

Spot the difference

The two images captured below are of an area just outside of Proserpine in Queensland, a rural area whose agriculture is dominated by sugarcane. The images were taken approximately 10 days apart by the Sentinel 2A satellite, which is a part of the European Space Administration's Copernicus program, and provide imagery with a 10m per pixel resolution.



The above images have clear differences, representing the harvest of sugarcane. By analyzing these types of images, it should be possible to determine which fields have been harvested on a week by week basis. Similarly, the different shades of green should provide some level of evidence as to how mature the crop is, as well as an indicator of the quality of the crop.

The Task

The task is to build an interactive App (e.g. browser based) which utilizes satellite data to provide insight into the production of sugarcane over time. You may use any publicly available dataset that you see fit, but it must incorporate the analysis of satellite imagery.

The task is deliberately open ended to provide you with the opportunity to think about the different ways that this kind of data may help different kinds of people.

We encourage you to participate in this task in teams of up to 5 people, although you may also participate as an individual.

Example use cases

- As a Farmer I want to know how to improve the yield of my crop, including:
 - How I can improve my soil (e.g. pH, nitrogen / phosphorus)
 - How I can better time my planting / harvesting
- As a Sugar Mill, I want better forecasts for when sugarcane is likely to arrive and how much
- As a farmer, I want better visibility into how my crop compares to other farmers in the area as it grows
- As a farmer, I want to know what my yield is likely to be at the end of the season
- As a banker, I want to be able to offer loans to farmers on the basis of the value of their crop when harvested
- As a government, I want to identify regions and people most effected by extreme weather events
- As a government, I want to identify new plots of land suitable for Sugarcane, so as to support a new BioFuels industry
- As a Fertilizer vendor, I want to know which farms I should be targeting to sell my products to maximize the benefit to the farmer

Assessment

Your final submission for the Data2App category will be a 5-minute video presentation whereby you pitch us on the value of your App to your target customer. This pitch should be in a similar format to that to Shark Tank (where hopeful entrepreneurs pitch their business for investment).

From the video submissions we will select a shortlist of 5 teams who will present their App at the award ceremony in November.

Project phases

We understand that this project may seem daunting, so we suggest that you approach the analysis of the satellite data in phases. You are of course free to jump straight in if you feel confident.

Phase 1 (15th of August)

We have pre-processed a time series of tiles, along with masks for a small region of Proserpine in Queensland.

- Tiles are 512 pixels in height and width, and therefore corresponds to an area of approximately 25Ha.
- Tiles are time stamped by the date of capture (approximately each 14 days)
- Tiles have one image per capture date, per sensing band
- Tiles have a json file providing meta data about the conditions of the capture, and its location in lat/long.
- Tiles also have a "mask" which corresponds to the Sugarcane regions, so you can more easily identify whether a pixel is likely to be growing sugarcane versus another crop

We suggest that you use this data to try to detect when crops have been harvested and build up a time series of when each field is harvested and how big it is. The average yield of Sugarcane in Queensland can be found online and thus you can identify how much sugarcane is harvested per week / month.

Phase 2 (29th of August)

Increase the size of the area you analyze to be the greater Proserpine area.

- Tiles and masks will be provided for you in Week 2 of a much larger area
- Code for generating tile masks from GeoJSON will be provided
- The larger area will mean you will need to be thinking about how you will handle clouds
- You may wish to jump ahead and develop a pipeline to pull data from SARA on your own
- You may wish to jump ahead and develop a pipeline to build masks (e.g. sugarcane regions) on your own

Phase 3 (12th of September)

Increase the size of your target area to all of Australia (or Queensland / Northern New South Wales)

- Code to generate tiles and masks will be provided for you in Week 4
- You are on your own in terms of data access & tooling

Additional data sets available

There is a wealth of additional data sources available, including:

- High resolution maps of soil quality (including: water carrying capacity, pH, Nitrogen, Phosphorus and Carbon) ⁱ
- High resolution maps (ESRI format) of regions which are zoned for Sugarcane production
- Weekly local sugarcane production reports (Wilmar mills) ⁱⁱ
- Weekly sugarcane production reports 2018 (National) ⁱⁱⁱ
- Weekly sugarcane production reports 2019 (National) ^{iv}

Sentinel 2A satellite band information

The Sentinel 2A satellite is a remote sensing platform and is capable of image capture at wavelengths outside of what is human visible. The platform therefore generates imagery across different bands. A table of these bands and their usage is defined below:

Sentinel-2 Bands	Central Wavelength (µm)	Resolution (m)
Band 1 - Coastal aerosol	0.443	60
Band 2 - Blue	0.490	10
Band 3 - Green	0.560	10
Band 4 - Red	0.665	10
Band 5 - Vegetation Red Edge	0.705	20
Band 6 - Vegetation Red Edge	0.740	20
Band 7 - Vegetation Red Edge	0.783	20
Band 8 - NIR	0.842	10
Band 8A - Vegetation Red Edge	0.865	20
Band 9 - Water vapour	0.945	60
Band 10 - SWIR - Cirrus	1.375	60
Band 11 - SWIR	1.610	20
Band 12 - SWIR	2.190	20

Clouds

One of the largest challenges in satellite imagery analysis is cloud cover, and the shadows that they task. This is an ongoing area of research (e.g. enhancing regions of cloud cover with data captured on previous days). Bands 9 (clouds) and 10 (Cirrus clouds, thin and wispy clouds) captured by Sentinel 2A are dedicated to sensing clouds, providing you with a reference source for whether a particular location is cloud effected.

Sugar in Australia

Sugar is Australia's second largest agricultural export (after wheat) totaling about \$2B per year and employing 48,000 people. Over 95% of sugarcane production is localized to Queensland (the remainder is in northern New South Wales).

Sugarcane also has an increasing role to play in the migration away from fossil fuels. Ethanol is a common byproduct from sugar production and can be used as a biofuel alternative to petrol. In Brazil (the world's largest producer of sugar), petroleum must contain at least 22% ethanol, sourced exclusively from Brazil's sugarcane crop.

Getting started

Demo script

The demo script contains a number of simple methods to help get you started with the analysis of imagery. The script shows how you can find the color of a pixel (in Red, Green, Blue, Alpha format) as well as how to track whether a certain pixel is set in a mask. The demo script prints an ascii version of the mask tile, with X's for regions in the mask, and nothing for others.

The demo script should be run using Python3. Dependencies can be installed by executing:

```
pip install -r requirements.txt
```

The demo script can be executed with:

```
python demo.py
```

Files & folders

demo.py	An example Python script to show how to load images and get colour information for a specific pixel
guide.pdf	This document
requirements.py	A list of the python3 requirements
data	
./sentinel-2a-tile-7680x-10240y	Data from the sentinel 2a satellite (you may wish to also add imagery from Sentinel 2b at a later date). The 7680 / 10240 indicate the position of this tile in a much larger image. All the tiles in the larger image will be released in Phase 2.
./geometry/file-x7680-y10240.geojson	A geometry file containing the geo-json regions which relate to areas where sugarcane is grown. This file was used to create the "sugarcane-region-mask.png" file
./masks/sugarcane-region-mask.png	A rasterized image which calculates which pixels are in a sugarcane area (black) and which are not (transparent)
./metadata/{yyyy-mm-dd}.json	Additional meta data relating to the capture of the larger image which this tile comes from.
./timeseries/*.png	An image for each band, for each day that an image could be captured by the satellite. The TCI band is the most interesting, as it represents a colorized version of the images which is intuitively easier to work with.

Helpful Resources

NationalMap.gov.au

<https://nationalmap.gov.au/>

This is a brilliant resource for viewing a diverse group of datasets, many of which will be useful for your project. Most notable is the ability for you to add your own “layers” to the map by uploading your own GeoJSON files. This is very useful to help verify that files you generate map back to the positions that they are supposed to.

Copernicus SARA

<https://copernicus.nci.org.au/sara.client/#/explore?collection=S1>

The Sentinel Australasia Regional Access portal provides access to the raw Sentinel satellite data. As your techniques progress, you will want to process the raw satellite data on your own. Scripts will be released to assist you with this, but you will need to register (free) with SARA to get API access. It typically only takes an hour or two for you to be provided with access. Please note that the raw data is in JPG2 format, which will require downstream processing.

Australian Collaborative Land Evaluation Program

<http://www.clw.csiro.au/aclep/>

ACLEP provides a high resolution grid of Australia, where each position on the grid contains a number of metrics relating to soil quality. These metrics are provided for:

- Water Capacity of soil
- Carbon content
- Nitrogen content
- Phosphorous content

These metrics are available for varying depths (e.g. 0-5cm, 5-15cm, 15-30cm, etc). This information could be utilized to understand how different soil quality effects yield, or to identify other areas which might be suitable for growing sugarcane.

Links and references

i <http://www.clw.csiro.au/aclep/index.htm>

ii <https://www.wilmarsugarmills.com.au/media-centre>

iii <https://asmc.com.au/industry-overview/2018-weekly-crushing-statistics/>

iv <https://asmc.com.au/industry-overview/statistics/weekly-crush-statistics-2019/>