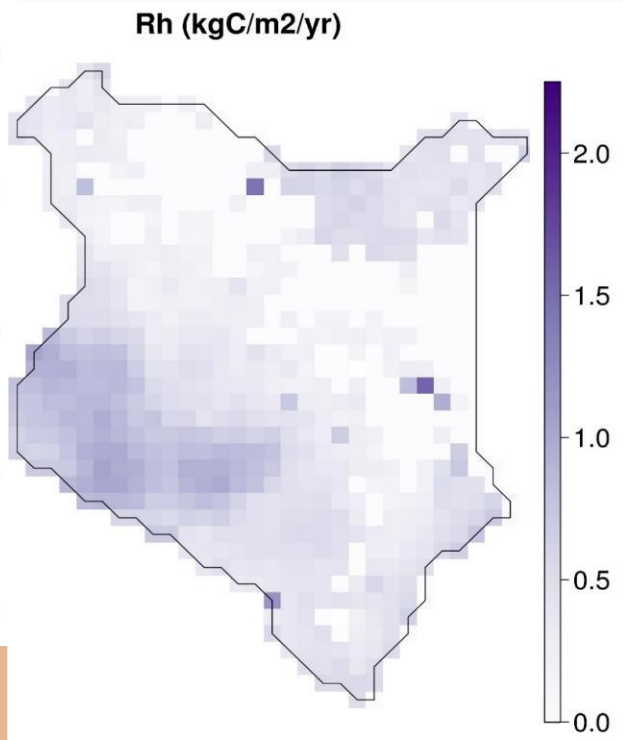
The NCEO ODA used the CARDAMOM model-data fusion framework combining the latest NCEO created - Earth Observations based maps of forest land cover and above ground biomass, and time varying meteorological information to create a state-of-the-art carbon cycle model for forest areas over Kenya.

Our analysis estimates that for the period 2015-2017 Kenya's forests were a net sink of 11.5 (3.8 / 20.2) TgC/yr or 2.1 (0.7 / 3.7) MgC/ha/yr. The areas of greatest biomass accumulation (indicated by NPP) are found is the mangrove forests along the south east coast and central western areas towards Lake Victoria

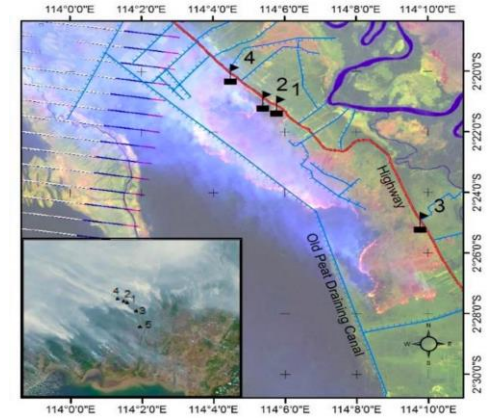


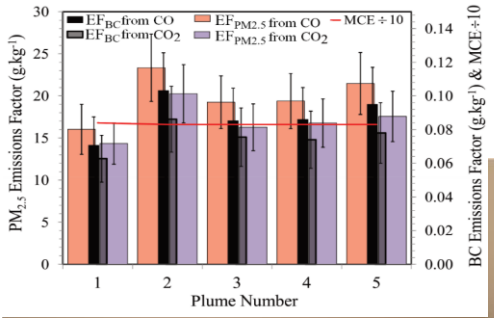
NPP = Net Primary Productivity, i.e. photosynthesis less autotrophic (plant originating) respiration

Rh = Heterotrophic (non-plant originating) respiration, result of soil and litter decomposition



Emission Factors for SE Asian Peatlands



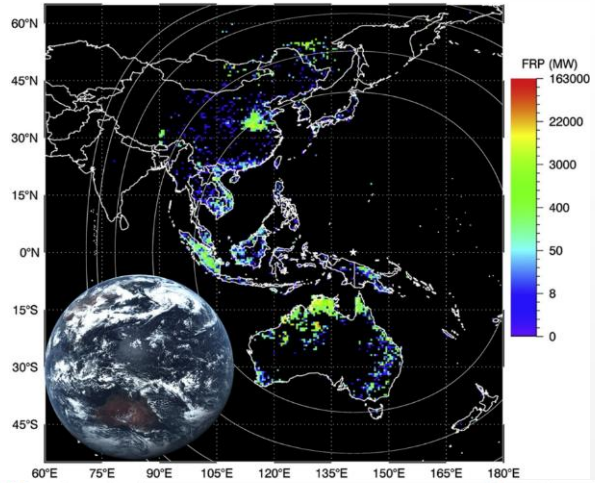


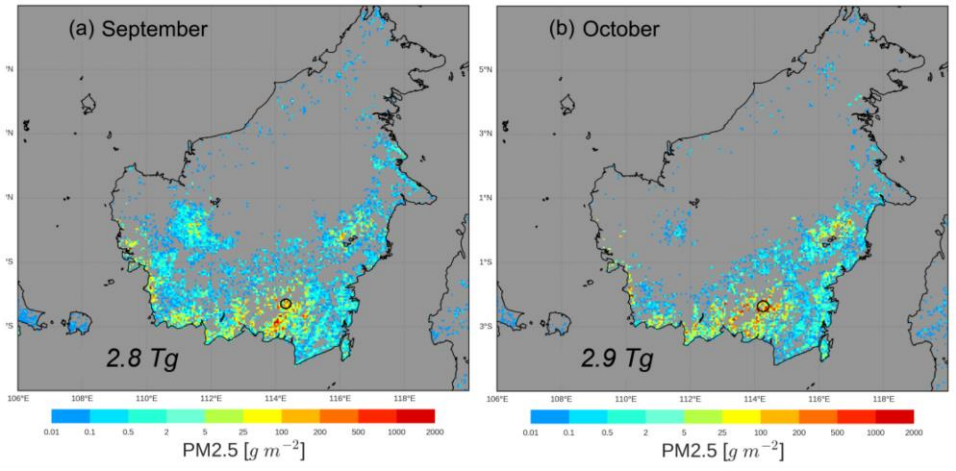
- **30 peat fire plumes sampled** in-situ with FTIR Spectrometer in 2015 across 5 locations on Kalimantan, Indonesia.
- **New carbonaceous trace gas emission factors** (CO_2, CH_4, CO) generated for tropical peat fires.
- **New particulate emission factors** ($PM_{2.5}$ and black carbon) generated for tropical peat fires.
- **Update emission and fuel consumption totals** generated for 2015 SE Asian peatland fires.

KING'S College LONDON
National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL
CIFOR
KNMI
INSTITUT PERTANIAN BOGOR
BMKG
RUB
ATB Leibniz-Institut für Agrartechnik und Biokonomie

Geostationary FRP Emission Source Mapping

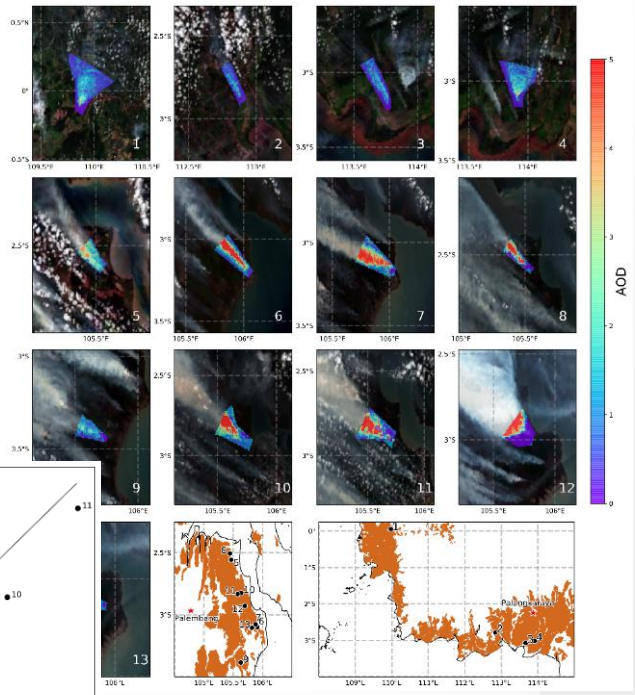
- **Very high temporal resolution** (10 minute) fire radiative power data from Himawari geostationary sensor produced.
- **Emission source maps** generated for Kalimantan at 2km resolution. Far higher resolution than in other emission inventories.

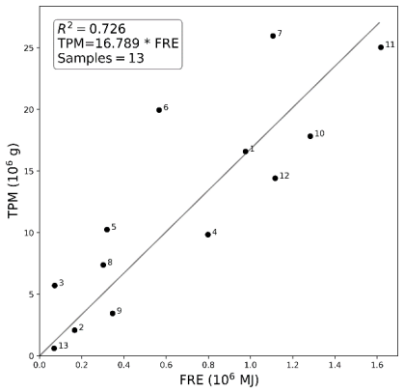




New ‘Top-Down’ Emission Coefficient

- **'Top-down’ particulate matter** emission coefficient generated for surficial peat fires in SE Asia.
- **Appropriate for flaming peat fires** this new coefficient has demonstrated the importance of fire phase when estimating emissions.





The Way Forward ?

- Survey of Agencies
  - CapD targeted datasets and associated notebooks
  - Key Datasets for CapD and associated notebooks
- Input of relevant in country capabilities and needs
- Technical discussion at WGISS 49
- Notebook standards and CEOS endorsement

Discuss !