Advanced GIS

Raphaëlle ROFFO

Sciences Po - Urban School

Today's plan

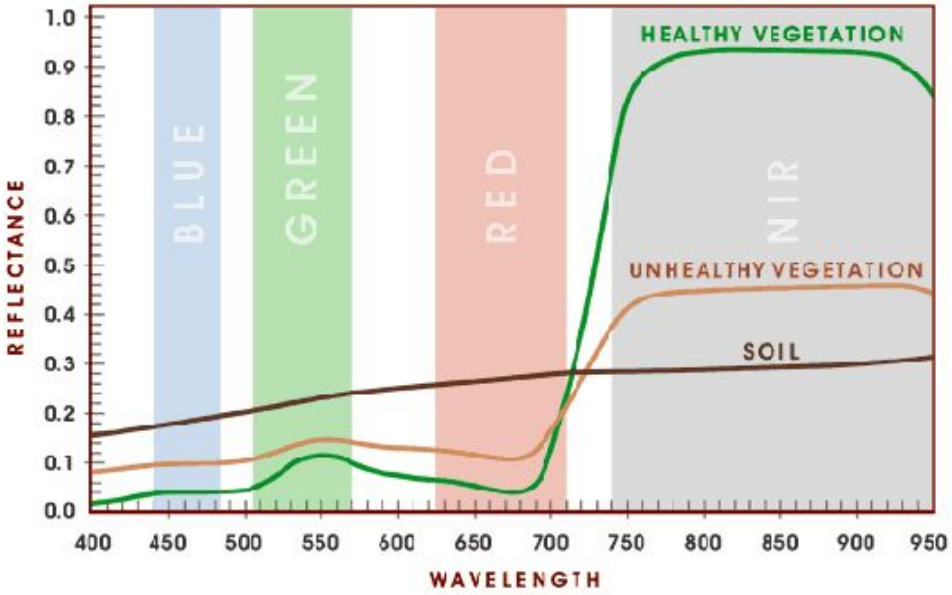
1. Tutorial feedback & questions
2. Raster data - examples
3. Georeferencing & Digitizing

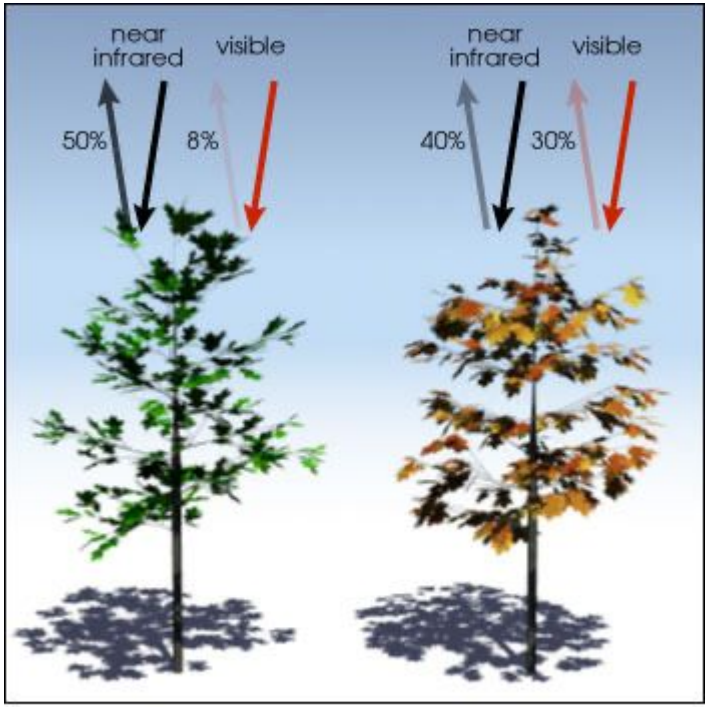
1. Tutorial feedback & questions

Take-away from the tutorial

1. Downloading data from public satellite (choosing the satellite, filtering out for clouds, picking the date etc)
2. Importing satellite data into QGIS
3. Working with multiple raster bands
4. Using the raster calculator to build the NDVI
5. Reclassifying data

The NDVI

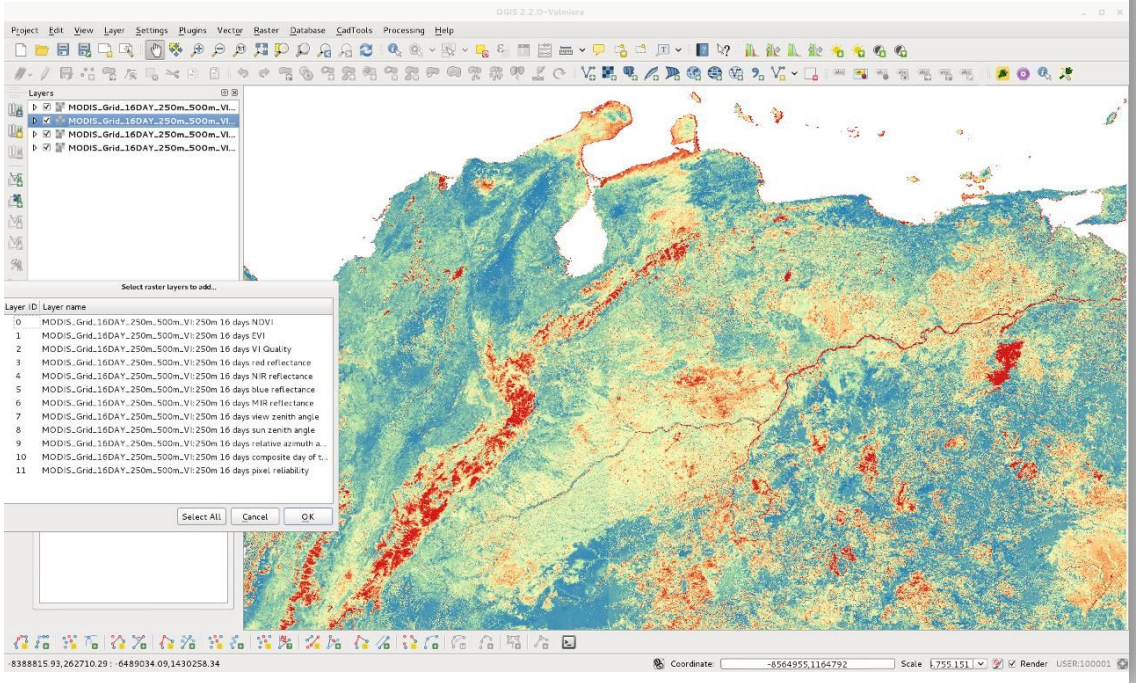

REFLECTANCE
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
BLUE
GREEN
RED
NIR
HEALTHY VEGETATION
UNHEALTHY VEGETATION
SOIL
400
450
500
550
600
650
700
750
800
850
900
950
WAVELENGTH


near infrared
visible
50%
8%
near infrared
visible
40%
30%
(0.50 - 0.08) / (0.50 + 0.08) = 0.72
(0.4 - 0.30) / (0.4 + 0.30) = 0.14

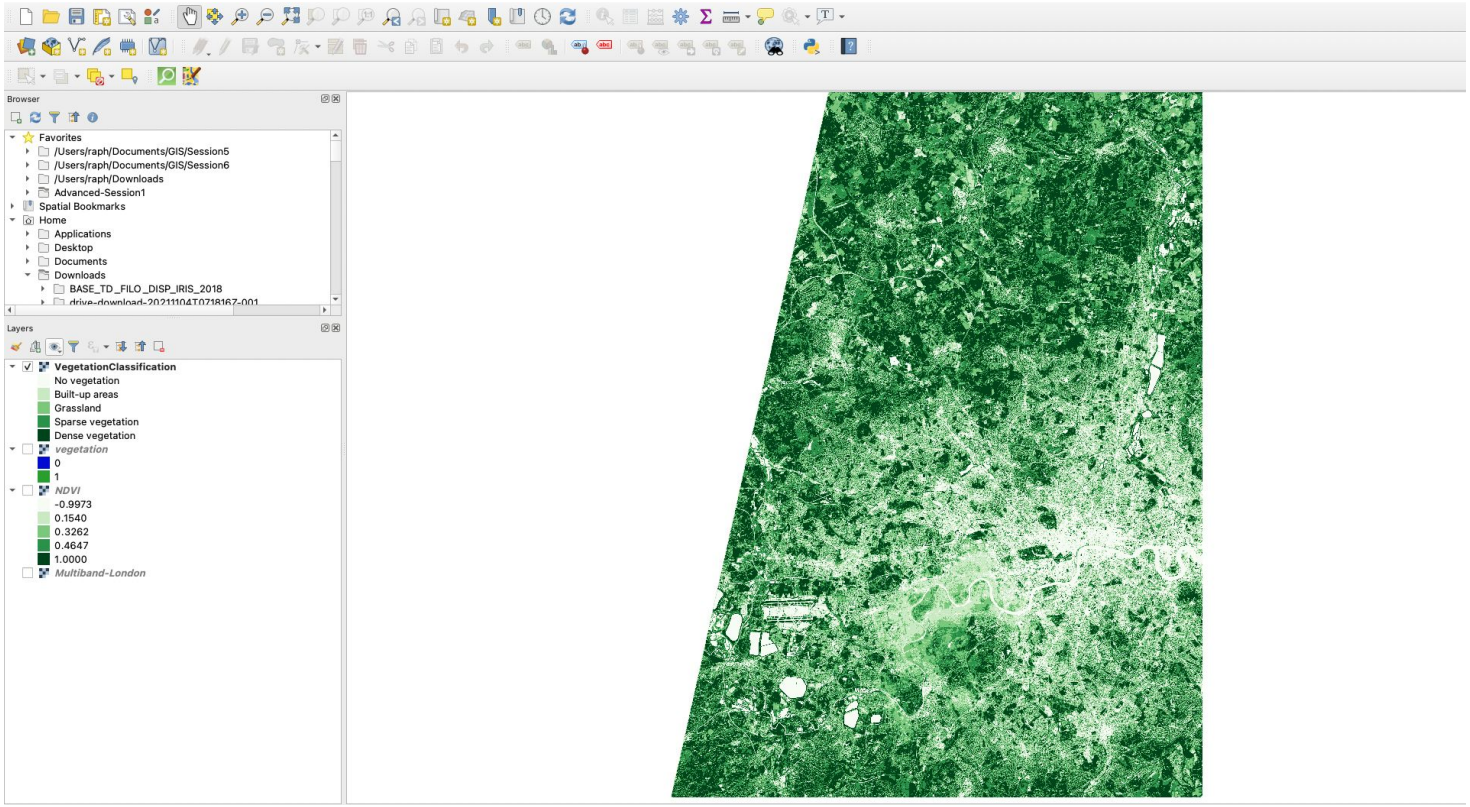
The NDVI

The normalized difference vegetation index (NDVI) allows us to infer the presence of green (healthy) vegetation based on the amount of light and near infrared wavelengths reflected from the surface of the Earth.

This is an index you can calculate in QGIS from satellite data (Sentinel-2, Landsat, SPOT etc)



Questions?


Browser
Favorites
/Users/raph/Documents/GIS/Session5
/Users/raph/Documents/GIS/Session6
/Users/raph/Downloads
Advanced-Session1
Spatial Bookmarks
Home
Applications
Desktop
Documents
Downloads
BASE_TD_FILO_DISP_IRIS_2018
Layers
VegetationClassification
No vegetation
Built-up areas
Grassland
Sparse vegetation
Dense vegetation
vegetation
0
1
NDVI
-0.9973
0.1540
0.3262
0.4647
1.0000
Multiband-London

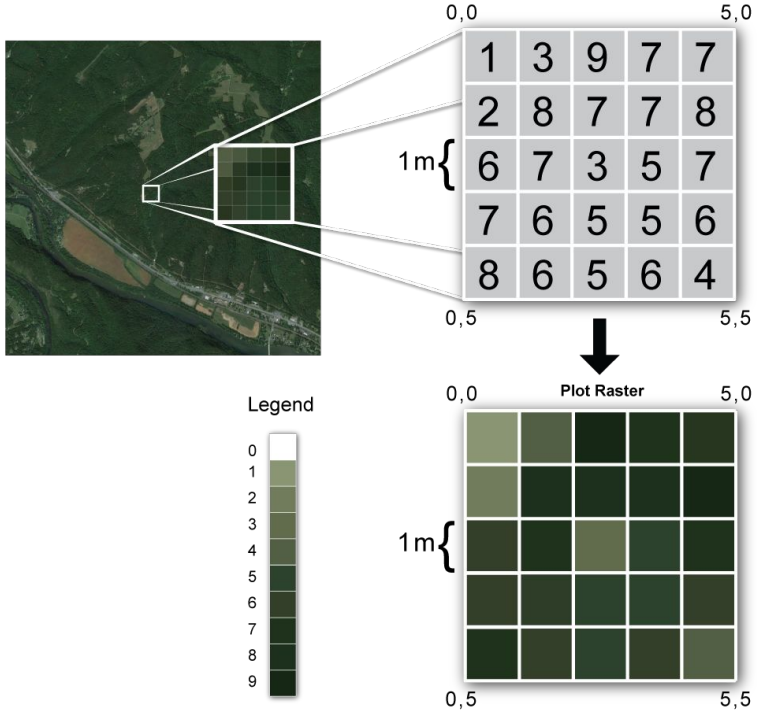
2. Raster data - use cases

1. What is raster data?

Raster vs vector

Spatial data can be represented in two ways:

- **Vector**: geometries: Point, Line, Polygon
- **Raster** : pixels, like a photo (each pixel is assigned a value)



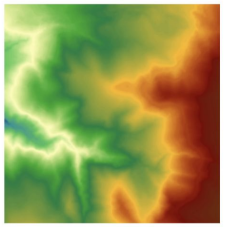
2. Continuous vs thematic raster data

Raster data models are used to represent continuous data like surface temperature or elevation...

But also thematic (= discrete = categorical) data such as land-use or soil class data.

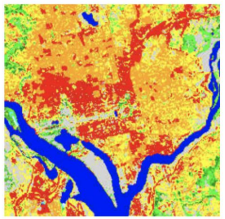
Continuous Data:

- Elevation
- Temperatures
- Precipitation
- Reflectance (light)



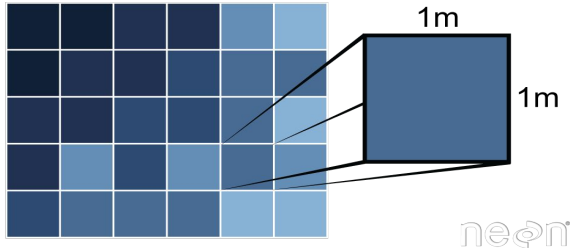
Categorical Data:

- Land Cover Type
- Soil Type
- Vegetation Type

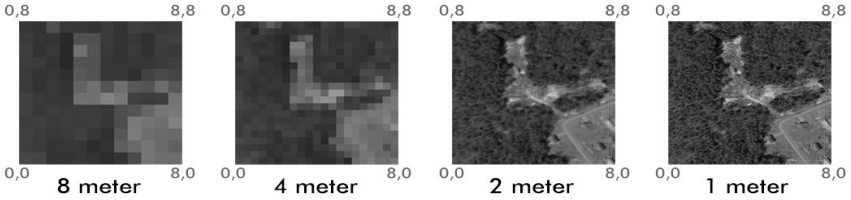


3. Spatial Resolution

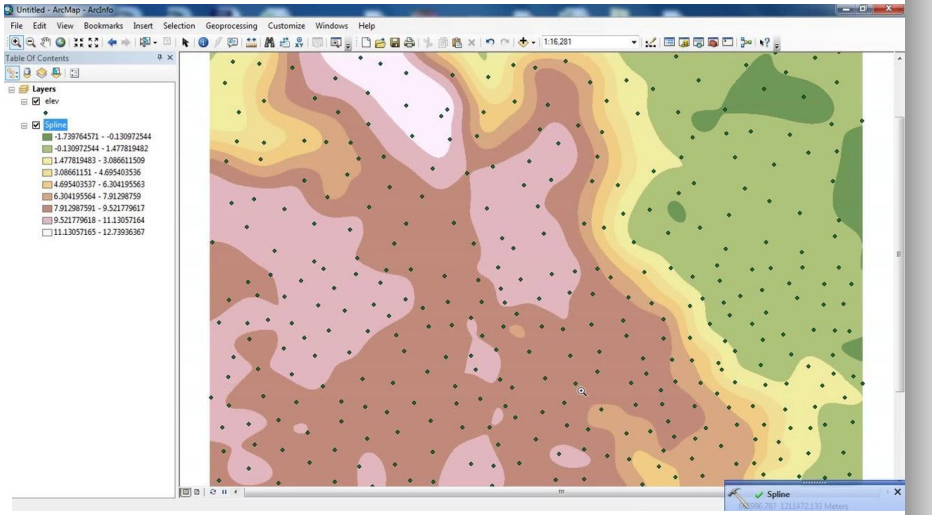
For remote sensing data, spatial resolution depends on the sensor used to take an image.



Raster over the same extent, at 4 different resolutions



For raster data computed by spatial analysis (e.g. rainfall), resolution will depend on your sampling scheme (e.g. number of weather stations).

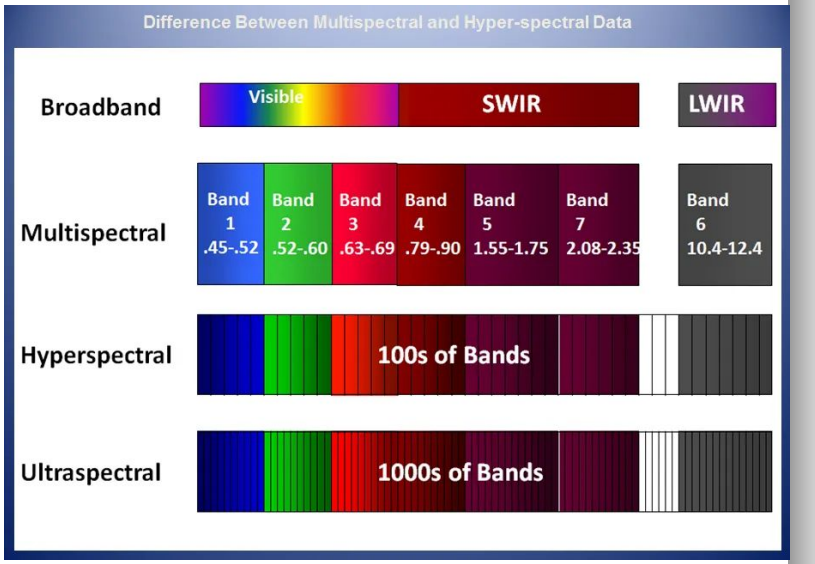


4. Remote Sensing: Spectral Resolution

With digital pictures, the camera uses electronic sensors to detect red, green and blue light. The red, green and blue (RGB) information is combined to show you an image that your eyes can interpret, but in fact RGB information is stored in **separate colour bands.**

This is the case for satellite images.

But remote sensing devices can also capture other wavelengths. They can record infra-red for instance (very useful in environmental science to detect bodies of water).



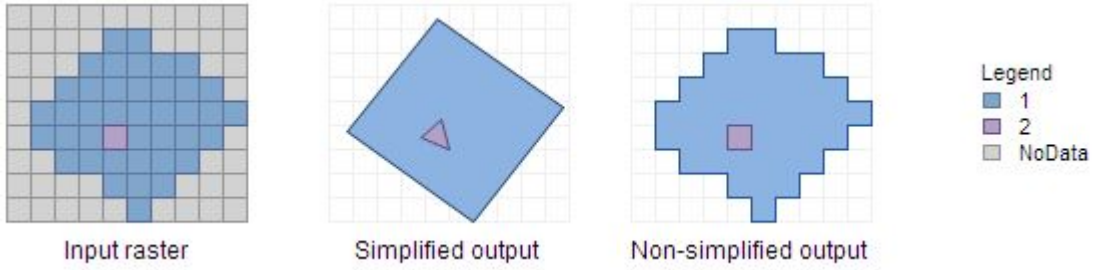
5. Raster can be converted to vector

(and vice versa)

You may want to convert your raster to vector:

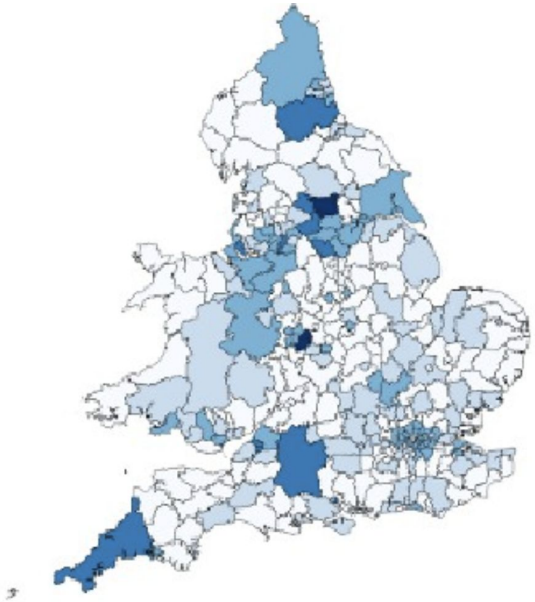
- Because some operations can only be carried out on vector data
- Because raster data uses a lot of storage space

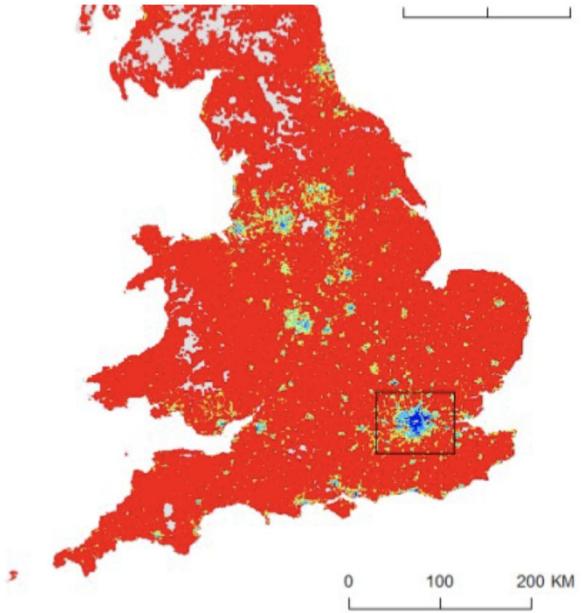
There are various options such as simplifying or not the output



6. Examples: Population grids

These two maps represent population density. On the left, the unit is output areas. On the right, it's a 1kmx1km grid. You can notice that these two maps depict a very different image. Rural areas look especially bad when aggregating data into output areas.





6. Examples: Population grids

Population is often represented as **vector**, where population is grouped in **output areas** (census, etc)

But there are many advantages to representing population as raster grids:

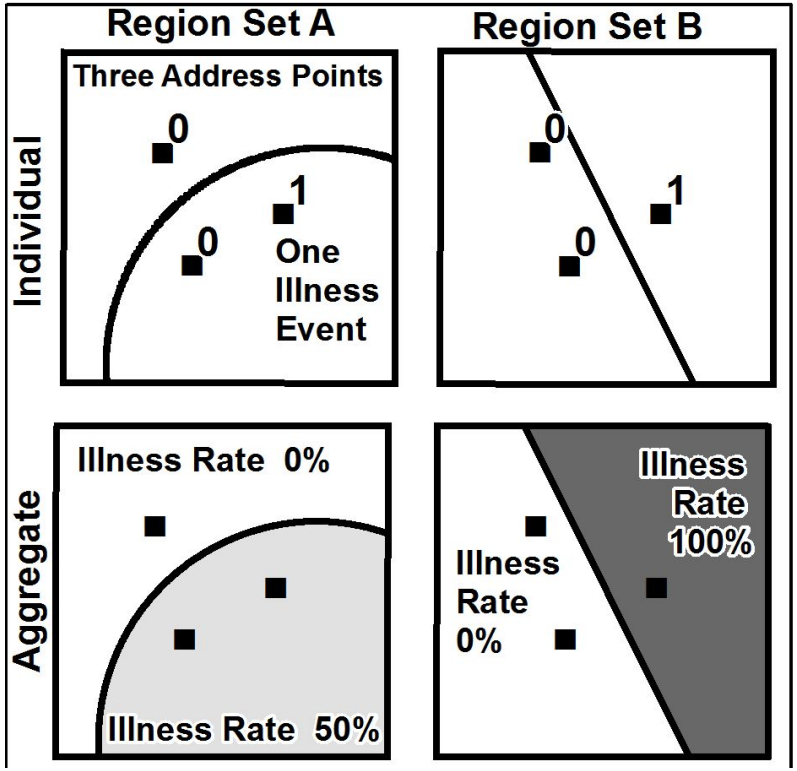
- They **minimize ecological fallacy and MAUP** (see Martin et al (2001) The Application of Zone-Design Methodology in the 2001 UK Census, Environment and Planning A, 33:11, pp. 1949 – 1962, https://doi.org/10.1068/a3497)

- We can clearly see where **nobody lives**

- They allow us to easily **calculate change over time**. On the contrary, output area boundaries are redefined frequently.

A famous population grid dataset is the one released by Facebook; you can read up on the methodology here.

Reminder: Modifiable Area Unit Problem - MAUP



6. Examples: Population grids

https://popchange.liverpool.ac.uk/

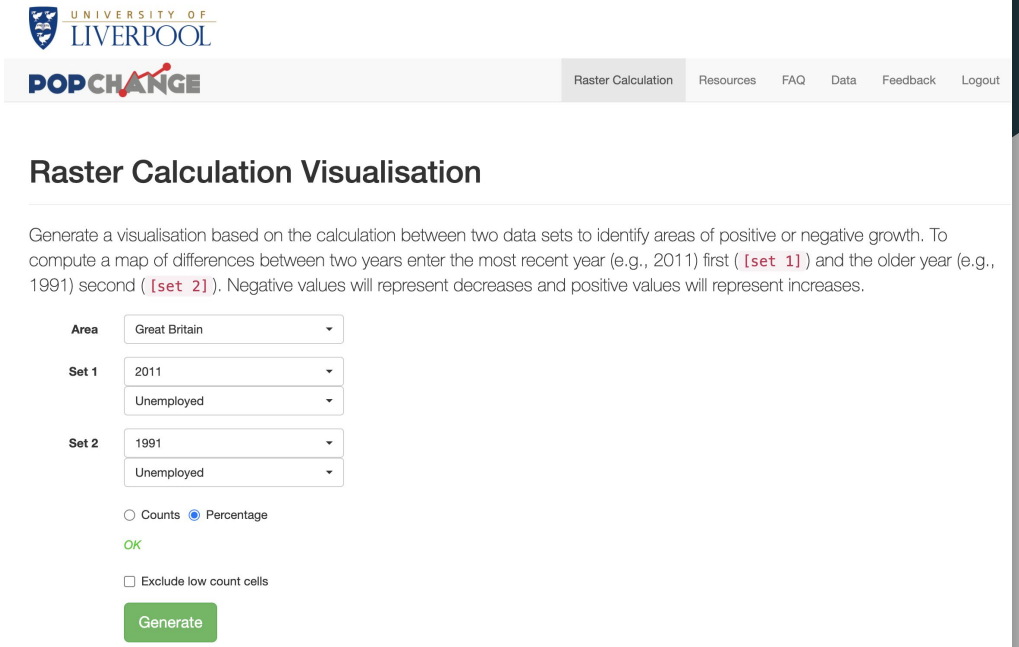
(you need to register)

Researchers at the University of Liverpool have taken census data since 1971 and rasterized it into 1km grid. The purpose is to be able to compare **the evolution of a variable across time**, regardless of how much the admin area boundaries changed.

6. Examples: Population grids

You can select an area (here I picked all of Great Britain), a variable you're interested in, and two dates to compare.

You can choose to see absolute values or percentage.

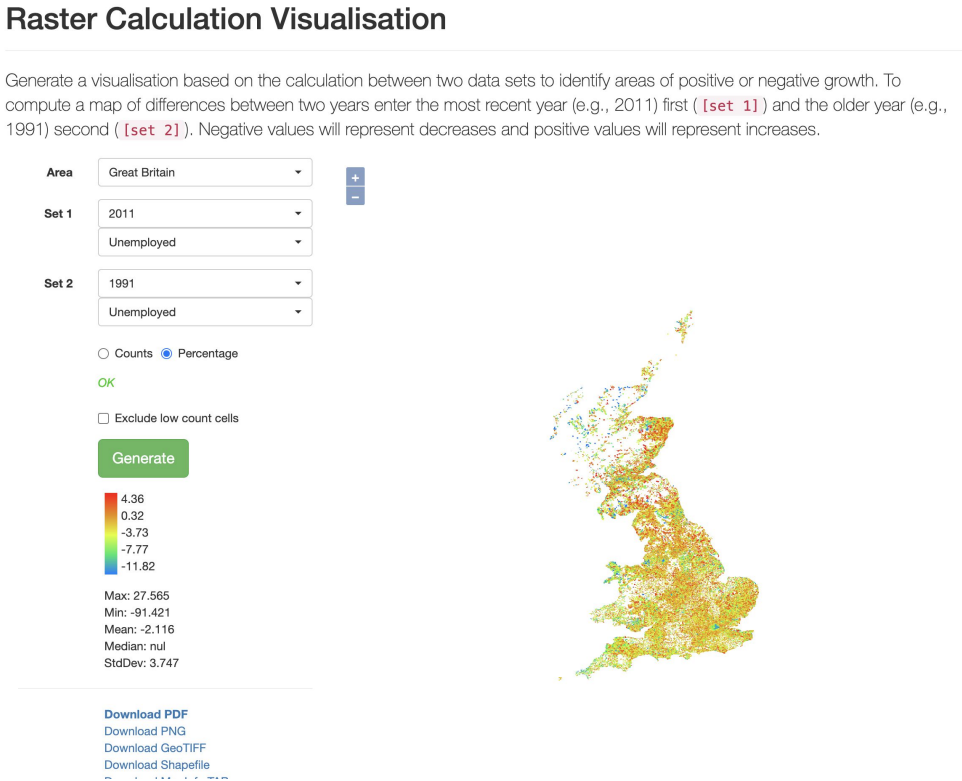


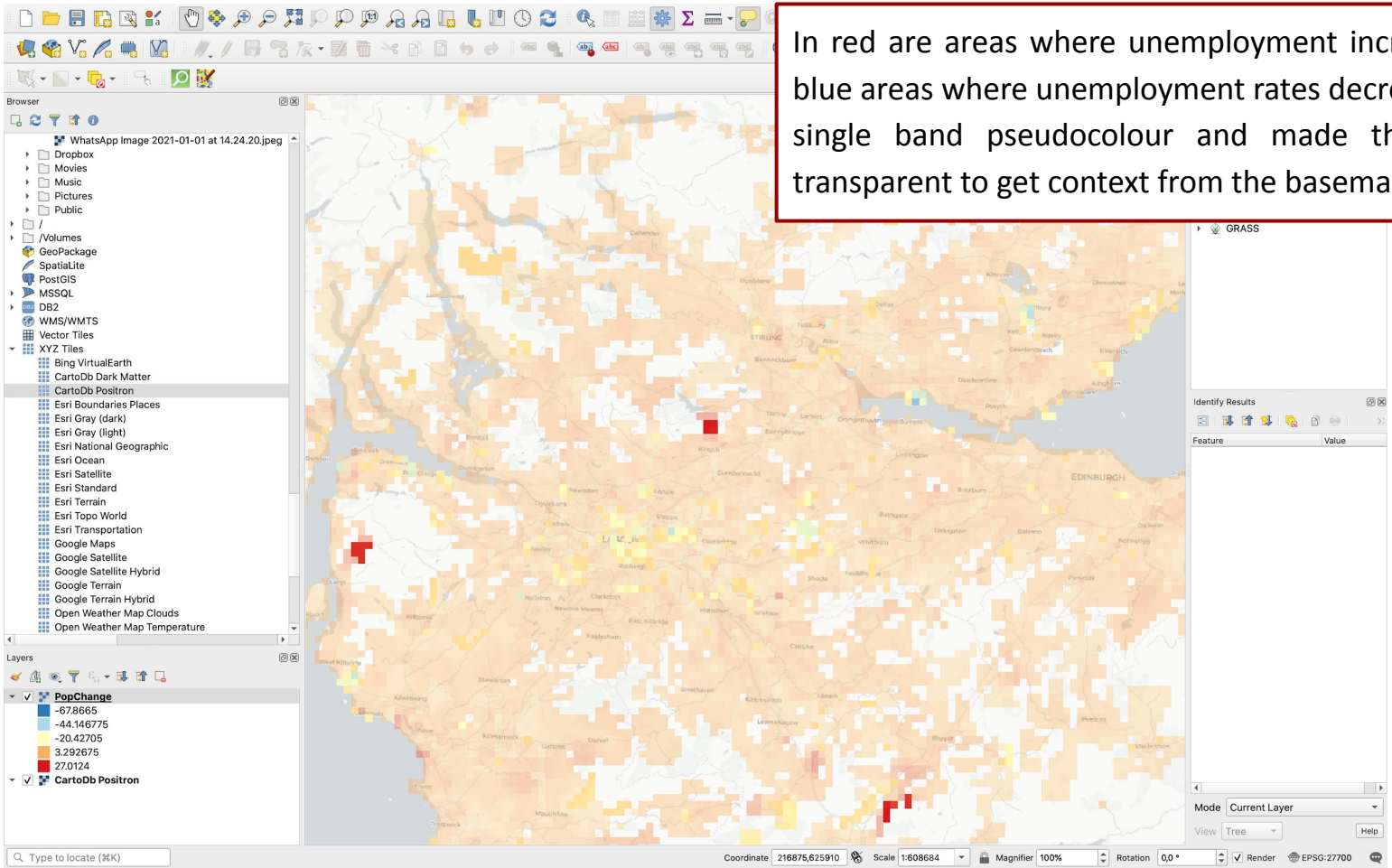
E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL
SDAI
PopChange was funded by the Economic and Social Research Council (ESRC) under the Secondary Data Analysis Initiative (SDAI), Phase 2 (project ES/L014769/1) and this support is acknowledged gratefully.
CENTRE FOR SPATIAL DEMOGRAPHIC RESEARCH

6. Examples: Population grids

The output is a layer that contains the DIFFERENCE between those two dates (i.e. how the percentage of unemployed people in the UK changed between 1991 and 2011.

You can also download the layer as GeoTIFF and use them in QGIS.





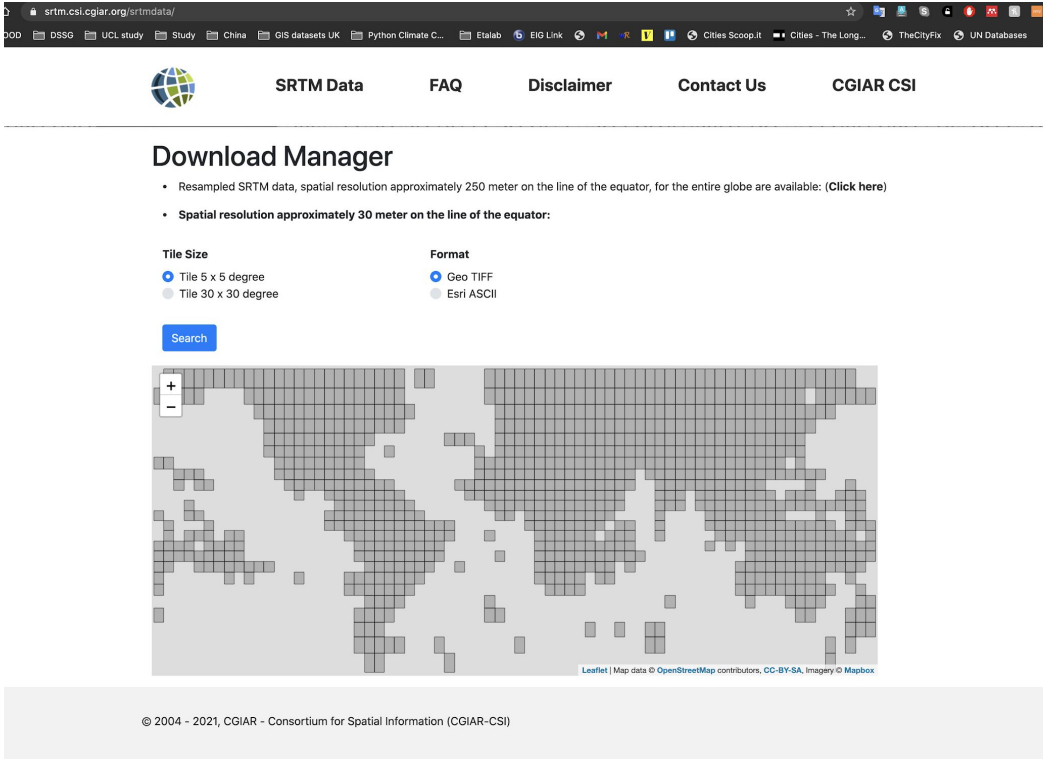
In red are areas where unemployment increased a lot, in blue areas where unemployment rates decreased. We used single band pseudocolour and made the layer a bit transparent to get context from the basemap.

7. Example: elevation raster from SRTM

Elevation data: the SRTM dataset.

In 2000, NASA's **S**huttle **R**adar **T**opography **M**ission collected the most complete high-resolution digital topographic database of Earth (later topped by ASTER GDEM in 2009). It covers most of the planet, with a resolution of roughly 30 meters.

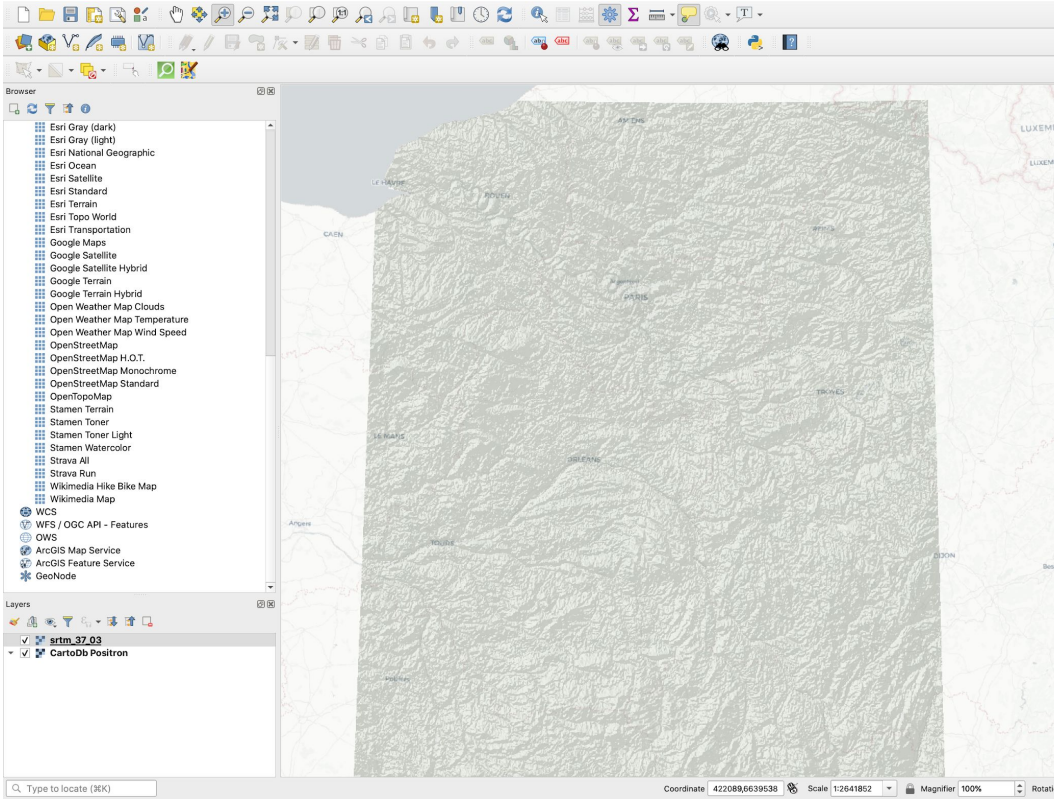
You can download tiles here: https://srtm.csi.cgiar.org/srtmdata/



7. Example: elevation raster from SRTM

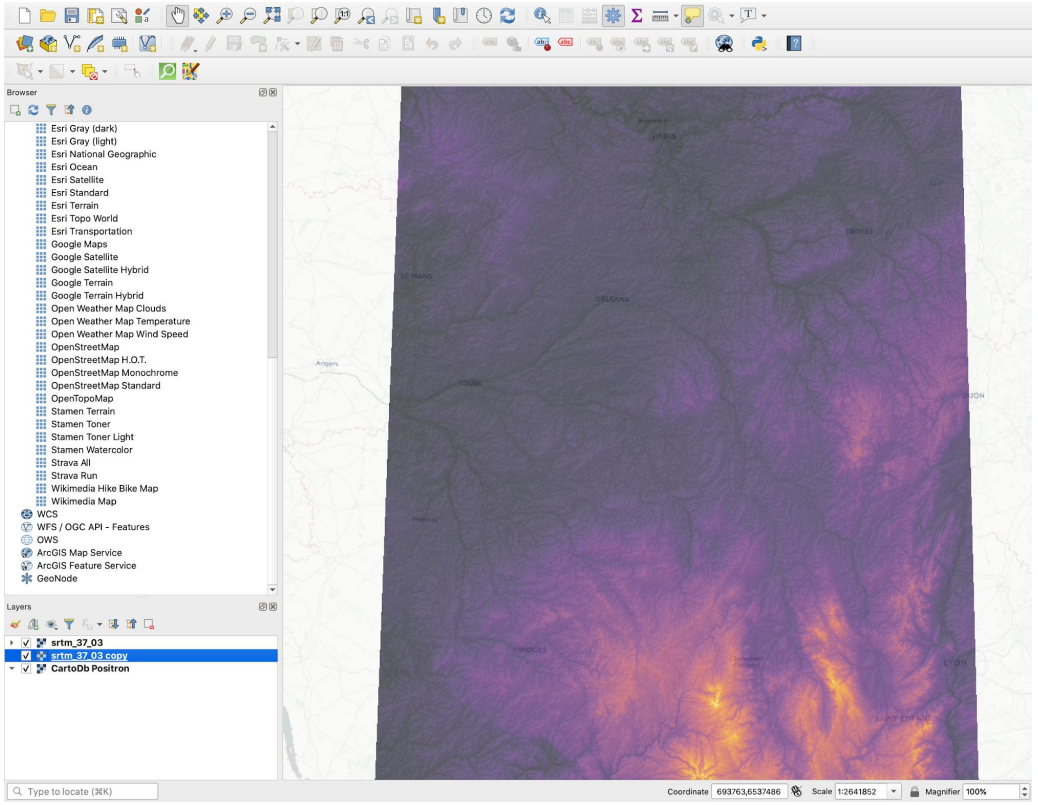
Several symbology options for elevation data.

This is hillshade ; it simulates the shadows of the sun on the relief.



7. Example: elevation raster from SRTM

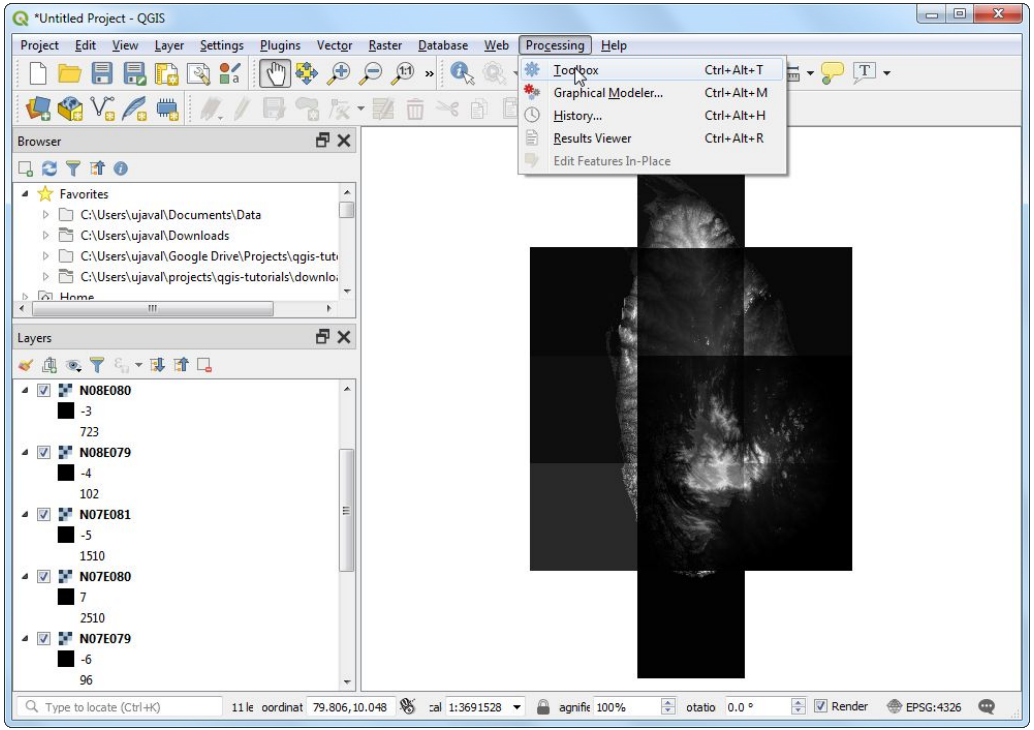
You can also duplicate the layer and combine two different symbologies. For instance here I used singleband pseudo-colour and hillshade, and played with transparency of each layer.



7. Example: elevation raster from SRTM

If you are interested in using topography data, you can give this tutorial a try and learn how to build a raster mosaic to cover a larger extent, then clip it to the shape of your area of interest.

http://www.qgistutorials.com/en/docs/3/raster_mosaicing_and_clipping.html



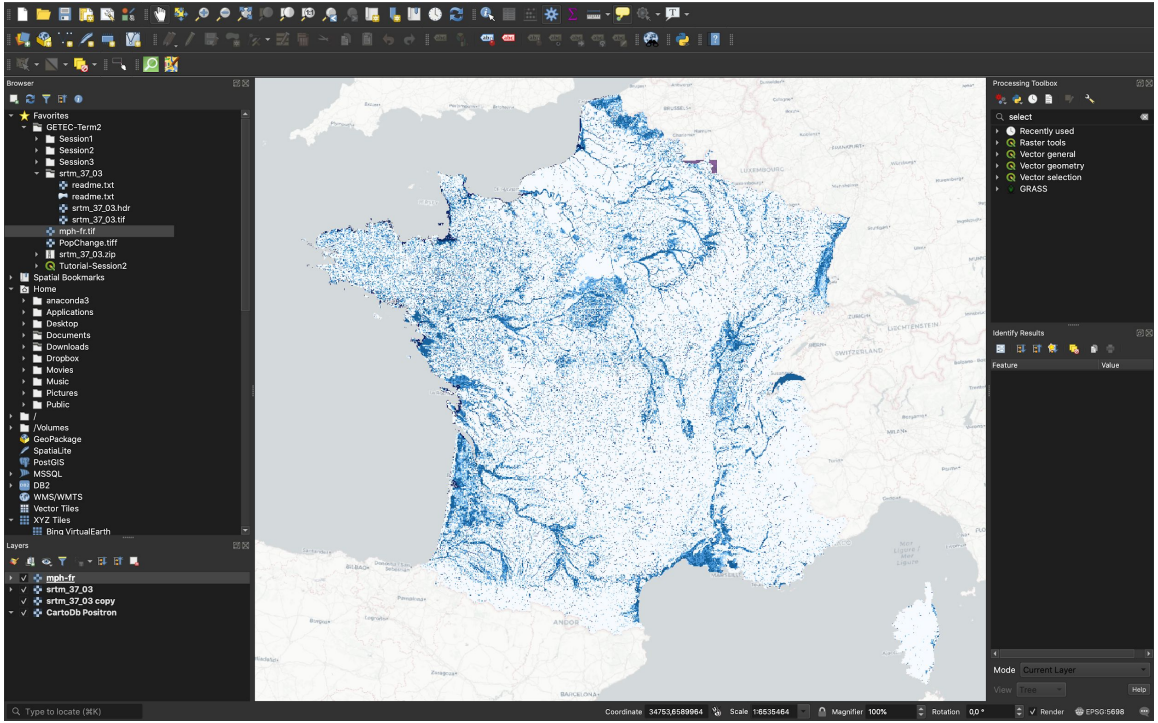
8. Example: Humid zones in France

Humid zones in France

https://geo.data.gouv.fr/fr/datasets/c2645280e85133c68fa7179b6b1feda6bbba63dc

Each pixel holds a score of 0, 1, 2 or 3 depending on the probability that the soil is rich in water.

(0= very low probability that the soil contains water
3= highest probability that the soil contains water)

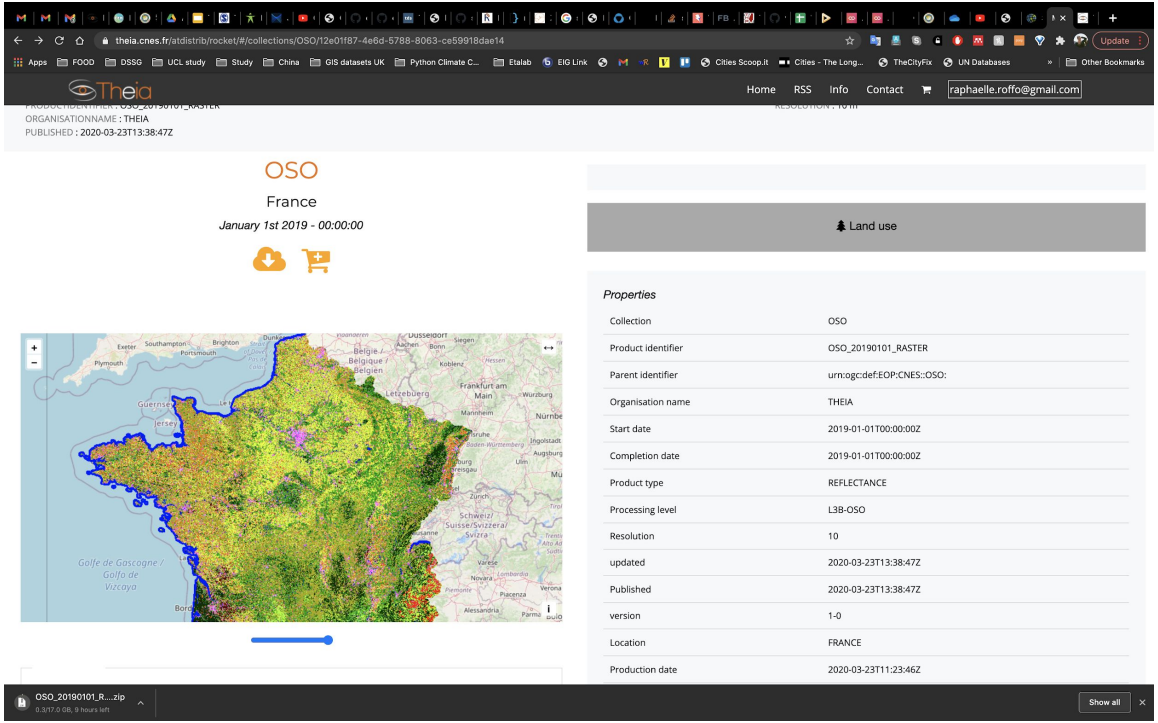


9. Example: land cover

You could download the Land Cover dataset on Theia (https://theia.cnes.fr/).

It uses reflectance to infer the nature of the surface (vegetation, concrete, water, etc)

Careful, due to their high resolution, these are VERY large datasets (17GB for France alone)



3. Georeferencing & Digitizing

Georeferencing

Many real life GIS projects require georeferencing some raster data. For example you may need to update the information for a forestry area for which you only have a paper inventory of the area. If you want to be able to look at the way the forest evolves over time, you will want to digitize that map and update it in the future. You may also have a satellite image that hasn't been georeferenced.

***Georeferencing* = assigning real-world coordinates to each pixel of a raster image, allowing it to be viewed, queried and analyzed with your other geographic data.**

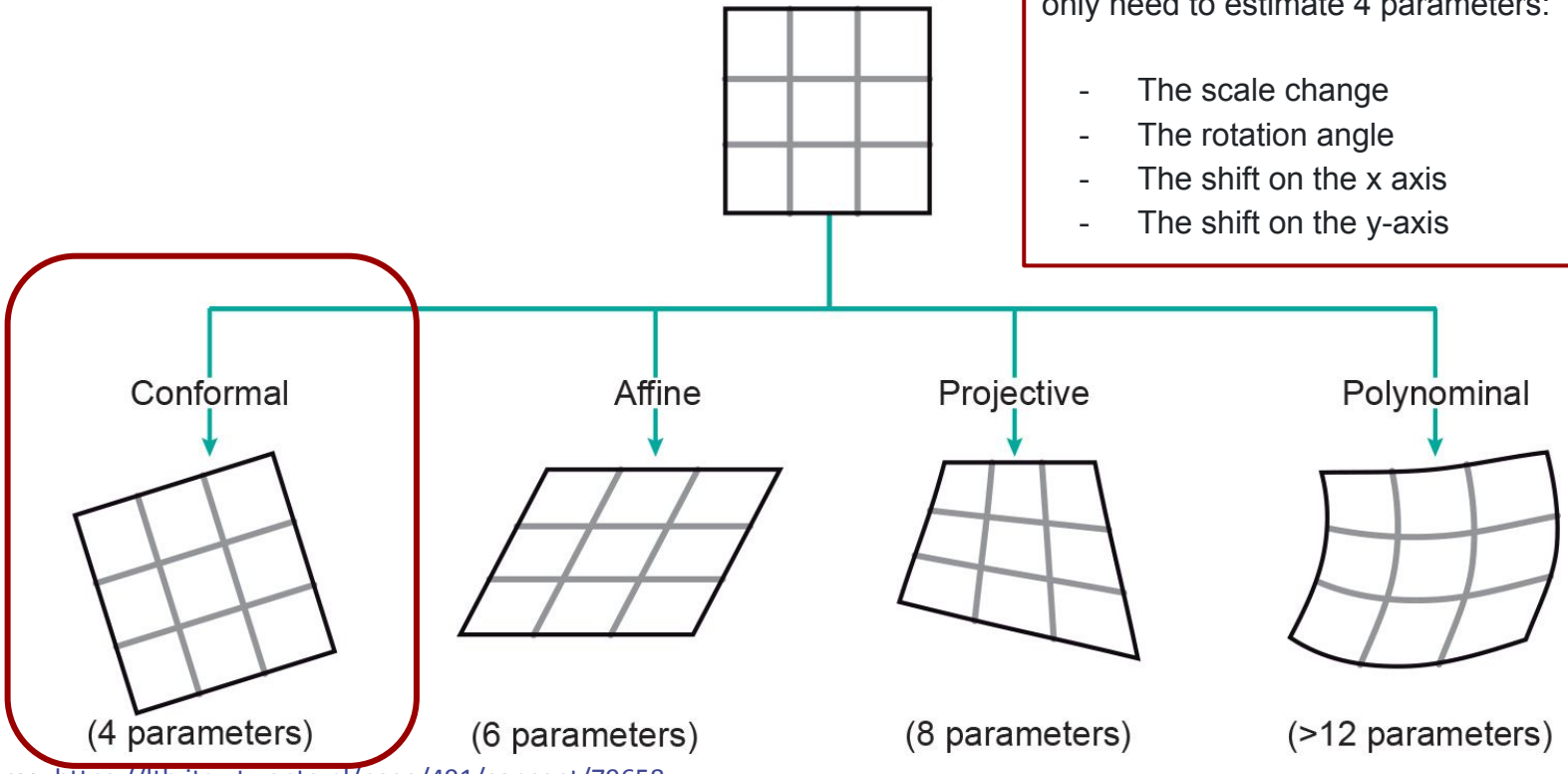
To do so, you will first scan that paper map (or get a digital picture of a map of your choice). Then, you will identify a few points (at least 3) for which you have precise coordinates (because you have those coordinates, or because you can identify those points on a basemap).

Finally, you can "pin" those Ground Control Points (GCP), and choose the way you want to warp the image to fit your coordinate reference system.

Georeferencing

In a **conformal or Helmert transformation**, the shape and angles of your image are retained; to describe your transformation in an equation you only need to estimate 4 parameters:

- The scale change
- The rotation angle
- The shift on the x axis
- The shift on the y-axis

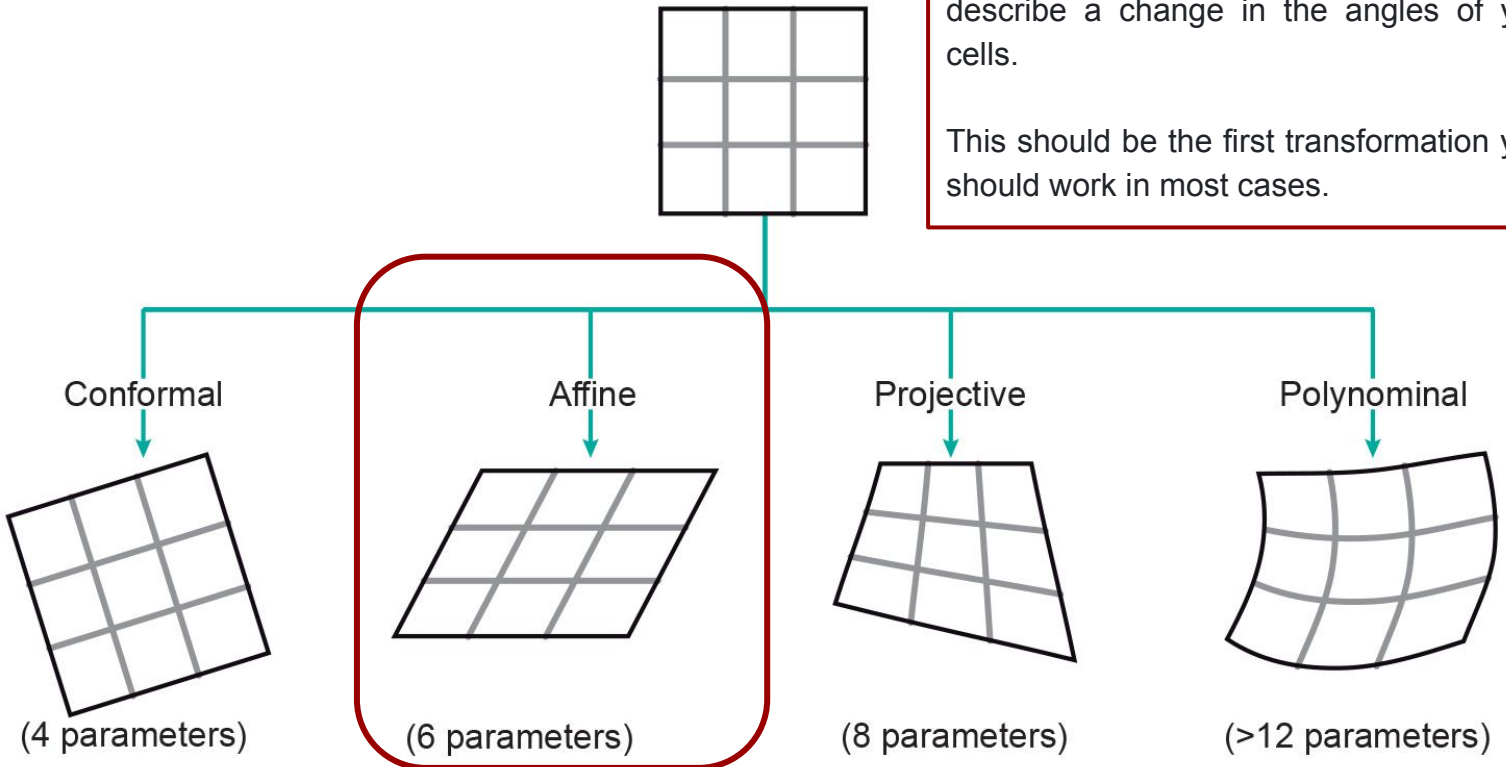


Source: https://itb.itc.utwente.nl/page/491/concept/79658

Georeferencing

With **affine (= Polynomial 1) transformation**, you introduce another set of parameters to describe a change in the angles of your raster cells.

This should be the first transformation you try and should work in most cases.

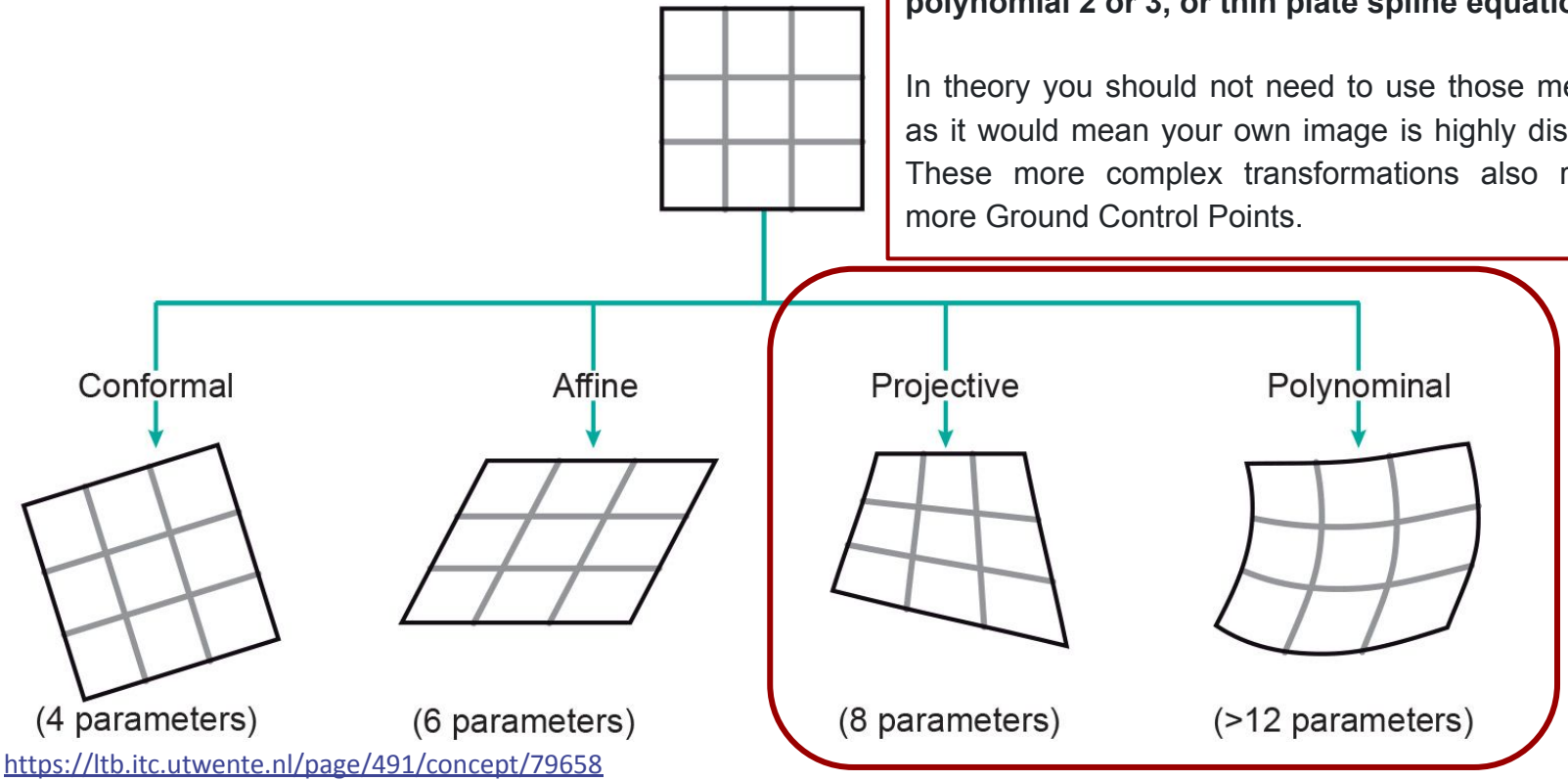


Source: https://ltb.itc.utwente.nl/page/491/concept/79658

Georeferencing

But if you need to warp or stretch your image to make it fit onto your basemap, then you will need a more complex transformation, described by **projective, polynomial 2 or 3, or thin plate spline equations.**

In theory you should not need to use those methods as it would mean your own image is highly distorted. These more complex transformations also require more Ground Control Points.

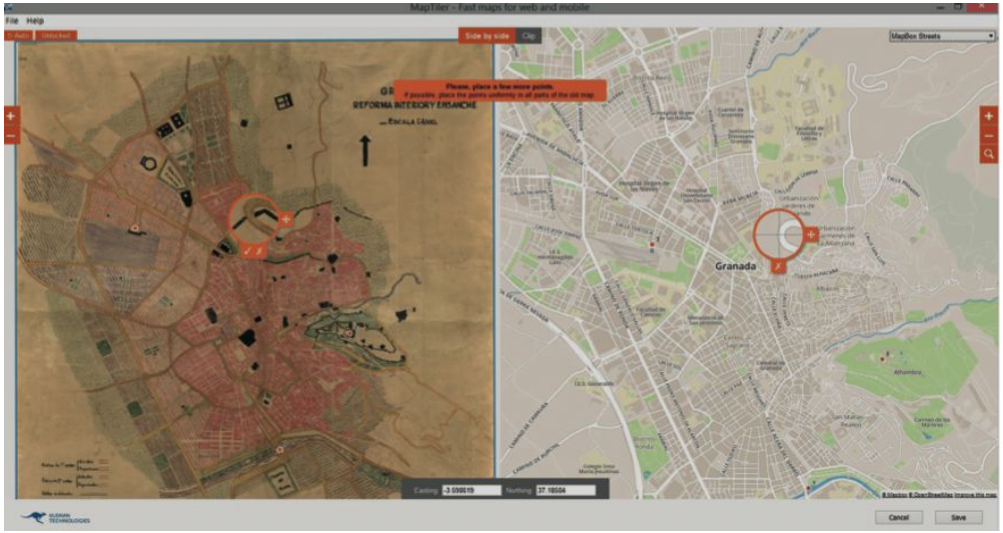


Source: https://ltb.itc.utwente.nl/page/491/concept/79658

Georeferencing: use cases

Georeferencing allows to revitalise and give new purpose to ancient cartography that remains in archive deposits, libraries and cultural centres.

This facilitates access for non-expert users to geographical, heritage, urban or environmental information of interest. Making those maps machine-readable and queryable also allows for interactivity: web mapping of photographs, mobile geolocation on old maps etc.



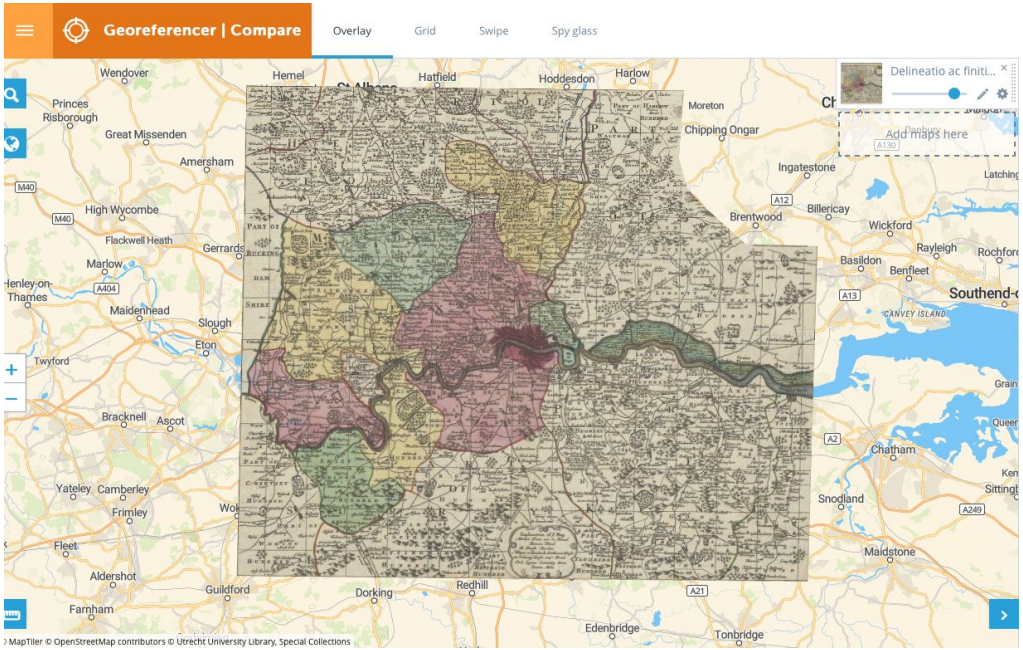
Source: https://revista.profesionaldelainformacion.com/index.php/EPI/article/viewFile/epi.2018.ene.19/38573

Georeferencing: use cases

See: https://www.oldmapsonline.org

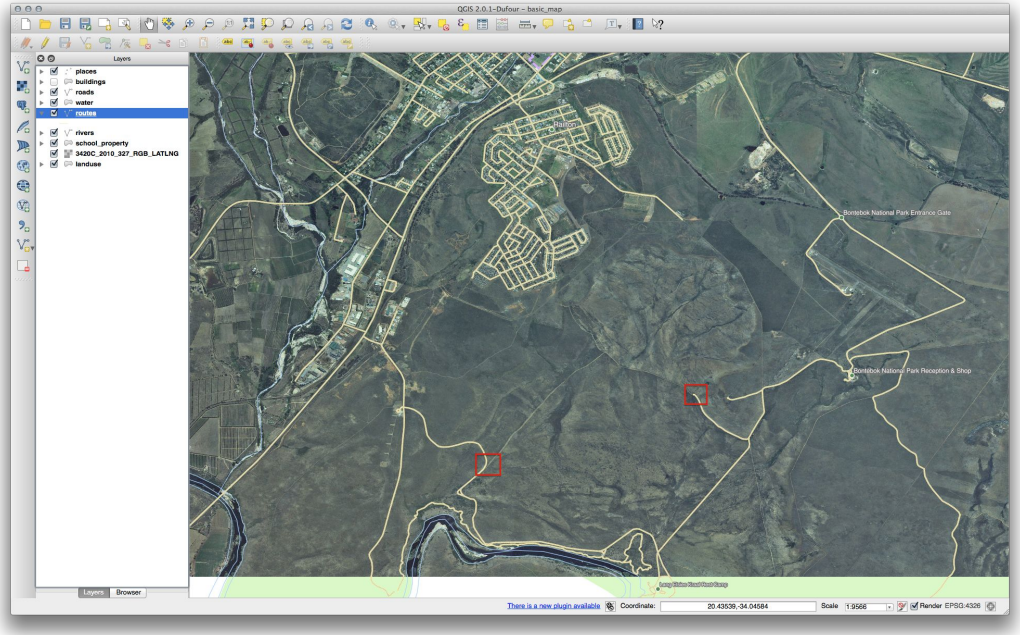
In this example, you might then be interested in extracting the boundaries of the boroughs in this ancient map in a new vector dataset.

In order to do so, you would use digitizing to draw the vertices of the polygons and save them in a new vector dataset (or update an existing one)



Digitizing

Digitization is an even more common task in GIS workflows. When working with environmental data in particular, it happens that you do not have access to recent data. A way to create your own data is to import satellite images into QGIS, and draw the boundaries of the features you're interested in (Points, Lines or Polygons). You can thereby create your own dataset and populate the attribute table with the information you want.



Next week

Homework

- Do the Session 2 Tutorial
- If you want to explore raster symbology, you can try and import some of the example datasets listed on these slides (Popchange, HumidZones, land cover) into QGIS and adjust the symbology.
- You can also try the SRTM tutorial, with an area of your choice!