

The Open Data Cube (ODC)

A tool to increase the value and impact of global Earth observation satellite data

We work with
Innovate UK

CATAPULT
Satellite Applications



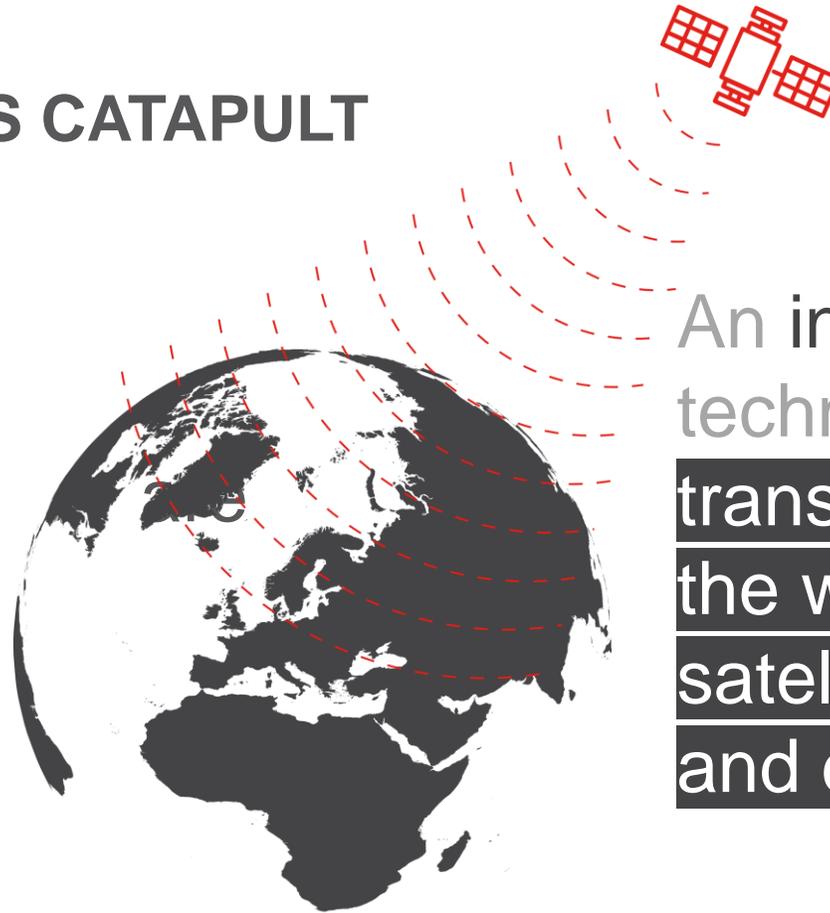
SATELLITE APPLICATIONS CATAPULT

Our Mission

// To innovate for a better world, empowered by satellites.

SATELLITE APPLICATIONS CATAPULT

Who are we?



An innovation and technology company transforming the way the world uses satellite technology and data.



WE HELP ORGANISATIONS GROW THEIR BUSINESS

We help organisations to use satellite applications to grow their business in the UK and internationally.



WE ARE INDEPENDENT

We bring together industry, researchers, end-users and government to explore and develop new ideas.

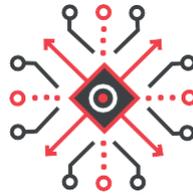


WE ARE GOVERNMENT BACKED

We are partly-funded by the Government and work closely with Innovate UK, UK Space Agency, UK Science & Innovation Network, and other public bodies.

SATELLITE APPLICATIONS CATAPULT

Our focus



UBIQUITOUS CONNECTIVITY

Super-fast, always on connectivity for the digital age



GEOSPATIAL INTELLIGENCE

An ever growing demand for information that is accurate, personalised, localised and available globally in real-time

APPLICATIONS

INTEGRATION FOR A CONNECTED AND INFORMATION CENTRIC WORLD

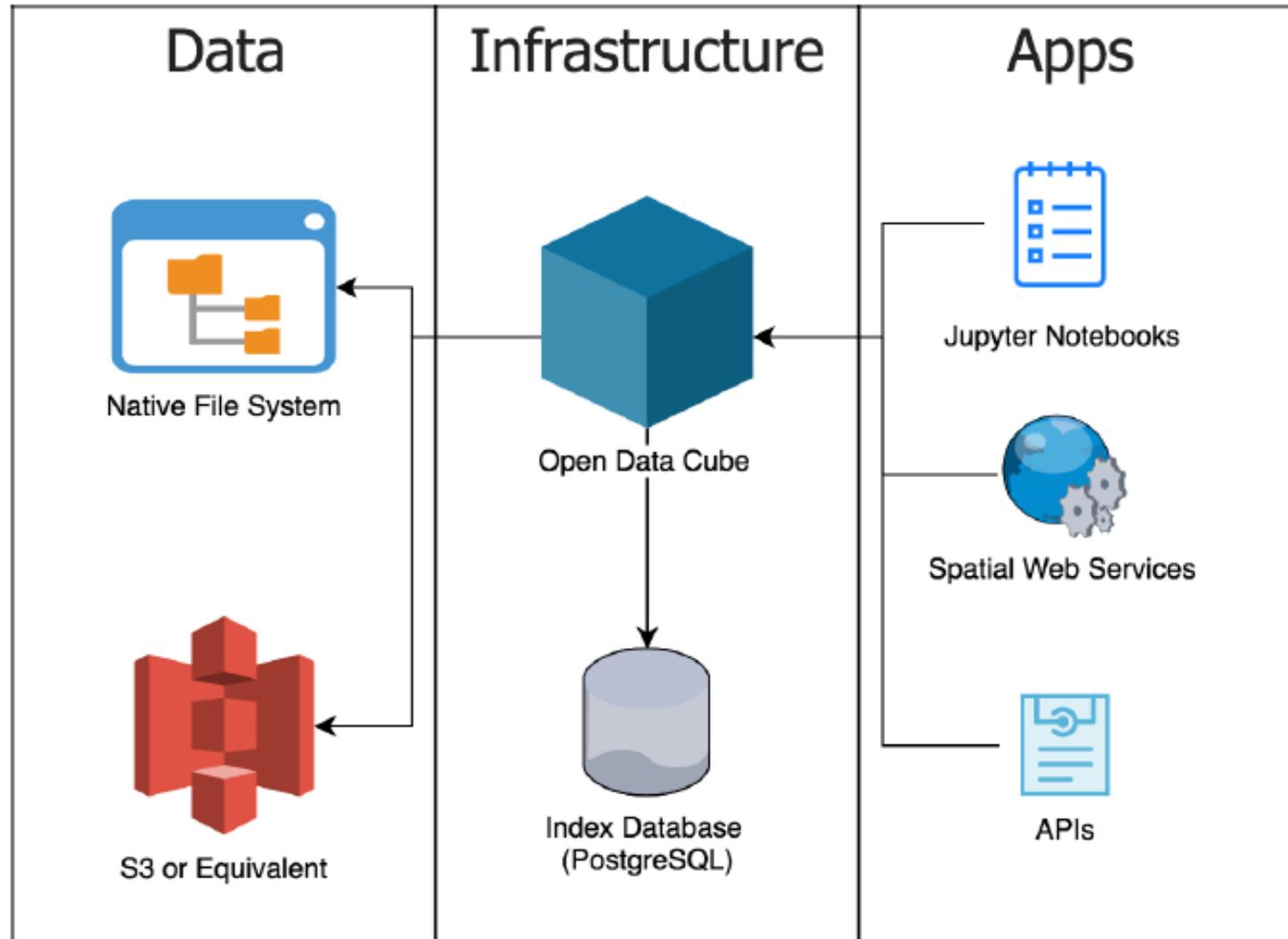


WHAT IS THE OPEN DATA CUBE?

**A Python Library
that facilitates
working with raster
data**



THE TECHNICAL COMPONENTS



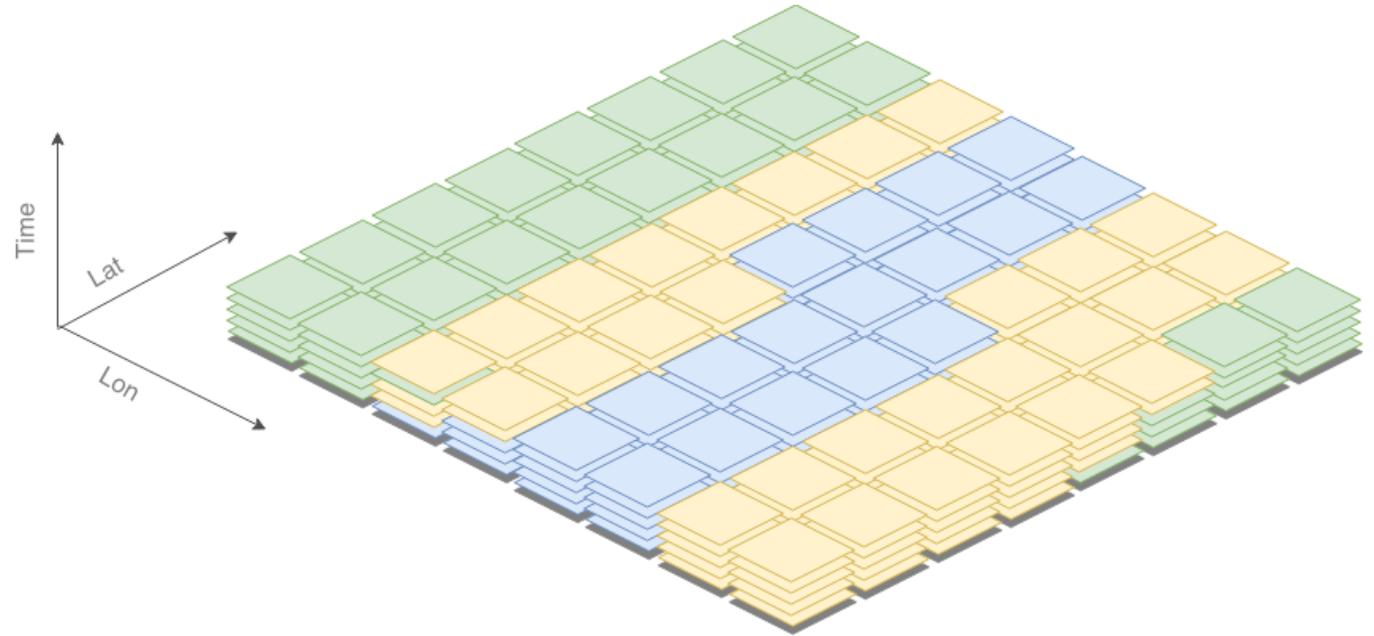
IN A NUTSHELL

- Data
- An Index
- Software



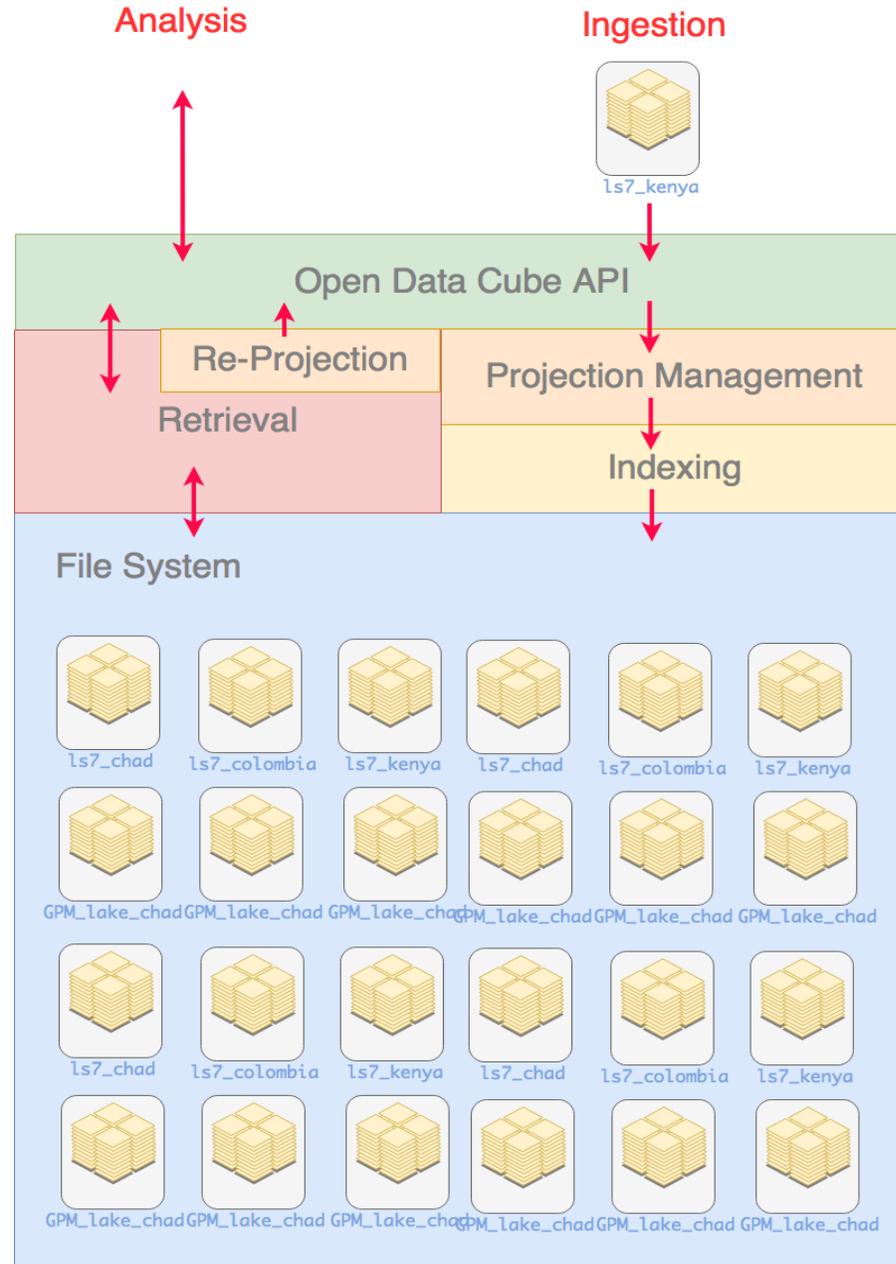
DENSE DATA

- **A versatile representation**
- **All points map to a value**
- **Easy to select/index**
- **Efficient representation in memory**
- **Called “Raster Data”**
- **Data Cube analysis case is spatially and temporally dense**



OPEN DATA CUBE: STORAGE

- When someone asks you about DataCube, you might say:
 - “It’s an efficiently indexed file system for satellite imagery”
- Open Data Cube focuses on the storage of dense data
- Open Data Cube’s storage system can handle and automate:
 - Storage
 - Projections
 - Retrieval



WHY PHYTON?

- **Numpy**
- **Scipy**
- **Scikit-learn**
- **Matplotlib**

I have used a combination of Perl, Fortran, NCL, Matlab, R and others for routine research, but found out this general-purpose language, Python, can handle almost all in an efficient way from requesting data from remote online sites to statistics, and graphics.



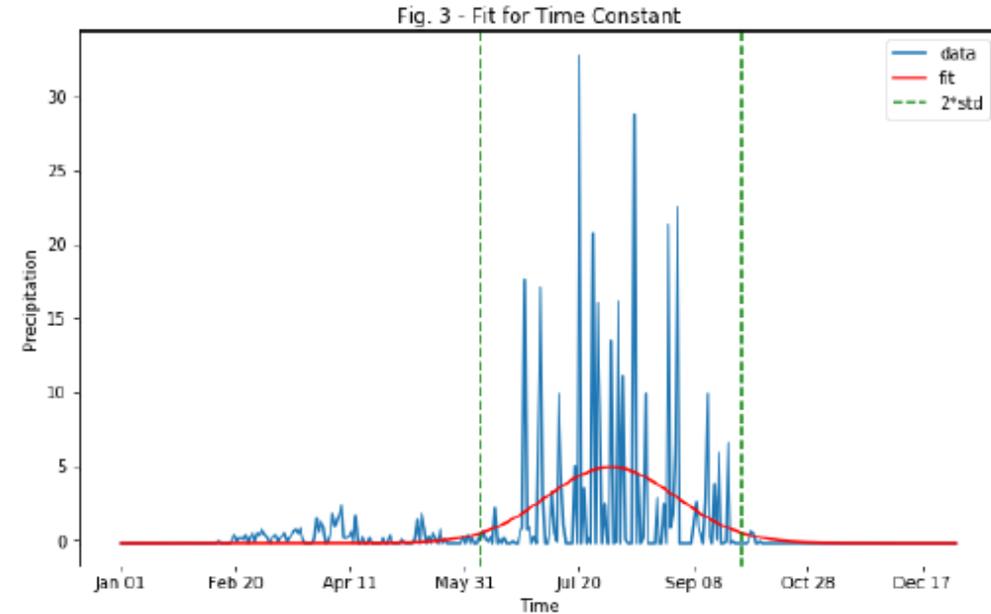
JUPYTER NOTEBOOK

The Jupyter notebook allows developers to parse their code into blocks which can be run independently of each other, with variables stored in the background.

Dividing the code up in this way can save vast amounts of time while developing a program, as it allows developers to test their code a few lines at a time, without running other lengthy processes included in the program.

```
In [34]: #Code for this algorithm is out of scope for this tutorial on datacube and is
         abstracted away for demonstration purposes.
         import demo.curve_fit_gaussian as curve_fit

         curve_fit.plot_fit(times, values, standard_deviations = 2)
```



We pick two points that are equidistant from the center/peak of this curve to act as our bounding points for the rainy season.

```
In [36]: curve_fit.get_bounds(times, values, standard_deviations = 2)
Out[36]: (numpy.datetime64('2015-09-28T11:59:59.500000000'),
         numpy.datetime64('2015-06-06T11:59:59.500000000'))
```

It appears that **JUNE** and **OCTOBER** should be adequate bounds for the rainy season.



ANALYSIS READY DATA

Sentinel-1:
Calibrated and
orthorectified absolute
backscatter products



CUBE IN A BOX

- Can be deployed locally or on the cloud
- Enables:
 - **Test environment**
 - **Development environment**
 - **Disposable workspace rapid evaluation**
- **Running Cube in a box on the cloud:**
 - **Template based AWS' CloudFormation**
 - **Deploys in 3 minutes**
 - **5 minutes to index data**
 - **Ready to run with Landsat 8 data from AWS auto-indexed**

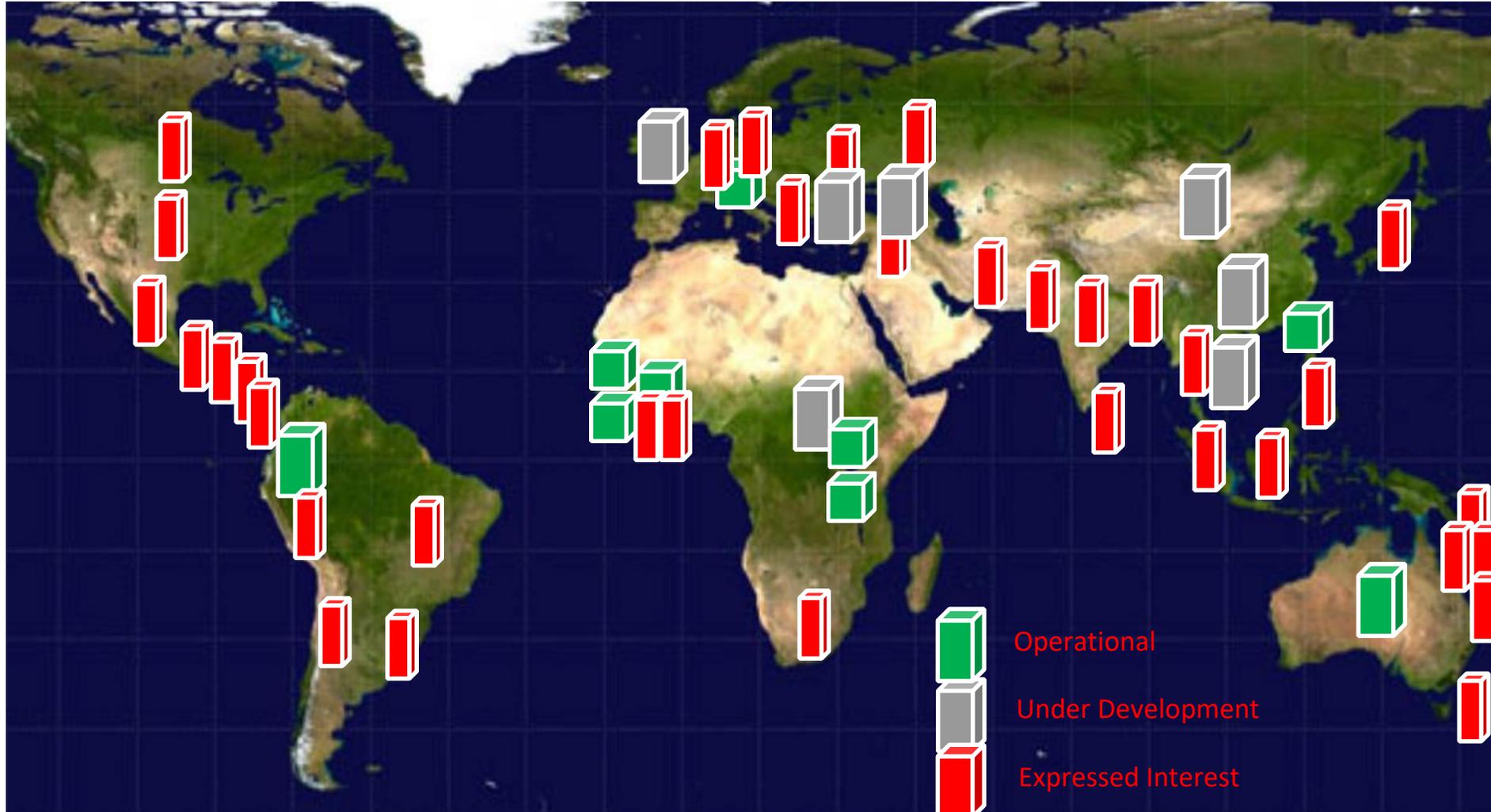


WHY DO WE NEED DATA CUBES?

- **Makes it easier to use satellite data by reducing preparation time**
- **Makes it easier to query data in time and space for powerful application products**
- **Provides free and open source software and algorithms for the cloud or a local computer**
- **Allows community development, shared capacity building and organized algorithms**
- **Enables efficient time series analyses and data interoperability**



GLOBAL IMPACT



More than 50 countries are interested with 9 in operation and 7 coming soon!



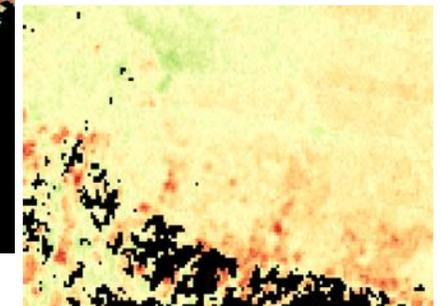
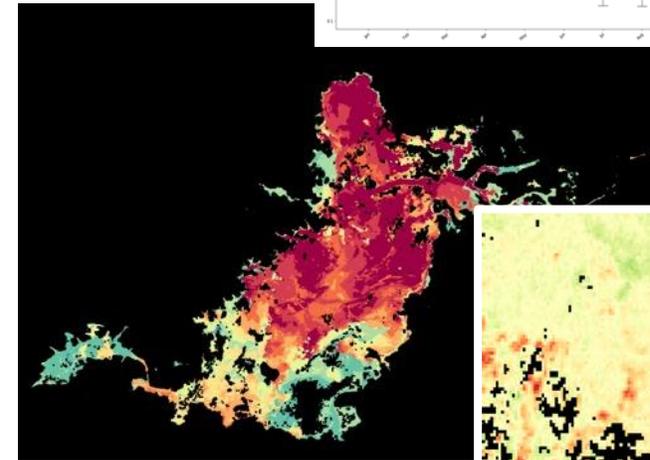
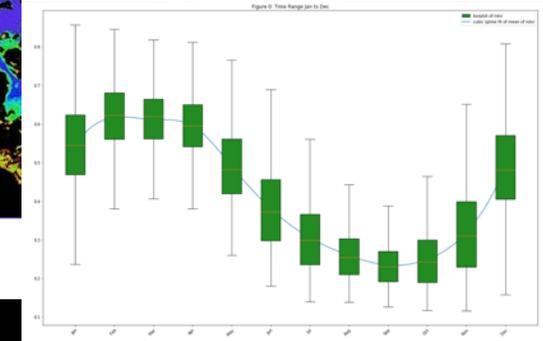
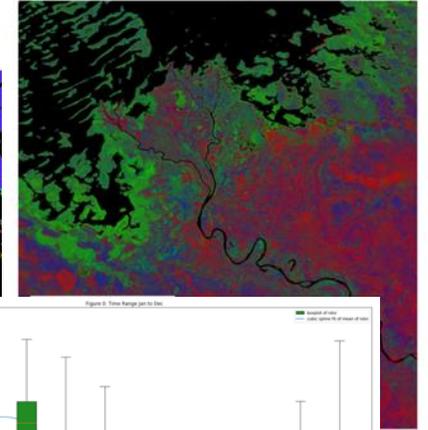
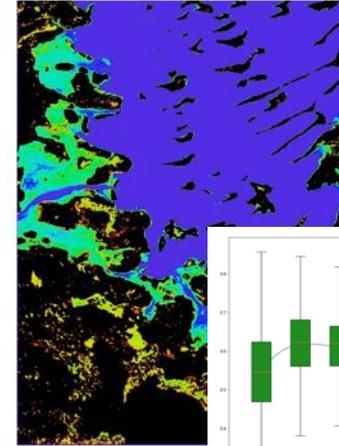
GLOBAL IMPACT

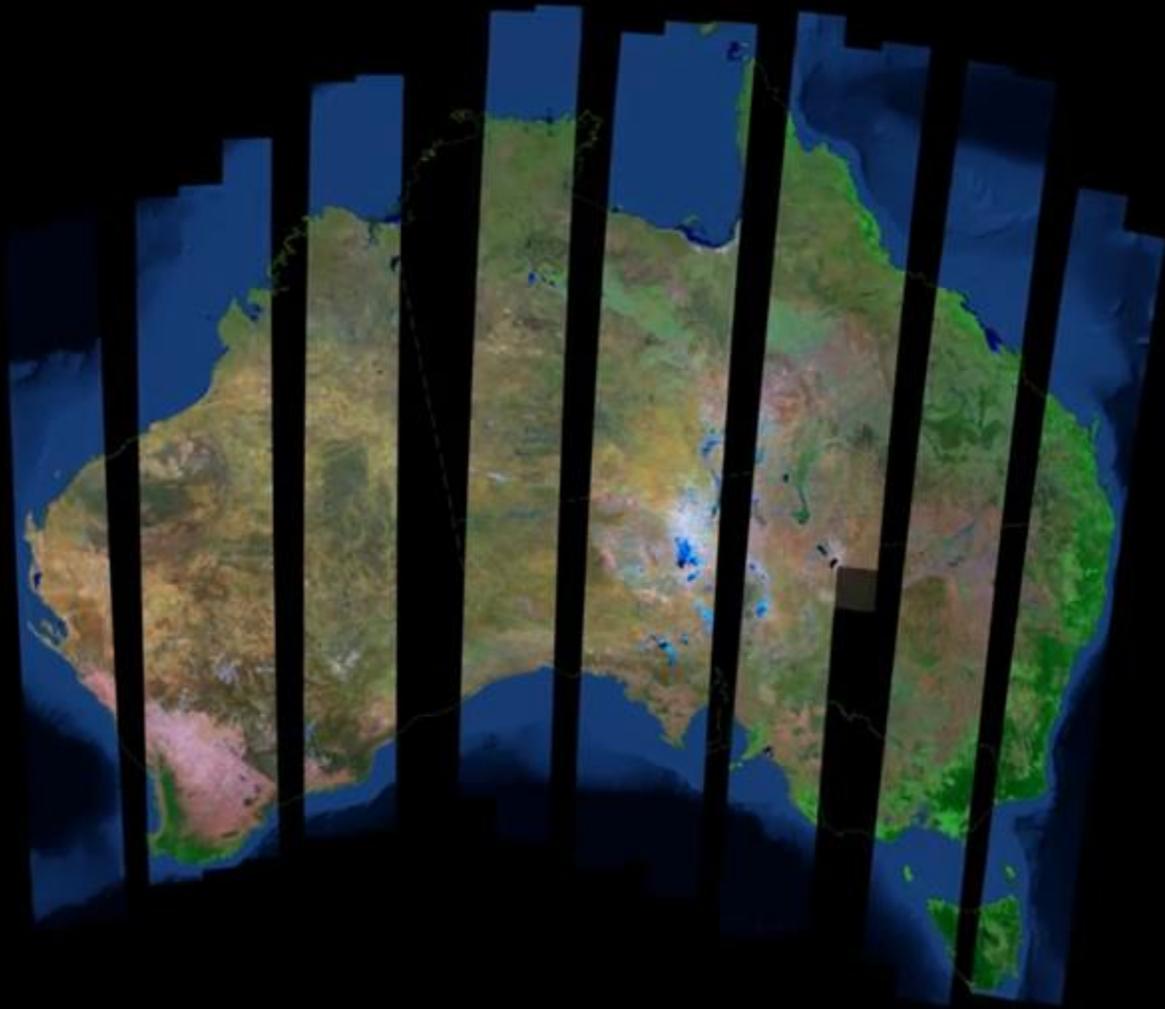
- **Switzerland** is using their Data Cube to develop new snow detection algorithms and develop time series snow coverage maps
- **UK** is using the Data Cube to develop a drought monitoring system for Uganda
- **UK** is using a Data Cube to monitor climate change resilience in the Pacific Islands
- **Taiwan** is using the Data Cube to monitor forest and vegetation restoration in landslide areas.
- **Uruguay** is using the Data Cube to adjust their water sampling approach to coincide with satellite overpasses in order to improve water quality data in their reservoirs.
- **The Gates Foundation** is using water extent time series data from the Lake Chad Data Cube to strategically supply polio vaccines to small villages that become "isolated" during the rainy season

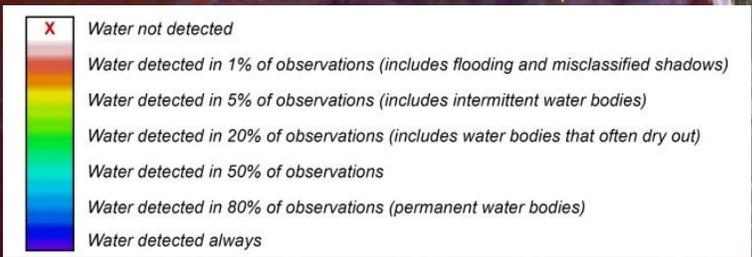
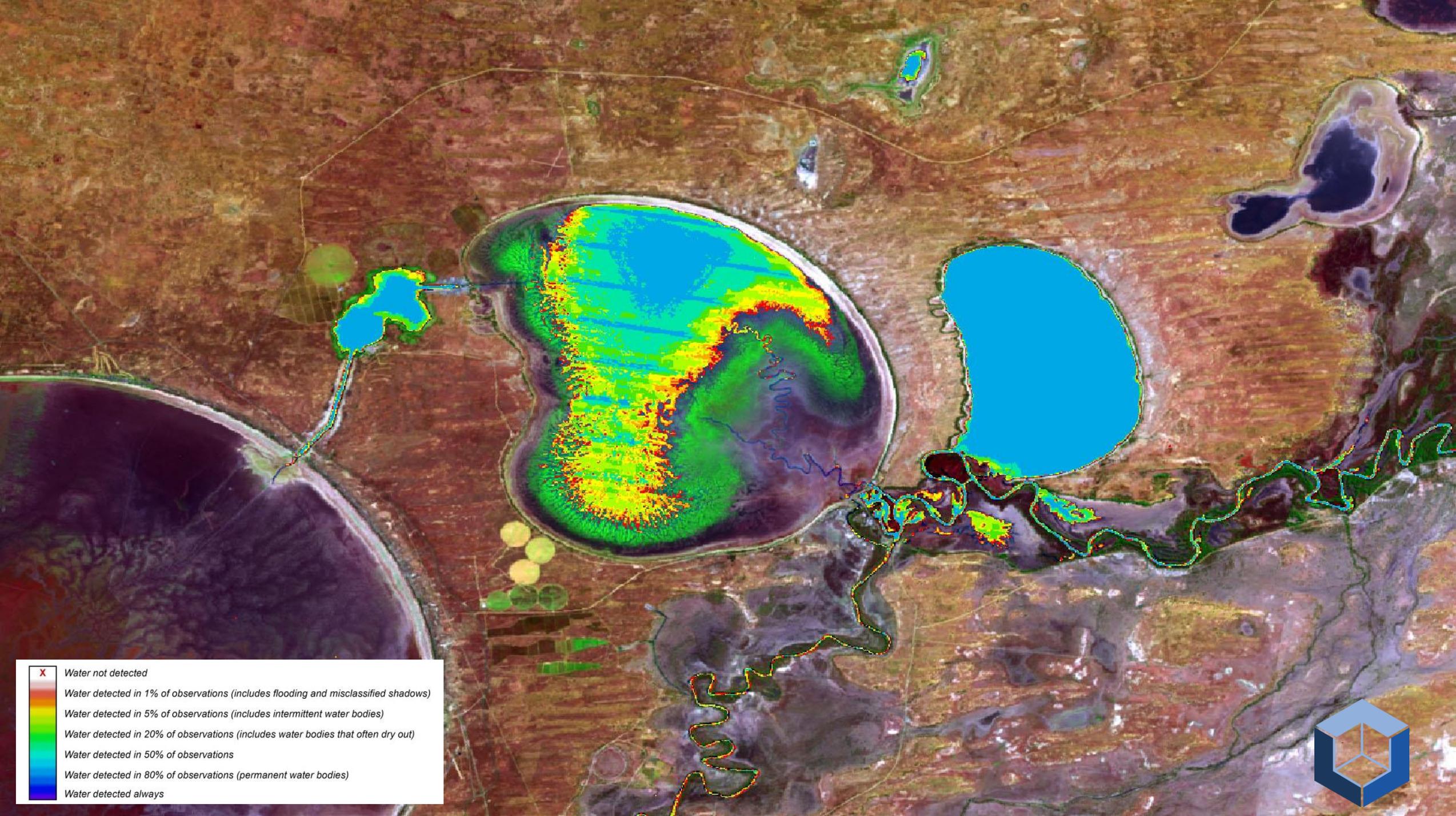


HOW CAN WE USE DATA CUBES?

- **Cloud-free Mosaics:** Recent Pixel, Median, Geomedian, Max-NDVI
- **Spectral Indices:** NDVI, EVI, NDBI, NDSI, NDWI, Fractional Cover, Tasseled Cap
- **Land Classification:** K-Means, Random Forest
- **Water:** WOFS (Australia), Sentinel-1 WASARD, Total Suspended Matter
- **Land Change:** PyCCD (USGS), PCA (Colombia), NDVI Anomaly, Sentinel-1 Radar (Deutscher), Coastal Change, Landslide Risk







ODC COMMUNITY

- **Founded by GA, CSIRO, NASA-SEO, USGS, and Satellite Applications Catapult in Jan 2017 as an open source data management solution with a community of contributors and users**
 - *www.opendatacube.org*
- **Driving ODC roadmap and managing technical repository**
 - *<https://github.com/opendatacube>*
 - *<http://slack.opendatacube.org/>*
- **CEOS is extending the ODC goals to demonstrate data cubes around the world and maximize the impact of CEOS satellite data with a focus on developing countries**
- **Focused on ARD, understanding user needs, and global capacity building**



THE FUTURE

What next?



Thank you



We work with
Innovate UK

