

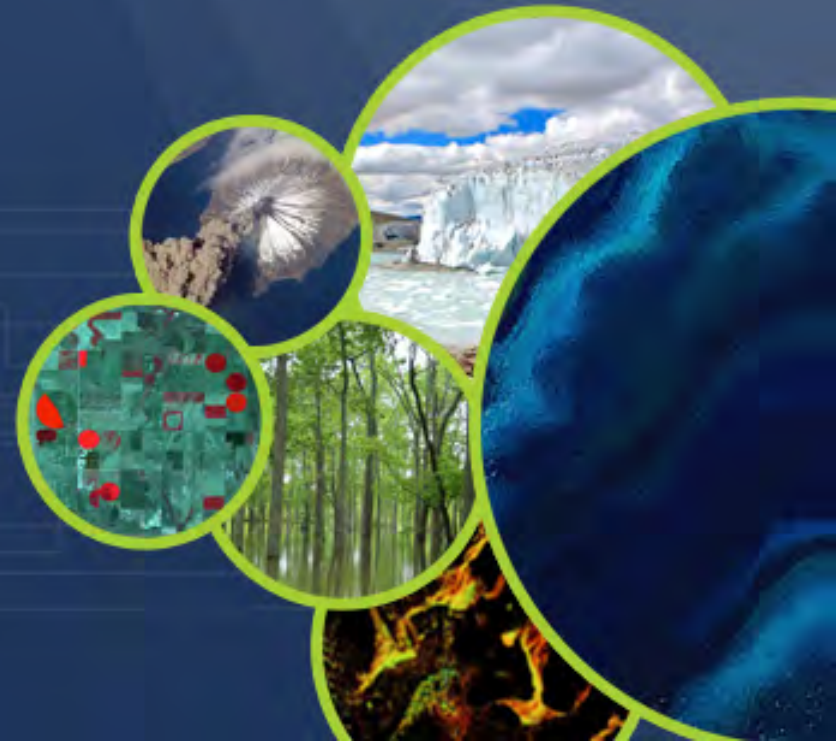


Open Data Cube (ODC) Provisions for the GEO-AWS Cloud Credits Programme

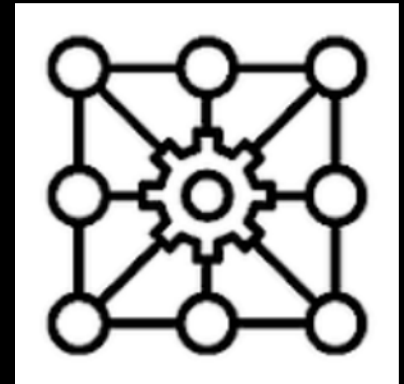
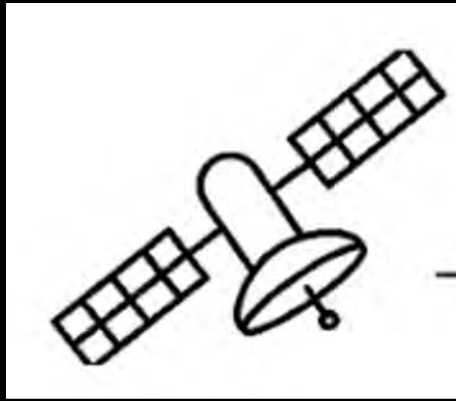
2nd GEO-AWS Earth Observation Cloud Credits Programme Webinar
March 19, 2019

Dr. Brian Killough

CEOS Systems Engineering Office
NASA Langley Research Center



What is the Open Data Cube?



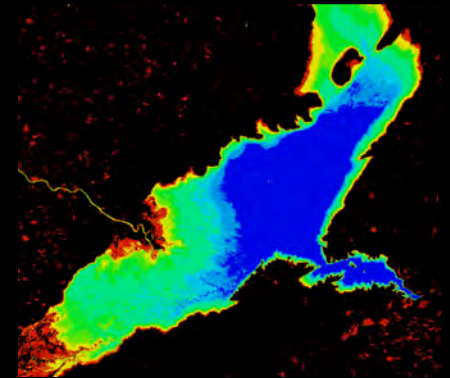
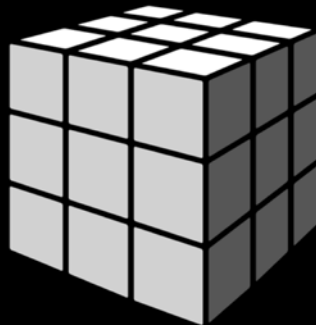
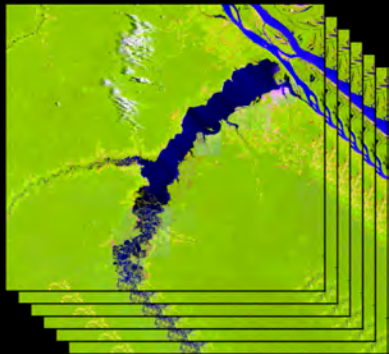
**Satellite
Analysis-Ready Data**

**Open Data Cube
Infrastructure**

Applications and Tools

*Data Cube Core Code,
API and Database*

*Application Library
User Interface*

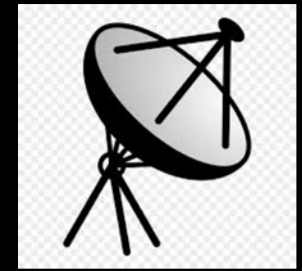


What are the benefits of the Open Data Cube?



- Uses analysis ready data to reduce data preparation time and complexity
- Based on free and open source software
- Enables data interoperability and efficient time series analyses
- Easily deployed in the Amazon Cloud
- Existing customizable web-based user interface and a library of application algorithms
- Proven concept ... Working in Australia, Colombia, Switzerland, Vietnam and 5 countries in Africa!

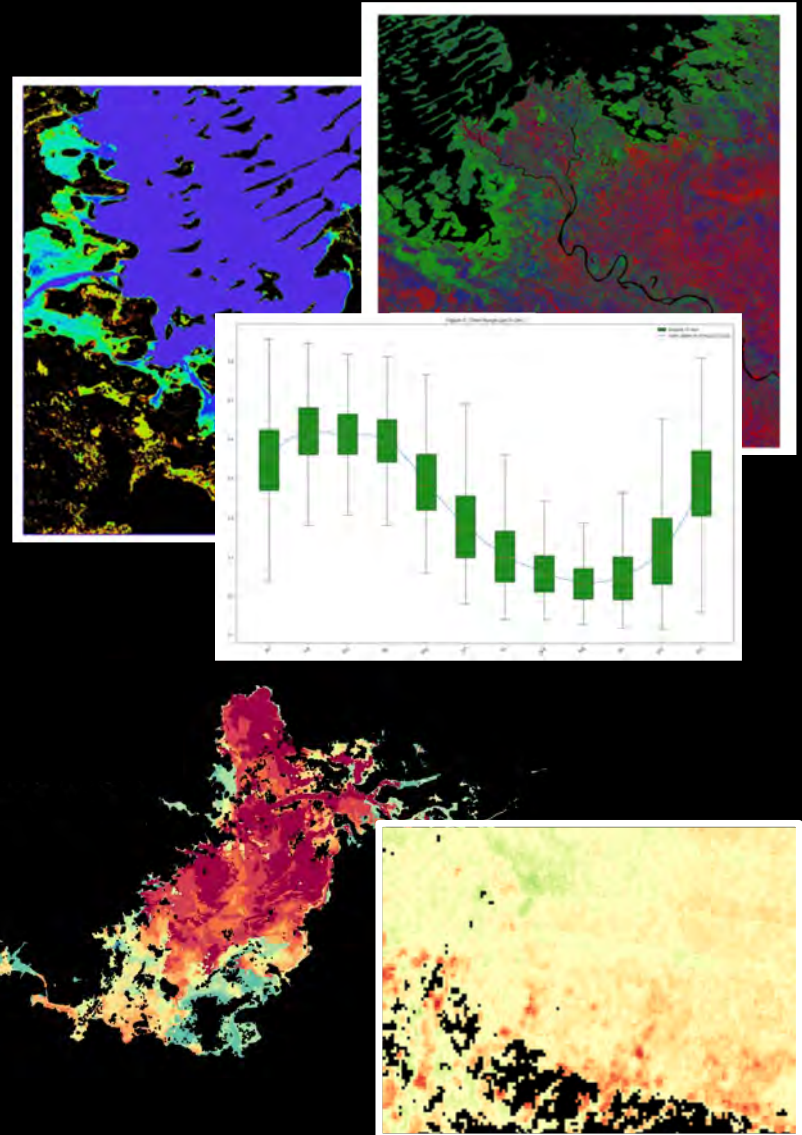
How do we “feed” the Open Data Cube with satellite data?



- We promote the use of pre-processed analysis-ready data (ARD) from Landsat, Sentinel-1 (S1), and Sentinel-2 (S2).
- **Landsat** data must be ordered and downloaded from USGS (e.g. surface reflectance) and then uploaded, ingested and indexed in the cloud. By the end of 2019, USGS will have all of their data on AWS.
- **Sentinel** data has been more challenging for our users, as downloads outside of Europe are less reliable. We are working with e-geos (Italian company) to develop a Sentinel data workflow that will allow users to create S1 or S2 ARD and move those products directly to the Amazon cloud without local downloads.
- “Automated” flow of new satellite data into cubes has been demonstrated by Australia and Switzerland.

How can we use the Open Data Cube?

- **Cloud-free Mosaics:** Recent Pixel, Median, Geomedian, Max-NDVI
- **Spectral Indices:** NDVI, NDBI, NDSI, NDWI, SAVI, EVI, Fractional Cover
- **Land Classification:** K-Means, Random Forest, FAO 8-class decision tree
- **Water:** Landsat WOFS (Australia), Sentinel-1 WASARD (NASA), Landsat Water Quality (Australia) - Total Suspended Matter
- **Land Change:** Spectral Threshold Anomaly, Coastal Change, PyCCD (USGS)



Web-based User Interface



Welcome to the Open Data Cube

CEOS is using the power of the Open Data Cube to help address the needs of satellite data users, giving them a better picture of their land resources and land change.

- Ease of use and access to satellite-based data
- Multiple dataset interoperability and spatial consistency
- Use of "Analysis Ready" Data Products
- A Shift in Paradigm from Scenes to Pixels

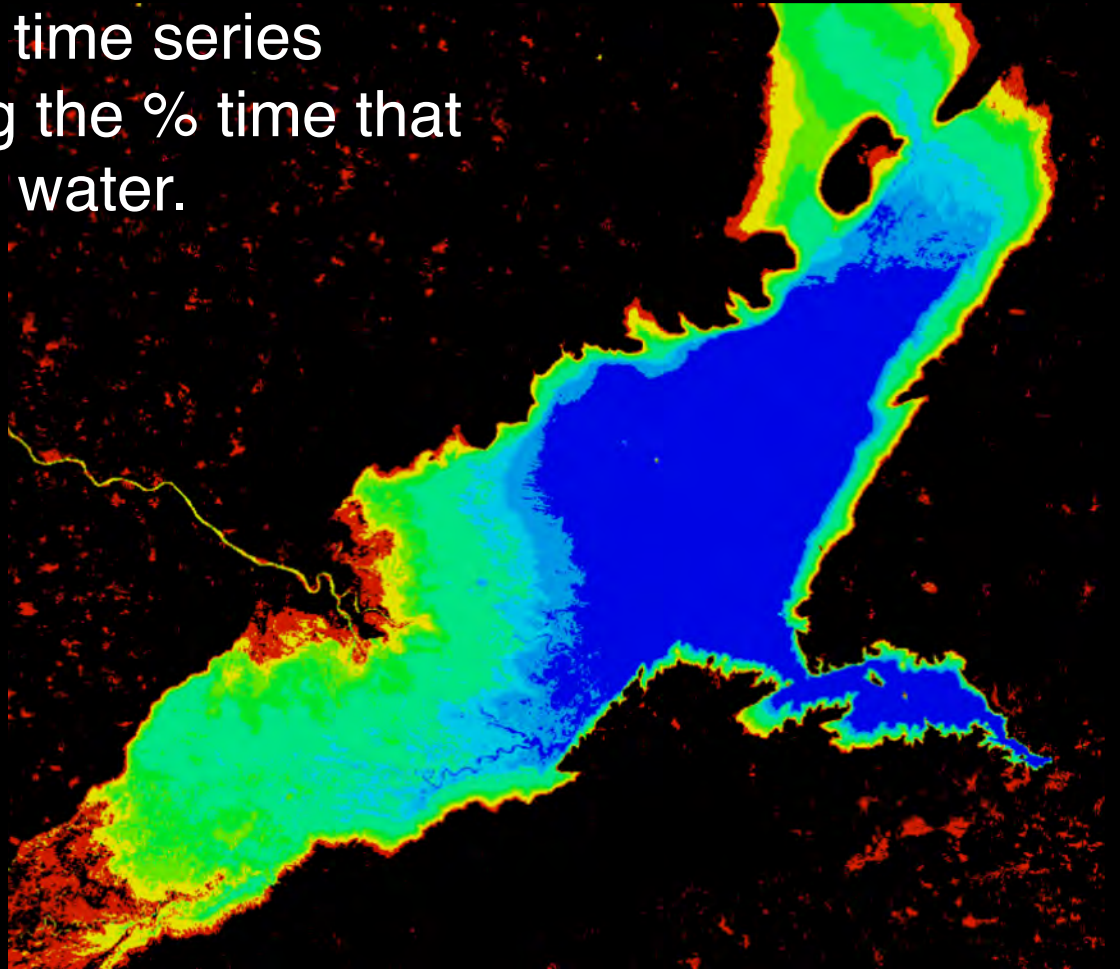
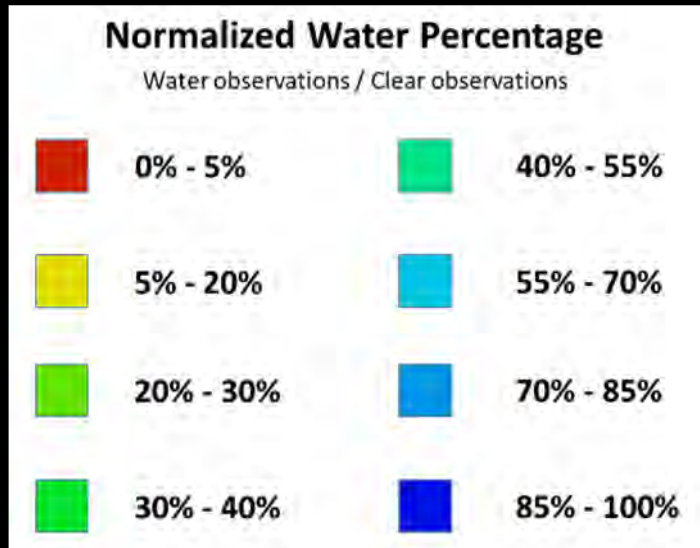
<http://tinyurl.com/datacubeui>

- 16 sample data cubes + 10 common applications
- Free and Open, easy-to-use, menu-driven, GIS tool
- Applications focused on Landsat data
- GeoTIFF outputs, animations, pixel drill

Water Extent

Mtera Reservoir in Tanzania

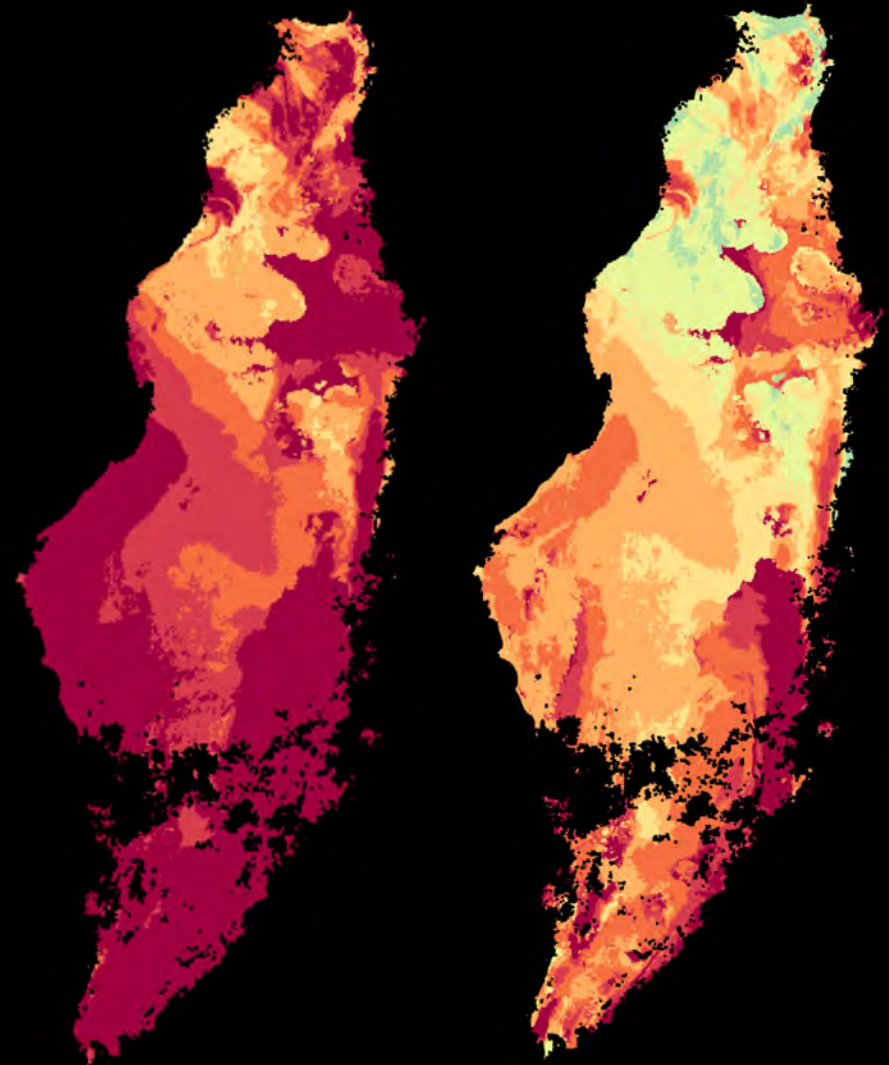
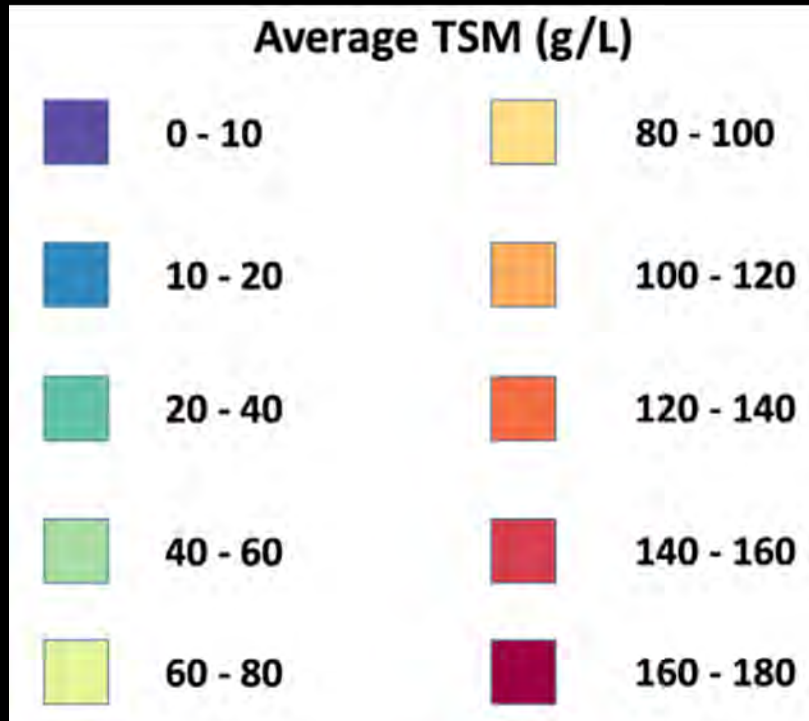
This product is a 17-year time series (1999 thru 2016) showing the % time that every pixel has observed water.



Water Quality

Lake Manyara in Tanzania
Feb 2013 to Dec 2016

Total Suspended Matter (TSM)

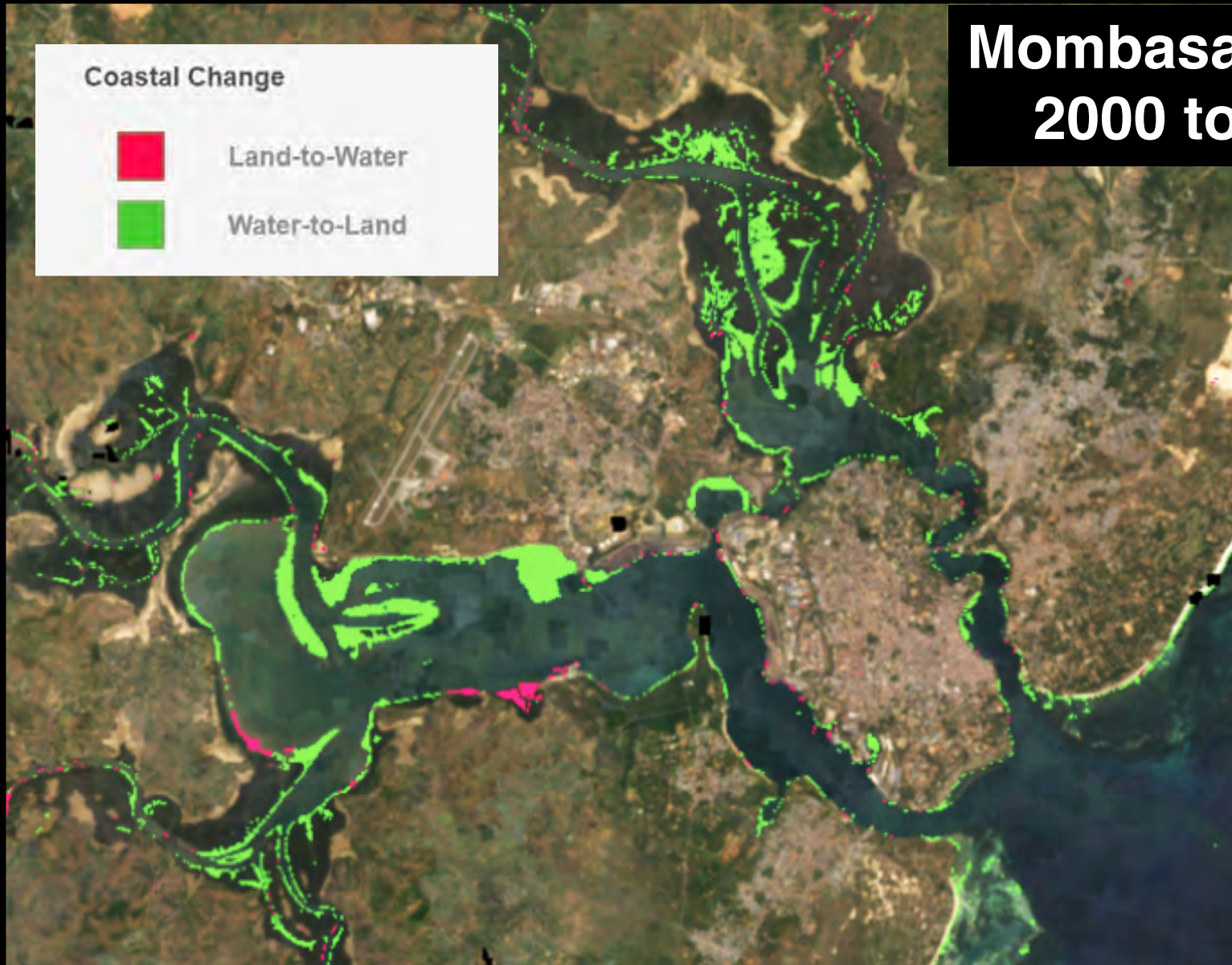


Maximum

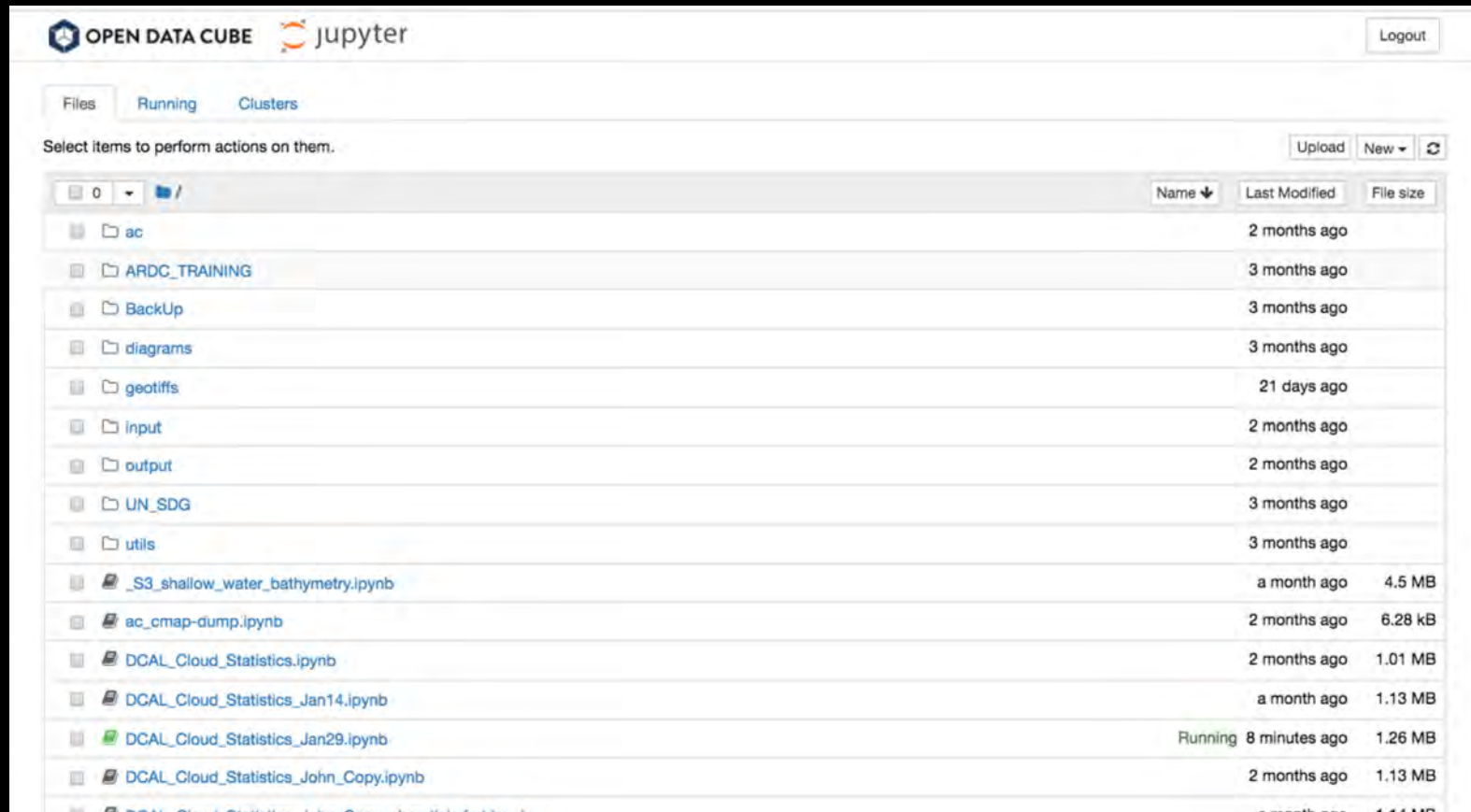
Variability

Coastal Change

**Mombasa, Kenya
2000 to 2015**



Jupyter Python Notebook Hub

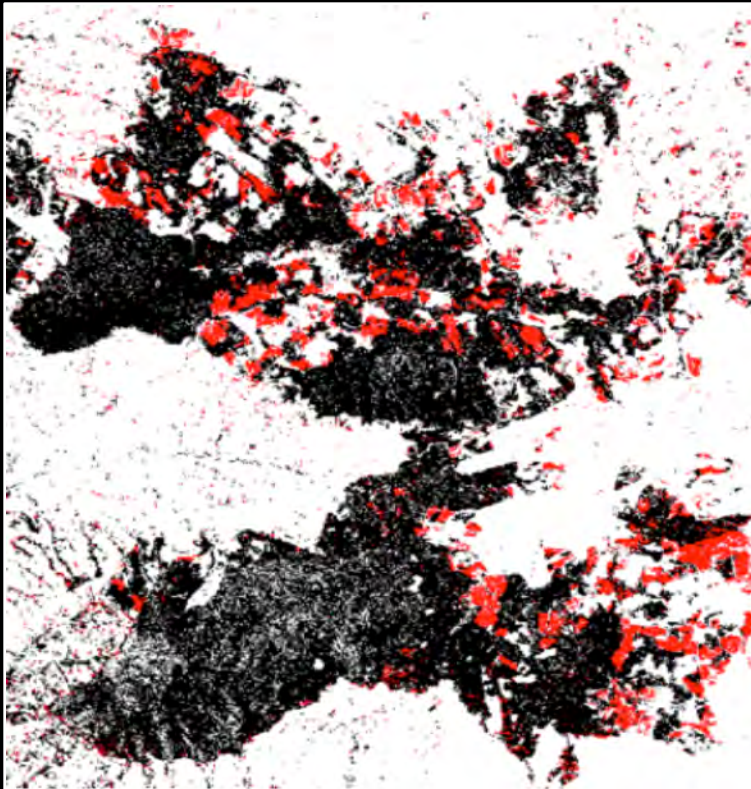


The screenshot displays the Jupyter Python Notebook Hub interface. At the top, there are logos for 'OPEN DATA CUBE' and 'jupyter', along with a 'Logout' button. Below the logos, there are tabs for 'Files', 'Running', and 'Clusters'. A message says 'Select items to perform actions on them.' with 'Upload', 'New', and a refresh icon. The main area shows a file browser view with columns for 'Name', 'Last Modified', and 'File size'. The files listed include folders like 'ac', 'ARDC_TRAINING', 'BackUp', 'diagrams', 'geotiffs', 'input', 'output', 'UN_SDG', and 'utils', and notebooks like '_S3_shallow_water_bathymetry.ipynb', 'ac_cmap-dump.ipynb', 'DCAL_Cloud_Statistics.ipynb', 'DCAL_Cloud_Statistics_Jan14.ipynb', 'DCAL_Cloud_Statistics_Jan29.ipynb', and 'DCAL_Cloud_Statistics_John_Copy.ipynb'. The 'DCAL_Cloud_Statistics_Jan29.ipynb' notebook is shown as 'Running'.

Name	Last Modified	File size
ac	2 months ago	
ARDC_TRAINING	3 months ago	
BackUp	3 months ago	
diagrams	3 months ago	
geotiffs	21 days ago	
input	2 months ago	
output	2 months ago	
UN_SDG	3 months ago	
utils	3 months ago	
_S3_shallow_water_bathymetry.ipynb	a month ago	4.5 MB
ac_cmap-dump.ipynb	2 months ago	6.28 kB
DCAL_Cloud_Statistics.ipynb	2 months ago	1.01 MB
DCAL_Cloud_Statistics_Jan14.ipynb	a month ago	1.13 MB
DCAL_Cloud_Statistics_Jan29.ipynb	Running 8 minutes ago	1.26 MB
DCAL_Cloud_Statistics_John_Copy.ipynb	2 months ago	1.13 MB
DCAL_Cloud_Statistics_John_Copy_what_link_for_Link	a month ago	1.14 MB

There are 6 core notebooks (**cloud statistics, custom mosaics, water extent, spectral products, land change, vegetation phenology**) with You-Tube videos. Many more algorithms are under development and testing.

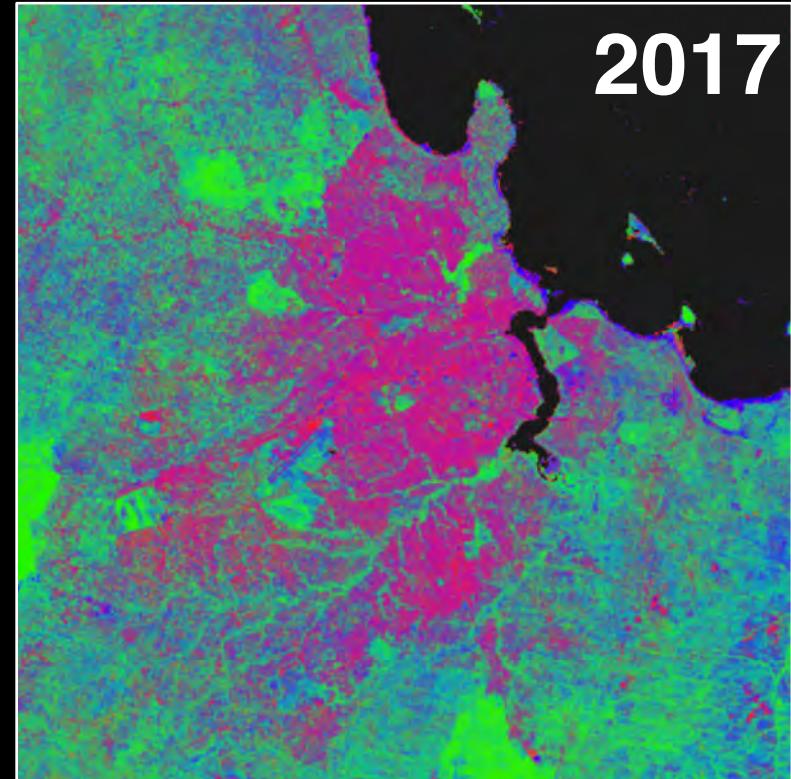
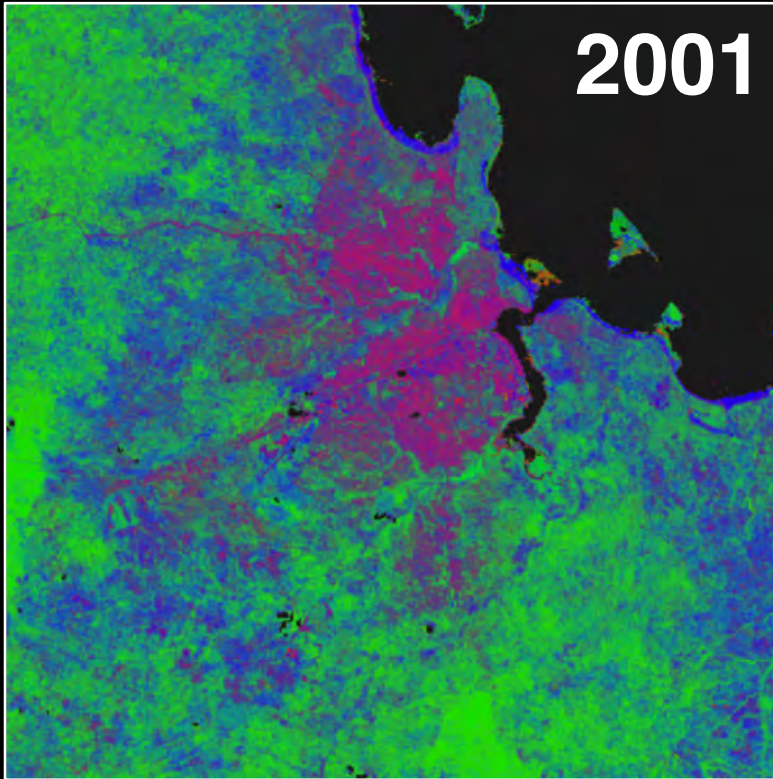
Land Change: Deforestation



Mau Forest in western Kenya

Data Cube results (left) closely compare with the Global Forest Watch results (right). Using an EVI threshold approach, there is a **loss of 5.7%** of the land to devegetation (likely deforestation) from 2002 to 2017.

Spectral Products: Urbanization



Urbanization in Dar es Salaam, Tanzania – 2001 to 2017 (16 years)

Urban Area Growth = 7% per year (Fractional Cover BS, 0.4 threshold)

Population Growth = 5 % per year (reference: web sources)

Where can I find the Open Data Cube?

- ODC Website: <https://opendatacube.org>
- ODC Installation: [~ / installation](#)
- Data Cube Application Library: [~ / dcal](#)
- ODC Documentation:
<https://opendatacube.readthedocs.io/en/latest/>
- ODC GitHub: <https://github.com/opendatacube>
- User Interface:
<http://tinyurl.com/datacubeui>

Thanks!