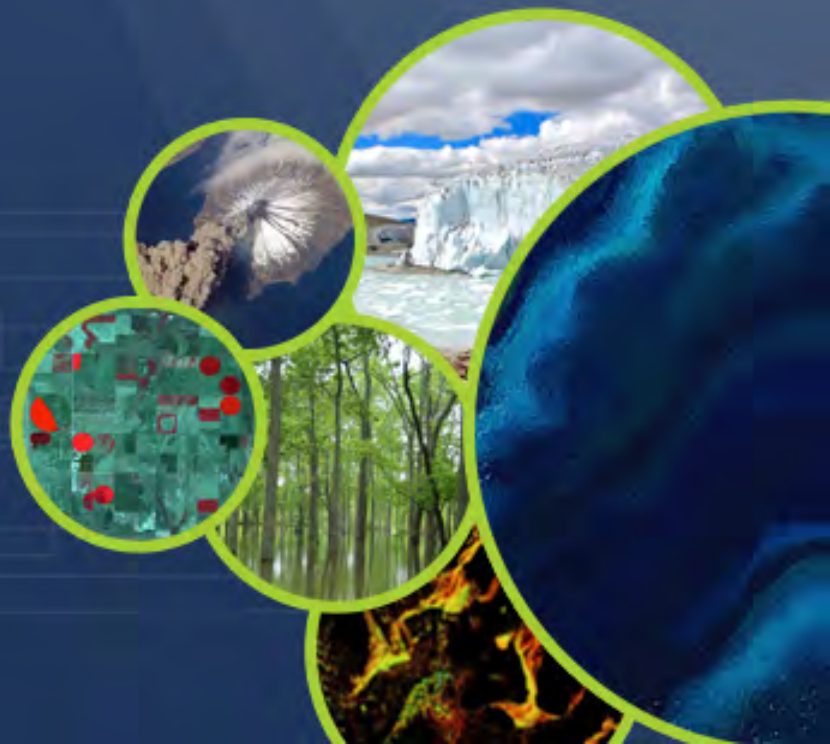
The Open Data Cube

A Big Data Solution for Global Capacity Building and Monitoring Environment

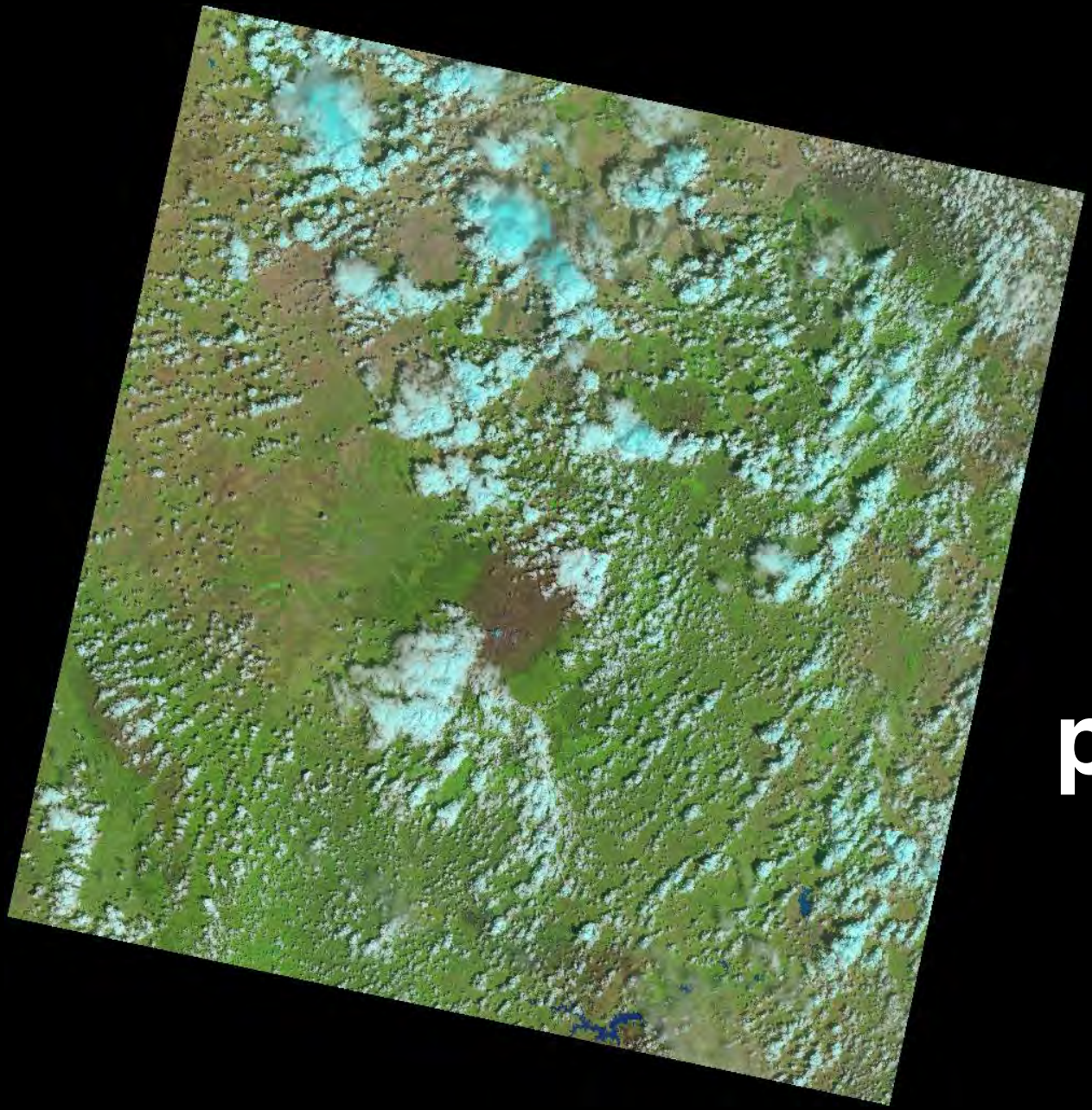
Africa Regional Data Cube
Training Workshop
Nairobi, Kenya
May 9, 2018 (Day #1)

Dr. Brian Killough
CEOS Systems Engineering Office
NASA Langley Research Center



Committee on Earth Observation Satellites (CEOS)

International coordination of satellite data …
22 countries, 60 members, 137 active satellites



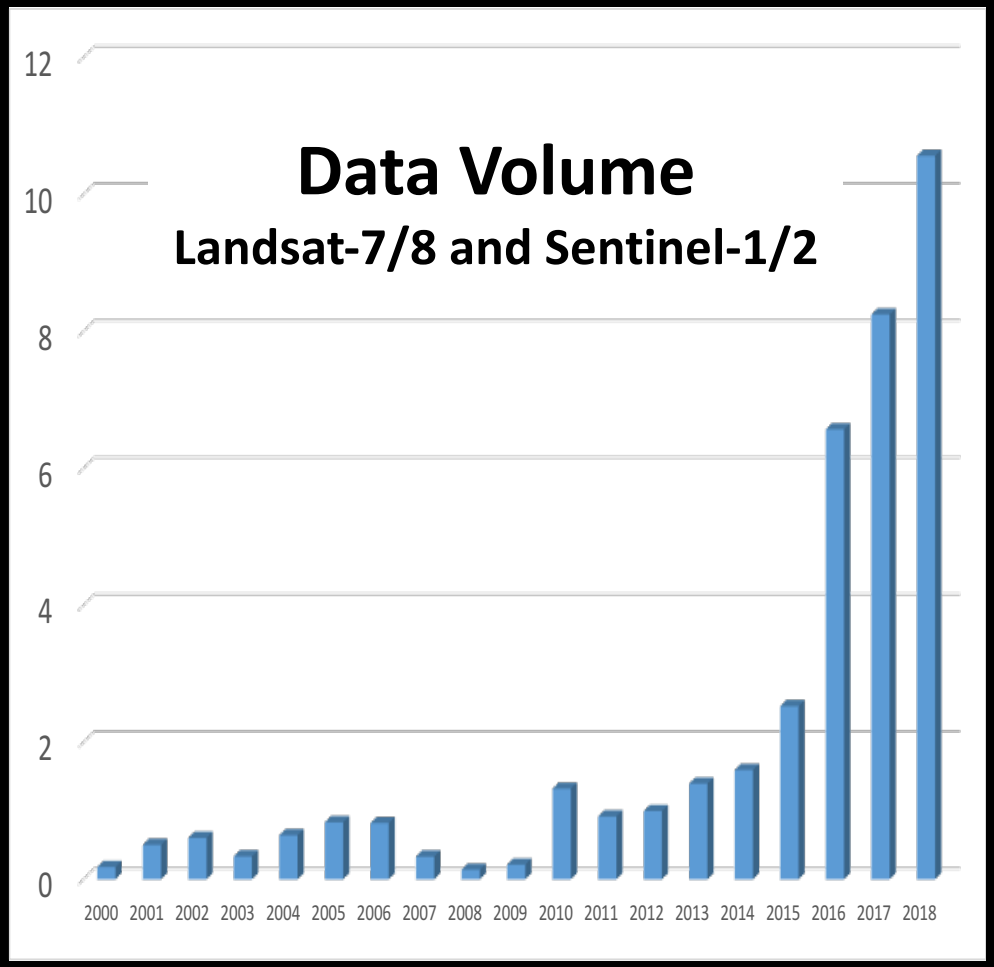
Landsat
$600
per scene
in 2007

34 scenes to cover Kenya

$918,000 per year

Now it is FREE!

The Big Data Problem



- Free data has **increased data volumes** by 10x in the last 5 years.
- Many countries **lack the expertise, infrastructure, and resources** to access and use the data to create products.
- Countries have requested help ...

The Latest Trends

Free and Open Resources

- Abundant Satellite Data
- Open Source Software and Tools

Global Engagement

- Improved communication
- Increased cooperation and collaboration

Global Philanthropy

- Google Earth Engine, Earth on AWS
- Governments, World Bank, Others (GPSDD)

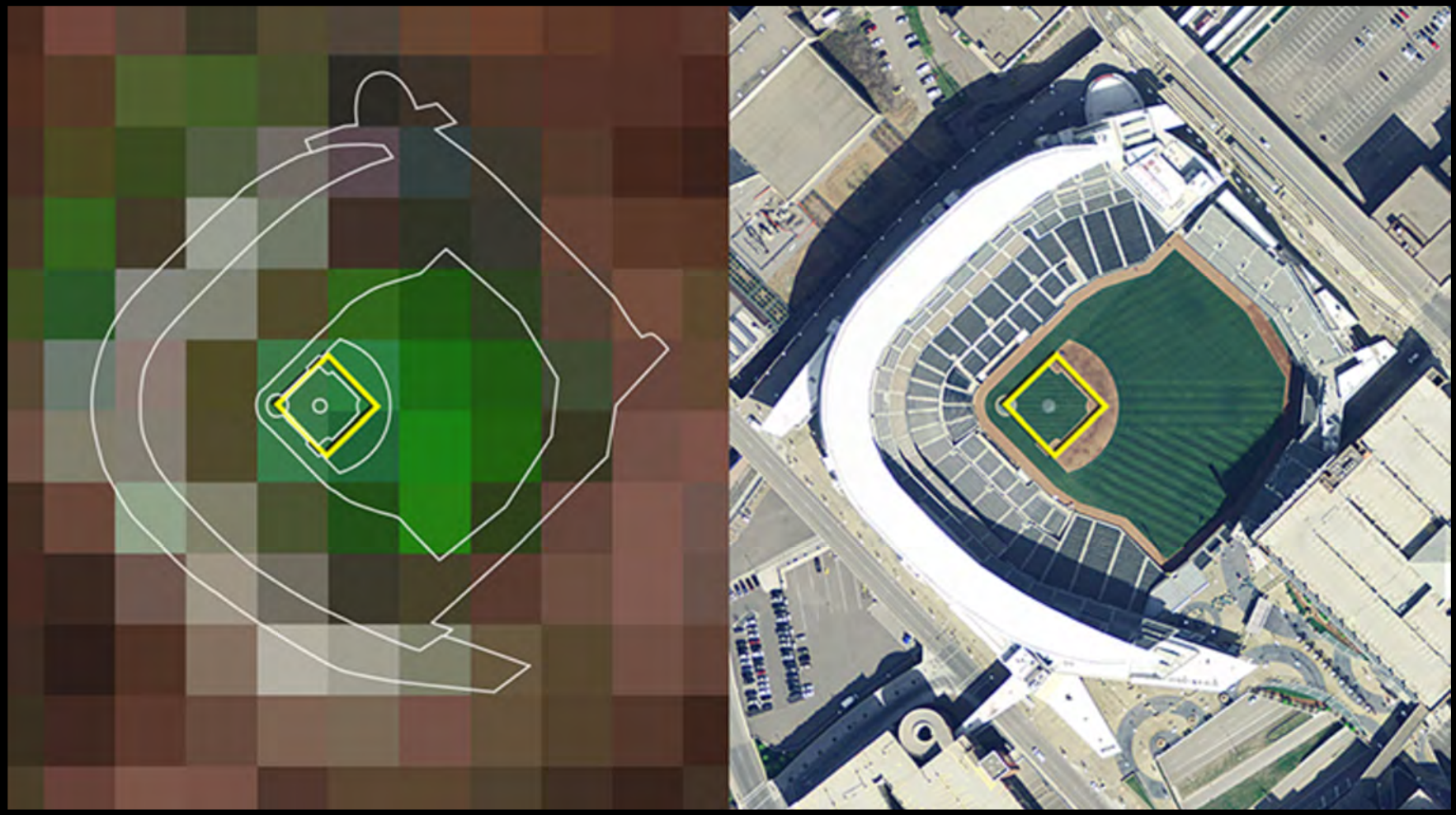
Improved Technology

- Cloud Storage and Computing
- **Data Cubes**

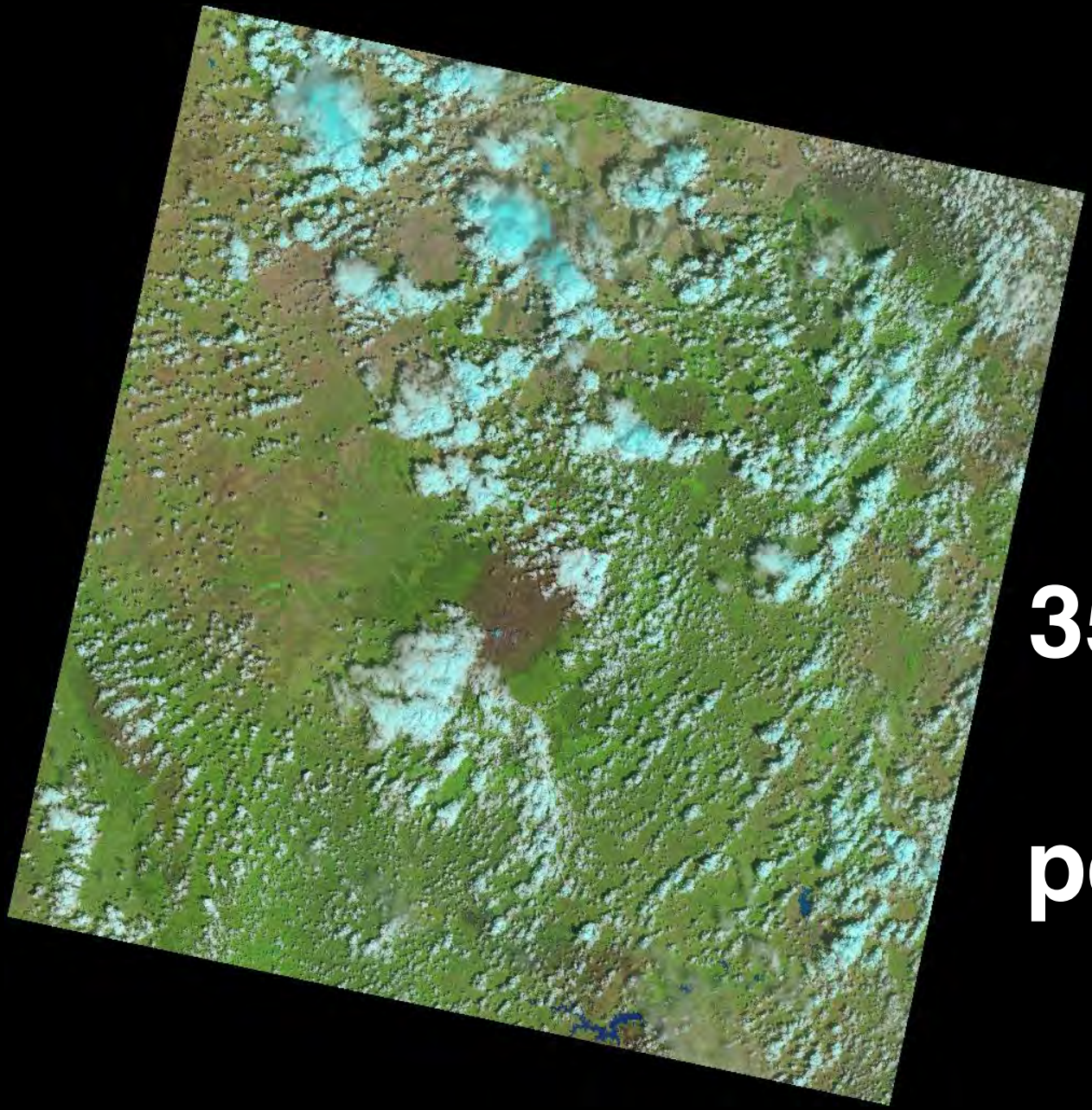
Everything is in place ...
Why doesn't the world use satellite data?

- It requires science knowledge to understand what data is needed … optical or radar
- It is hard to access and download
- It is hard to prepare … atmospheric correction, grid formats, pixel alignment, cloud filtering
- It requires capacity building and training

A New Solution ... **Data Cubes**

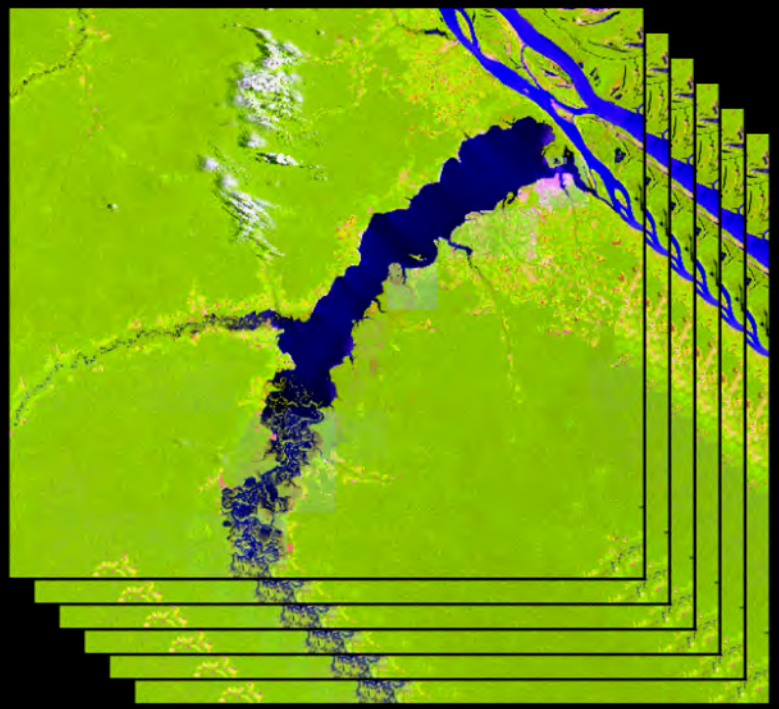


1 pixel = 30 meters square

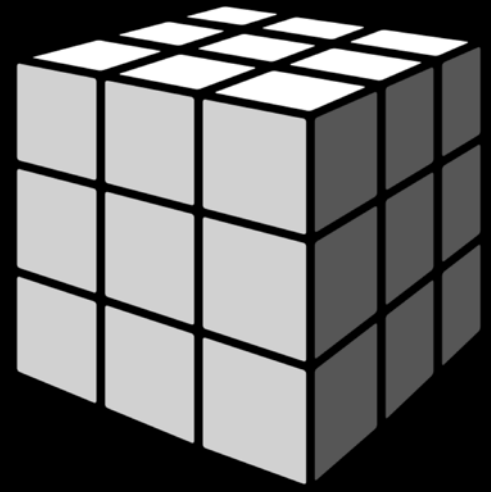


35 million pixels per scene

What is a Data Cube?



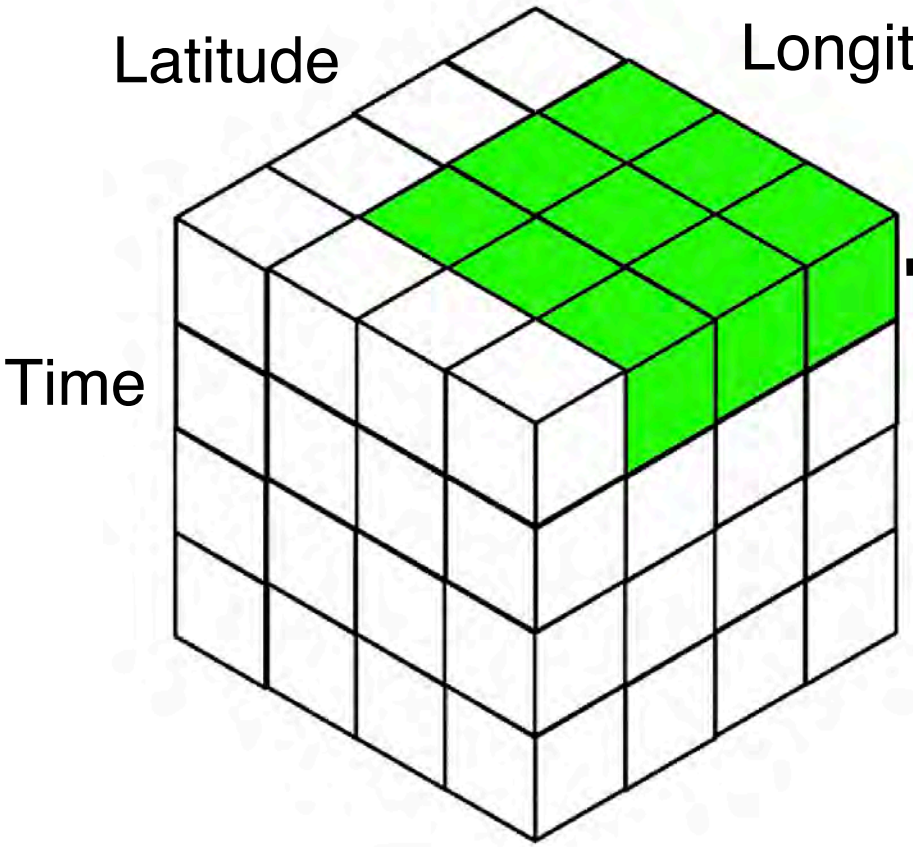
8000 Landsat scenes
(17 years of data)



1000 Data Cube storage units (1° x 1° x 1 yr)

8x data compression

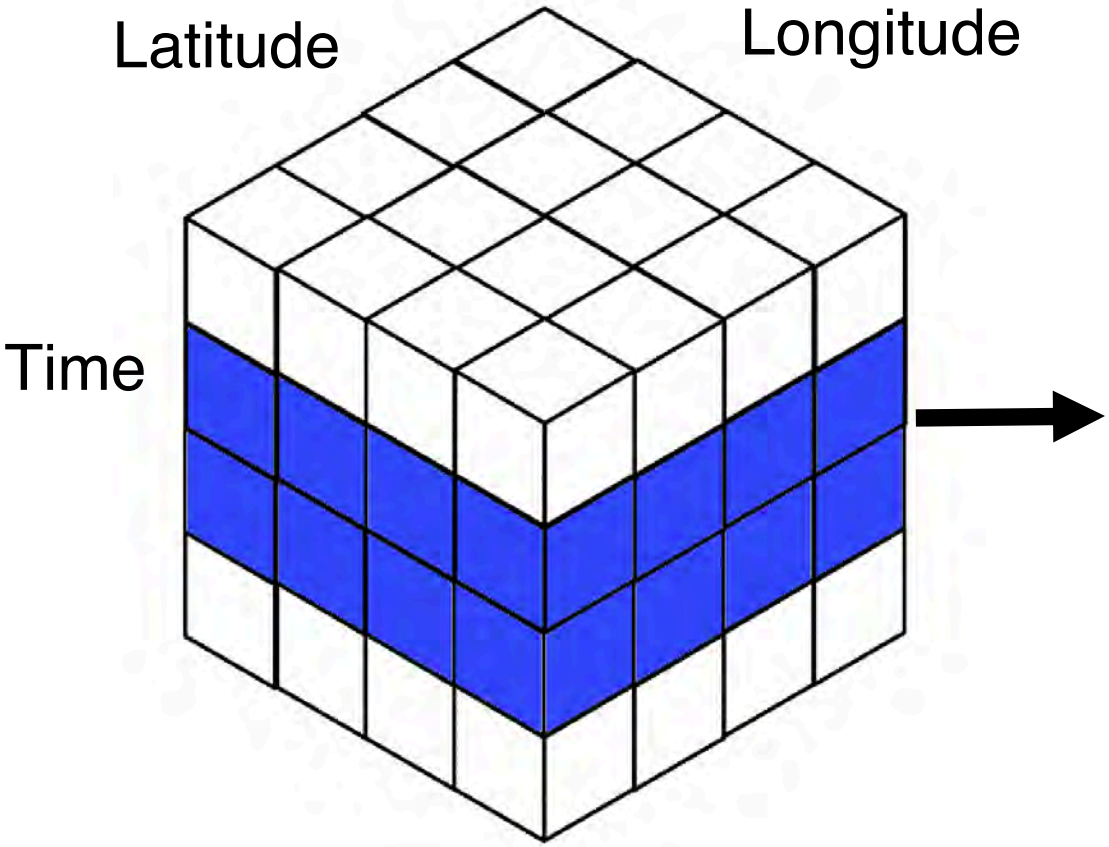
Sampling a Data Cube



A single time slice, similar to a standard “scene” can be used to assess a single point in time

Pixels in the Data Cube are processed, aligned, and compressed and ready for data analysis

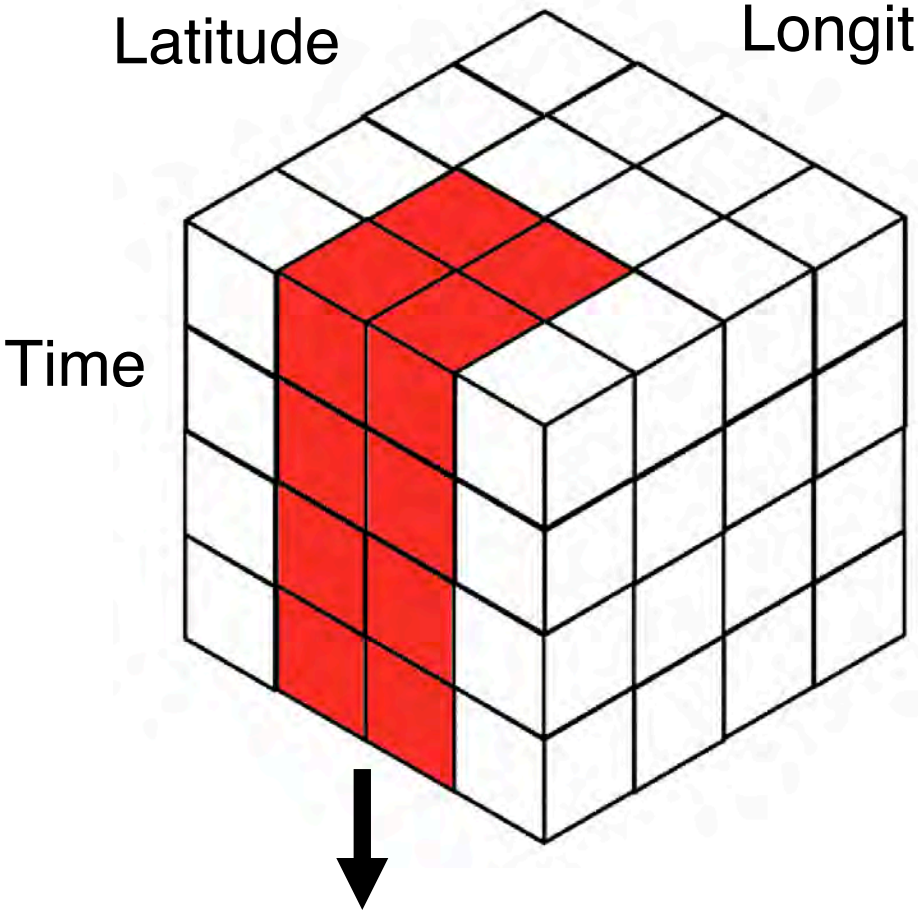
Sampling a Data Cube



Several time slices can be combined into one to form a "**Mosaic**". This is often used to reduce clouds or create seasonal or annual images.

Typical Mosaics ... Most/Least Recent Pixel, Mean/Median, Geomedian, Min/Max NDVI

Sampling a Data Cube



Examples of **Time Series** analyses include: Land Change (PyCCD), Water Change (WOFS), Parameter variation along a transect (Hovmoller plot)

Time Series analyses consider the variation of data over time to assess change

Benefits of Data Cubes

- Expanded use of satellite data
- Reduced data preparation
- Enables data interoperability and efficient time series analyses
- Free and open access
- Flexible deployment (local or cloud) using a common architecture
- Community development and sharing
- Proven concept ... Australia, Colombia, Switzerland, Taiwan **and now Africa!**

The Data Cube Vision

A solution supporting priority objectives …

- Build capacity of users to apply CEOS satellite data
- Support GEO and United Nations agendas

Customer focused …

- Easy to install and maintain with training materials
- A brand that people know and trust

Scalable solution …

- Operational Data Cubes in **20 countries by 2022**
- Partnerships with … Google, Amazon

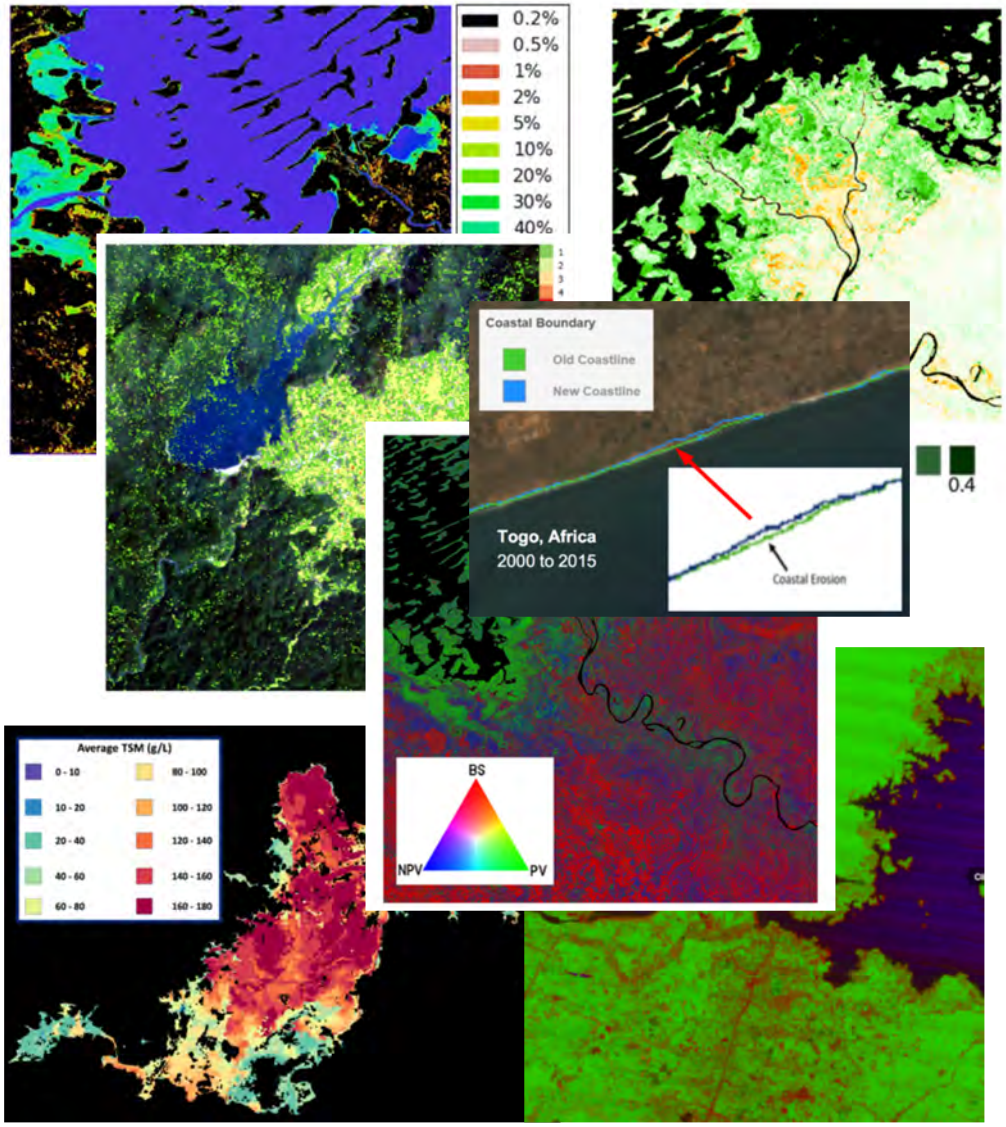
The “Road to 20”

43 countries in the 17 months!

Application Products

30 algorithms

- **Cloud-free Mosaics:** Most/Least Recent Pixel, Median, Geomedian (in testing) and MVC (coming soon)
- **Spectral Indices:** NDVI, EVI, NDBI, NDSI, NBR, NDWI, FC, TCT, SVVI (coming soon)
- **Land Change Classification:** K-Means, Random Forest (in testing), FNF (from Colombia), SOFS (from Switzerland)
- **Water:** WOFS, WASARD (radar), TSM, WLUT (coming soon)
- **Land Change:** PyCCD, PCA (from Colombia), NDVI Trend (in testing), NDVI Anomaly, Deutscher (radar), Coastal Change, SLIP (landslides), Hovmoller/Transects



The Future

- Data Cube deployments: Vietnam, U.K., Uganda and the Africa Regional Data Cube
- Progress collaborations with Google and Amazon
- IGARSS Conference in Valencia, Spain (July 2018) … dedicated paper session and training course
- New technical additions: Jupyter (Python) Notebooks, Web-based User Interface tools
- New user applications and algorithms: Custom Cloud-filtered Mosaics, Land Classification, Water Quality, Land Change

THANK YOU

Web: opendatacube.org
Twitter: @opendatacube