Digital Earth
AFRICA

Introduction to Digital Earth Africa

Chad Burton | chad.burton@ga.gov.au

Contents

1. **Introduction** to the Digital Earth Africa program, platforms, and services.

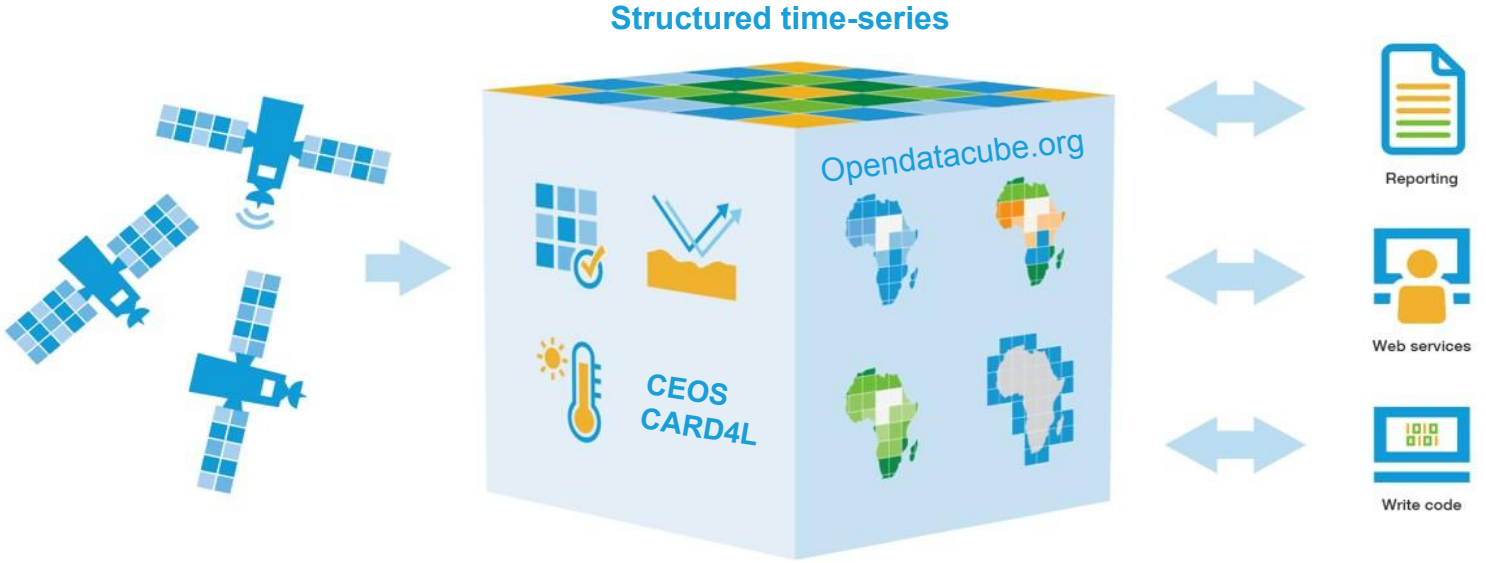
2. **Use Case:** Building a continental cropland extent mask for Africa using ML

3. **Showcase:** Introduce a notebook *series* exploring supervised learning in the Open Data Cube

4. **Exercise:** Run a simple ML model to classify urban areas in Kenya

What is Digital Earth Africa?

Digital Earth AFRICA



Observations

Open data from government observing systems

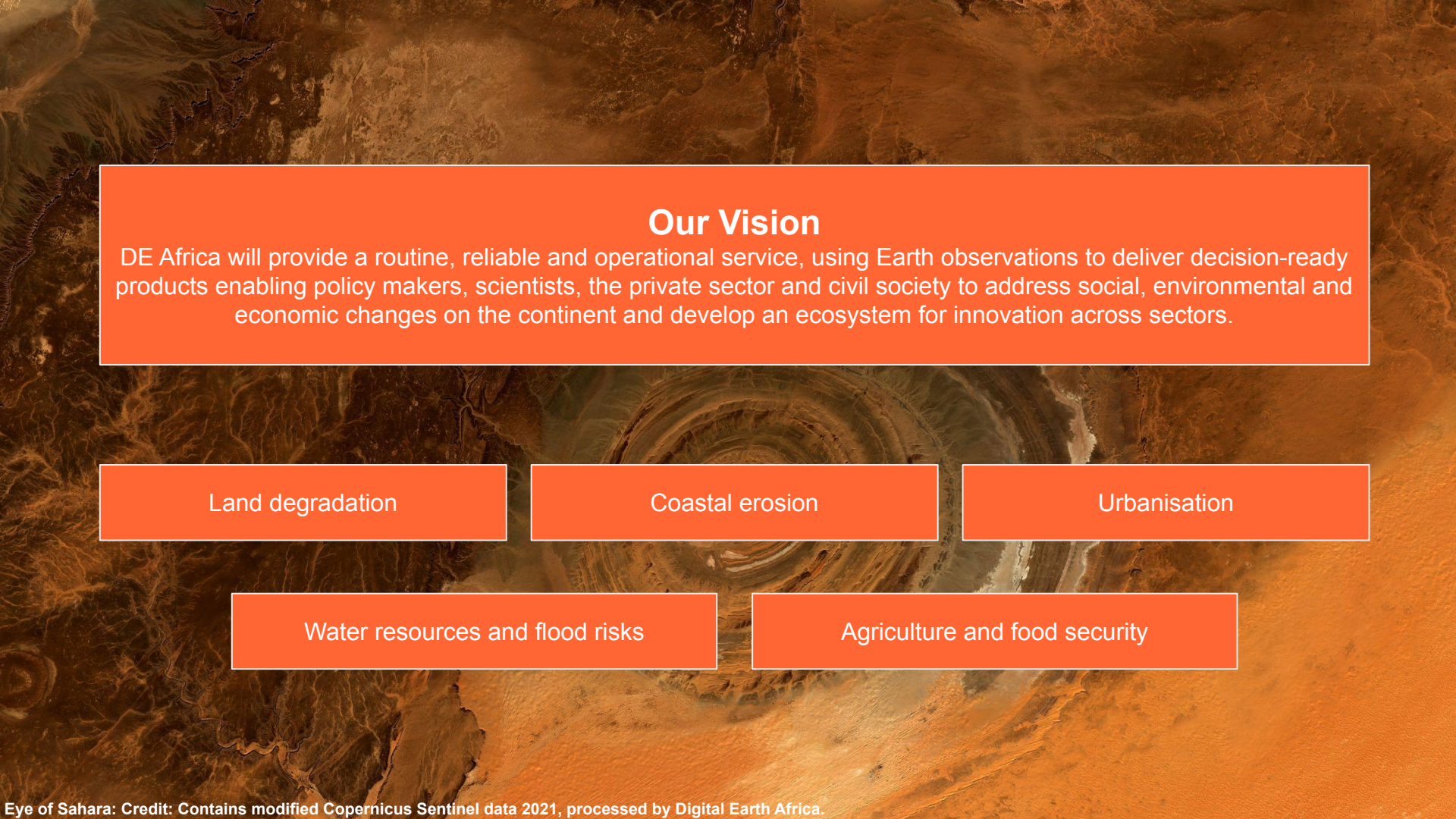
Analysis Ready Data

Products

EO products that are comprehensive, routine, user-orientated & CC

Information for decisions

Specific information products, including simple facts, numbers, visualisations



Our Vision

DE Africa will provide a routine, reliable and operational service, using Earth observations to deliver decision-ready products enabling policy makers, scientists, the private sector and civil society to address social, environmental and economic changes on the continent and develop an ecosystem for innovation across sectors.

- Land degradation
- Coastal erosion
- Urbanisation
- Water resources and flood risks
- Agriculture and food security

Eye of Sahara: Credit: Contains modified Copernicus Sentinel data 2021, processed by Digital Earth Africa.

Analysis Ready Satellite Datasets

Digital Earth AFRICA

All datasets cover Africa, hosted in AWS Cape Town,

Freely accessible and optimized for analysis

Datasets	Type & Specs	Coverage	Source	Timeline
Sentinel-2 level-2A	Multispectral optical, up to 10 m, every 5 days	2017 - current	ESA, reformatted	Now
Landsat collection 2 level 2	Multispectral optical and surface temperature, 30 m, every 16 days	1984 - current	USGS	Now
Sentinel-1 backscatter	Terrain corrected, 20 m, every 12 days	2017 - current	Sinergise (generated from ESA data)	~ May/June 2021
SRTM Digital Elevation Model	Static, 30 m	-	NASA	Now
ALOS annual mosaic	radar backscatter, 30 m	Selected years	JAXA	Now

Products

Water Observation from Space(WOfS)

Sentinel-2 Annual GeoMAD composites (semi-annuals in the pipeline)

Crop Mask - Prototype

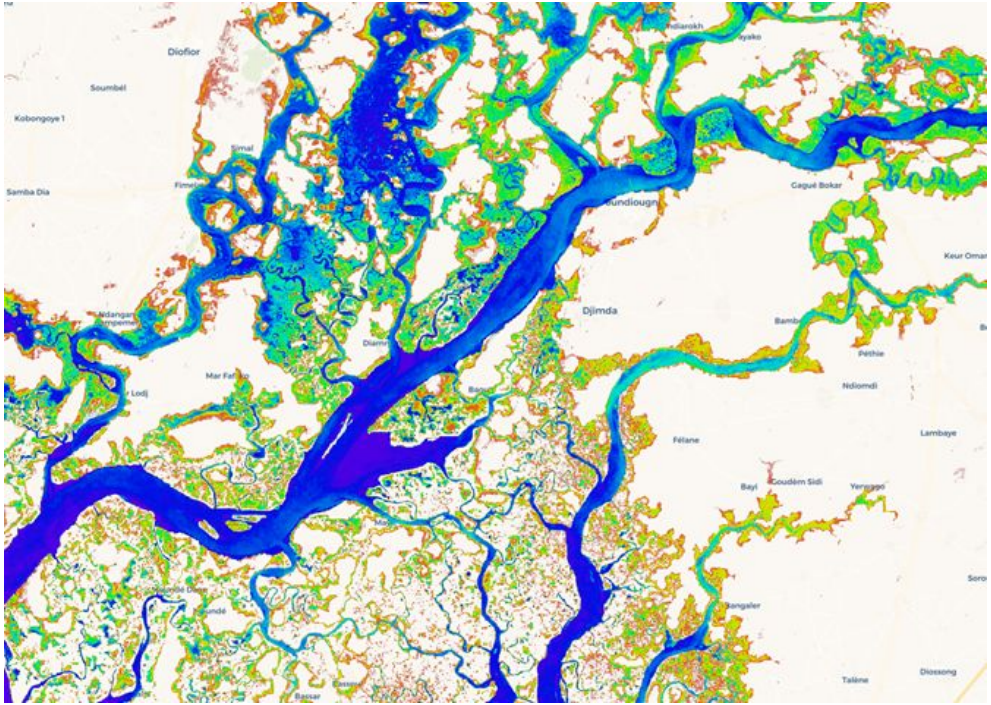
Water Observations from Space

Digital Earth AFRICA

Mapping surface water dynamics for the continent

- 30 m resolution
- 1980s to now
- Validated across Africa

https://docs.digitalearthafrica.org/en/latest/data_specs/Landsat_WOfS_specs.html



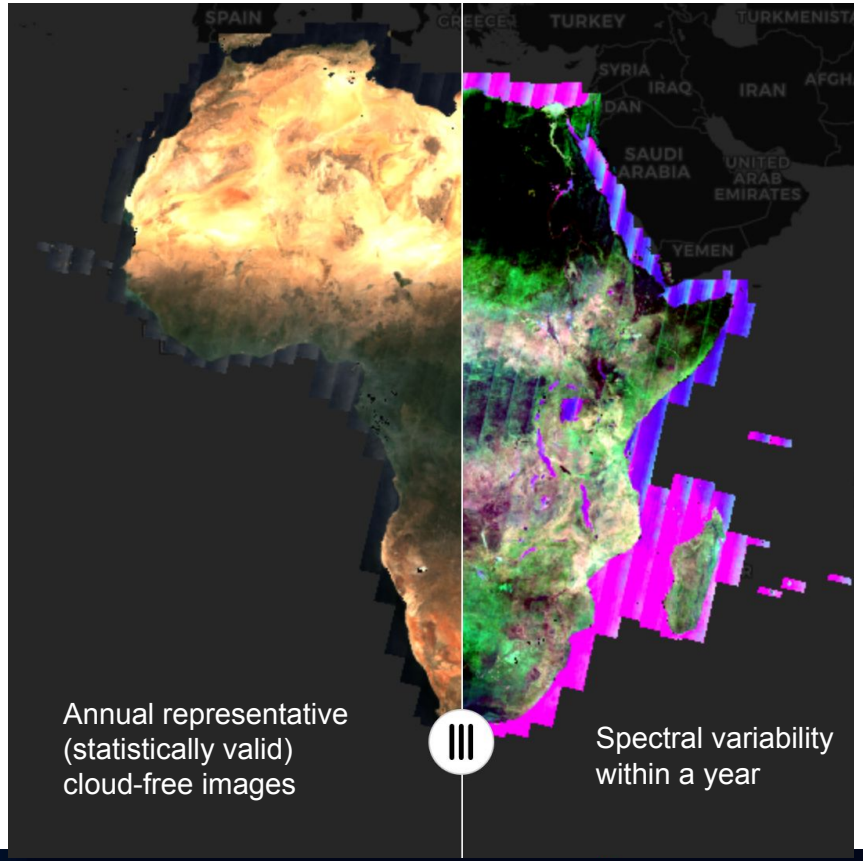
Summary for an area in Senegal from 2013 to 2019. Colours indicate the percentage of times water was detected (Red = rarely water, green = often water, blue = always water).

Sentinel-2 (10m) Annual GEOMAD

Digital Earth AFRICA

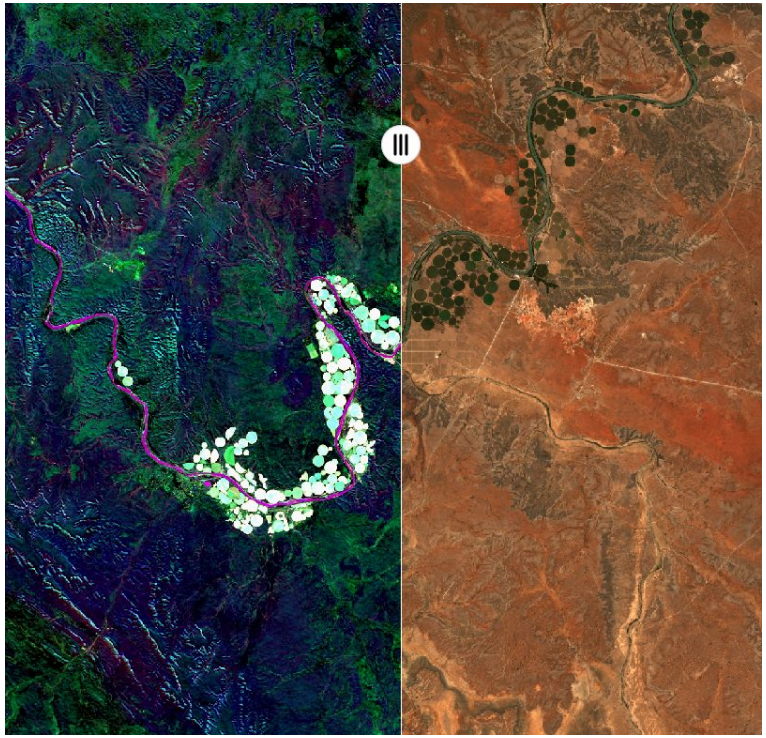
Annual composite images for the Sentinel-2 archive.

- Geomedians are statistically valid high-dimensional composites → relationship between bands are preserved. Band indices and other typical RS analysis all work as expected
- Includes three measures of median absolute deviations (MADs) - captures annual variability
- https://docs.digitalearthafrica.org/en/latest/data_specs/GeoMAD_specs.html



Sentinel -2 Annual GEOMAD

Digital Earth AFRICA



Irrigation along the Orange River, South Africa

Border of Serengeti Protected area, Kenya

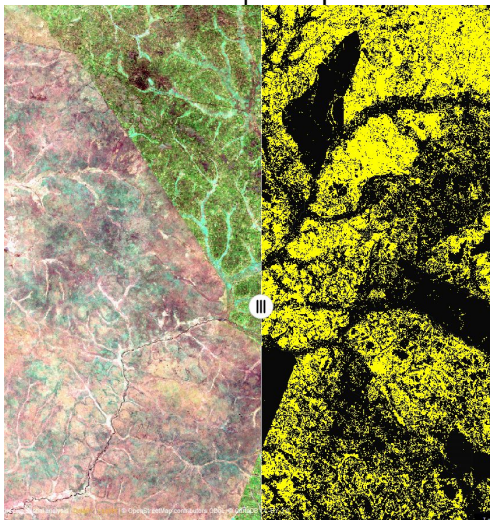
Crop Mask (upcoming)

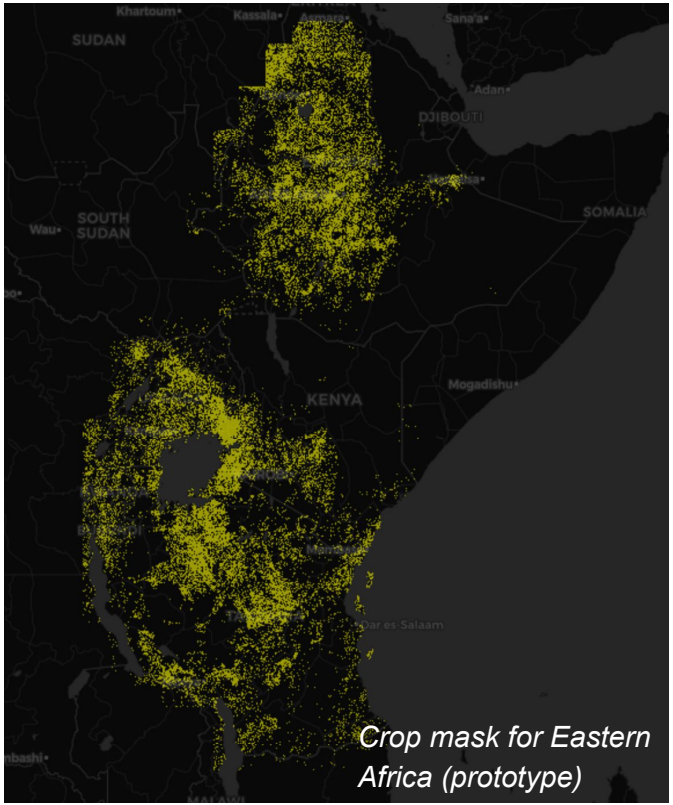
Digital Earth AFRICA

Cropland extent map.

- Created using a supervised ML workflow
- Co-produced with experts across Africa

GeoMAD | Cropmask





Crop mask for Eastern Africa (prototype)

Platforms/Resources

Maps Portal	http://maps.digitalearth.africa/
Jupyter Lab Sandbox	https://sandbox.digitalearth.africa/
Jupyter-Notebook Repository:	https://github.com/digitalearthafrica/deafrica-sandbox-notebooks
Jupyter-Notebook Docs:	https://docs.digitalearthafrica.org
Training Modules:	https://training.digitalearthafrica.org/en/latest/index.html
OWS Map Services	https://ows.digitalearth.africa/
Help Desk	helpdesk@digitalearthafrica.org

Digital Earth
AFRICA

ML4EO Use Case: Building a Continental Cropland Mask for Africa

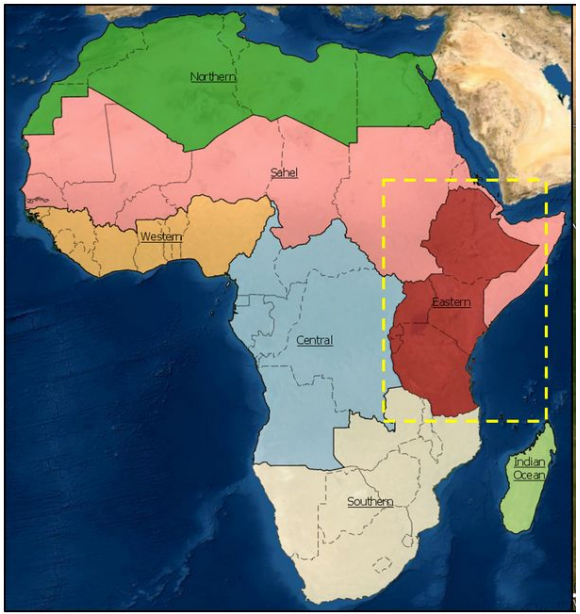
Chad Burton | chad.burton@ga.gov.au

Background

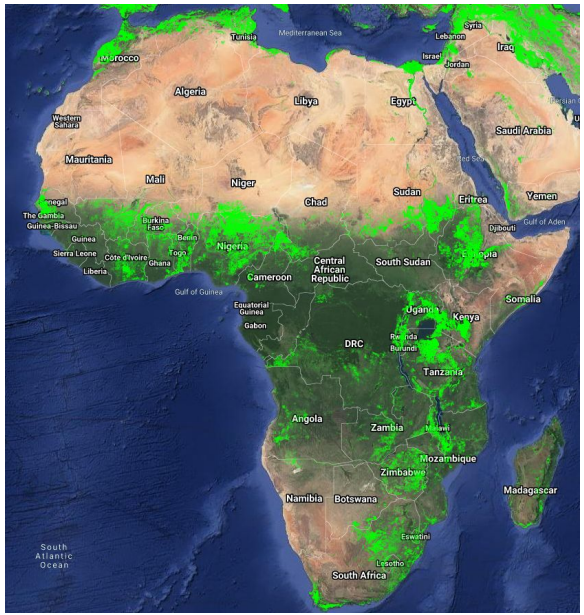
- Technical Advisory Committee mandate to produce a crop mask for Africa. Fulfill a need for fundamental agriculture datasets to support crop monitoring at scale.
  - Requested a crop mask that is:
    - Continental scale at 10m resolution
    - Updated annually
    - Initially for 2019

- Guided by *co-production* philosophy. Established a ‘crop mask task-team’ consisting of members from six regional institutions:
  - Task team members help with:
    - Defining product specs through interviews with stakeholders
    - Developing conceptual framework for algorithm
    - Reference data collection
    - Running classifications (not yet)
    - Developing use cases around the crop mask (just started this)

Reference Data - Sample Design

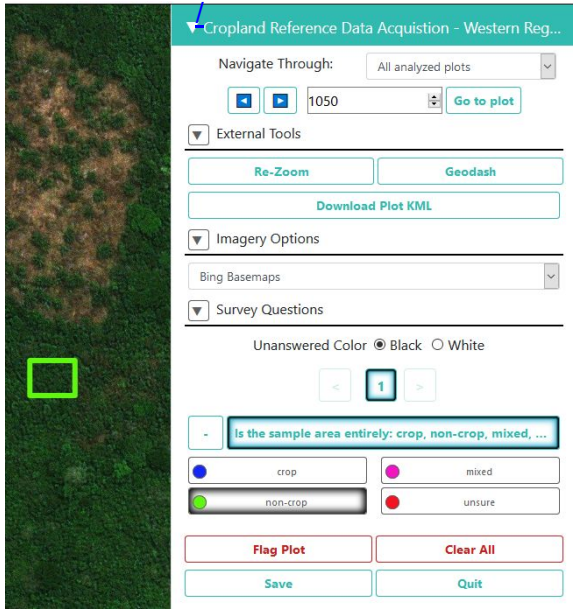


Stratify by simple AEZs,both for reference data collection and model building. Improves classifications, and allowed for working on Eastern crop mask while data was collected for other regions.



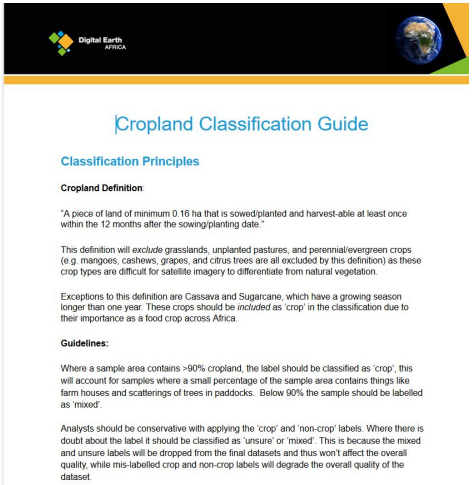
Within each AEZ, samples stratified by a pre-existing, crop mask (GFSAD2015). 1000 samples randomly distributed in non-crop class, 1000 in crop class (half those numbers for smaller regions)

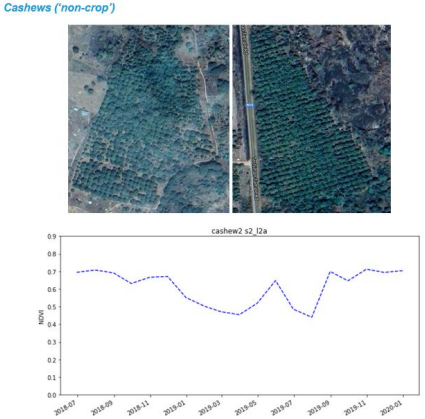
https://collect.earth



Samples uploaded to Collect Earth. Image interpretation used to classify each 40 x 40m sample/polygon. Labelled as either crop, non-crop, mixed, or unsure.

Reference Data - quality control measures





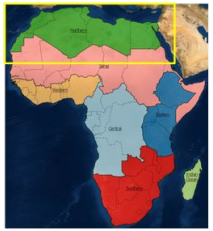
To assist analysts with classifications, and ensure data quality standards, analysts had access too:

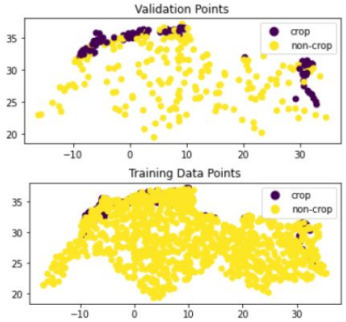
Classification guidelines:

- Instructions
- Cropping definitions, with examples
- Image library of examples

QA was controlled/measured in three ways:

1. Analysts underwent training. Classified ~100 'known' samples.
2. ~100 'known' labels were randomly seeded throughout the 2000 reference samples.
3. 100 samples randomly selected per AEZ and sent to Radiant Earth for labelling with their access to high-res data. Provides external accuracy assessment.

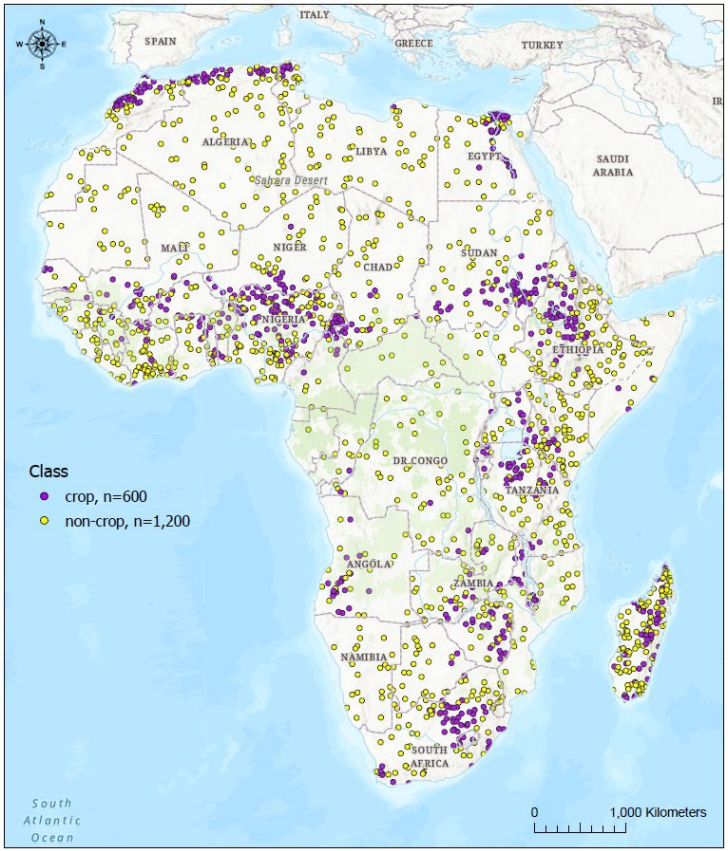




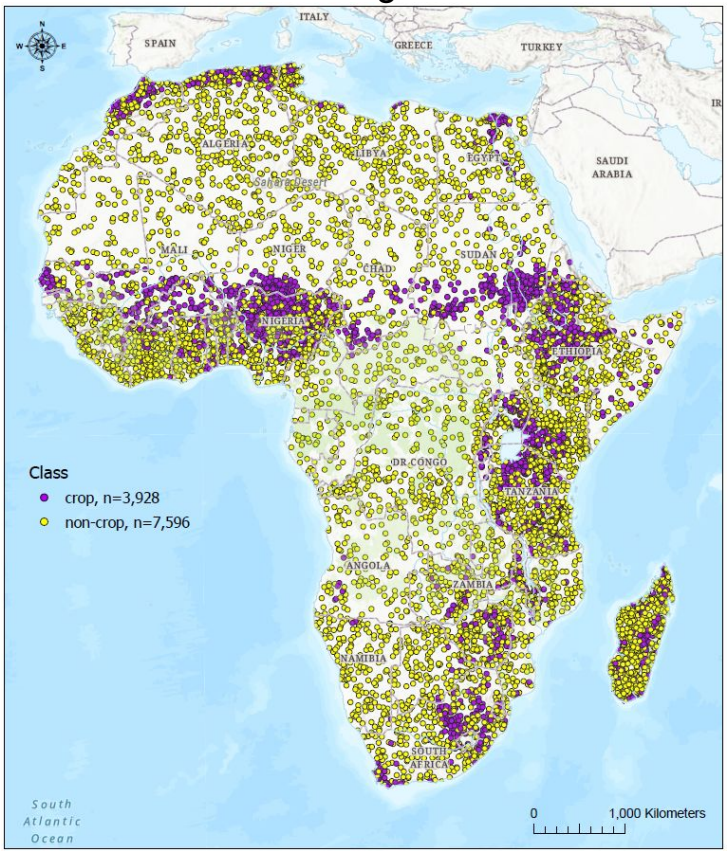
Prediction Actual	Non-crop	Crop	Total	Producer's
Non-crop	49.00	2.00	51	96.08
Crop	0.00	43.00	43	100
Total	49.00	45.00	94	--
User's	100.00	95.56	--	97.87
F-score	0.98	0.98	--	--

Reference Data - results

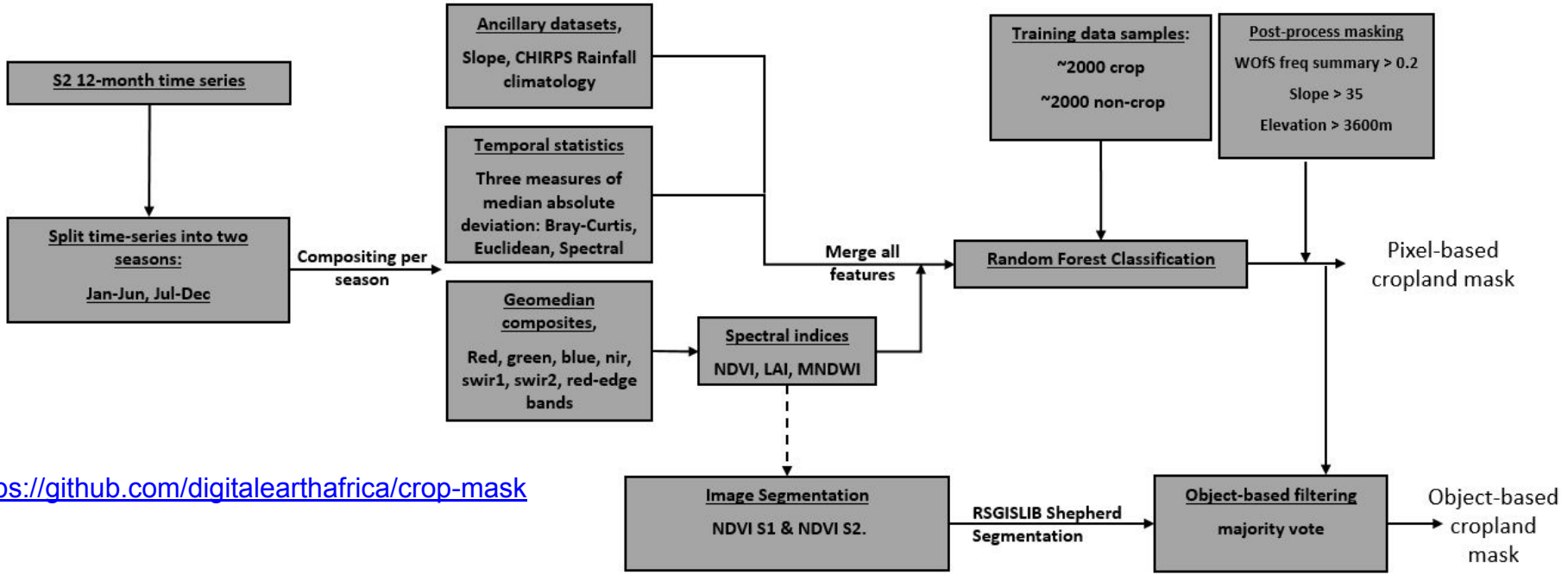
Validation Data



Training Data



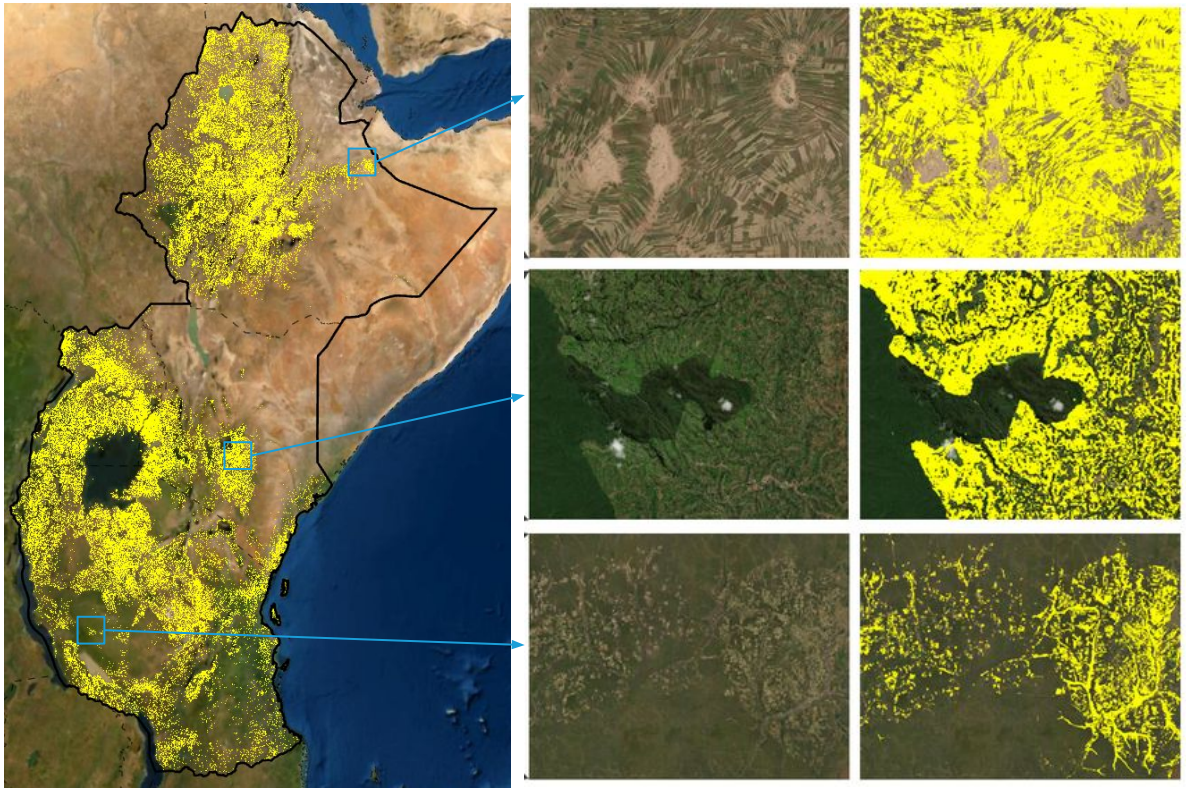
Algorithm



https://github.com/digitalearthafrica/crop-mask

1. Split annual Sentinel-2 time-series into two six-month seasons
2. Generate composite statistics, band indices, add ancillary datasets
3. Random Forest classification produces pixel-based mask
4. Image segmentation and majority vote produces object-based mask

Results (prototype)



Preliminary Accuracy

Overall: 89 %

Crop:

- Commission: 87 %
- Omission: 79 %
- F1: 0.83

Non-crop:

- Commission: 90 %
- Omission: 94 %
- F1: 0.92

AEZ	Country	Total cropped area (Mha)		
		GFSAD (30m)	GLC30 (30m)	DEA (20m)
Eastern	Ethiopia	25.70	21.72	22.54
	Tanzania	22.57	18.25	20.49
	Kenya	9.23	8.38	7.40
	Uganda	7.19	6.24	9.14
	Rwanda	1.42	1.20	1.10
	Burundi	0.85	0.67	1.20
Total		66.96	56.46	61.87

What have I learned?

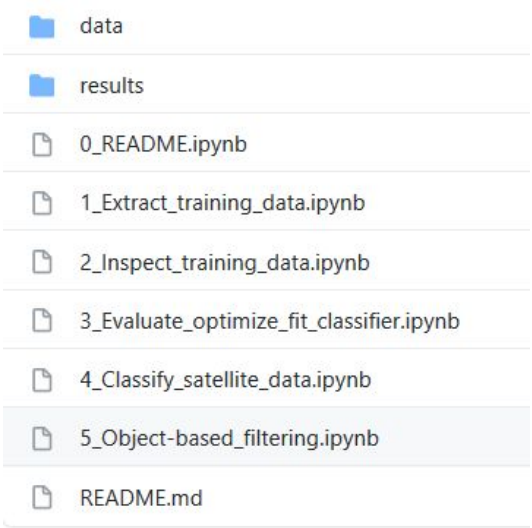
- To worry *less* about the engineering (hyperparameter optimization, feature optimization).
- To worry *more* about the training data.
- To worry *more* about computational limitations. Scaling is very hard when you have trillions of observations to summarise.
- To think carefully in the early stages about the biophysics of the land classes when deciding what features to include in a model.
- It’s easier to get stakeholders excited for the end product when they’re substantively involved in the process.

Digital Earth
AFRICA

Showcase:
Mapping Cropland Extents with Random Forests

Chad Burton | chad.burton@ga.gov.au

Scalable Supervised ML on the ODC Notebook Series



A *series* of Jupyter Notebooks (and accompanying python modules) for guiding a user through supervised ML on the ODC. These notebooks go into much more depth than the urban extent ML notebook we just explored. Each critical step in ML workflow is explored in its own notebook.

https://github.com/digitalearthafrica/deafrica-sandbox-notebooks/tree/master/Real_world_examples/Scalable_machine_learning

Digital Earth
AFRICA

Exercise:
Mapping Urban Extent with Decision Trees

Chad Burton | chad.burton@ga.gov.au

ML in the ODC notebook example

Using the DE Africa Jupyter Lab Sandbox, we will together run the ML in the ODC example notebook. This notebook maps the urban extent of Kampala, Uganda