Call for Code Spot Challenge for Wildfires:

Fighting Wildfires with Big Data and AI

Hendrik Hamann
Chief Scientist
AI Applications

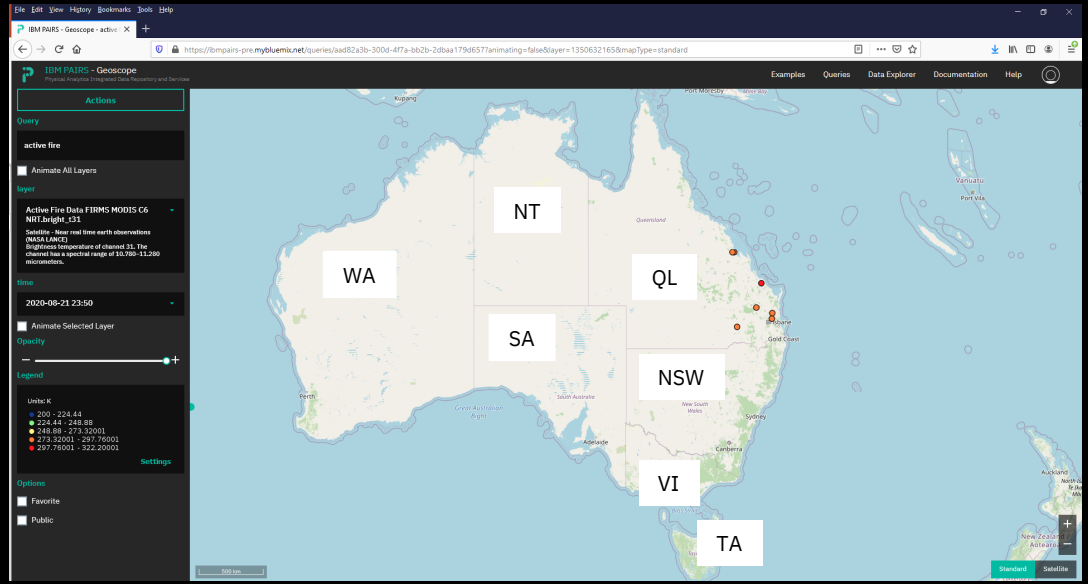
IBM

What will be forecasted/predicted in this contest?

Forecast/predict the daily total estimated fire area for 7 states in Australia?

NSW=New South Wales*
NT=Northern Territory
QL=Queensland
SA=Southern Australia
TA=Tasmania
VI=Victoria
WA=Western Australia

*excluded the capital region



What will be forecasted/predicted?

Date	Region	Total estimated fire area [km²]
2/1/2021	NSW	15
2/2/2021	NSW	327
…	…	…
2/28/2021	NSW	234
…	…	…
2/1/2021	WA	50
2/2/2021	WA	800
…	…	…
2/28/2021	WA	785

7x28 rows = 196 rows

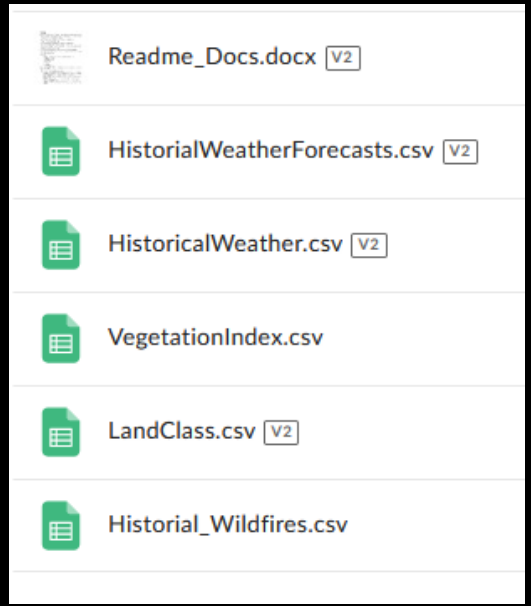
IBM

Why is it important to forecast wildfires?

1. To prepare and respond
2. To understand the root causes
3. To help to mitigate them in the future

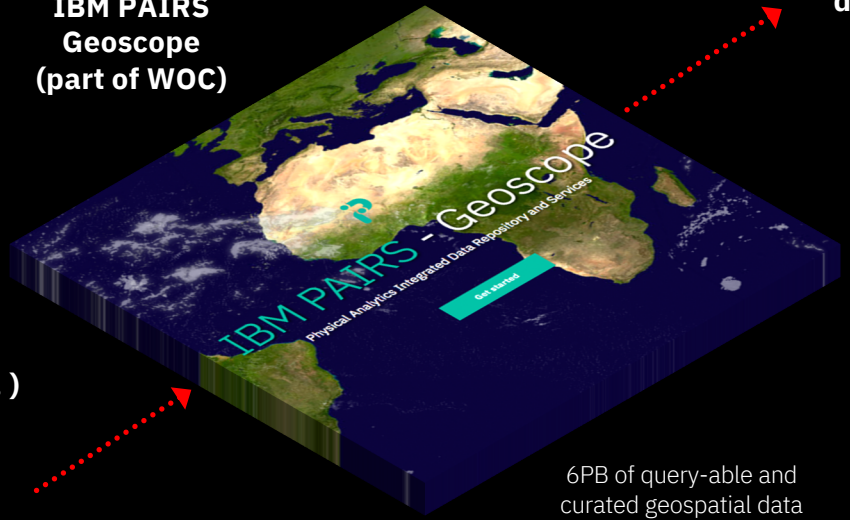
The data for this competition

1. Historical Wildfires
2. Historical Weather
3. Historical Weather Forecast
4. Land Class
5. Normalized Vegetation Index



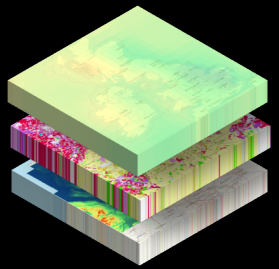
Where does the data set comes from?

IBM PAIRS Geoscope (part of WOC)



Extract for data science contest

Raw data (NASA, Copernicus, NOAA etc.)

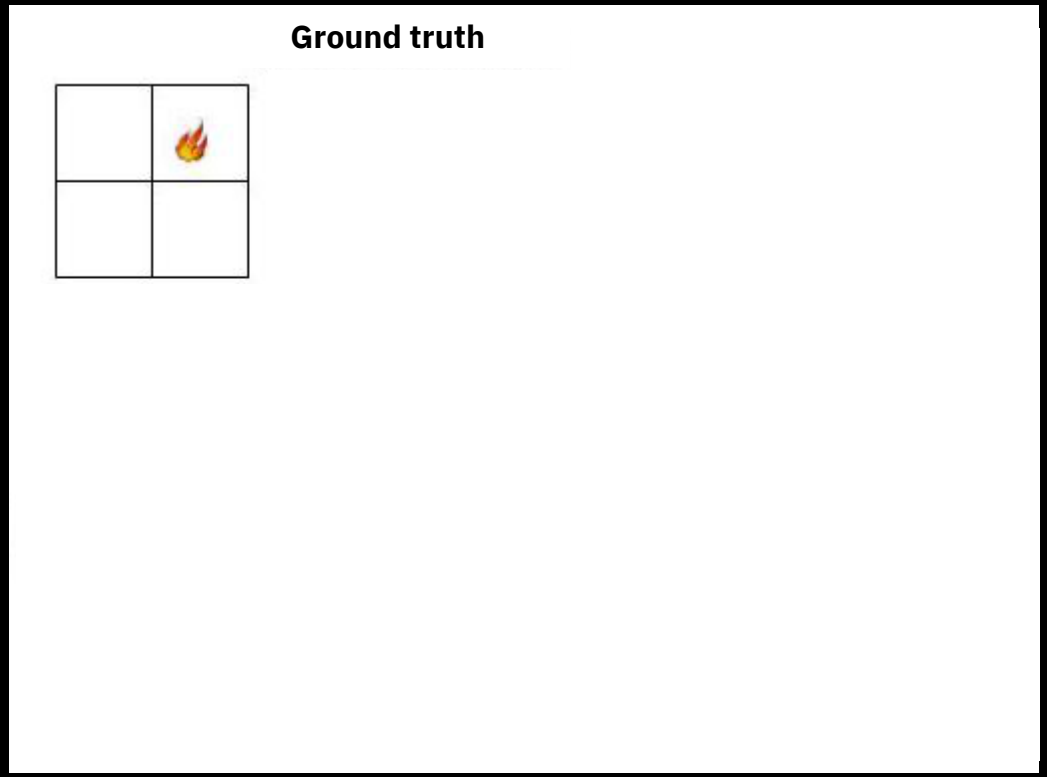


6PB of query-able and curated geospatial data

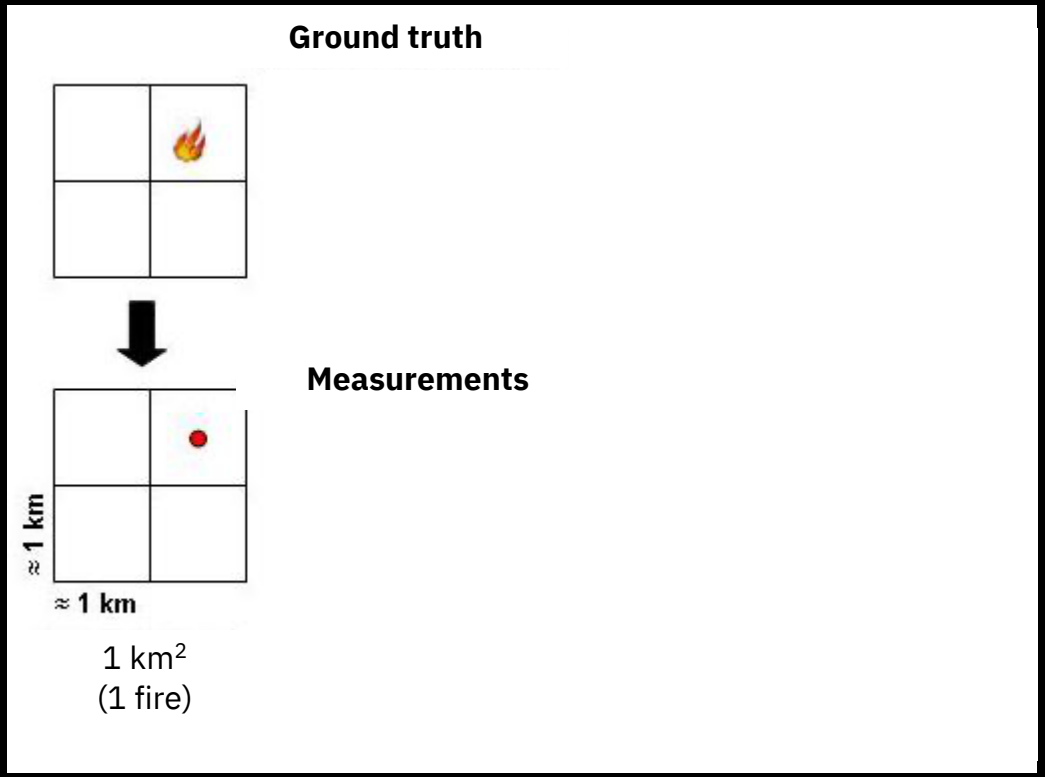
IBM

Historical fire data from a NASA satellite

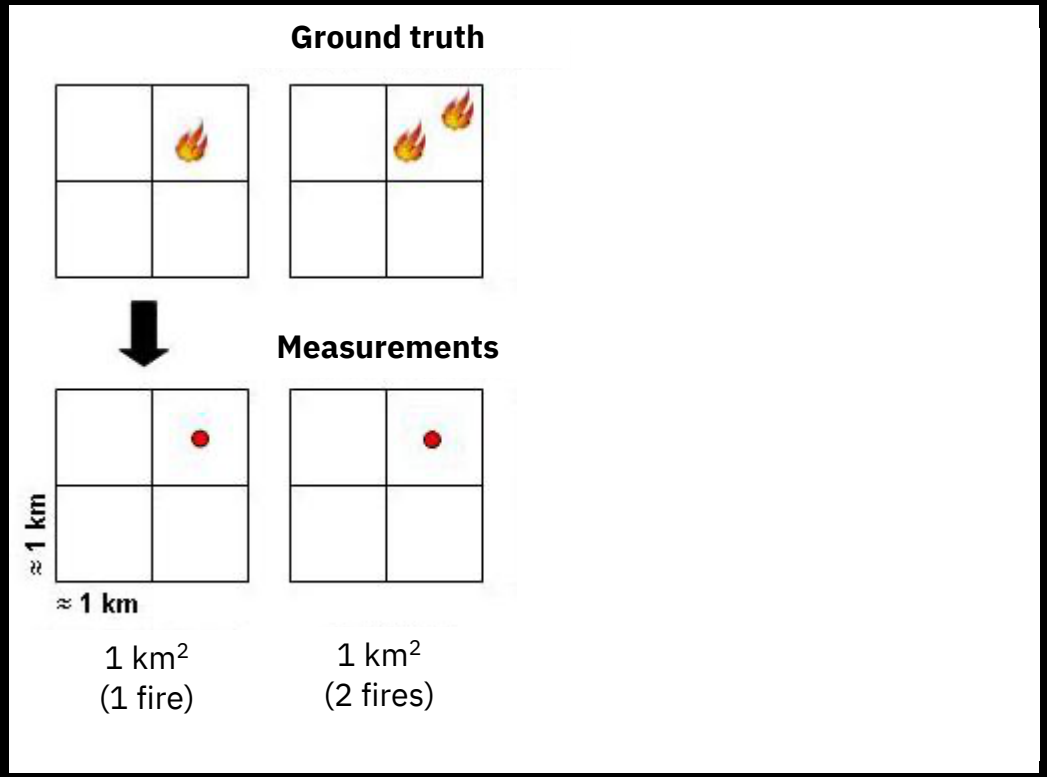
Historical fire data – from hotspots to pixels to estimated areas



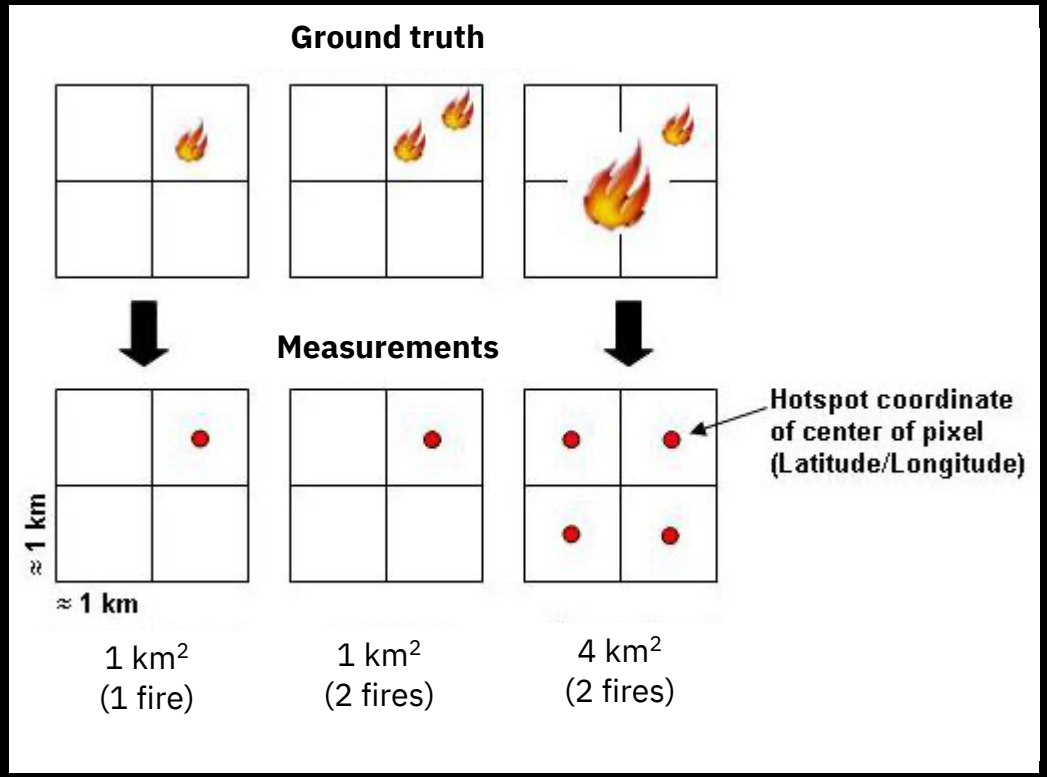
Historical fire data – from hotspots to pixels to estimated areas



Historical fire data – from hotspots to pixels to estimated areas

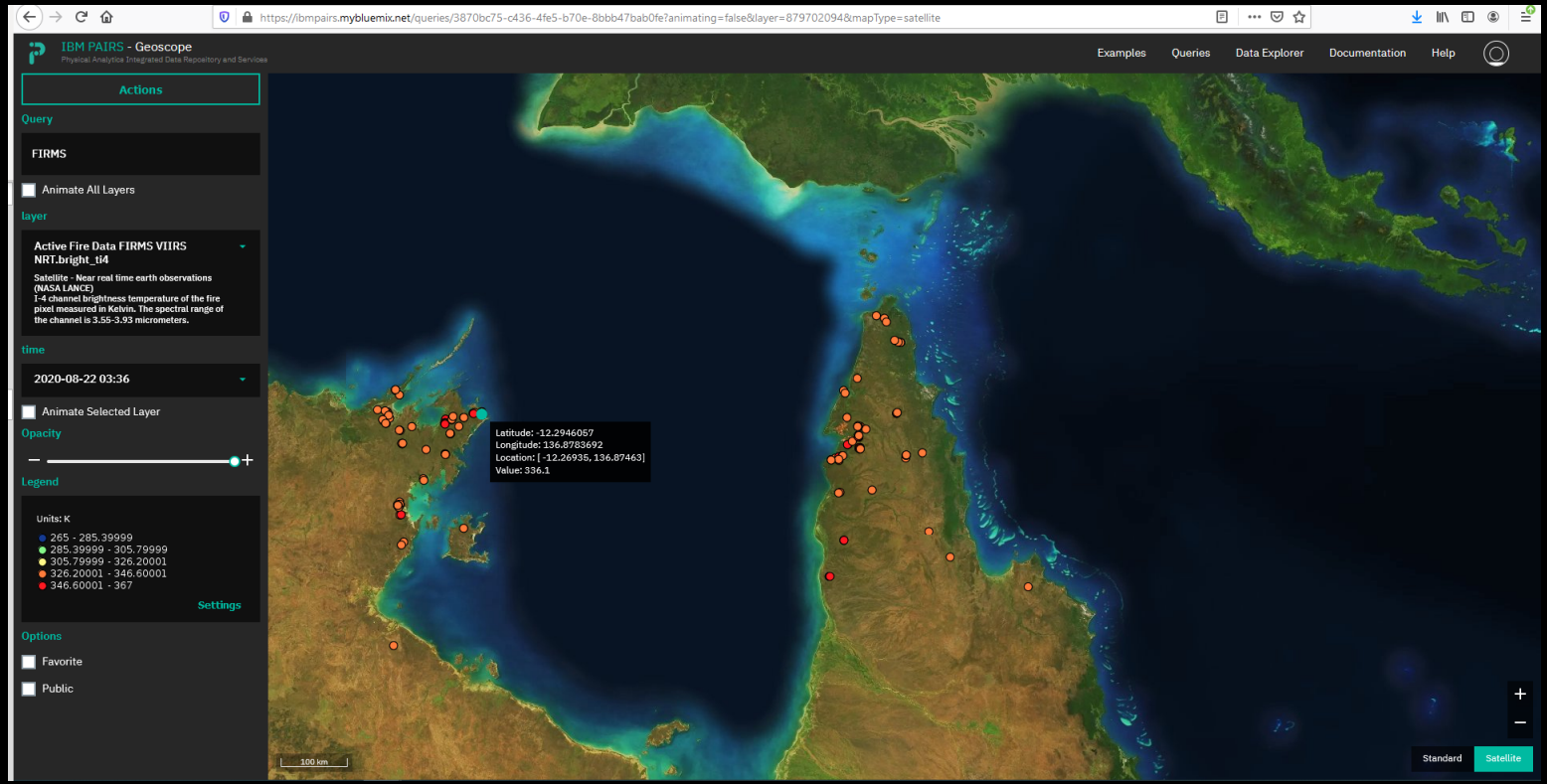


Historical fire data – from hotspots to pixels to estimated areas

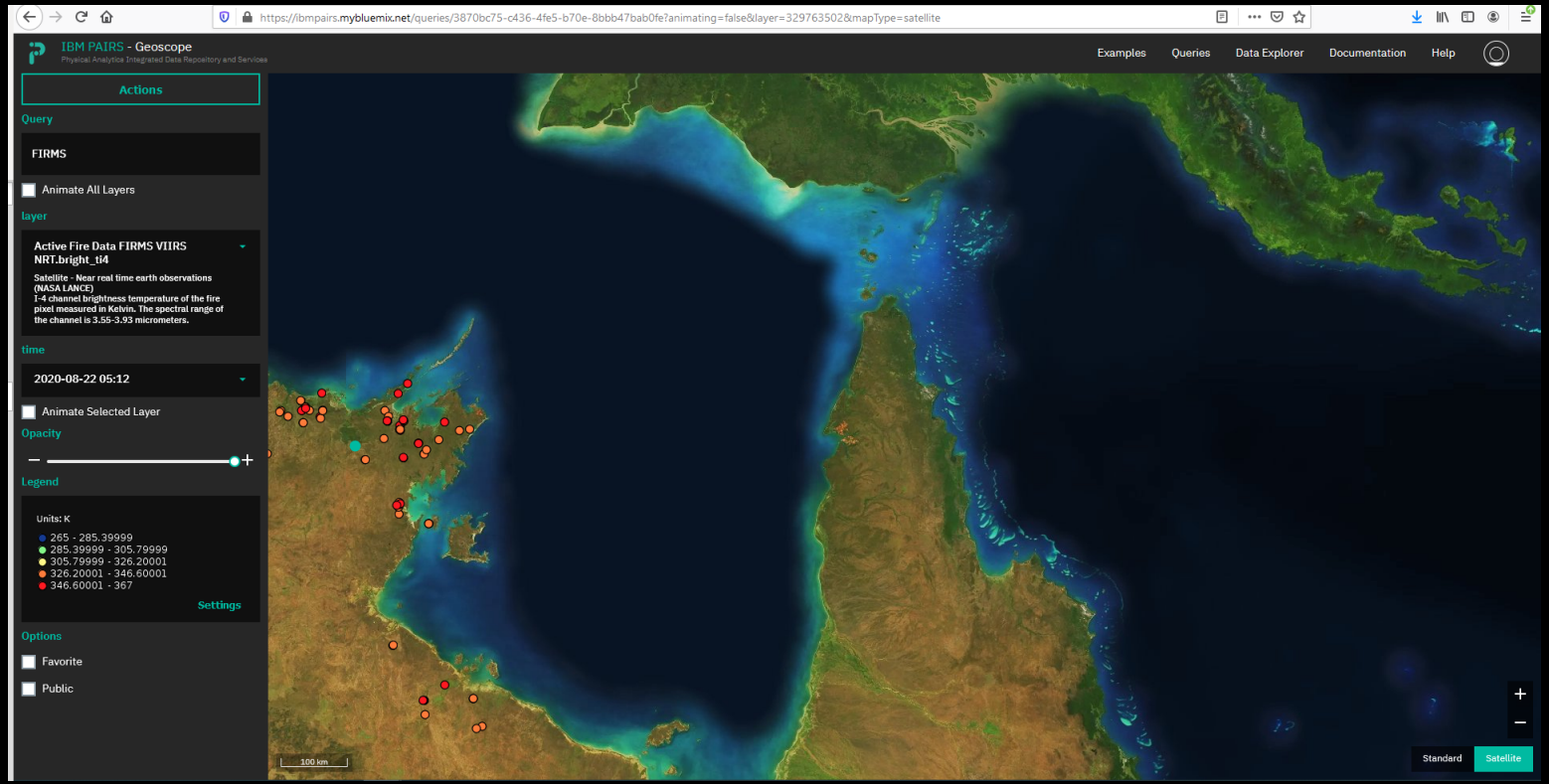


A hotspot plotted using the MODIS thermal anomalies algorithm represents the center of an approximately one-square-kilometer pixel flagged as containing one or more thermal anomalies, which may indicate a fire (upper half of image). The hotspot "location" is the center point of the pixel, which is an approximation of the actual thermal anomaly (lower half of image). Illustration courtesy of NASA FIRMS.

Historical fire data – Raw data



Historical fire data – Raw data



Historical fire data – PAIRS processing

1. Daily averaged
2. Spatially aggregated over 7 regions
3. Confidence of >75%
4. Inferred hotspot type = 0 meaning a presumed vegetation fire
5. Area estimated by multiplying the along scan pixel size by the along track pixel size.
6. Brightness estimated by averaging the means of both the brightness temperature 21 (obtained from channel 21/22) and brightness temperature 31 (obtained from channel 31).

Historical fire data – Columns

1. Region
2. Date
3. Estimated_fire_area [km^2]
4. Mean_estimated_brightness [K]
5. Mean_estimated_fire_radiative_power [MW]
6. Mean_confidence [%]
7. Std_confidence [%]
8. Var_confidence [%]
9. Count
10. Replaced [Y/N]

Historical weather data – PAIRS processing

1. Daily averaged
2. Spatially aggregated over 7 regions

Historical weather data – Columns

1. Region
2. Date
3. Parameter
   - ✓ Precipitation [mm/day]
   - ✓ Relative humidity [%]
   - ✓ Soil water content [m3 m3]
   - ✓ Solar radiation [MJ/day]
   - ✓ Temperature [C]
   - ✓ Wind speed [m/s]
4. count() [km2]
5. min(): Minimum value of the spatial aggregation.
6. max(): Maximum value of the spatial aggregation.
7. mean(): Average of the spatial aggregation.
8. variance: 2nd moment of the spatial aggregation

Historical weather forecasts – PAIRS processing

1. Daily averaged
2. Spatially aggregated over 7 regions

Historical weather forecasts – Columns

1. Region
2. Date
3. Parameter
   - ✓ Precipitation [mm/day]
   - ✓ Relative humidity [%]
   - ✓ Solar radiation [MJ/day]
   - ✓ Temperature [C]
   - ✓ Wind speed [m/s]
4. Lead time [day]
5. count() [km^2]
6. min(): Minimum value of the spatial aggregation.
7. max(): Maximum value of the spatial aggregation.
8. mean(): Average of the spatial aggregation.
9. variance: 2nd moment of the spatial aggregation

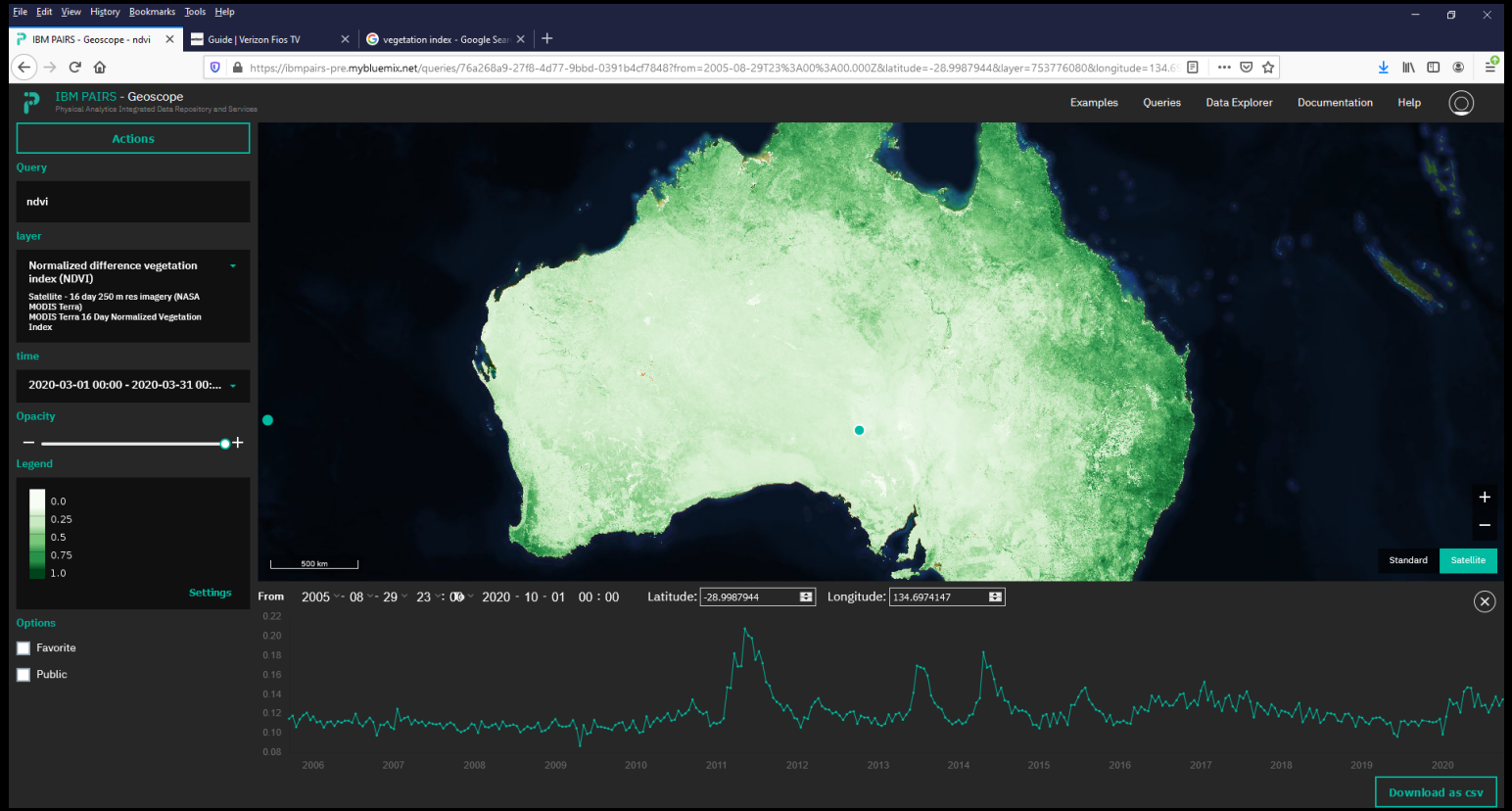
Land class – PAIRS processing

1. Spatially aggregated over 7 regions
2. Data is normalized to 100 [%]

Land class – Columns

1. Region
2. Shrubs [%]
3. Herbaceous vegetation [%]
4. Cultivated and managed vegetation/agriculture (cropland) [%]
5. Urban / built up [%]
6. Bare / sparse vegetation [%]
7. Permanent water bodies [%]
8. Herbaceous wetland [%]
9. Closed forest, evergreen, broad leaf [%]
10. Closed forest, deciduous broad leaf [%]
11. Closed forest, unknown [%]
12. Open forest, evergreen broad leaf [%]
13. Open forest, deciduous broad leaf [%]
14. Open forest, unknown definitions [%]
15. Open sea [%]

Vegetation index from a NASA satellite



Vegetation index – PAIRS processing

1. Monthly aggregated
2. Spatially aggregated over 7 regions

Vegetation index – Columns

1. Region
2. Date
3. Vegetation_index_mean
4. Vegetation_index_max
5. Vegetation_index_min
6. Vegetation_index_variance


IBM