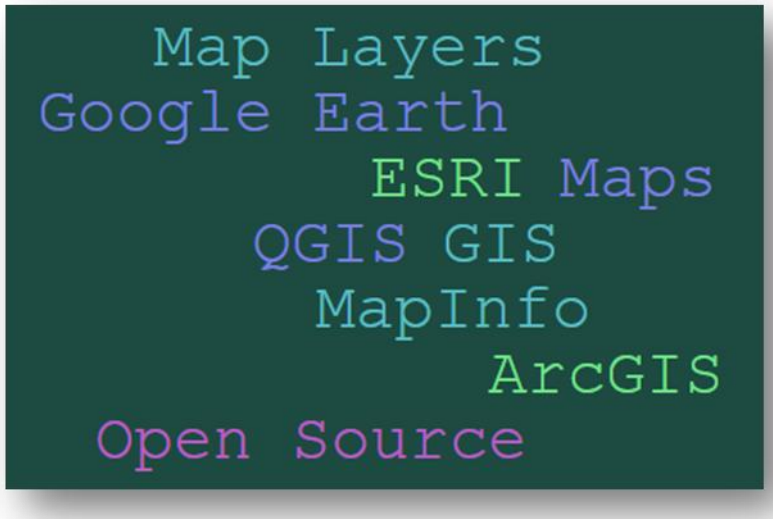
QGIS: An introductory workshop for biological recorders

Led by ***Rich Burkmar***

FSC Preston Montford

Thursday 28th May 2015



Terminology & concepts

GIS Software

GIS stands for Geographic Information System – a class of computer software which deals with spatial data, or to put it plainly, maps on computers.

Among the commercial GIS products that you have probably come across are MapInfo and ESRI's ArcGIS (sometimes people just refer to ArcGIS as 'ESRI'). These are probably the market-leading GIS software – particularly in the environment sector – and have been around for at least a couple of decades.

But in recent years the domination of these commercial products has been challenged by a new kid on the block – QGIS. The name QGIS evolved from Quantum GIS but the 'Quantum' part has officially been dropped. I say it like this - "kew-jiss" - but you can say tomato if you like!

Unlike MapInfo and ArcGIS, QGIS is not sold commercially – it is developed as a collaborative 'open source' product and is available free of charge. But don't make the mistake of thinking that means it is any less functional or 'less serious' than the others; QGIS is comparable to ArcGIS and MapInfo and, in many respects, surpasses them.

Google Earth – also a kind of GIS, but with much less analytical capability than the others we have mentioned – is also freely available. But unlike QGIS, it is *not* an open source development and so there's nowhere near as much innovation and collaborative development going on around it although it is a very popular piece of software because of its excellent visualisation capabilities.

A feature common to all GIS is that of working with 'map layers'. A single map can be composed of data from many different map layers – for example a layer showing roads, another with rivers and another with semi-natural habitat. It is the flexible ways in which GIS source, create, visualise, combine, analyse and navigate maps and map layers that make them what they are.

GIS File formats and georeferencing

Map layers in GIS are normally represented by files on a file system and these can occur in many different file formats.

The standard format for ArcGIS is the Shapefile (file extension '.shp'). The Shapfile is also the preferred format of QGIS; when you create new map layers in QGIS from scratch, you normally create them as Shapefiles.

MapInfo's standard data format is known as Tab File, or just TAB (file extension '.tab'). MapInfo also has a popular data exchange format know as MIF or MIF/MID (MIF stands for MapInfo Interchange File). QGIS can open Tab files and MIF/MID files. QGIS can also save layers as TAB or MIF/MID.

Most GIS, including MapInfo and ArcGIS, support more data formats than just their 'own' ones, but QGIS is remarkable for the large number of data formats that it supports.

The main data format for Google Earth is KML (stands for Keyhole Markup Language) and KMZ (which is a compressed form of KML). QGIS can open KML files and save layers in the KML format too.

GPS (Global Positioning Systems) can record tracks and routes (linear data) and waypoints (point data) and usually encode these in a standard format called GPX (which stands for GPS Exchange format). QGIS can both read and write layers in the GPX format.

CSV stands for Comma Separated Values and it is a widely used format in computing in general. In particular, spreadsheets – such as Excel – can normally export data in a CSV format. Most GIS can import CSV data and show records on a map as long as each line (each record) is georeferenced.

If a piece of data is georeferenced, it means that information is attached to it so indicate a location on the earth to which it is associated. This information could be an OS grid reference, easting & northing or latitude & longitude (often referred to as lat/lng or lat/long). A photo taken with a smartphone or modern camera that includes GPS will normally be georeferenced (also called geotagged). The process of assigning a georeference to an item (e.g. an address) is often called geocoding.

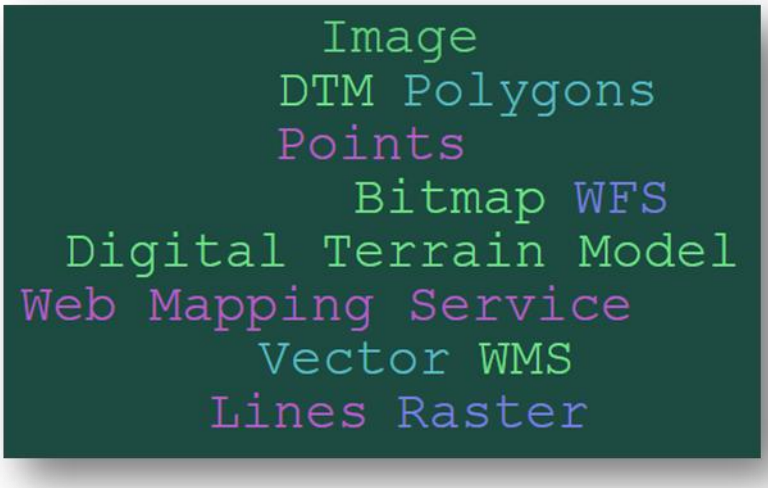
Classes of map layers

There are two major classes of map data common to all GIS: raster and vector.

Although it's an over-simplification, raster layers can be thought of as 'dumb images' – although the GIS can display them in the right place on earth, the lines, points polgyons and other features within a raster layer don't mean anything to the GIS – they are just pixels within an image (sometimes called bitmaps). They are very useful to us as viewers of them in GIS though, because our brains know how to interpret maps.

Vector layers are more meaningful to the GIS since it understands the location of every point which forms a line or the boundary of a polygon. It can, for example, intersect two polygons and construct a new polygon which is the area common to both. Or it can, for example, create a polygon by buffering a point or a line.

To these two main types we could add others. For example DTMs (Digital Terrain Models) which store information about landscape elevation. In fact these can be encoded either as raster or vector data, but the essential distinguishing factor is that each point represented in a DTM encodes not only its location on the earth (x and y), but its elevation too (z).

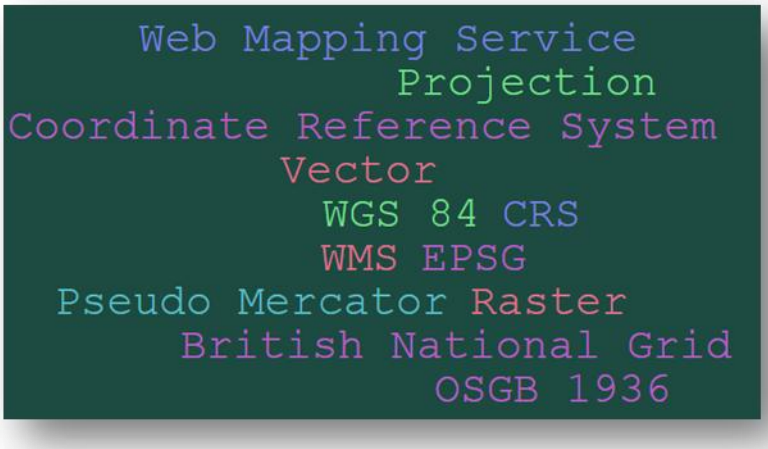


In recent years more and more GIS data is being made available over the internet via WMS (Web Mapping Services) and WFS (Web Feature Services). These serve up data into a GIS over a live internet connection. Every time you pan or zoom the map, your GIS sends the map server a request for the new area and this is sent into your GIS over the internet as a raster image. So you don't have the layers stored on your computer – you rely on a good internet connection to get them. The difference between WMS and WFS is essentially the same as the difference between raster and vector: WMS get raster data over the internet and WFS get vector data over the internet.

Projections

A problem in cartography which is as old as maps is that of representing a spherical (or rather near spherical) surface (that of the earth) on a flat surface – be it paper or a computer screen. This is known as projection. Map projections necessarily distort reality and depending on where you are in the world, and the purpose of your map, different projections are appropriate. That's why there are so many.

QGIS, projections are generally referred to by the acronym CRS (Coordinate Reference System). In the UK we commonly favour the CRS used by the Ordnance Survey which QGS knows as 'OSGB 1936 / British National Grid'. QGIS also uses an international referencing system known as EPSG (European Petroleum Survey Group) to refer to projections. The EPSG code for OSGB 1936 is EPSG:27700.



Because of the ubiquity of GPS, most of us are also very familiar with WGS84 (World Geodetic System 1984) which is used by GPS. The EPSG code for WGS 84 is EPSG:4326. You will also notice that layers from WMS like Bing and Google are projected in QGIS with WGS 84 / Pseudo Mercator (EPSG:3857).

The EPSG codes for these three projections that you are likely to commonly encounter are given below:

- OSGB 1936 / British National Grid – EPSG:27700
- WGS 84 – EPSG:4326
- WGS 84 / Pseudo Mercator – EPSG:3857

Sources of basemap and other data

The workshop covers access to National Biodiversity Network (NBN) Gateway data which is accessible in QGIS through the Tom.bio NBN plugin. We also cover the wealth of aerial photography and other basemap data available in QGIS through the 'OpenLayers' plugin. But there are many other sources of free data – a quick search of the internet for 'free GIS data' will turn up many such sources.

When downloading data, remember that QGIS can handle almost any format but it is especially well-geared up to use Shapefiles. So if presented with a choice of format, it's normally best to go with Shapefile. But if someone sends you data in another format, the chances are that QGIS will handle it.

Some other sources of data are described below.

Ordnance Survey

Many people don't realise the extent to which Ordnance Survey (OS) have opened up access to their data. Although their flagship products like 1:50,000 and 1:25,000 raster basemaps and the Mastermap vector data are *not* available for free, they do provide some other less-detailed datasets free of charge.

For full details of what is available in their 'OpenData' suite of products, and online form for ordering the free data, visit this website:

https://www.ordnancesurvey.co.uk/opendatadownload/products.html

A particularly useful product that can be used to provide good basemaps against which you can display your own data at quite a large-scale (small area) is the 'Vectormap District' set. Confusingly, this is available both in vector format (as the name would suggest) and raster format.

The vector 'Vectormap District' data is more flexible but it requires styling. (Some Vectormap District style files are available – I think Lutra Computing have done something.) The raster 'Vectormap District' tiles are ready to use 'straight out of the box' (as you've seen on this course) and can be more convenient.

Natural England, Environment Agency and NRW

Natural England (and other country agencies) make much of their GIS data available. To download GIS layers from NE, you must first register, but it is painless. See here:

http://www.gis.naturalengland.org.uk/pubs/gis/GIS_register.asp

This EA site offers a wealth of GIS data for download:

http://www.geostore.com/environment-agency/WebStore?xml=environment-agency/xml/ogcDataDownload.xml

The following CCW site is referred to from the NRW website:

http://www.ccw.gov.uk/landscape--wildlife/protecting-our-landscape/gis-download---welcome/gis-boundary-downloads.aspx

Many agency datasets are also available for download from the Magic website:

http://magic.defra.gov.uk/Dataset_Download_Summary.htm

Forestry Commission

GIS data from the Forestry Commission can be obtained here:

http://www.forestry.gov.uk/datadownload

British Geological Survey

GIS data for download can be obtained here:

http://www.bgs.ac.uk/downloads/browse.cfm?sec=6

An impressive list of Web Mapping Services (WMS) is available here:

http://www.bgs.ac.uk/data/services/wms.html

General catalogues of free GIS data

A few useful websites are indicated below, but you will find many more by using the internet:

http://freegisdata.rtwilson.com/

http://wiki.osgeo.org/wiki/Public_Geodata_for_the_UK

CSV data

Remember that you can use the Tom.bio 'Biological Records' QGIS plugin to display any CSV data geocoded with OS grid references or easting/northings (or lat/longs) – it doesn't have to be biological records.

After-course support and mentoring

A key aim of the FSC Tomorrow's Biodiversity project – under which this training is being provided – is to explore the utility of providing follow-up support and mentoring to learners after they leave the classroom. You should consider me as a QGIS mentor to you.

Therefore if, after your training, you have questions about using QGIS or want to discuss any ideas or your own learning needs, you should feel free to contact me, your trainer, Rich Burkmar. I can't always guarantee to get back to you right away, but I will do my best to respond as soon as possible. The best way to contact me is by email: richardb@field-studies-council.org.

For the Tom.bio QGIS plugin, there are a series of demonstration YouTube videos that you might find useful. You can find links to them all here: http://www.tombio.uk/qgis-plugin. Of course you can email me with specific questions about the tools (or even suggestions for enhancements!).

There is a lot of good QGIS training material on the web. But I think that one of the best places to start looking is on the official QGIS website itself: http://www2.qgis.org/en/docs/index.html

The QGIS Training Manual, which you can link to from the page above, is excellent. Note that the development path of QGIS is so rapid that the documentation often lags behind the latest version of the software. So, for example, at the time of writing QGIS is available at version 2.8 but the documentation on the website is associated with version 2.6. Even where the version of the software and documentation match, you will often find that the documentation is talking about features or showing pictures of an interface from an older version. Don't be put off by this, you can normally work things out with a bit of persistence.

QGIS doesn't have official forums, but it provides support through a number of email lists. To find out more about them, explore this page: https://www.qgis.org/en/site/forusers/support.html