

The background of the image is a photograph of a dense forest of tall evergreen trees. Overlaid on this image is a green digital matrix or binary code pattern, consisting of vertical lines and small green characters (letters and numbers) that appear to be falling or moving across the screen.

Is London a Forest?

Using Python and Open Data to Find Out

This talk: Independent research



Objective: Find tree coverage percentage of the Greater London Area



Vector image from
freepik.com

Structure

- **Existing research** (London tree cover percentage)
- The **Open Datasets** we're going to use
- Loading the **MODIS dataset** to db
- Loading the **shapefile dataset** to db
- Find London **tree cover percentage** with SQL



Structure

- **Existing research** (London tree cover percentage) context
- The **Open Datasets** we're going to use
- Loading the **MODIS dataset** to db
- Loading the **shapefile dataset** to db processing
- Find London **tree cover percentage** with SQL

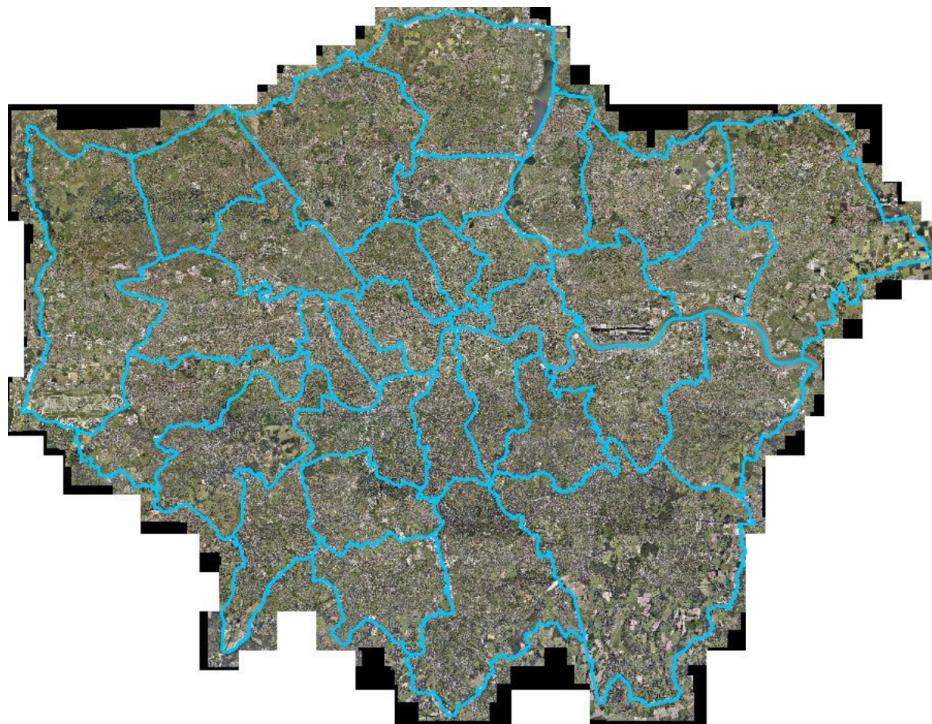


2015 study

- City Hall = the governance body for London & the Greater London Area
- Did a study on **tree canopy cover**
- Objective:
 - Measure impact of tree planting policies



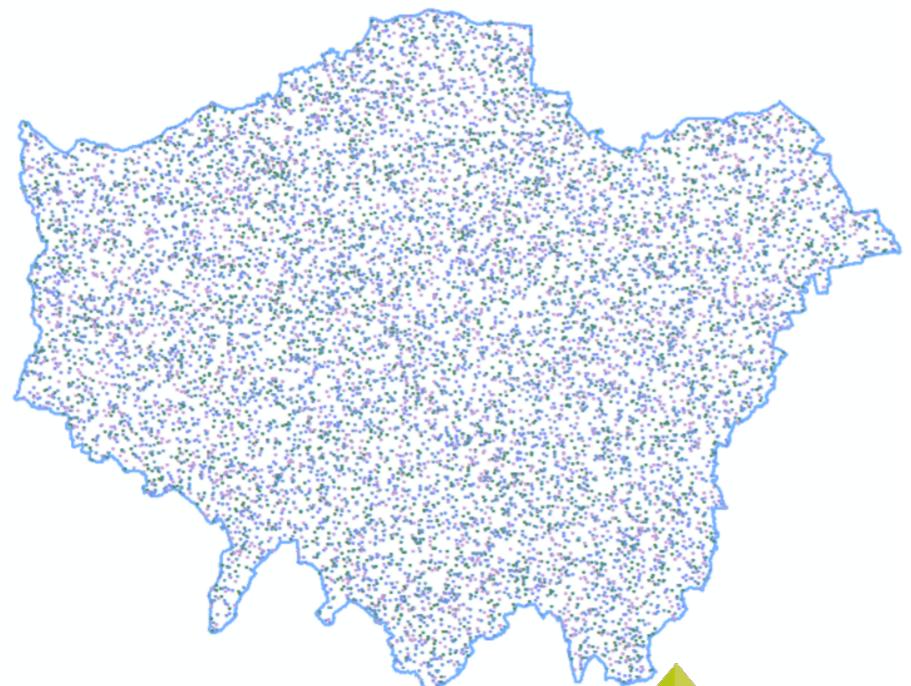
2015 study: Greater London Area



(Aerial image from The Geoinformation Group, 2011)

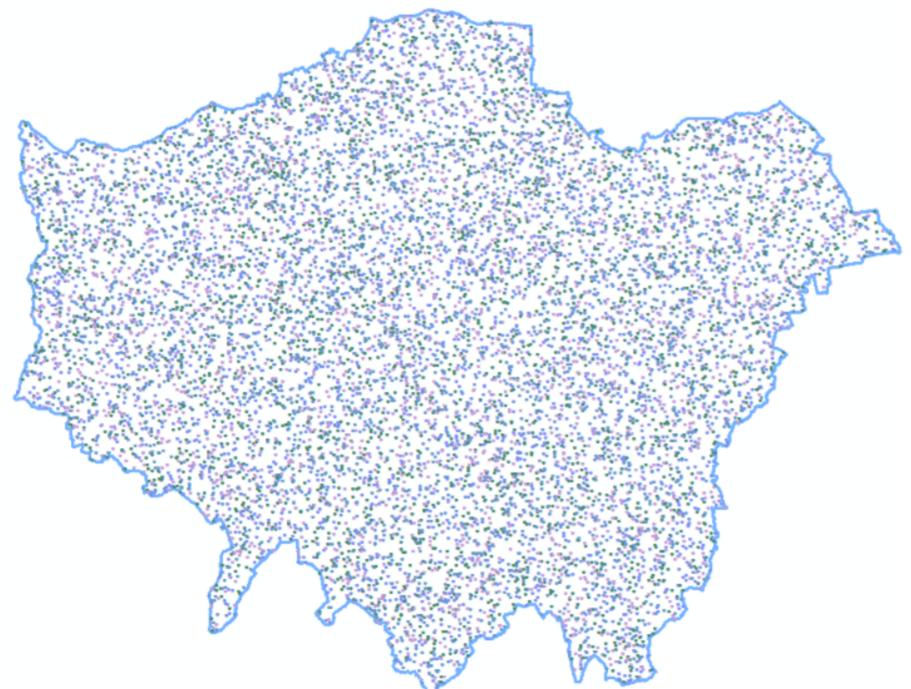
2015 study

- Aerial images (private dataset)
- Humans classify a sample of points as tree covered or not



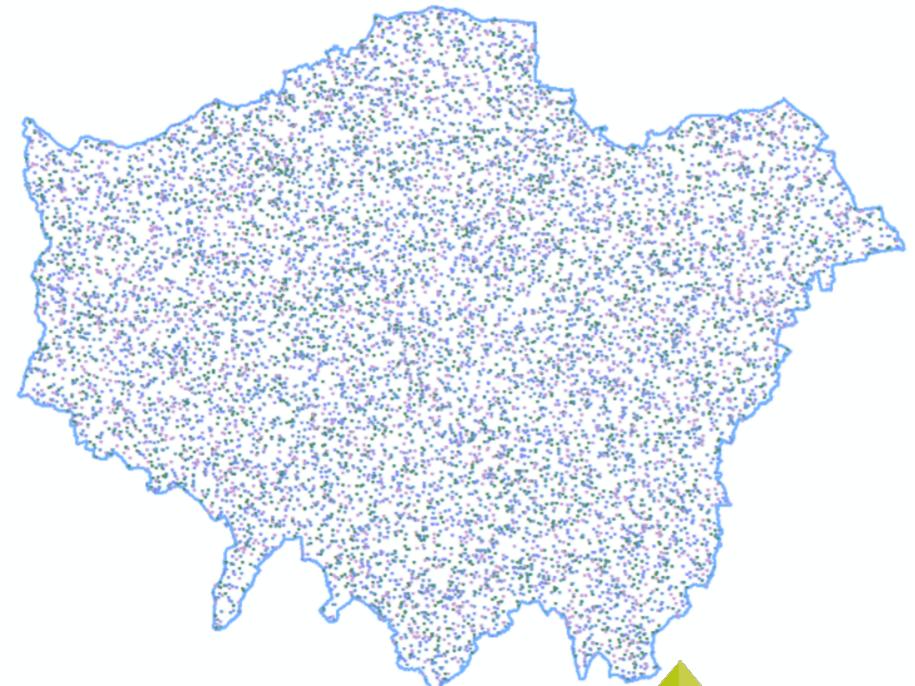
2015 study

- Aerial images (private dataset)
- Humans classify a sample of points as tree covered or not
- Multiple users classifying same image (quality check)
- Statistical inference of tree cover



Results

19.5 % tree cover



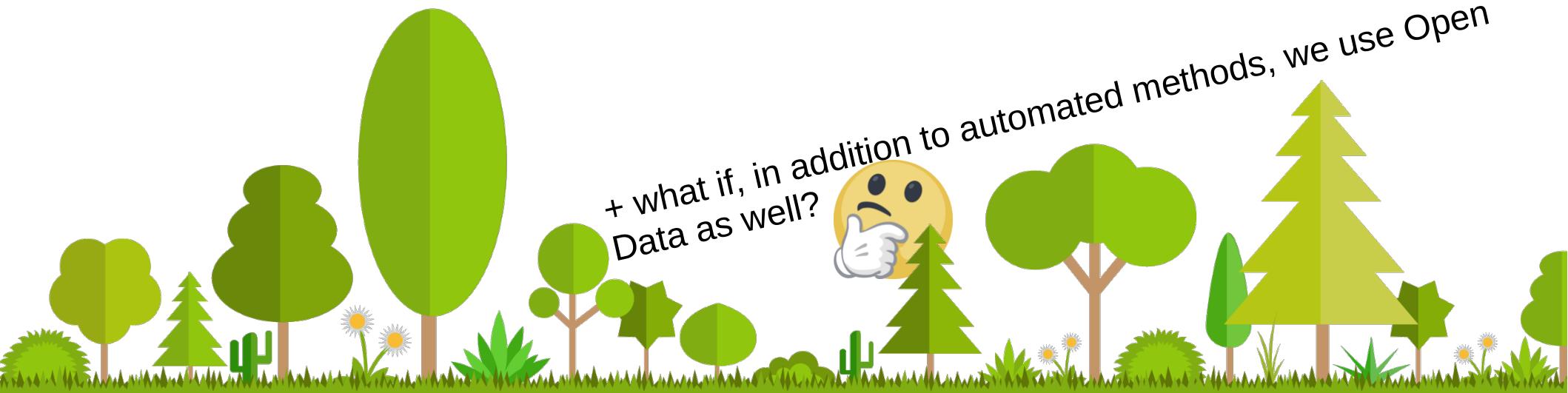
Idea for further research

Using aerial imagery is a time-consuming method of identifying land cover in an urban area. However, it is effective and accurate and does not require specialist processing or software. Few studies utilise aerial imagery to classify urban environments and this report contributes to the wider research on the best methods to analyse urban land cover. It would be useful to compare the findings produced by random point classification of aerial imagery to more automated methods.



Idea for further research

Using aerial imagery is a time-consuming method of identifying land cover in an urban area. However, it is effective and accurate and does not require specialist processing or software. Few studies utilise aerial imagery to classify urban environments and this report contributes to the wider research on the best methods to analyse urban land cover. It would be useful to compare the findings produced by random point classification of aerial imagery to more automated methods.



The Datasets



The MODIS dataset



- Data captured by satellite!
- Multiple layers with information (HDF format)
- One layer contains estimated* tree cover percentage (for a specific area)

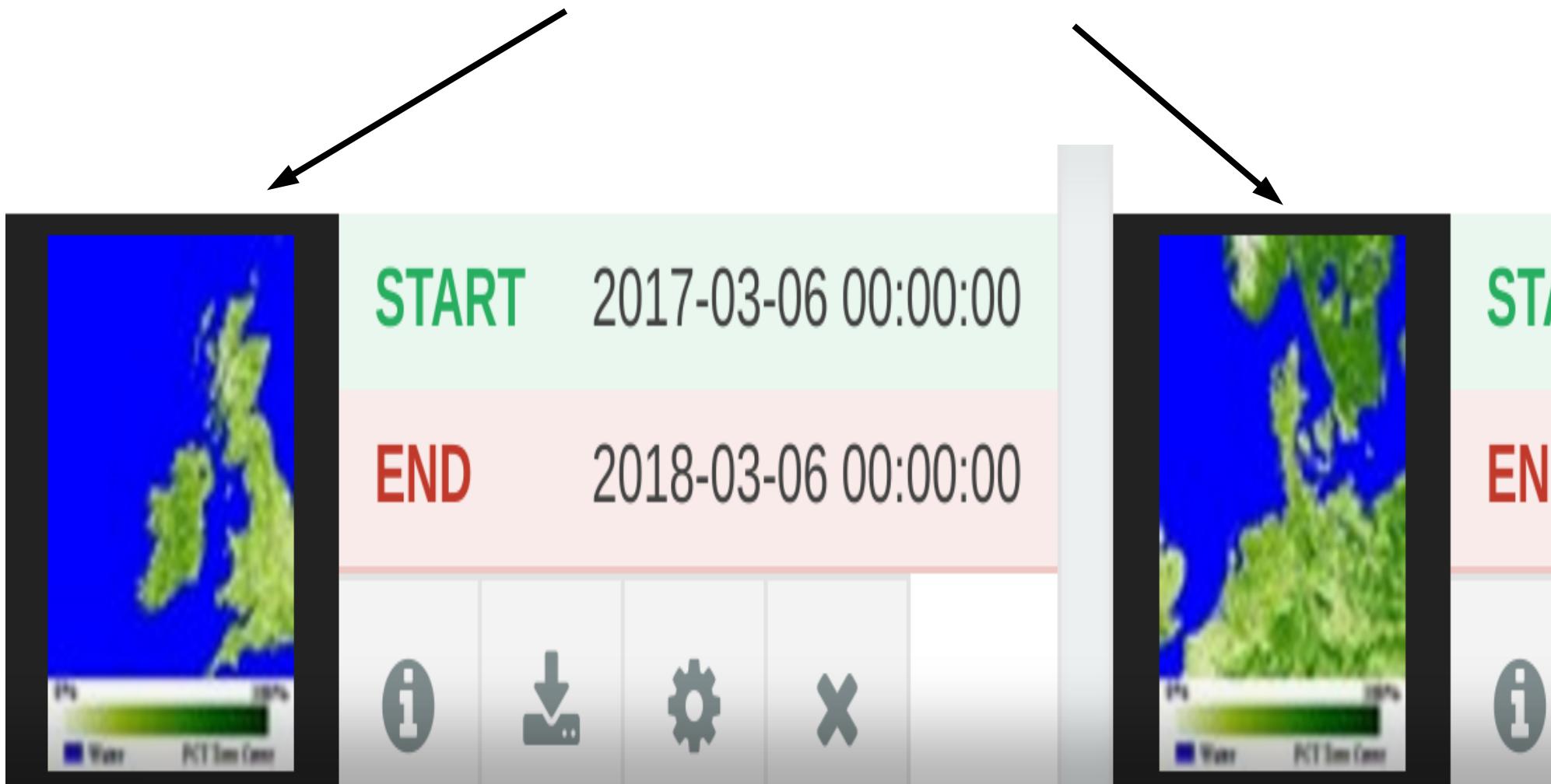
* using machine learning methods



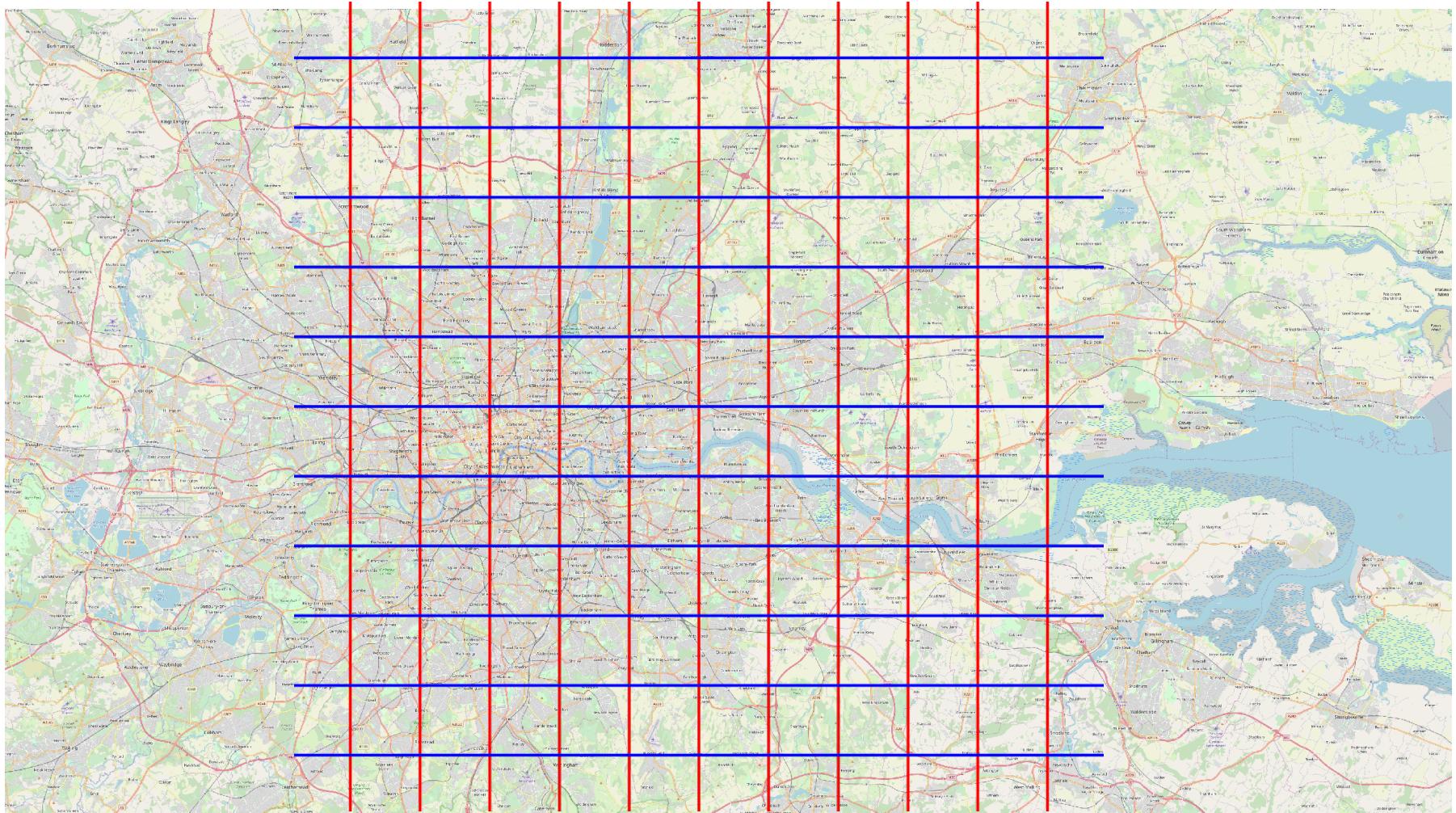
The MODIS dataset



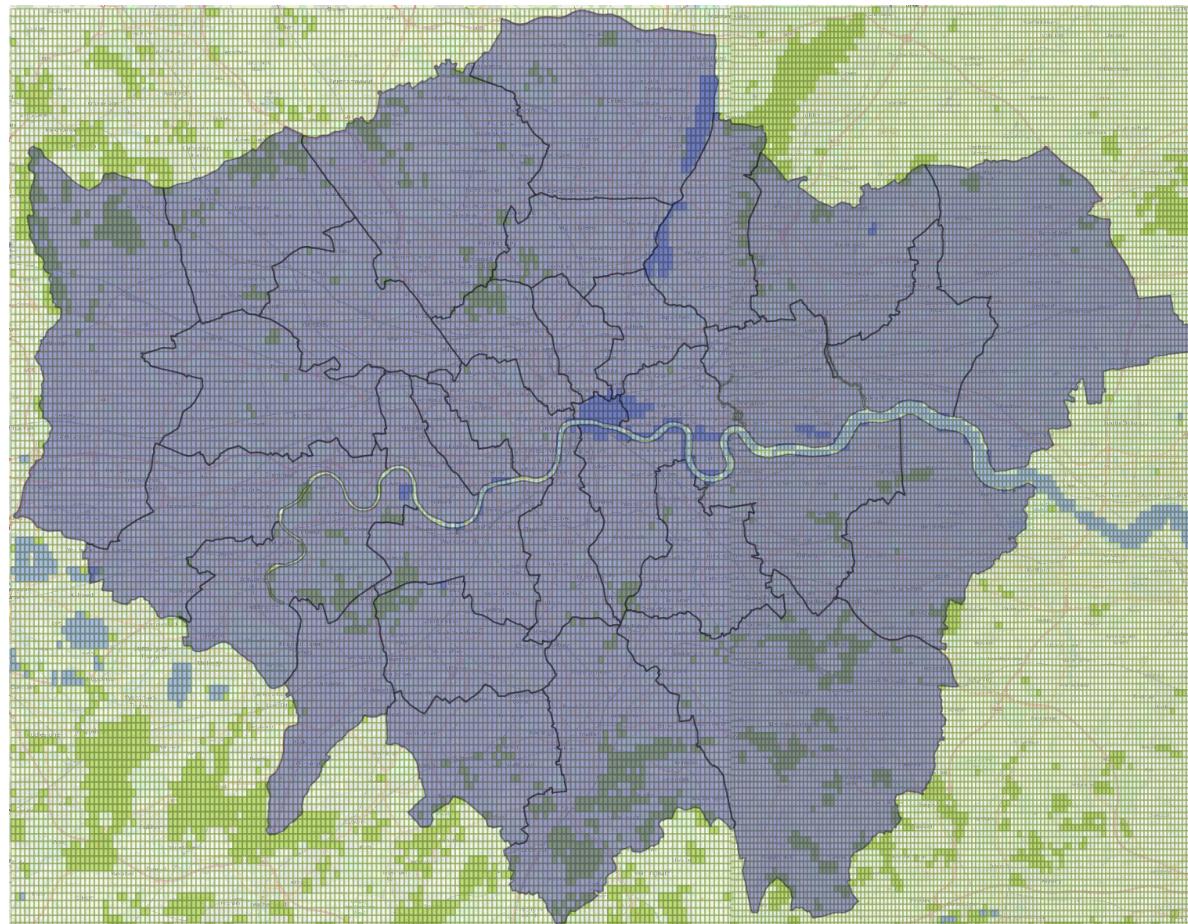
Earth divided into ~300 granules, each one has a local grid



Local grid



London shapefiles

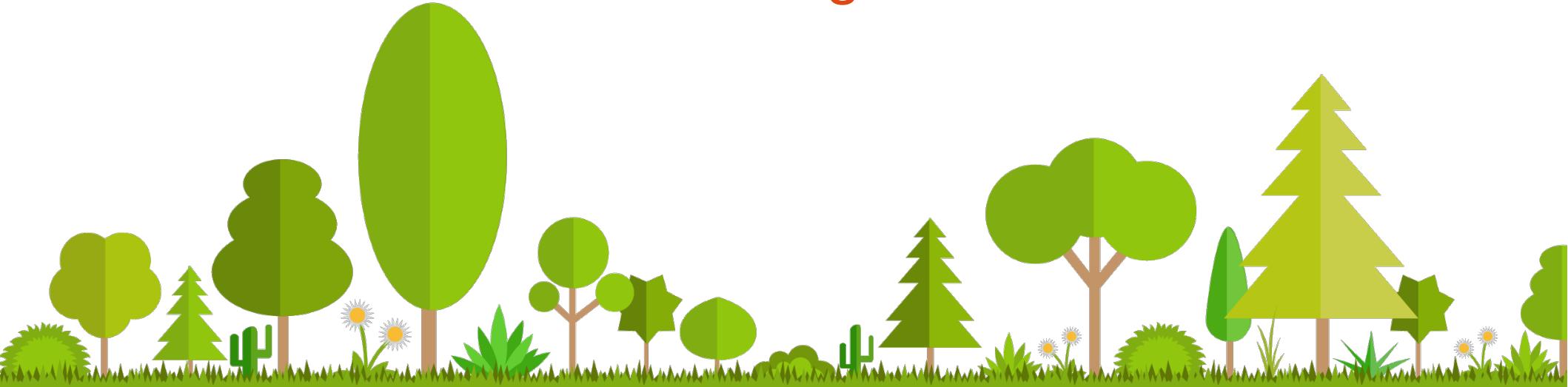


Processing

Main Idea:

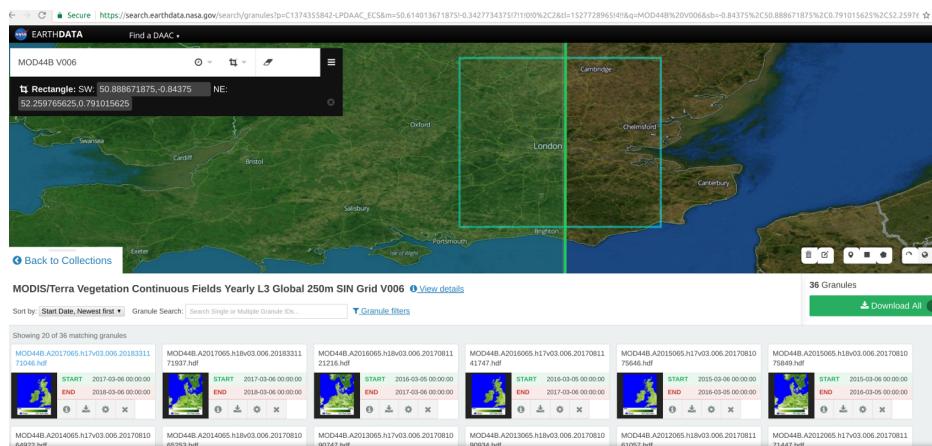
- Load both datasets
- Make them directly comparable
- Find tree cover area of London

Code on [github](#)



Processing

- Find the HDF files relevant for London (2 granules)
 - Bounding box query in the NASA website



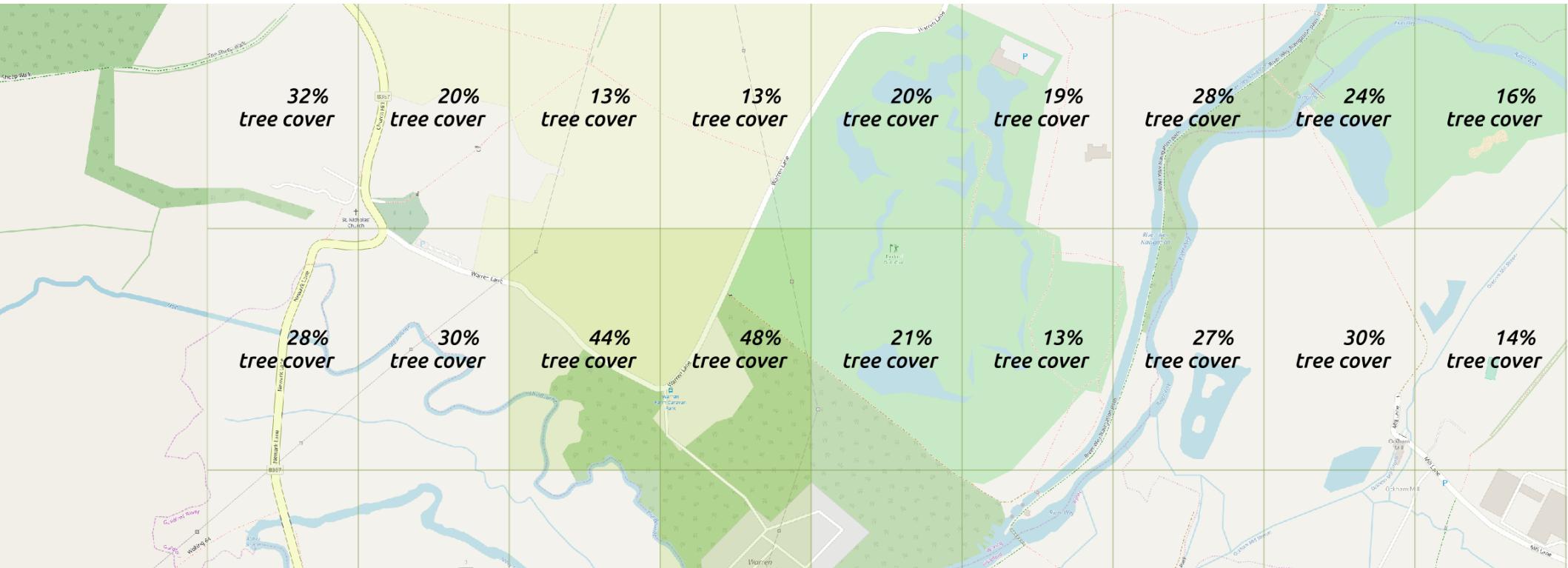
Processing

- Find the HDF files relevant for London (2 granules)
- **Load all rectangles + percentages in a PostGIS table**



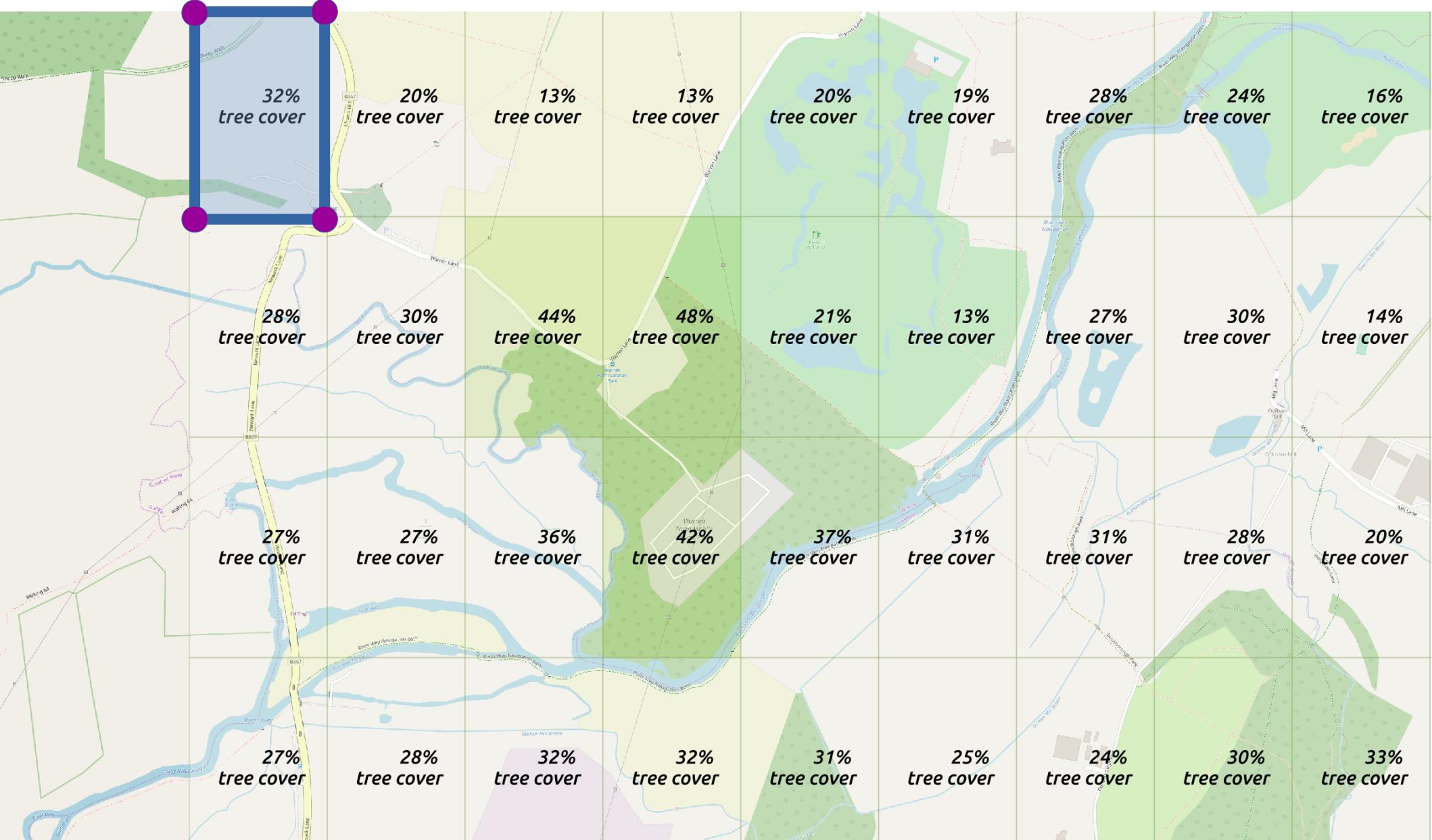
HDF file metadata

- Dimensions of grid (Grid_X, Grid_Y)
- Dimensions of each rectangle in grid
- Longitude and latitude of bounding box (of whole grid)
- Sinusoidal projection (reproject to lon/lat with **GDAL**)



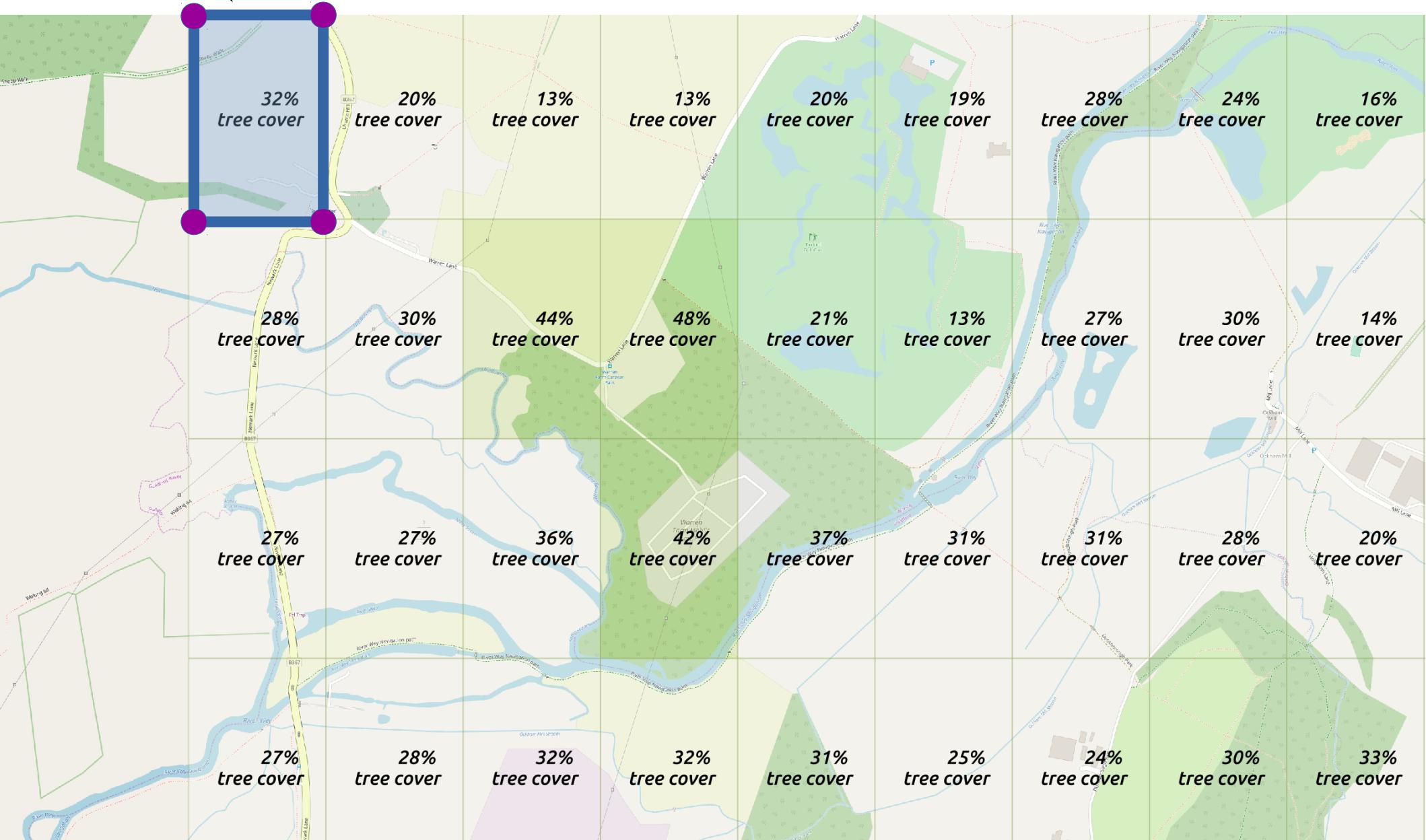
Can find out
lon/lat of 4
corners

Processing the HDF files



Load spatial object
and percentage
into PostGIS

Processing the HDF files



Resulting table

TABLE **tree_cover_rectangles**

(

rectangle **geometry**,

tree_cover_percentage **double precision**,

);



Processing

- Find the HDF files relevant for London (2 granules)
- Load all rectangles + percentages in a PostGIS table
- **Load the London shapefile in a PostGIS table**

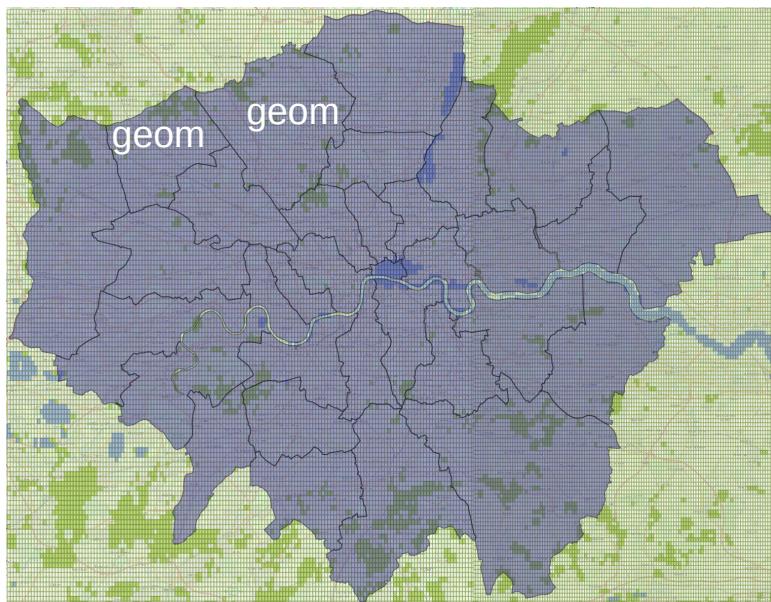


Loading the shapefile

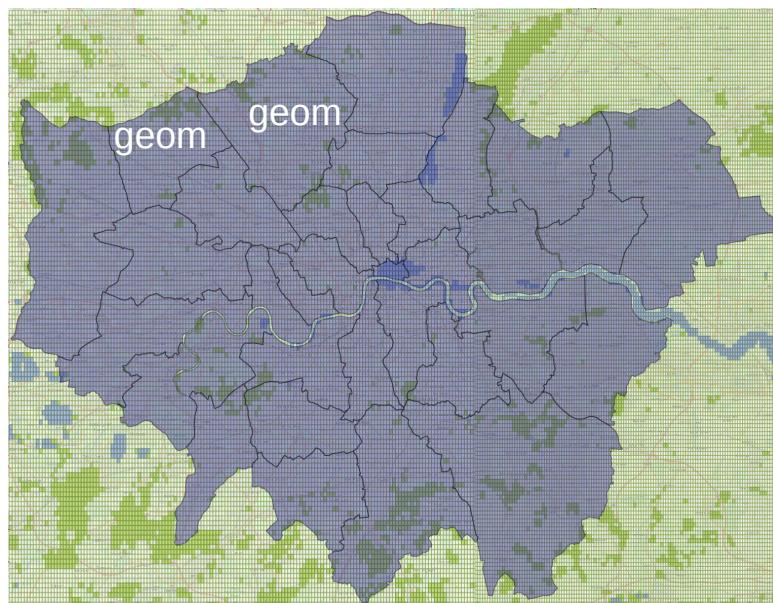
- Download from **official source**
- Load into Postgis table using ***shp2pgsql*** (*32 rows, one polygon per borough*)
- Transform from British National Grid projection to lon/lat



Unifying shapes



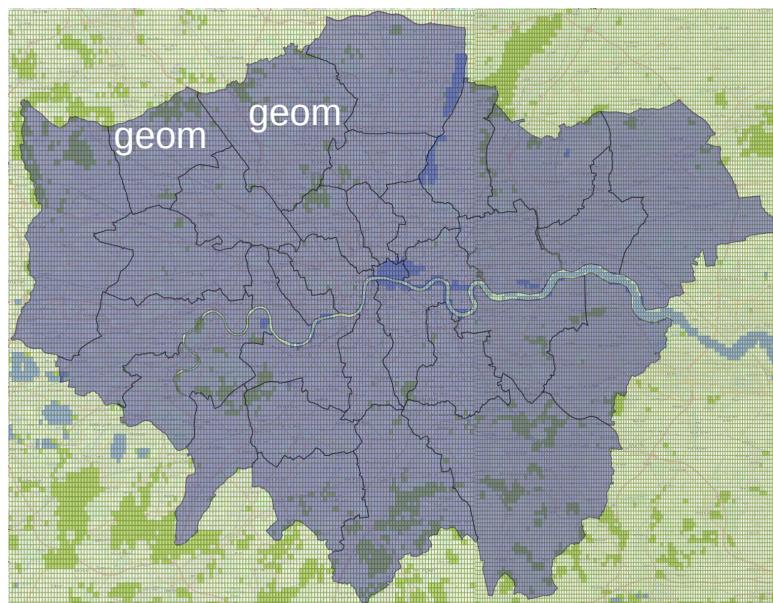
Unifying shapes



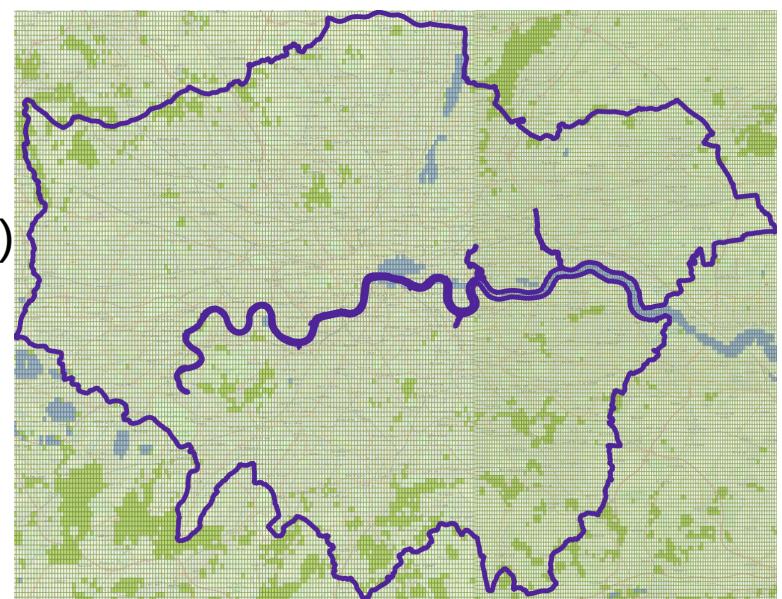
`ST_Union(geom)`



Unifying shapes



`ST_Union(geom)`



Resulting table

TABLE **greater_london**

(

geom **geometry** (just one row)

);



Processing

- Find the HDF files relevant for London (2 granules)
- Load all rectangles + percentages in a PostGIS table
- Load the London shapefile in a PostGIS table
- **Answer the question about tree cover using (spatial) SQL queries**



Queries

```
TABLE tree_cover_rectangles  
(  
rectangle geometry,  
tree_cover_percentage double precision,  
);
```

```
TABLE greater_london  
(  
geom geometry (just one row)  
);
```



```
TABLE tree_cover_rectangles
(
    rectangle geometry,
    tree_cover double precision,
);

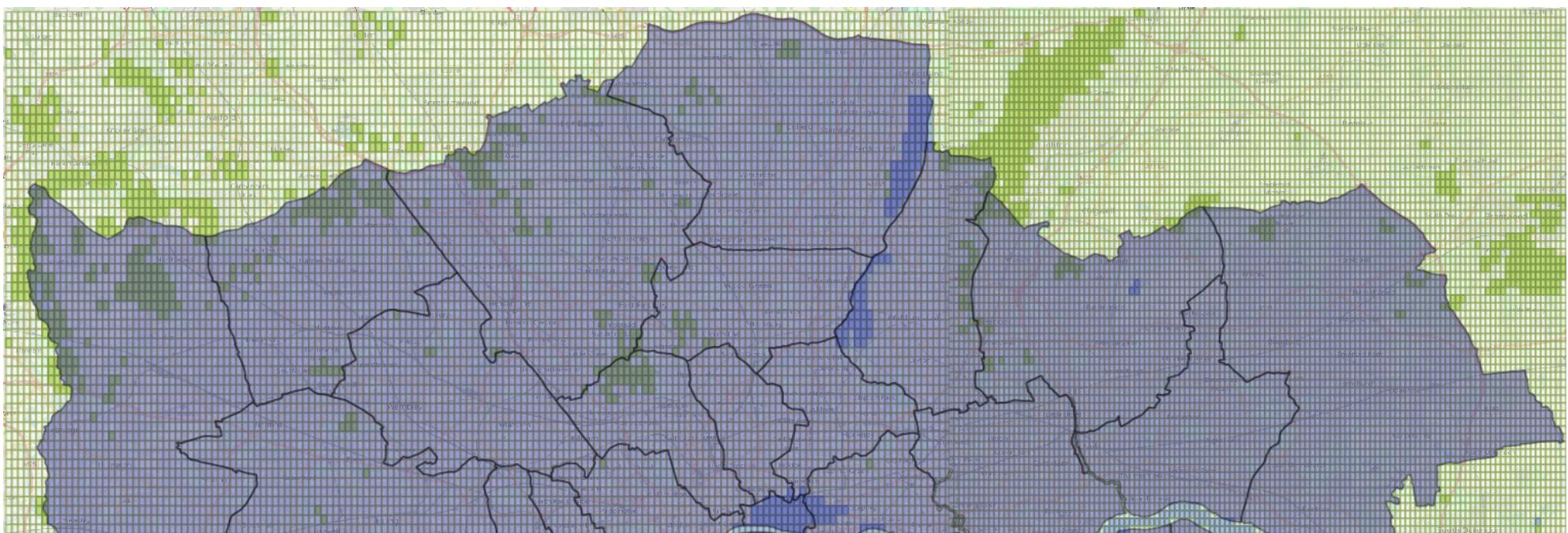
```

```
TABLE greater_london
(
    geom geometry (just one row)
);

```

- Which rectangles are contained in the Greater London area? →

**greater_london JOIN tree_cover_rectangles ON
ST_Intersects(geom, rectangle)**

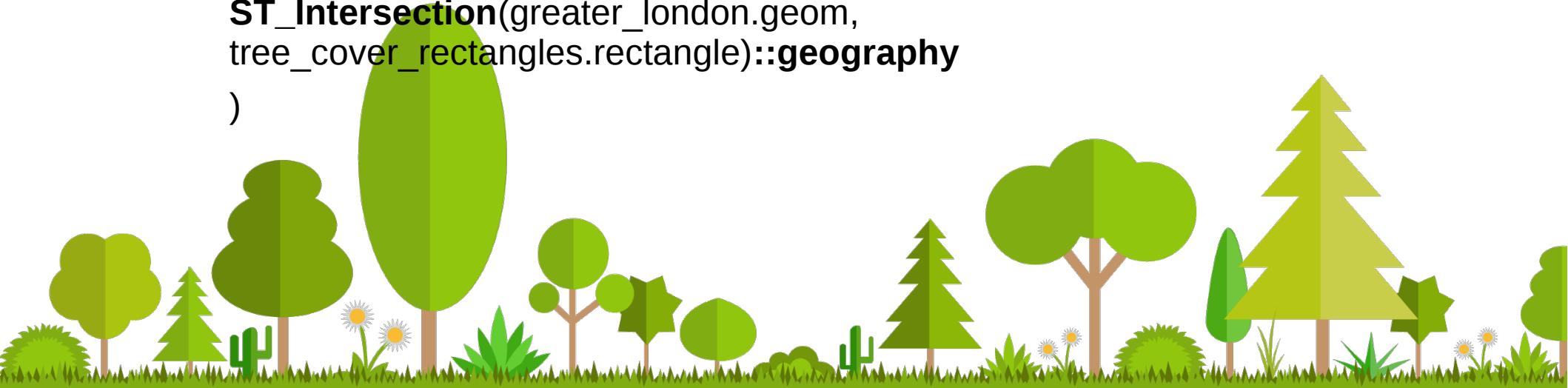


```
TABLE tree_cover_rectangles
(
    rectangle geometry,
    tree_cover double precision,
);
;
```

```
TABLE greater_london
(
    geom geometry (just one row)
);
```

- Which rectangles are contained in the Greater London area? →
greater_london **JOIN** tree_cover_rectangles **ON**
ST_Intersects(geom, rectangle)

- What is the tree cover area of each rectangle (total area of rectangle x tree cover percentage) →
(tree_cover_percentage/100) * **ST_Area**(
ST_Intersection(greater_london.geom,
tree_cover_rectangles.rectangle)::geography
)



```
TABLE tree_cover_rectangles
(
    rectangle geometry,
    tree_cover double precision,
);

```

```
TABLE greater_london
(
    geom geometry (just one row)
);
```

- Which rectangles are contained in the Greater London area? →
greater_london **JOIN** tree_cover_rectangles **ON**
ST_Intersects(geom, rectangle)
- What is the tree cover area of each rectangle (total area of rectangle x tree cover percentage) →
(tree_cover_percentage/100) * **ST_Area**(
 ST_Intersection(greater_london.geom, tree_cover_rectangles.rectangle)::geography
)
- Sum of tree cover area of each rectangle = tree cover area of London
(just plain SQL SUM)



Results

30,900 hectares or **19.4 %** tree cover (compared to 19.5 % in original study)



Results

Only possible with

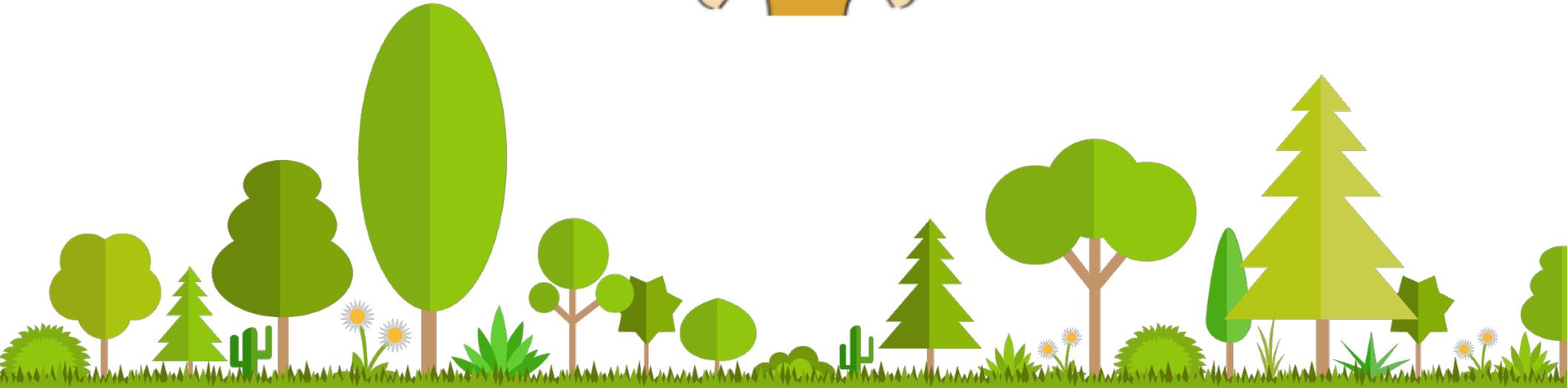
- **Open Data**
- **Open Tools** (QGIS, GDAL, PostGIS, shp2pgsql)



Thanks



- carolinegr@gmail.com
- github.com/carolinux/londonwald



Extra slides



Granules are not seamless



- Need to be careful 😅 with second granule
- Can use ST_Difference



Another methodology

- The City Hall applied machine learning techniques on aerial maps for tree canopy cover (2018)

Data freely available:

<https://www.london.gov.uk/what-we-do/environment/parks-green-spaces-and-biodiversity/trees-and-woodlands/tree-canopy-cover-map>

