

Identifying Wildfires: Geospatial Data and Machine Learning

Allison Lee

With gratitude to instructors, coaches, peers at Flatiron, and Forrest Wrenn for guidance

THE CONTEXT

Pollutants from the bush fires in Australia will circle the

Australian fires have incinerated the habitats of

ge for irreversible forest losses in Australia

n | AP

Fires exceeded human and computer predictions

st decade on record in Australia.

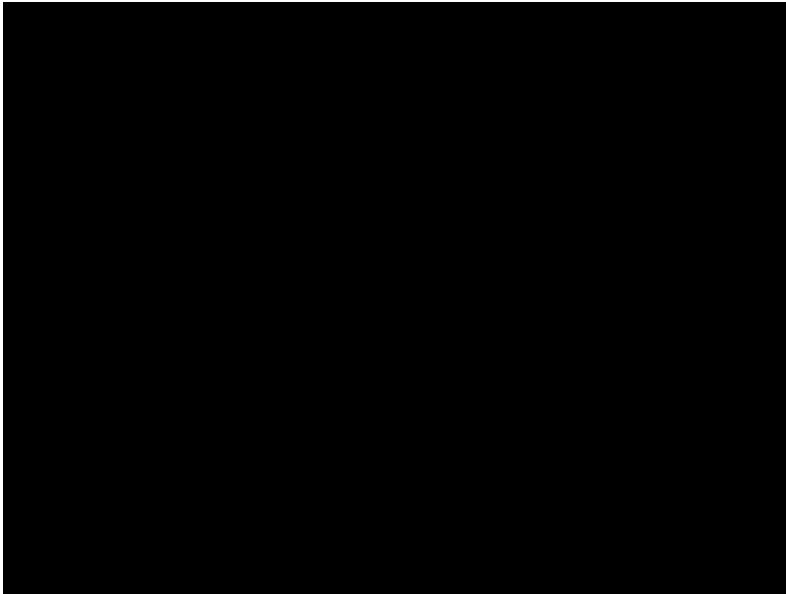
Why These Australia Fires Are Like Nothing We've Seen Before

More than 16 million acres have gone up in flames. And it has happened in populated areas, unlike most of the world's other blazes of this scale.



THE CHALLENGE

Visualization of Fire Spread in Shoalhaven Area



Guardian Graphic | Source: [Nasa Firms](#) VIIRS / Google Earth

Footage of Fires in Victoria



Guardian, Jan 2, 2020

THE OVERARCHING GOAL

Can we use machine learning to predict where a fire will spread?

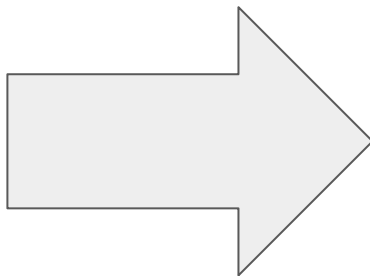
Data input



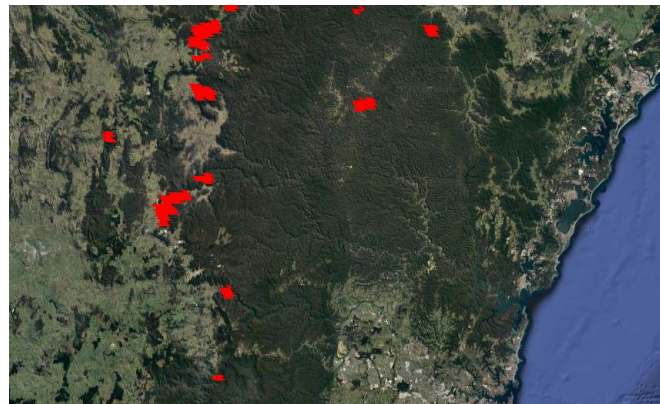
Google Earth Engine

Weather data (NASA, JAXA)
Free, publicly-available
Outputs every 3 hours
Global scale

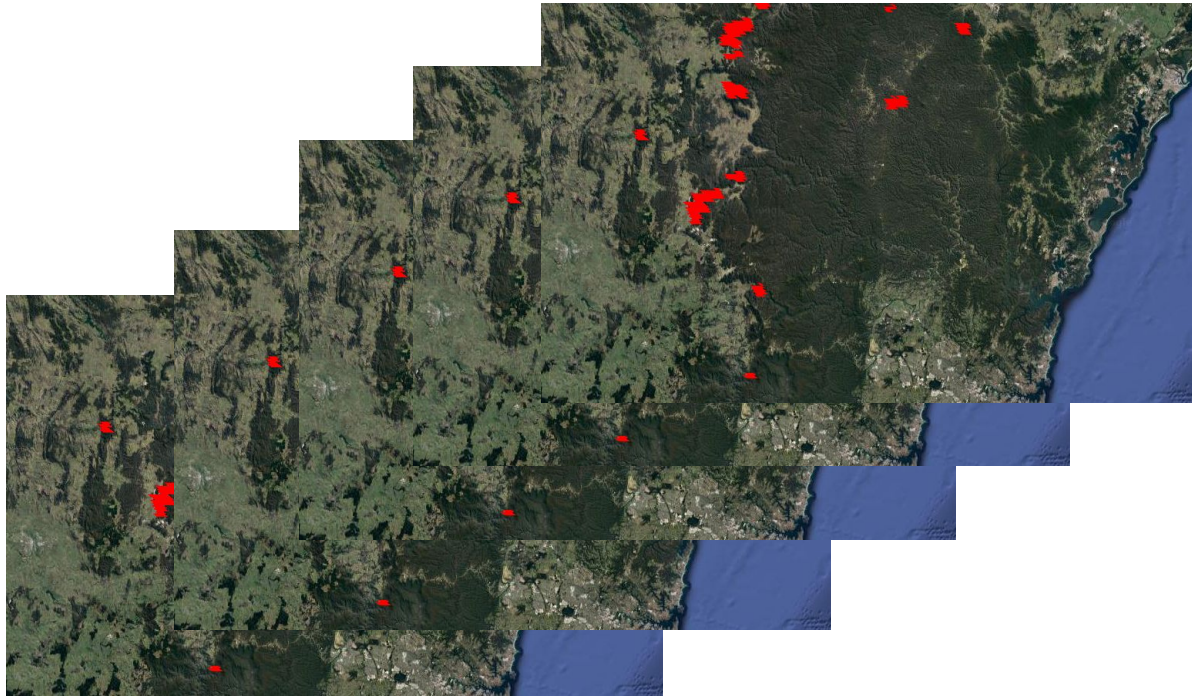
Model



Next-day Predictions of
Areas Likely to Catch Fire



THE OVERARCHING GOAL



THIS PROJECT SCOPE

For one point (day) in time:

**can a model distinguish between areas of active fire,
burned areas, land (other) and water**

based on weather data?

THE DATA COLLECTION

The screenshot displays the Google Earth Engine web interface. At the top, the search bar contains the text 'glda'. Below the search bar, a dropdown menu lists search results under three categories: PLACES, RASTERS, and TABLES. Under PLACES, 'Gldani-Nadzaladevi, Tbilisi, Georgia' is listed. Under RASTERS, 'GLDAS-2.1: Global Land Data Assimilation System' and 'Reprocessed GLDAS-2.0: Global Land Data Assimilation System' are listed. The 'import »' link is visible next to the first raster result. On the right side of the interface, a panel shows the details of the selected dataset, 'ImageCollection MODIS/006/MOD14A1 (1 elem...)'. It displays the 'Number of images' as '1'. The bottom of the interface shows a map of the world with a blue rectangle highlighting a region in Australia. The map controls on the left include a zoom in (+) and zoom out (-) button, and a status bar at the bottom indicates 'aus (1 poly)' and 'Point drawing.'.

Google Earth Engine

Scripts Docs Assets

Filter scr NEW

Owner (1)
users/allisonalee/te...
fire
air
burned
gsmmap_precip
mo...
original

mod_ac

1
2
3
4
5
6
7
8
9

PLACES

Gldani-Nadzaladevi, Tbilisi, Georgia

RASTERS

GLDAS-2.1: Global Land Data Assimilation System import »

Reprocessed GLDAS-2.0: Global Land Data Assimilation System

TABLES

View Data Catalog

Suggest a dataset

is console.

ImageCollection MODIS/006/MOD14A1 (1 elem...) JSON

Number of images JSON

1

var dataset = ee.ImageCollection("MODIS/006/MOD14A1")
.filterDate(start, end)

aus (1 poly)

Point drawing. Exit

Map Satellite

THE DATA SOURCES



MCD64A1.006 MODIS Burned Area Monthly Global 500m

MOD14A1.006: Terra Thermal Anomalies & Fire Daily Global 1km

Source: Land Processes Distributed Active Archive Center (LP-DAAC) within NASA's Earth Observing System Data and Information System

GSMaP Operational: Global Satellite Mapping of Precipitation

Source: Earth Observation Research Center, Japan Aerospace Exploration Agency

Global Land Data Assimilation System (GLDAS 2.1)

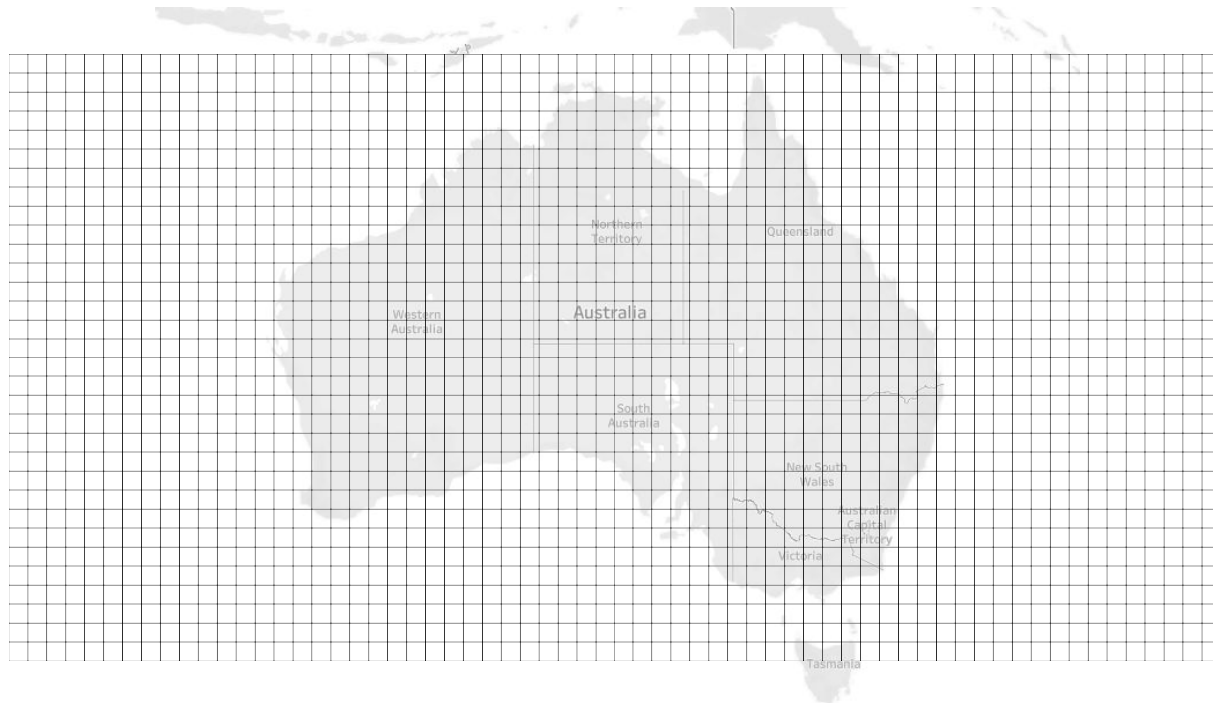
Source: NASA's Goddard Earth Sciences Data and Information Services Center

THE DATA INPUT

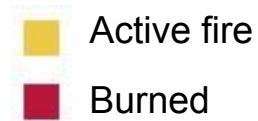
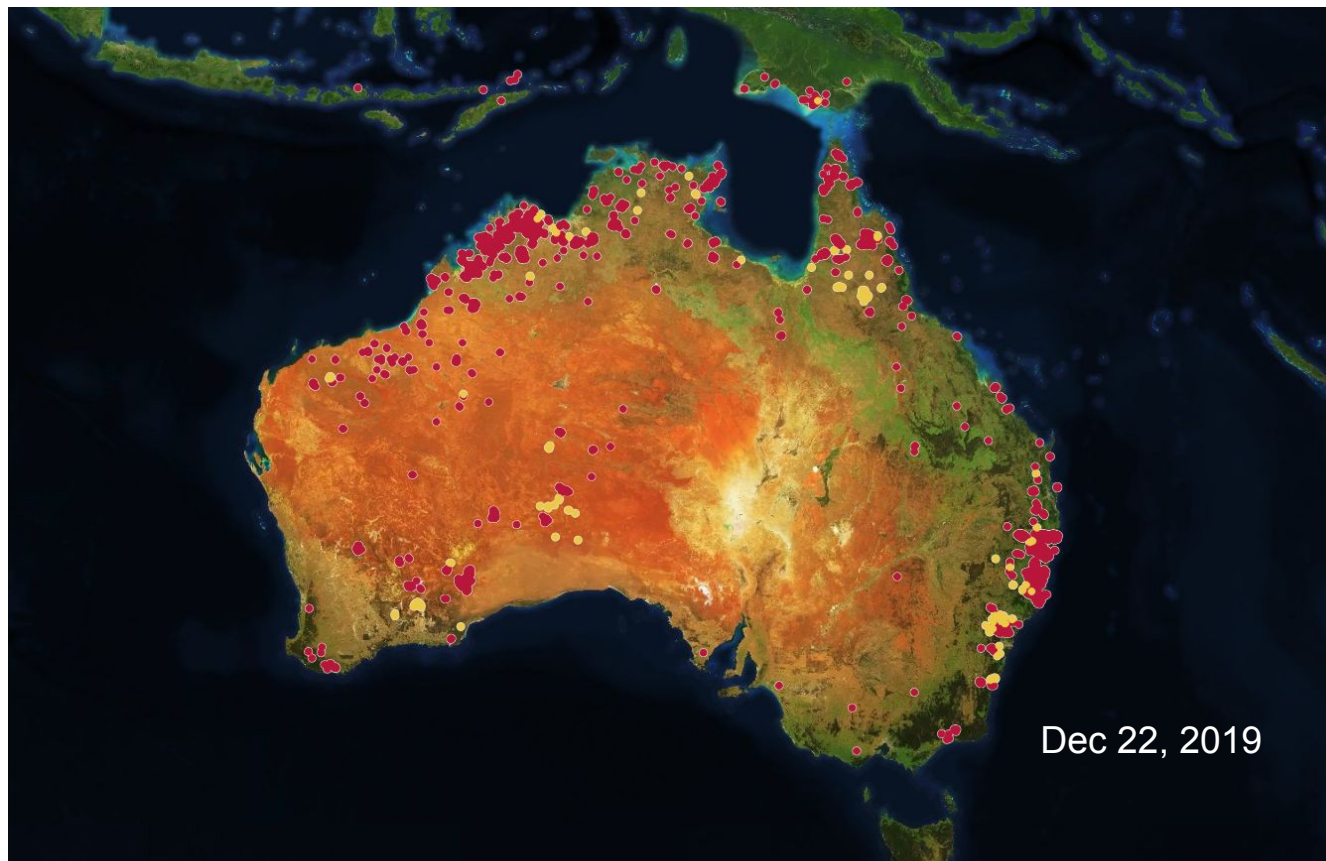
Soil moisture
Soil temperature
Baseflow-groundwater runoff
Root zone soil moisture
Average surface skin temperature
Latent heat net flux
Potential evaporation rate
Snapshot of hourly precipitation rate
adjusted to rain gauge
Total precipitation rate
Pressure
Air temperature

Wind speed
Plant canopy surface water
Albedo
Specific humidity
Downward long-wave radiation flux
Evapotranspiration
Net long-wave radiation flux
Net short wave radiation flux
Snapshot of hourly precipitation rate
Direct evaporation from bare soil

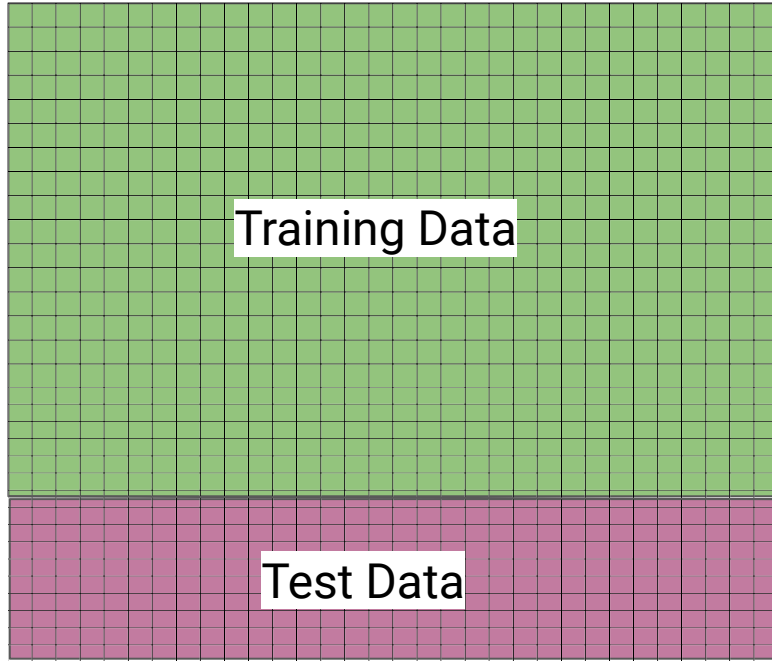
THE DATASET



THE LABELED DATASET



THE MODEL



Predicted Categories

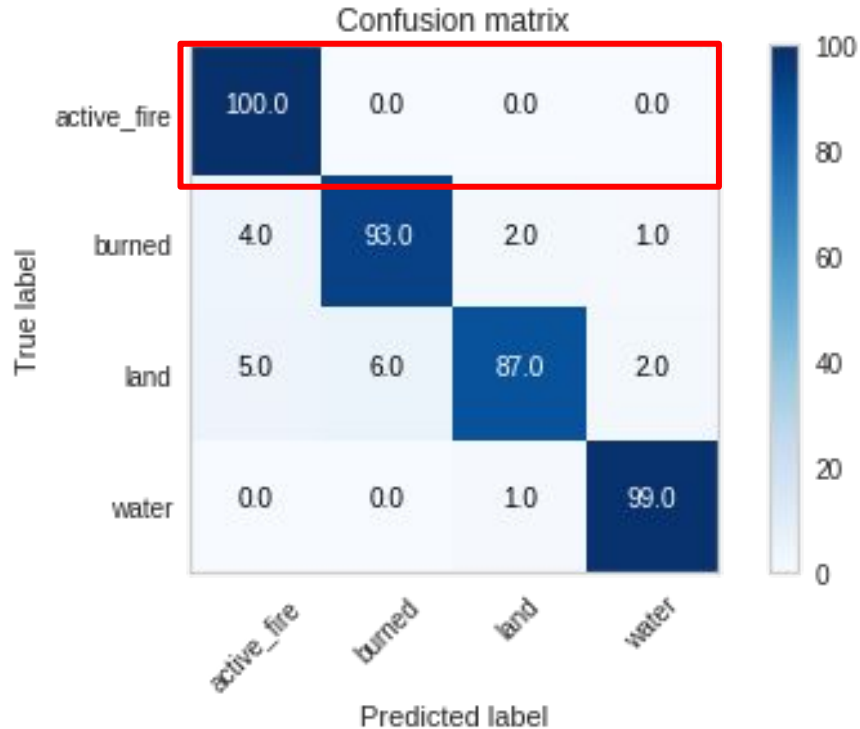
Active Fire

Burned Area

Land (Other)

Water

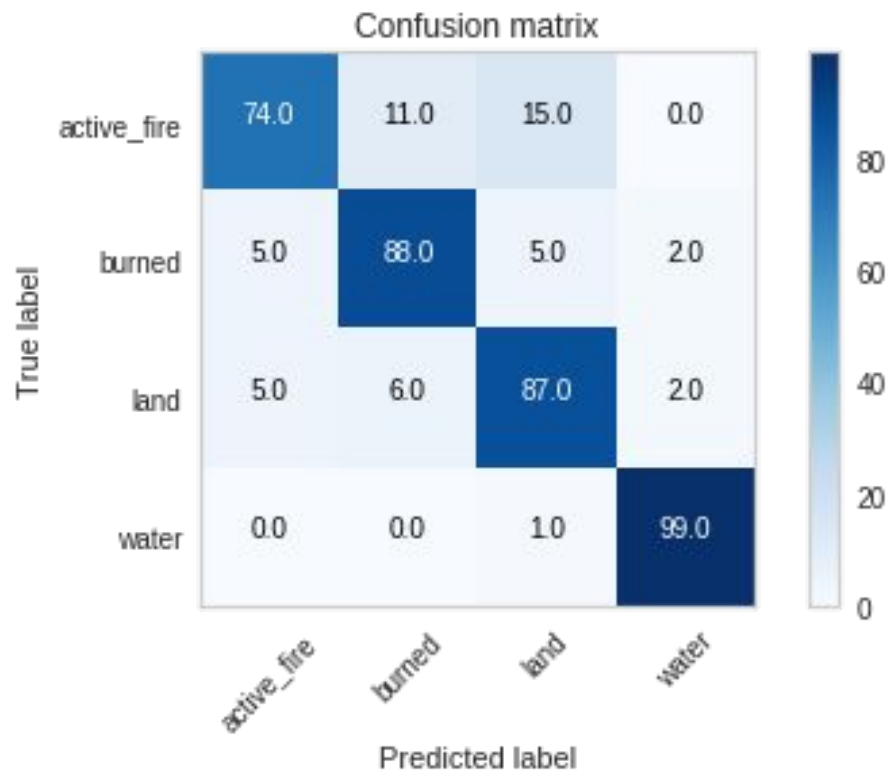
THE RESULTS: TRAINING DATA



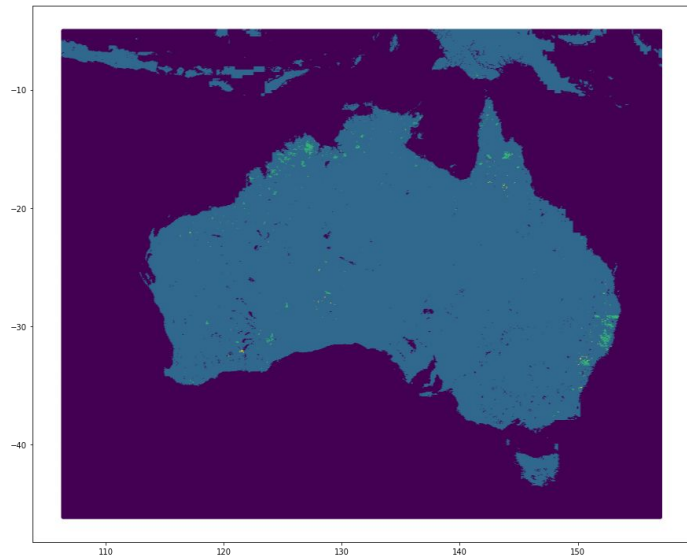
Top Most Important Features

- Specific humidity
- Wind speed
- Latent heat net flux
- Air temperature
- Albedo
- Soil moisture
- Potential evapotranspiration rate
- Evapotranspiration
- Hourly precipitation rate

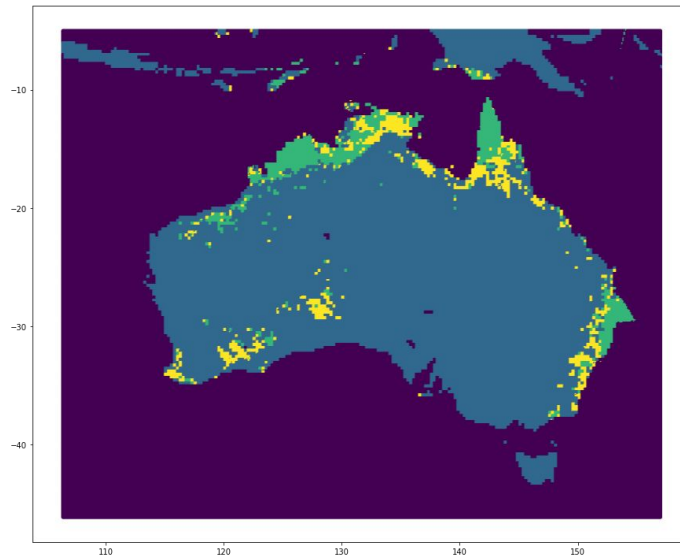
THE RESULTS: TEST DATA



THE RESULTS



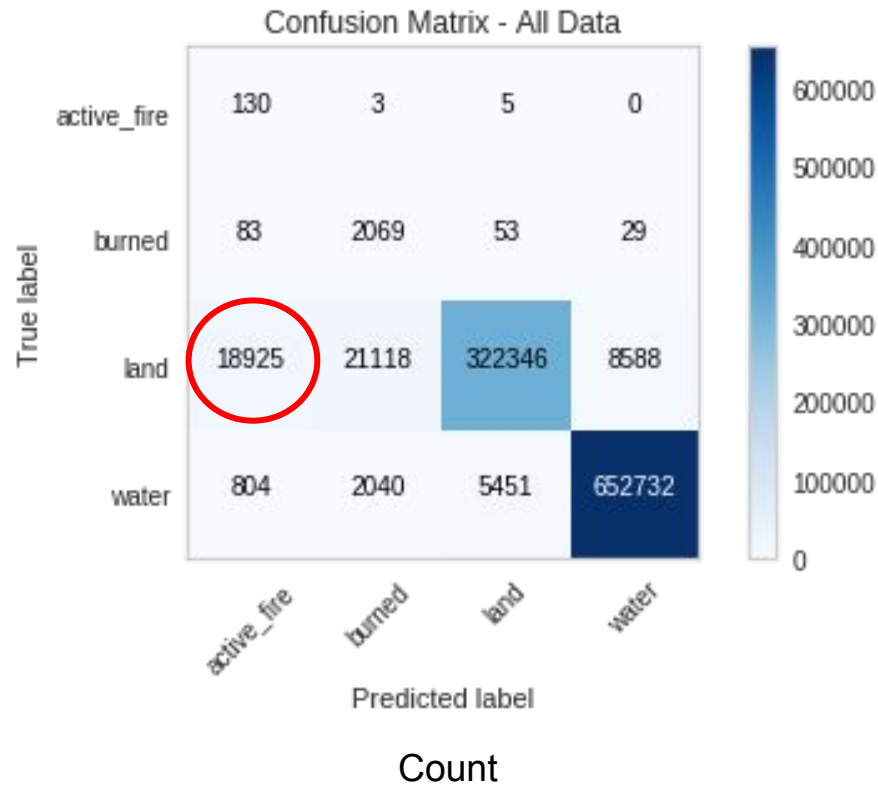
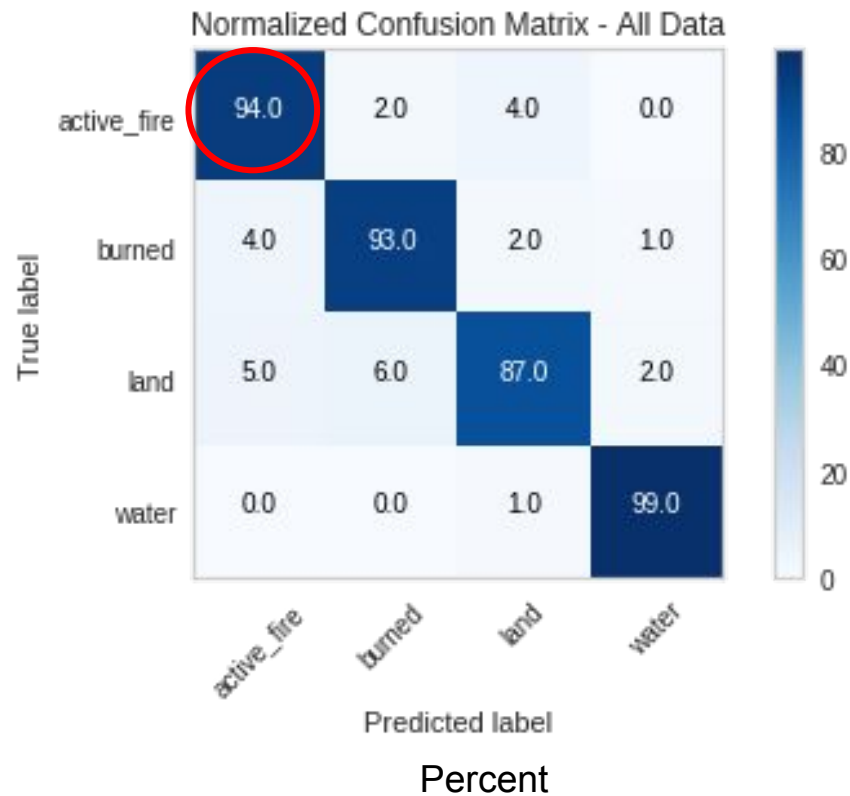
Actual
December 22, 2019



Predicted
December 22, 2019



THE RESULTS

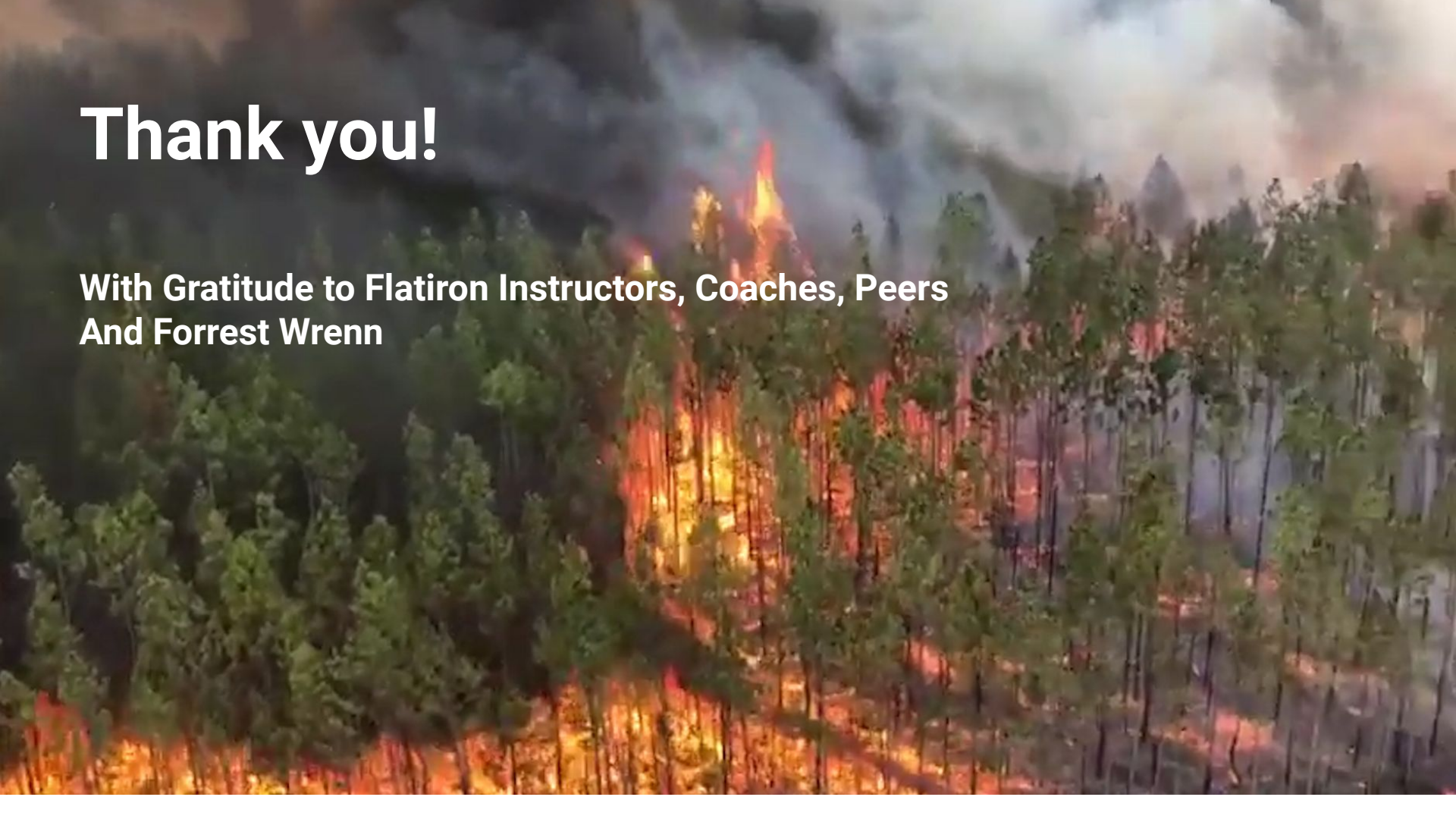


NEXT STEPS

- **Improve the model.**
 - Try different classification algorithms and hypertuning
- **Test the model on a different day.**
- **Explore different data inputs.**
 - Add data on terrain, slope, elevation, geolocation
- **Predict over time.**
 - Add time dimension and time series data
- **Explore over different geographies.**
 - Downscale to smaller geographic area, with higher resolution
 - Explore model application to different regions
- **Automate the model.**
 - Clean code and integration with Google Earth Engine Scripts

Thank you!

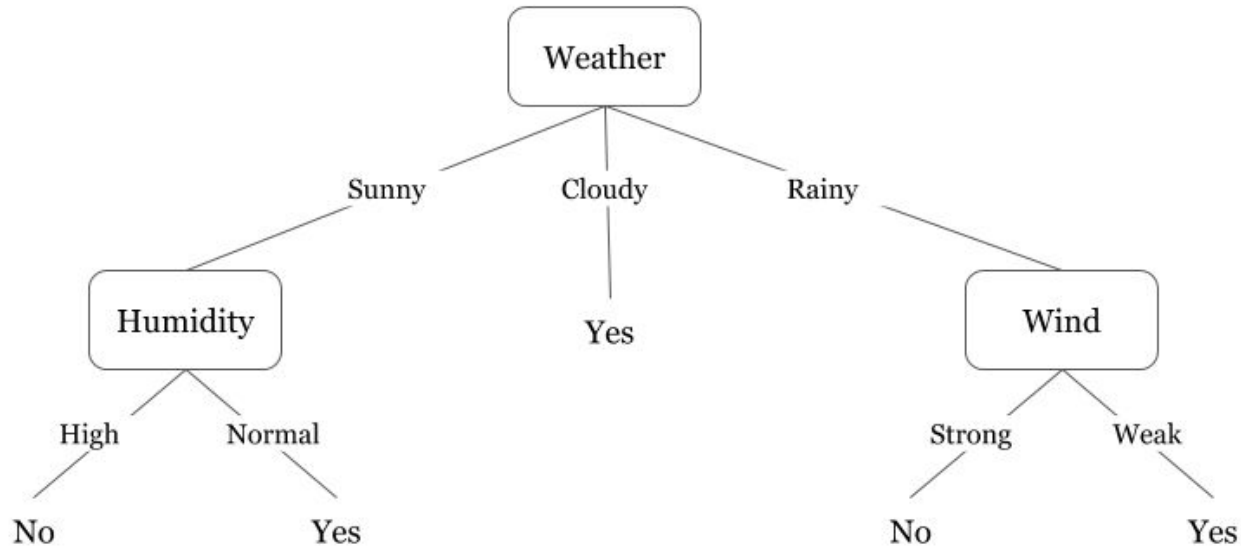
With Gratitude to Flatiron Instructors, Coaches, Peers
And Forrest Wrenn



APPENDIX

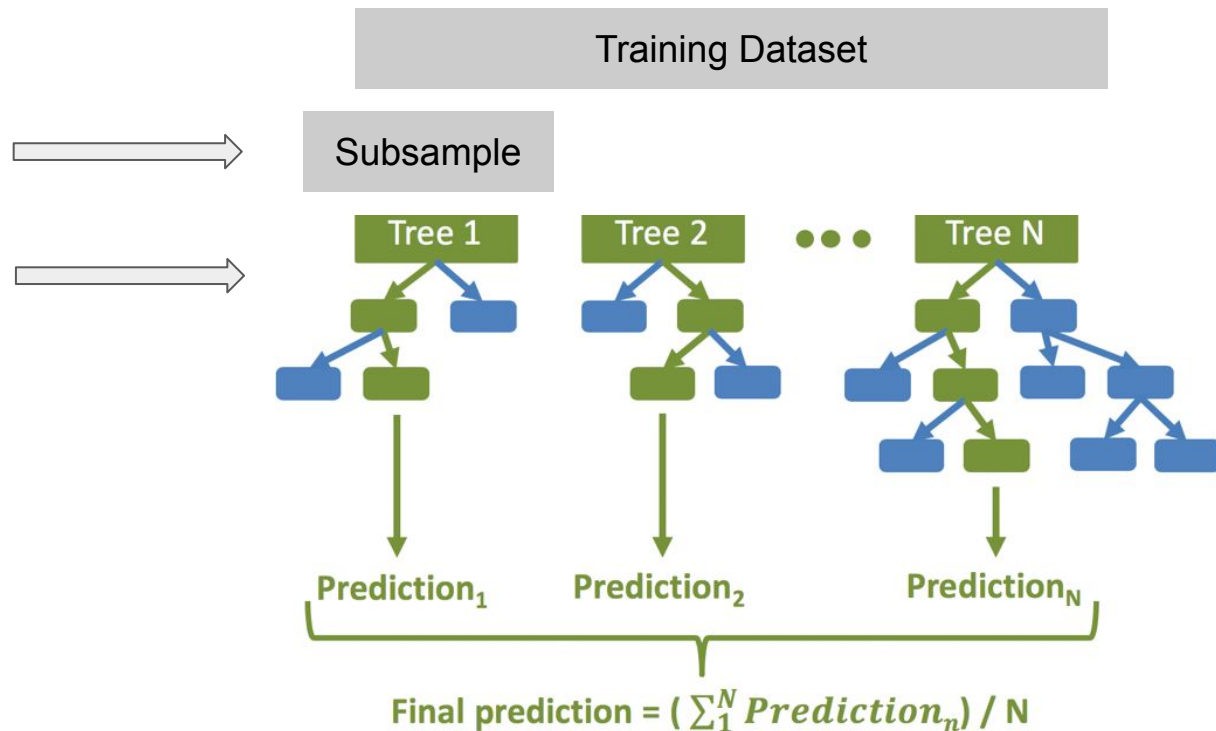
EXPLAINING THE MODEL

Binary Decision Tree Classifier: Should I play outside?



EXPLAINING THE MODEL

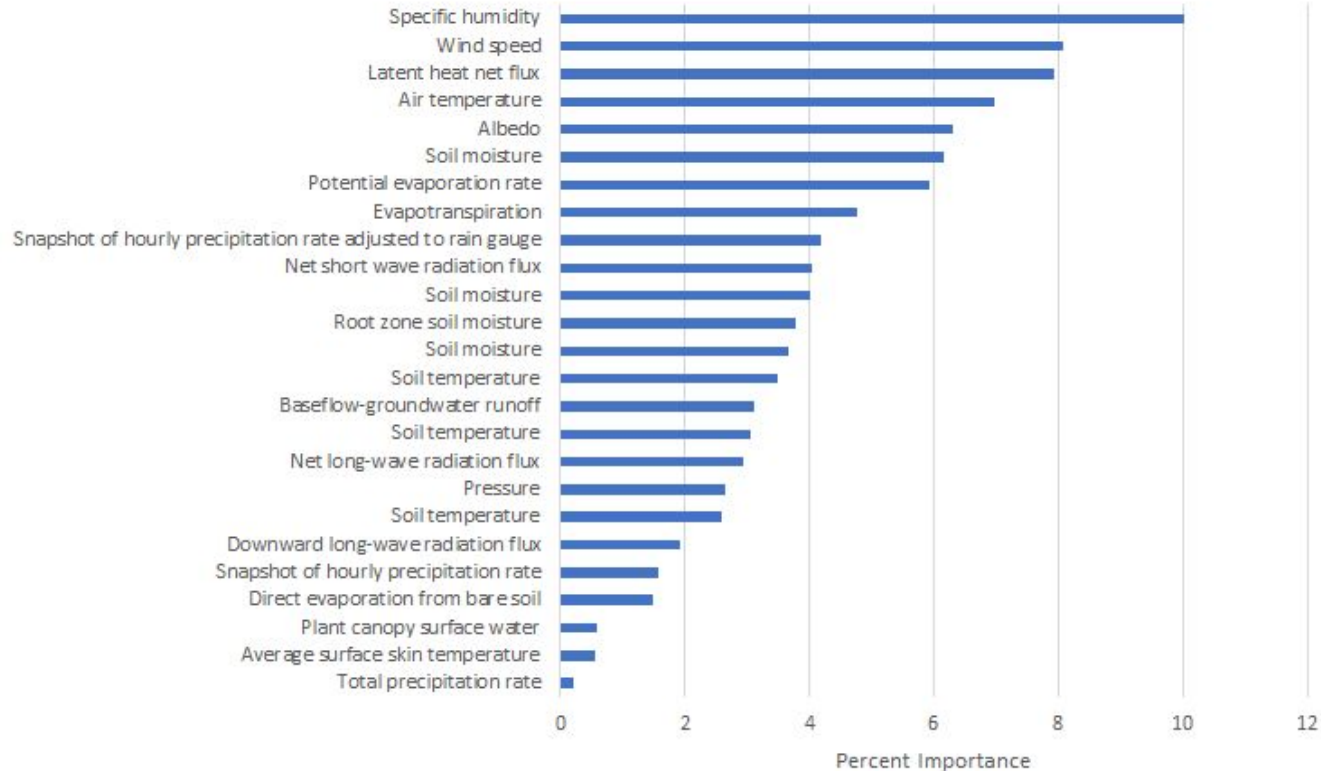
Random Forest Model



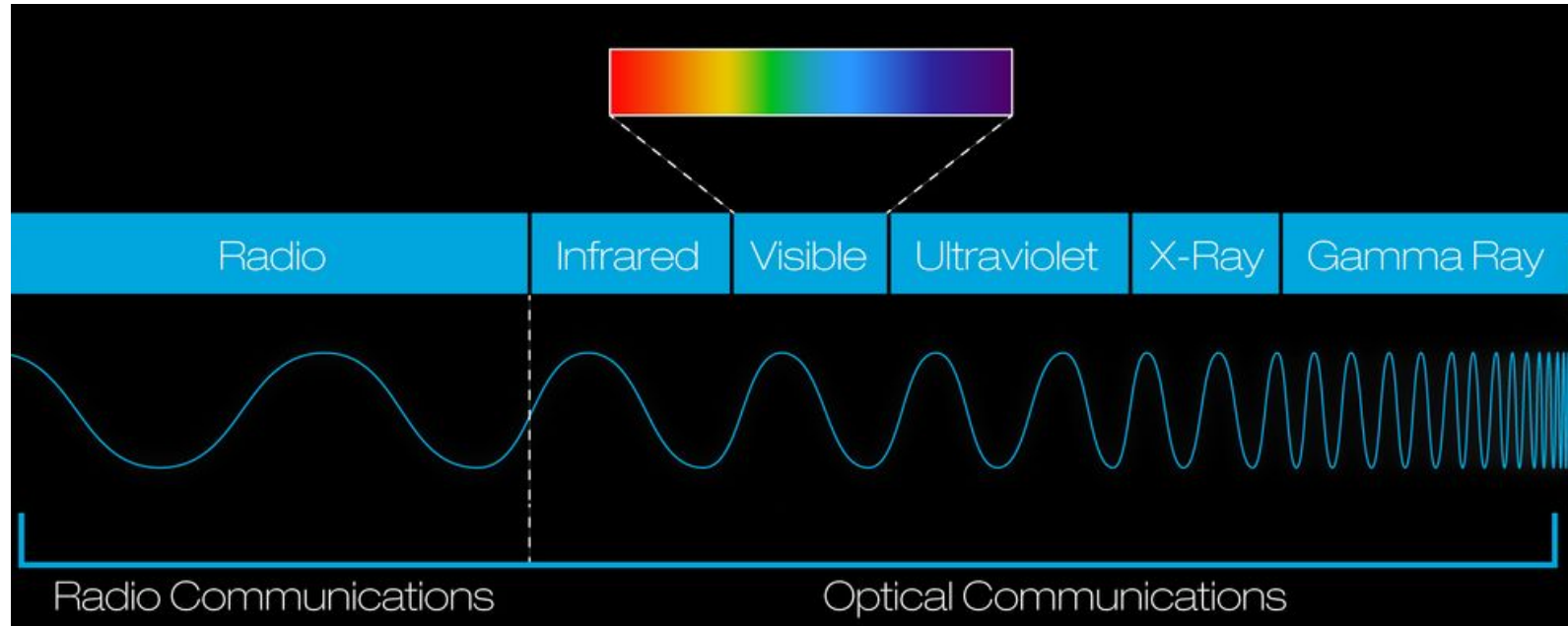
EXPLAINING THE MODEL

```
RandomForestClassifier(bootstrap=True, ccp_alpha=0.0, class_weight=None,  
                        criterion='gini', max_depth=10, max_features=5,  
                        max_leaf_nodes=None, max_samples=None,  
                        min_impurity_decrease=0.0, min_impurity_split=None,  
                        min_samples_leaf=1, min_samples_split=1000,  
                        min_weight_fraction_leaf=0.0, n_estimators=50, n_jobs=-1,  
                        oob_score=False, random_state=None, verbose=0,  
                        warm_start=False)
```

EXPLAINING THE MODEL

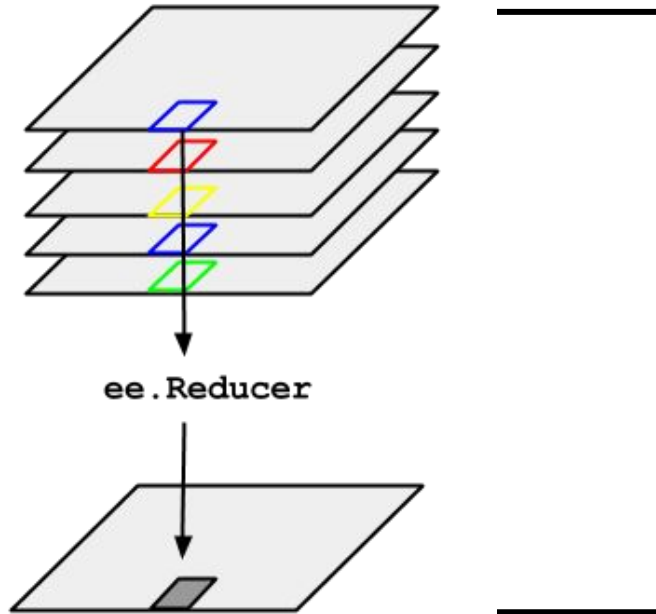


SATELLITE DATA



Source: <https://www.nasa.gov/directorates/heo/scan/spectrum/overview/index.html>

CONSTRUCTING THE DATASET

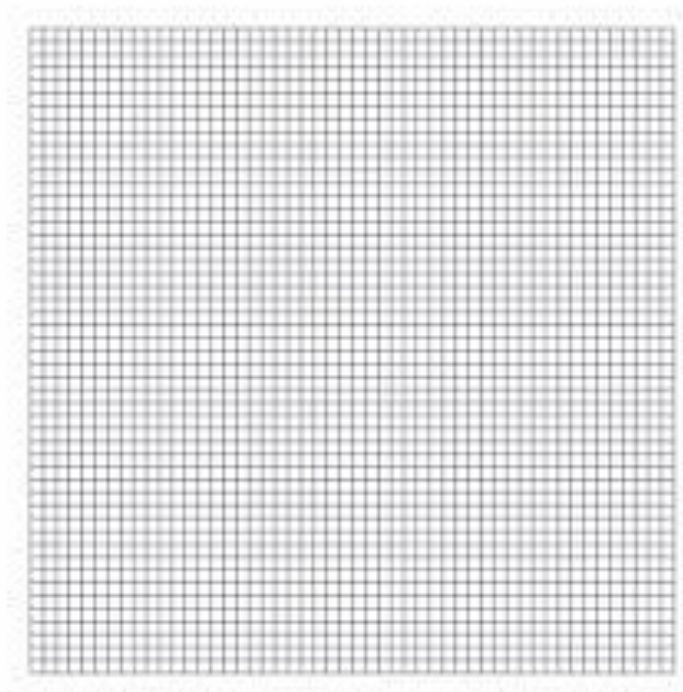
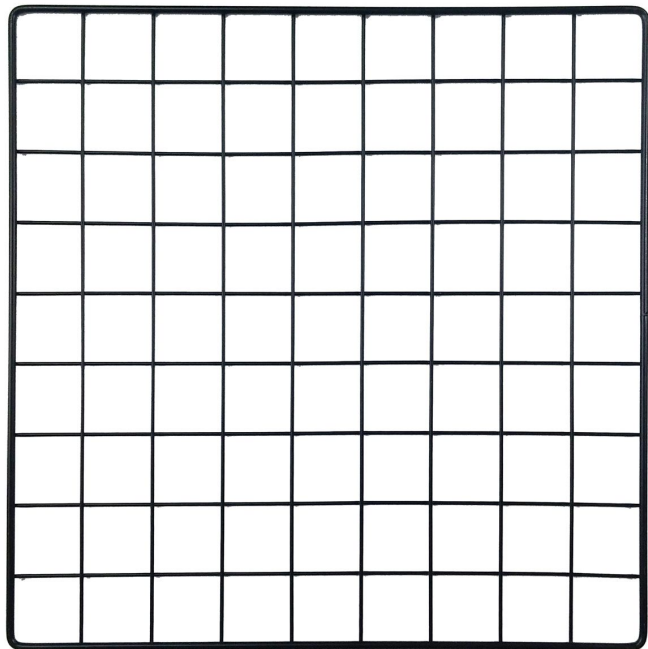


Temporal aggregation

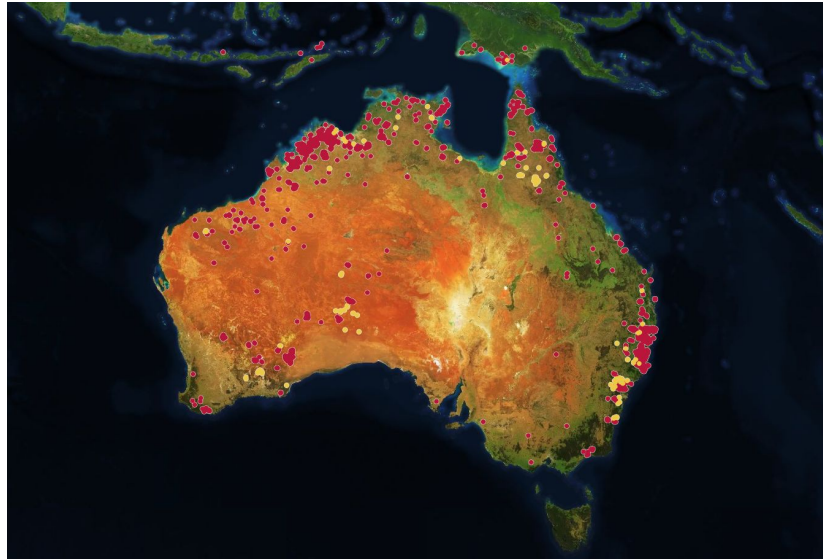
Taking several “images” over a time span and taking the median value per pixel range

Most recent data available as of December 22, 2019

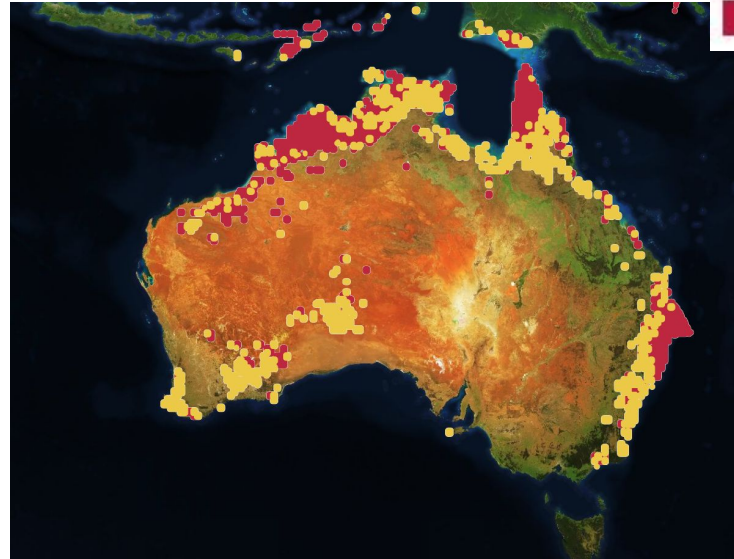
CONSTRUCTING THE DATASET



THE RESULTS



Actual
December 22, 2019



Predicted
December 22, 2019

