



Committee on Earth Observation Satellites

Using Satellite Data in the Cloud with the Open Data Cube

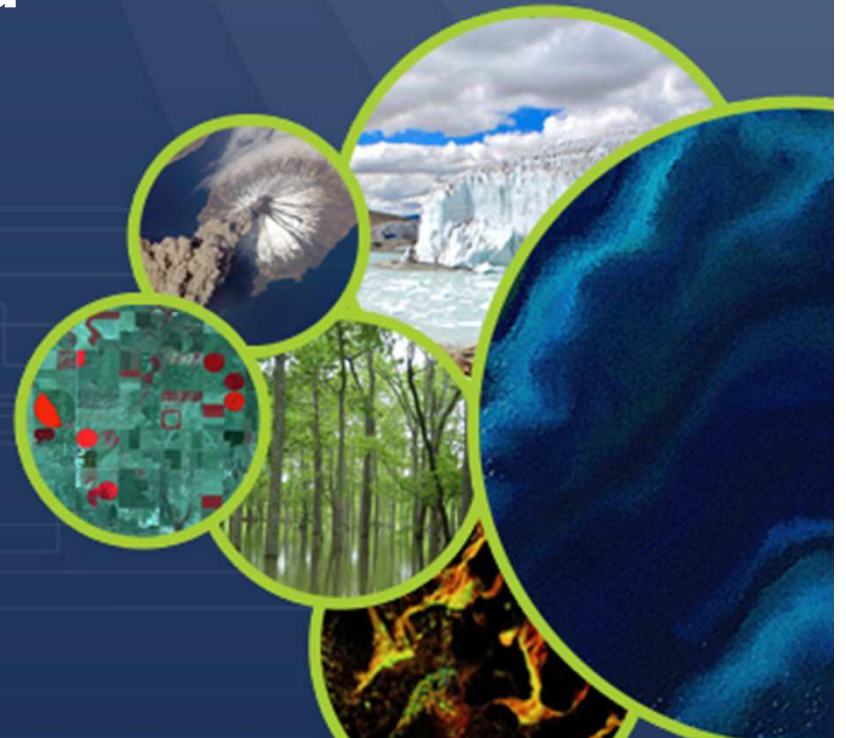
Brian Killough, PhD

NASA Langley Research Center
CEOS Systems Engineering Office (SEO)



Eyes on Nature Workshop

October 5, 2021

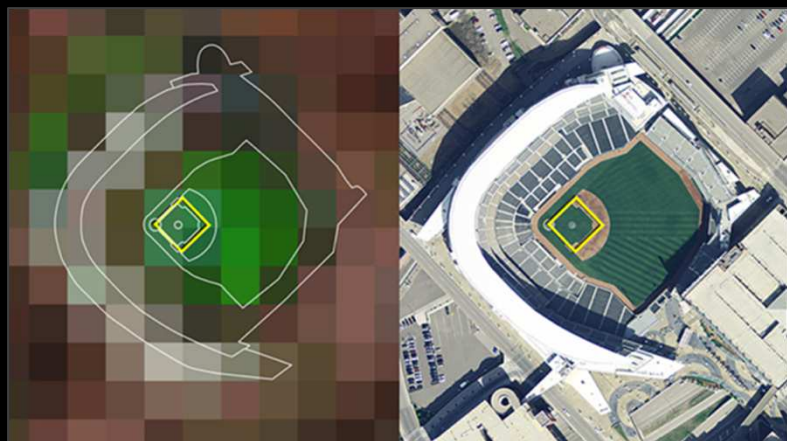
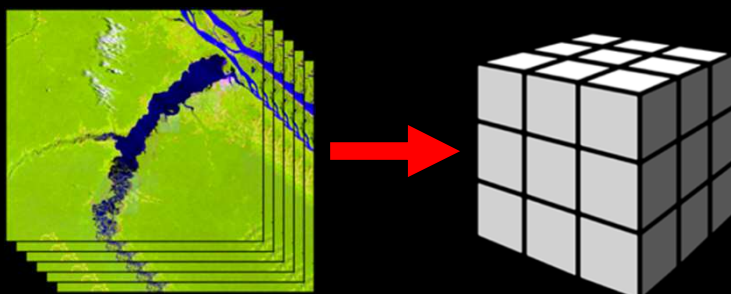


The Trends in Space Data

- Growing free/open satellite data available in the cloud ... no need to download large volumes of data
- Increasing cloud vendor options ... Google, Amazon, Microsoft
- Satellite data being pre-processed into analysis-ready formats
- New open sources tools (e.g. **Open Data Cube**) for using data and promoting open science
- Jupyter notebook programming environments using Python
- Increasing international collaboration and regional initiatives (e.g. Digital Earth Africa / Pacific / **Americas**)



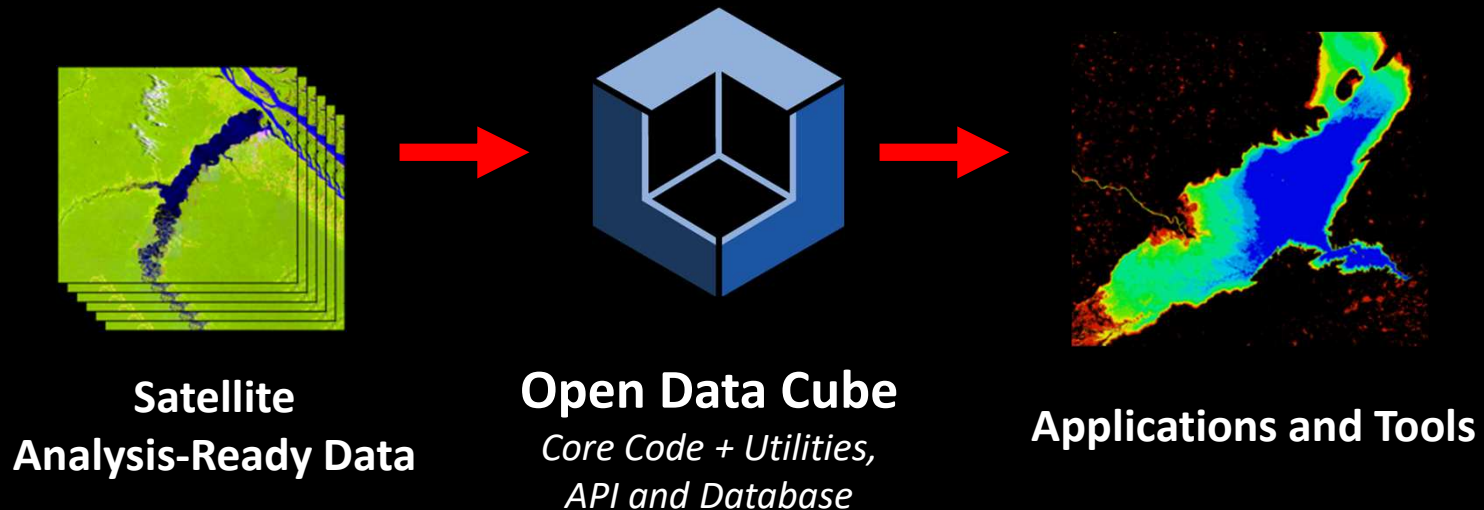
What is a Data Cube?



- **Satellite data** from many days and years are organized into a cube of space (latitude and longitude) and time.
- These cubes are made of small “**pixels**” that give us data at a scale of 30-meters (Landsat) ... about the size of a baseball diamond.
- When organized in a data cube, it is **much easier to analyze** the data and create valuable products

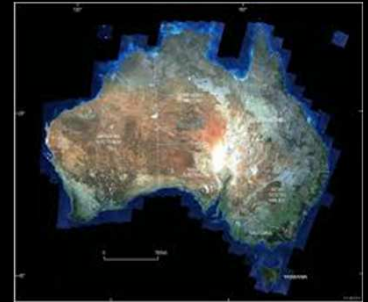
What is the Open Data Cube?

Open Data Cube (ODC) ... an open source geospatial data management and analysis framework for decision-making



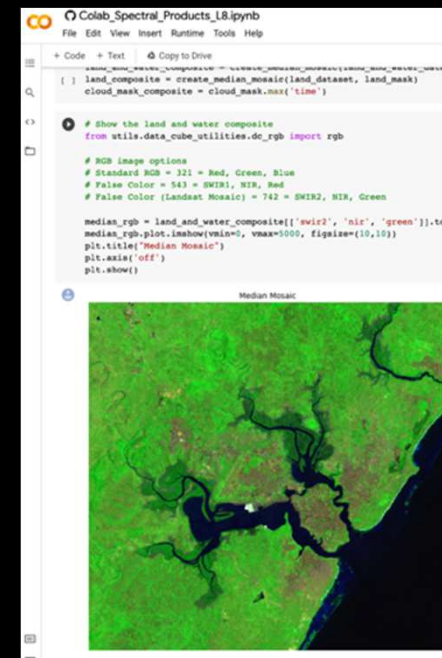
Past, Present, Future

- **Past ...** Initiated and proven in Australia as the Australian Geoscience Data Cube. The Open Data Cube (ODC) concept and brand was born within the Committee on Earth Observation Satellites (CEOS).
- **Present ...** Digital Earth Africa (operational) + Digital Earth Pacific (early planning) + **Digital Earth Americas** (early planning) + over 100 local or country-level data cubes. Improvements in core ODC code, application algorithms, and cloud computing methods.
- **Future ...** The goal is to achieve a global network of connected regional data cubes using ODC algorithms. We hope to grow the user community and share/test algorithms and methods to support SDGs and user needs.



Open Data Cube Sandbox

- The Committee on Earth Observation Satellites (CEOS) has developed a new Open Data Cube (ODC) Sandbox that runs on Google Colab. <http://openearthalliance.org/sandbox>
- This tool is a free/open Jupyter notebook interface connected to Google Earth Engine datasets that can create sample application products anywhere in the world.
- Sample **applications** include: Landsat and Sentinel-2 cloud statistics, spectral products, cloud-filtered mosaics, vegetation change, water extent, vegetation phenology, VIIRS nighttime lights, mission coincidences, and radar products.



```
Colab_Spectral_Products_18.ipynb
File Edit View Insert Runtime Tools Help

Code Text Copy to Drive

[] land_composite = create_median_mosaic(land_dataset, land_mask)
cloud_mask_composite = cloud_mask_msr('time')

# Show the land and water composite
from utils.data_cube_utilities.dc_rgb import rgb

# RGB image options
# Standard RGB = 321 = Red, Green, Blue
# False Color = 433 = SWIR1, NIR, Red
# False Color (Landsat Mosaic) = 742 = SWIR2, NIR, Green

median_rgb = land_and_water_composite[['swir2', 'nir', 'green']].to_array()
median_rgb.plot.imshow(vmin=0, vmax=1000, figsize=(10,10))
plt.title('Median Mosaic')
plt.axis('off')
plt.show()
```



Tabasco, Mexico – November 2020

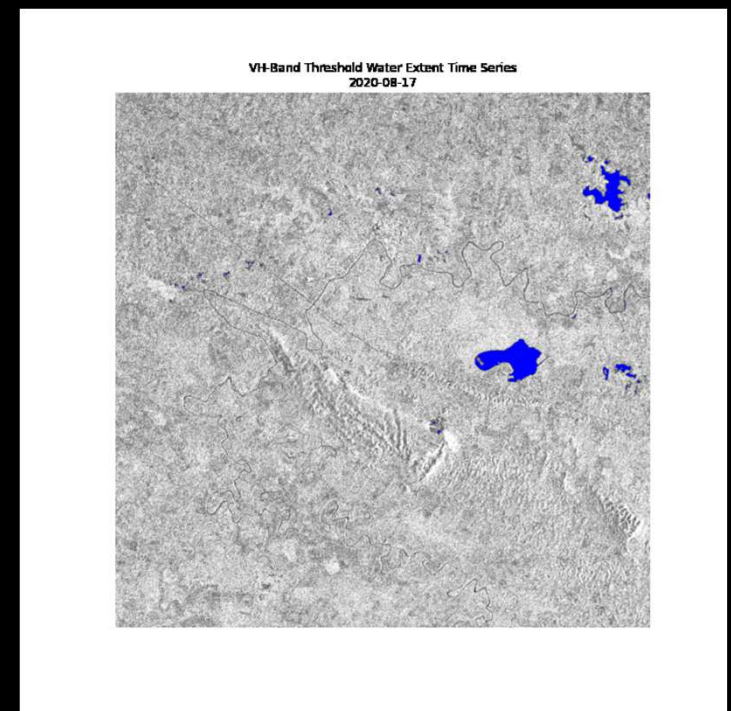
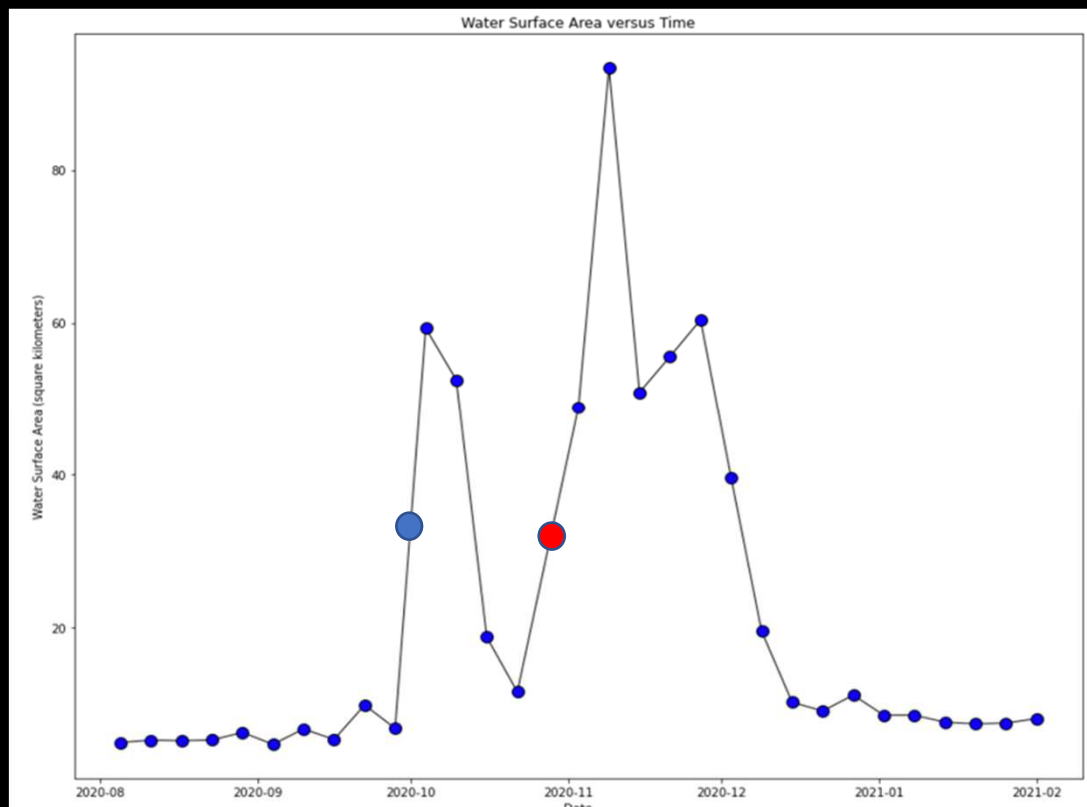
- The International Charter was activated on November 6, 2020 for Tabasco, Mexico due to flooding from Tropical Storm Eta.
- Heavy rainfall across south and southeast Mexico affecting over 100,000 people. At least 21 people have been killed and thousands of homes destroyed.
- The worst affected areas of Tabasco, Chiapas and Veracruz, received torrential rain triggering landslides which claimed 2000 homes.
- At least 10 rivers have burst their banks in the Gulf coast state, causing widespread flooding.



Macuspana, Tabasco, Mexico on Nov 9, 2020
after Tropical Storm Eta
Reference: atalayar.com

Flooding in Macuspana, Tabasco, Mexico

- Tropical Storm Gamma – October 2-3, 2020
- Tropical Storm Eta – November 1-6, 2020



Time Series Animation of
Flooding Extent

What do we see for the future?

- More satellite data available in more clouds
- More regional and local data cubes around the world ... broad adoption by governments, academia and industry
- Faster and more efficient ODC applications using parallel processing, and Machine Learning.
- More Python proficiency across the globe
- Combining satellite data with diverse datasets
Drones, Internet of Things (IoT)



THANK YOU

ODC Sandbox: openearthalliance.org/sandbox

ODC website: opendatacube.org

Twitter: [@opendatacube](https://twitter.com/opendatacube)