Mapping 1: displaying geographical data with QGIS

Hands-on at NICAR 2014, Baltimore, Mar 1

Peter Aldhous

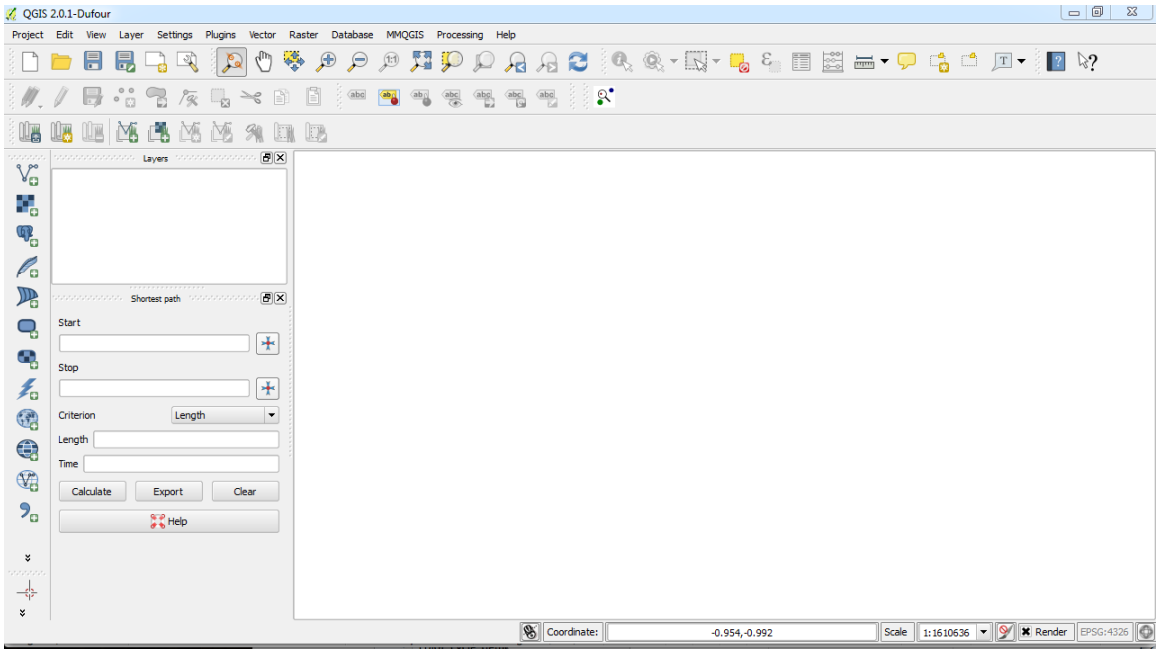
peter@peteraldhous.com

@paldhous

GQIS is the leading free, open source Geographic Information Systems (GIS) program. It is capable of sophisticated geodata processing and analysis, but you don't need to be a GIS expert to put it to effective use in displaying and processing geographic data for both print and online.

Here we'll learn how to use QGIS to make a simple thematic map, with areas colored according to data, plus a map of points, and how to export them as a vector graphic.

Launch **QGIS Desktop**, and you should see a screen like this:

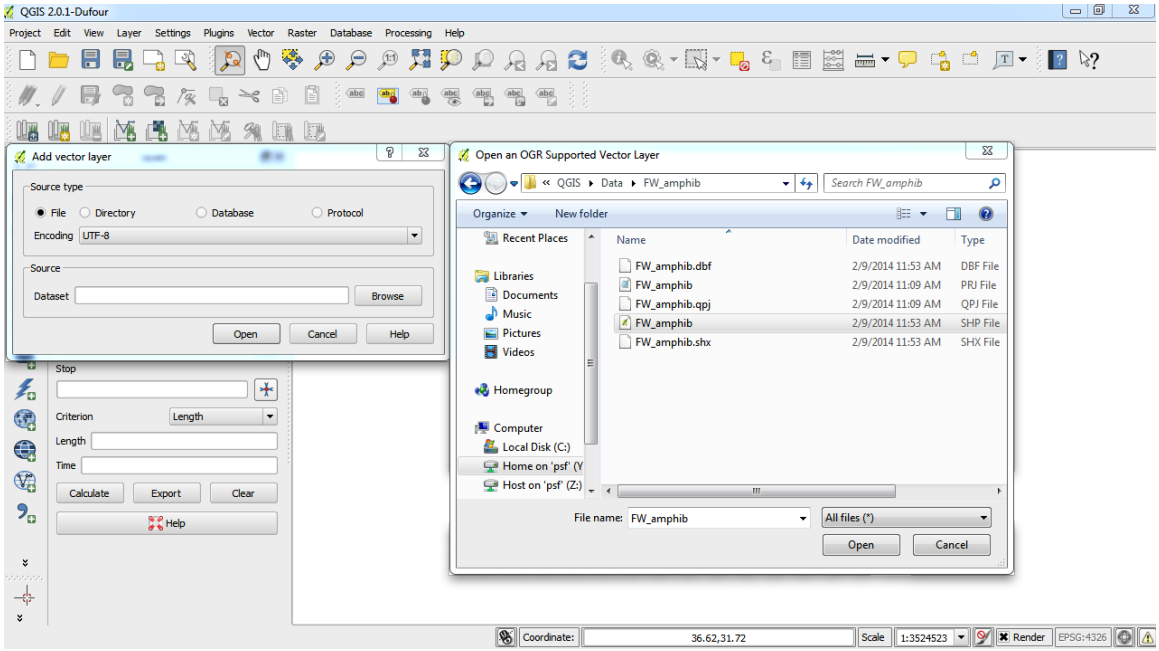


We're going to recreate a version of one of the maps from The Nature Conservancy's Atlas of Global Conservation – showing the number of globally threated amphibian species in each of the world's freshwater ecoregions.

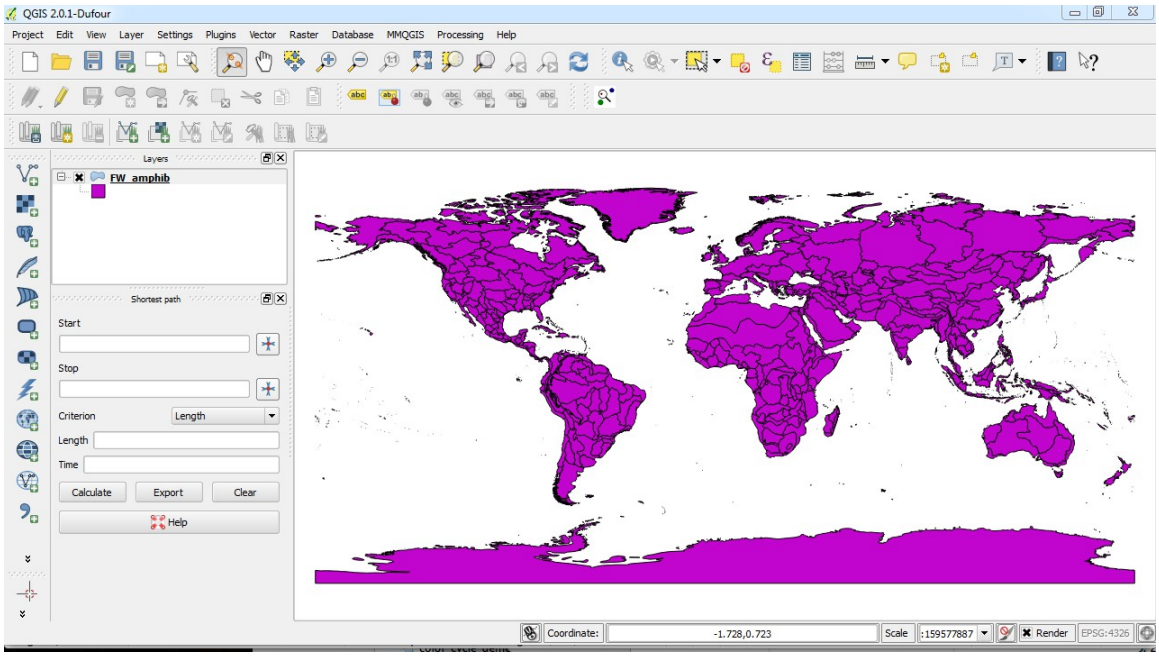
The data is here, in shapefile format, commonly used in GIS. A shapefile actually consists of a series of files, one of which (in a format called **DBF)** is a table of data. Many organizations, including the US Census Bureau, make their data available as shapefiles; Natural Earth also provides a great library of global shapefiles.

To import a shapefile into QGIS select **Layer>Add vector layer**, or click this icon: 

At the dialog box, click **Browse** and navigate to the **FW_amphib** shapefile, select the file with the type **SHP,** then **Open**:



Once the shapefile loads, you should have a screen like this:

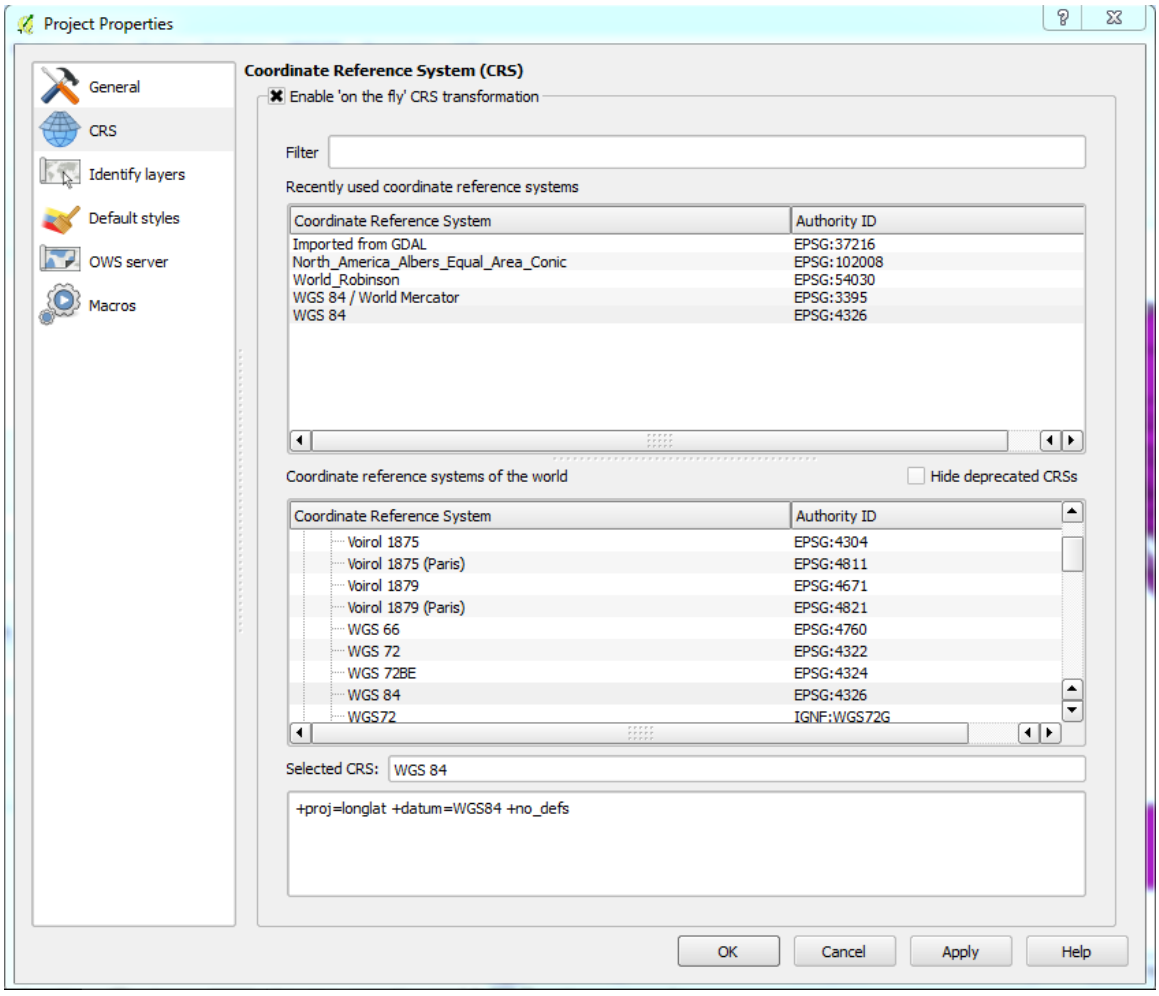


Notice how the name of the shapefile has appeared in the **Layers** panel, with a little square icon showing that it displays polygons – areas with boundaries. Now is a good time to save the project, so select **Project>Save** or use **Ctrl+S**.

This world map is plotted as rectangle with each degree of longitude and latitude given the same size. Usually, the first job in a GIS project is to set a specific map projection. Projections are important because any two-dimensional map is a distortion of reality: Just as you can't peel an orange and arrange the skin as a perfect rectangle, it's impossible to plot the Earth's surface in two dimensions and accurately represent distances, areas, shapes and directions.

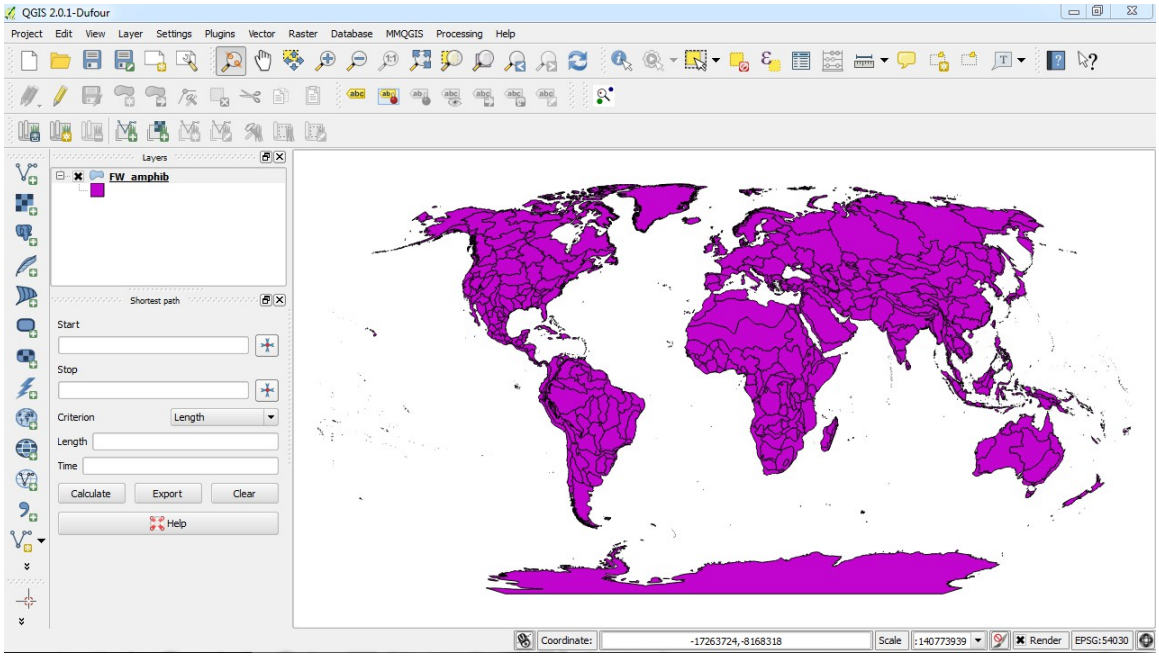
The best projection for your map will depend on which of these attributes is most important to conserve, and the area you're mapping. Your graphics department may be able to advise on which one to choose. (For more on projections, see this reference.)

To set the projection, select **Project>Project Properties>CRS** and check the box marked **Enable 'on the fly' CRS transformation**. (This will convert any data you import subsequently to the correct projection for the project.)



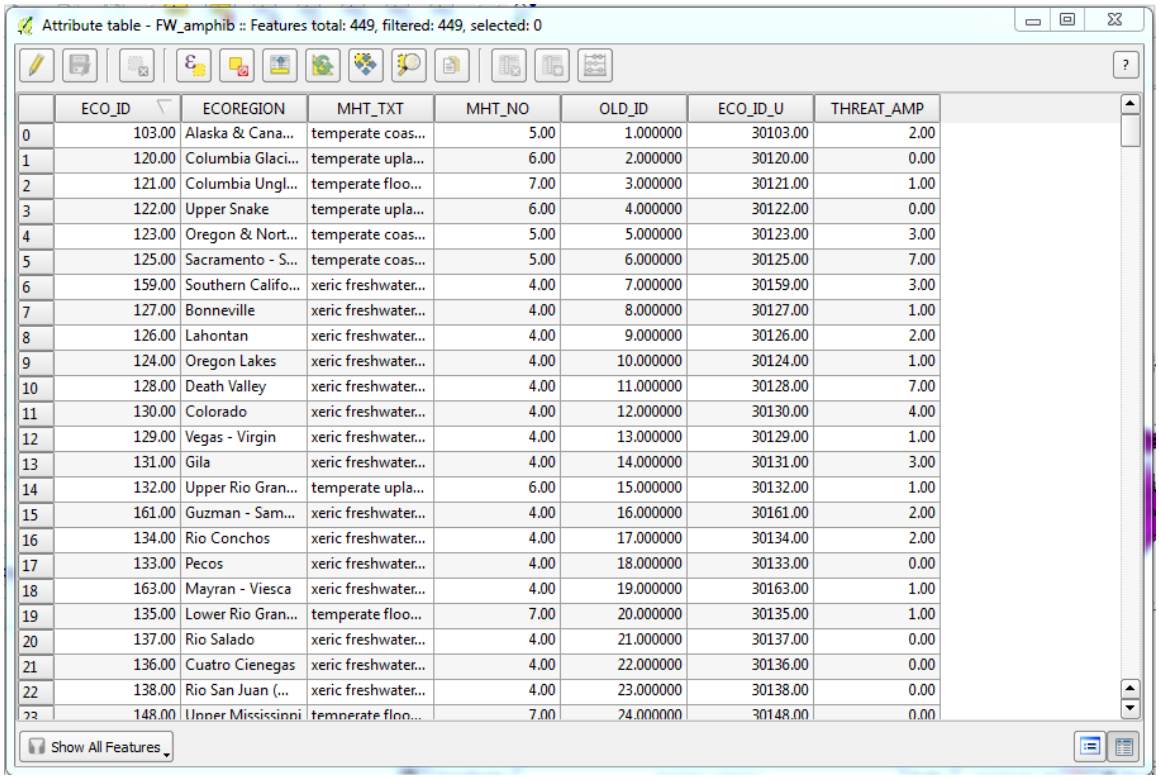
For this project, we'll use a Robinson projection. To find it, start typing Robinson into the box marked **Filter. World_Robinson** should appear in the box marked **Coordinate Reference System**: select it, then click **Apply** and **OK**. (You can also search for projections by their numerical codes, which you can find here.)

The map should now look like this:



Now let's look at the data available for us to map. Either right click on the name of the shapefile in the **Layers** panel and select **Open Attribute Table**, or click this icon: 

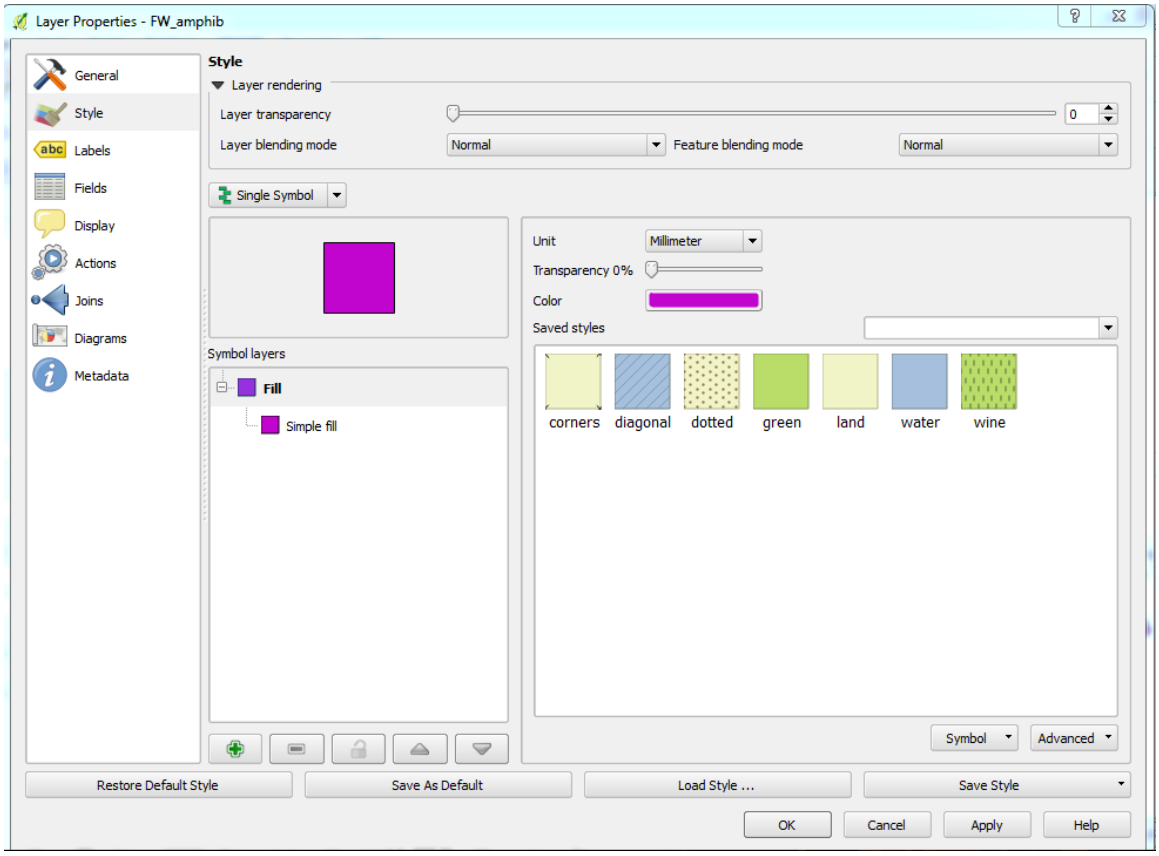
You should now see this data (from the shapefile's DBF):



	ECO_ID	ECOREGION	MHT_TXT	MHT_NO	OLD_ID	ECO_ID_U	THREAT_AMP
0	103.00	Alaska & Cana...	temperate coas...	5.00	1.000000	30103.00	2.00
1	120.00	Columbia Glaci...	temperate upla...	6.00	2.000000	30120.00	0.00
2	121.00	Columbia Ungl...	temperate floo...	7.00	3.000000	30121.00	1.00
3	122.00	Upper Snake	temperate upla...	6.00	4.000000	30122.00	0.00
4	123.00	Oregon & Nort...	temperate coas...	5.00	5.000000	30123.00	3.00
5	125.00	Sacramento - S...	temperate coas...	5.00	6.000000	30125.00	7.00
6	159.00	Southern Califo...	xeric freshwater...	4.00	7.000000	30159.00	3.00
7	127.00	Bonneville	xeric freshwater...	4.00	8.000000	30127.00	1.00
8	126.00	Lahontan	xeric freshwater...	4.00	9.000000	30126.00	2.00
9	124.00	Oregon Lakes	xeric freshwater...	4.00	10.000000	30124.00	1.00
10	128.00	Death Valley	xeric freshwater...	4.00	11.000000	30128.00	7.00
11	130.00	Colorado	xeric freshwater...	4.00	12.000000	30130.00	4.00
12	129.00	Vegas - Virgin	xeric freshwater...	4.00	13.000000	30129.00	1.00
13	131.00	Gila	xeric freshwater...	4.00	14.000000	30131.00	3.00
14	132.00	Upper Rio Gran...	temperate upla...	6.00	15.000000	30132.00	1.00
15	161.00	Guzman - Sam...	xeric freshwater...	4.00	16.000000	30161.00	2.00
16	134.00	Rio Conchos	xeric freshwater...	4.00	17.000000	30134.00	2.00
17	133.00	Pecos	xeric freshwater...	4.00	18.000000	30133.00	0.00
18	163.00	Mayran - Viesca	xeric freshwater...	4.00	19.000000	30163.00	1.00
19	135.00	Lower Rio Gran...	temperate floo...	7.00	20.000000	30135.00	1.00
20	137.00	Rio Salado	xeric freshwater...	4.00	21.000000	30137.00	0.00
21	136.00	Cuatro Cienegas	xeric freshwater...	4.00	22.000000	30136.00	0.00
22	138.00	Rio San Juan (...	xeric freshwater...	4.00	23.000000	30138.00	0.00
23	148.00	Upper Mississippi	temperate floo...	7.00	24.000000	30148.00	0.00

The data include various ID codes, the name of the ecoregion, and its type. The final column, marked **THREAT_AMP**, gives the number of threatened amphibians in each ecoregion. If you click on the column header it will sort the rows by that data, which ranges from 0 to 95. Close the attribute table.

Now we'll color the map according to the number of threatened amphibian species in each ecoregion. Click **Layer>Properties** or double click on the shapefile in the Layers panel, then select **Style**:

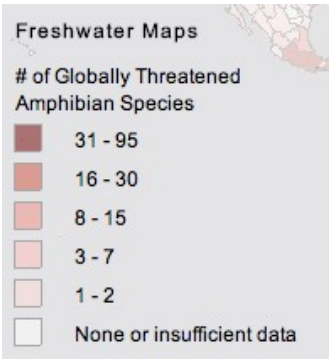


Click the little downward-pointing arrow next to **Single Symbol** and choose **Graduated**, which allows us to map by numerical data. (If we were mapping categories, such as the ecoregion type, we'd select **Categorized**.)

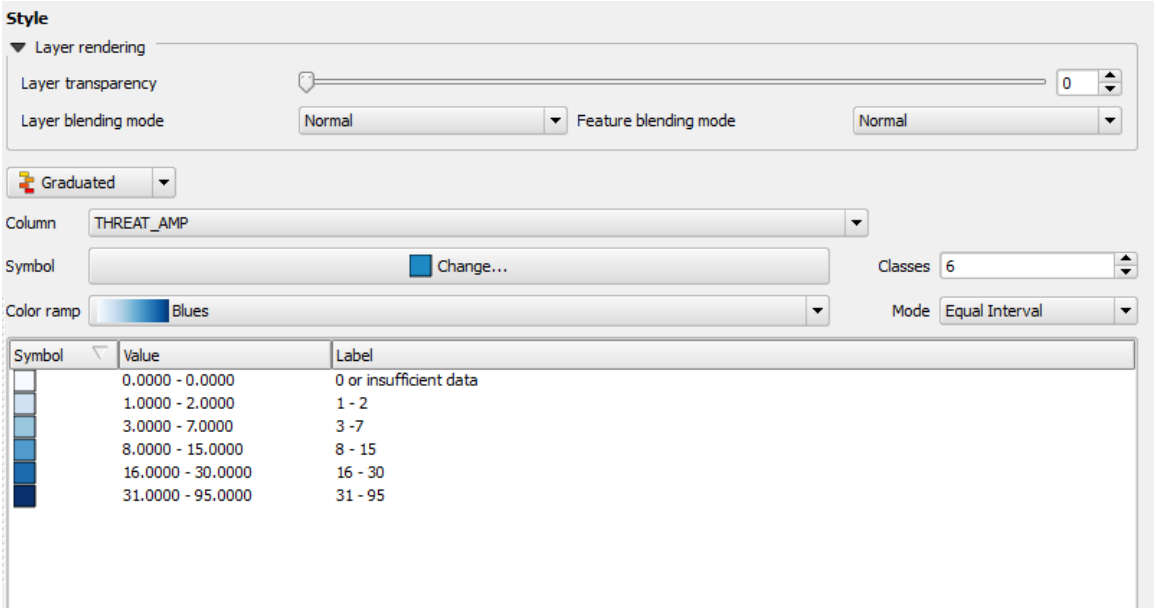
Then select **THREAT_AMP** under **Column**.

Using the **Mode** option instructs QGIS to divide the data in classes using some common defaults – for example, selecting 5 **Classes** and **Quantile** would put the lowest 20% of values in the first class, the next highest 20% in the second, and so on.

Here, however, we're going to follow the legend on the original map from The Nature Conservancy:

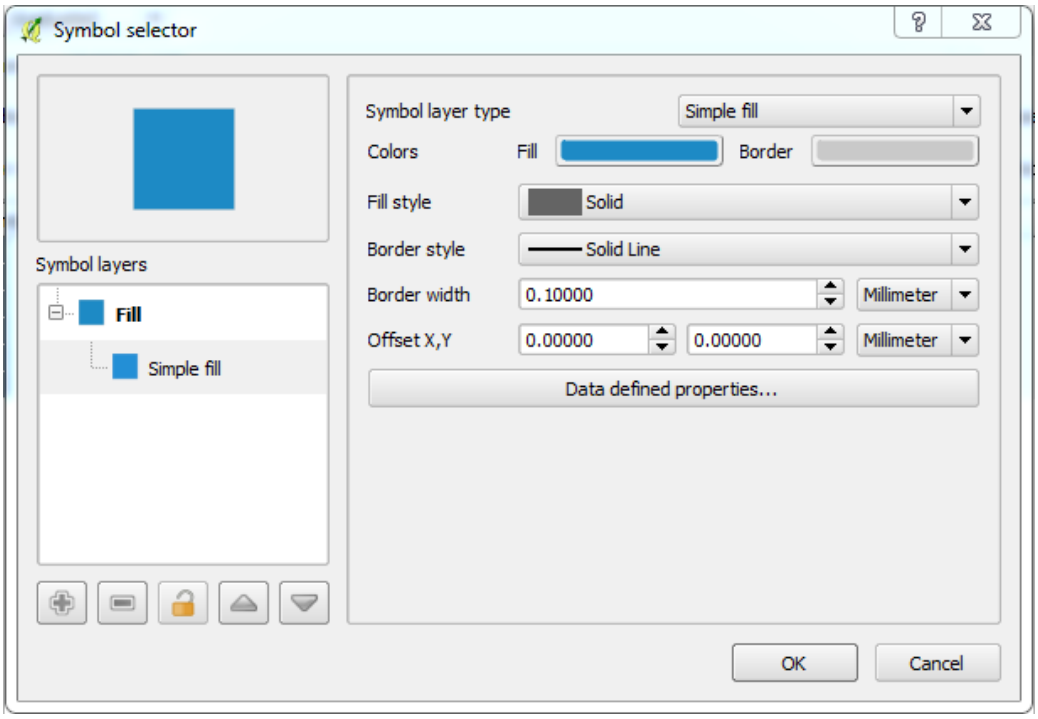


Select 6 **Classes**, and edit them manually – double-click on each **Value** and **Label** to edit.



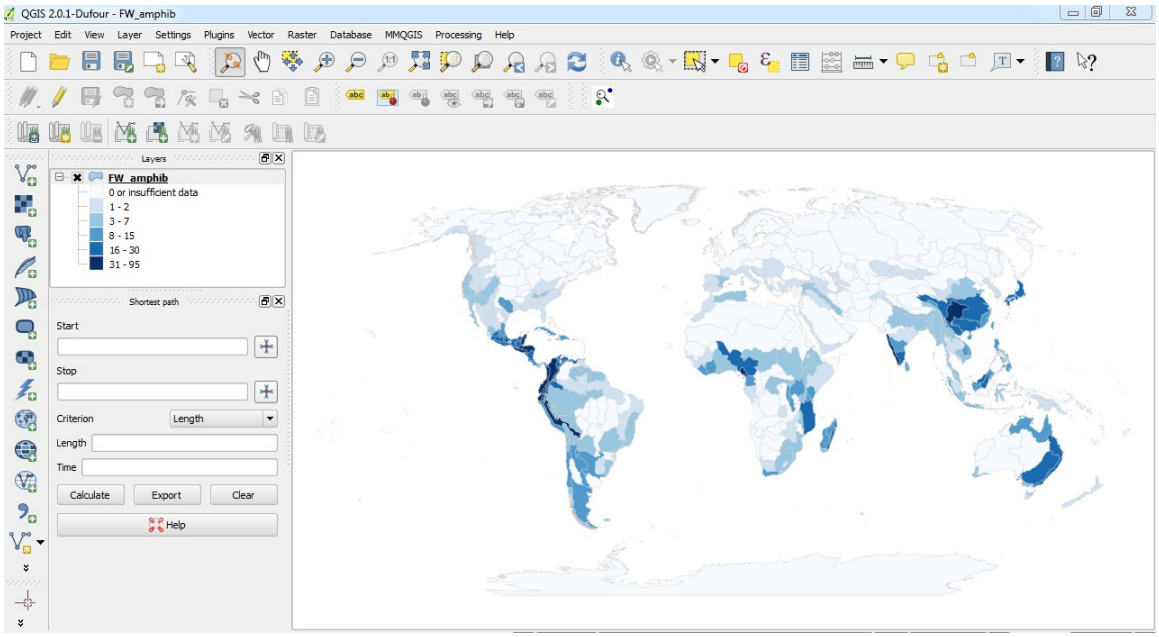
Now we can select a **Color ramp**. For data like this, a sequential color scheme with increasing intensity of a single color works well, so the default **Blues** option is fine. Misleading or confusing color schemes are a common problem in mapping, so I'd strongly recommend looking at ColorBrewer's suggestions to find one that fits your data. ColorBrewer schemes are available in QGIS – scroll to the bottom of the **Color ramp** options, select **New color ramp** and then select **ColorBrewer**.

Next we will edit the boundaries. Click **Change** under **Symbol**. At the next dialog box, click **Simple fill**, and then you can edit the color and the thickness of the boundary lines:



Click **OK**, then back in the main Properties dialog boc click **Apply** and **OK.**

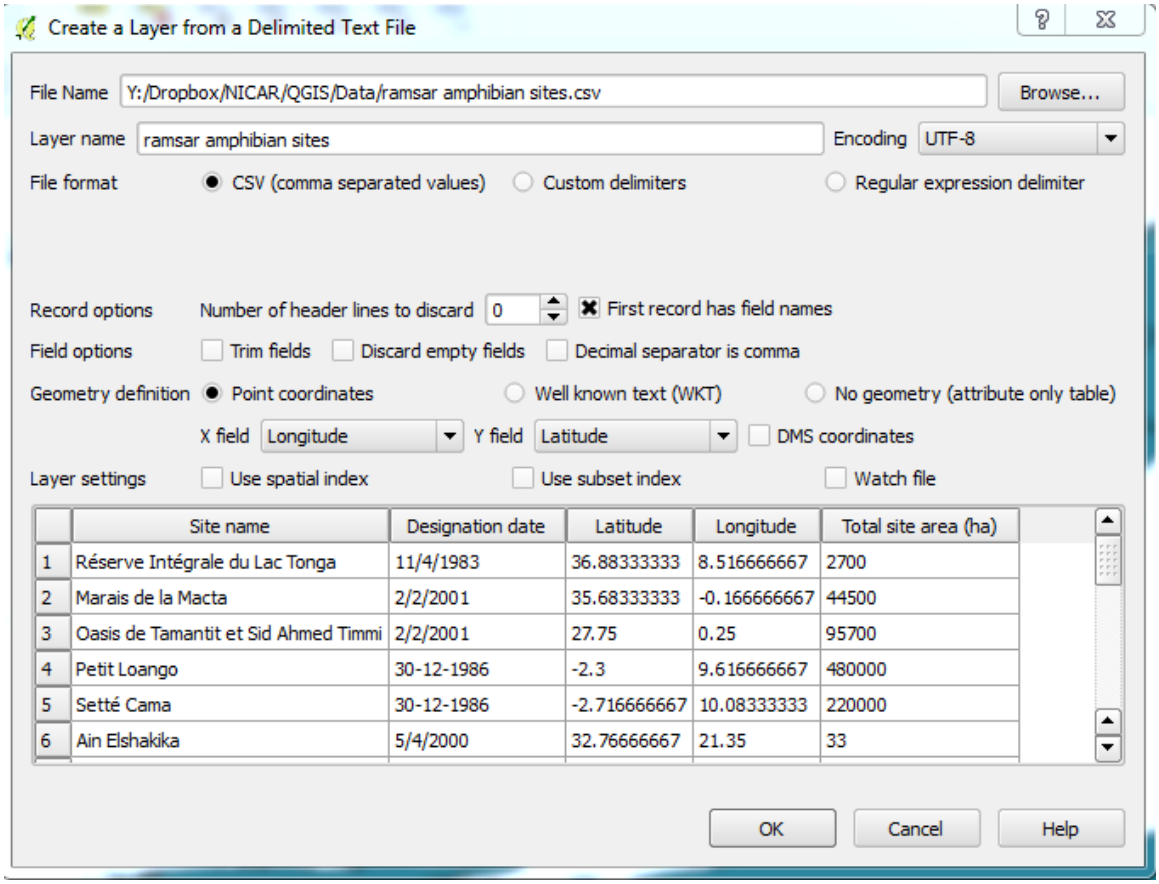
The map should now look like this:



Now we'll add a second map layer, showing the locations of sites deemed important for amphibian conservation, protected under the international Ramsar Convention on Wetlands, which is in this CSV file.

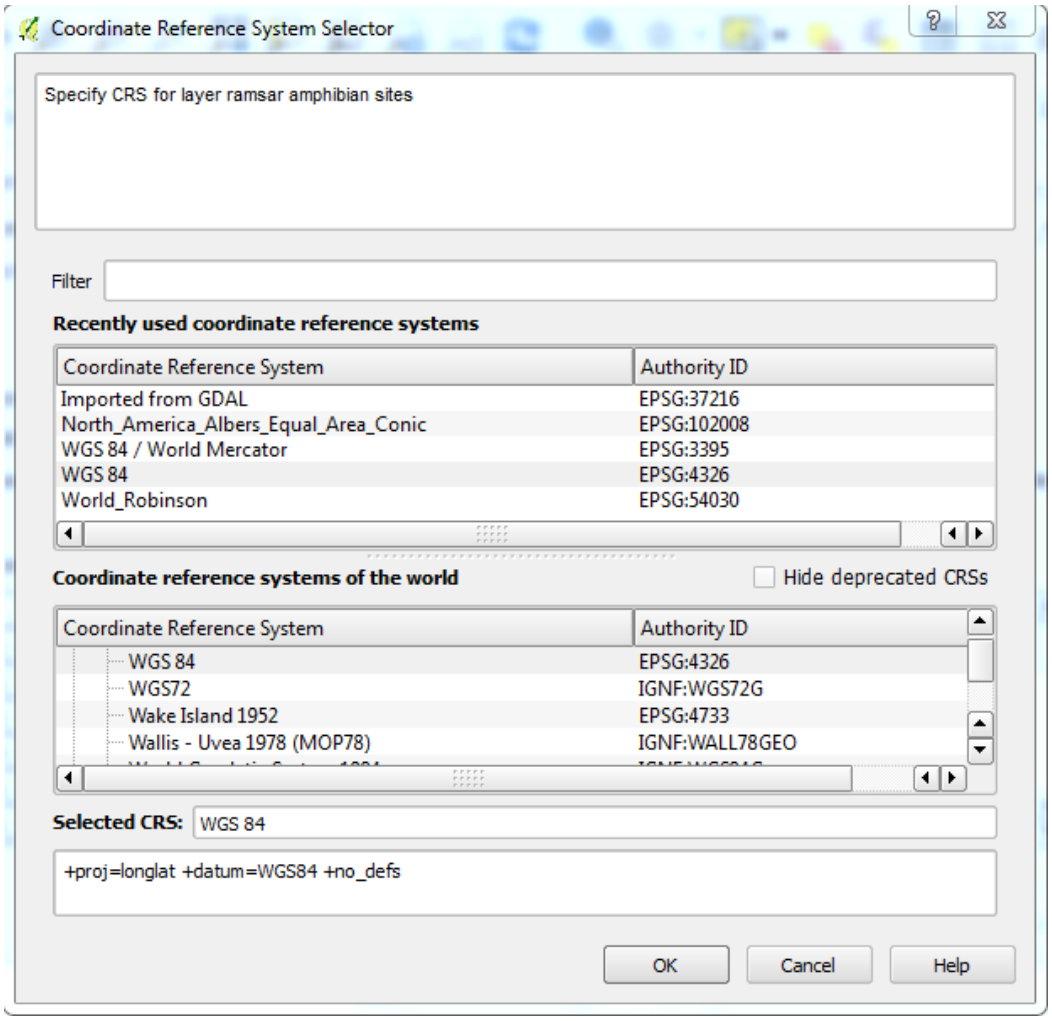
To import this data, select **Layer>Add Delimited Text Layer**, or click this icon: 

QGIS should recognize that **Longitude** and **Latitude** are the X and Y coordinates, and that the delimiters are commas. If not, you can select the correct values at this dialog box:

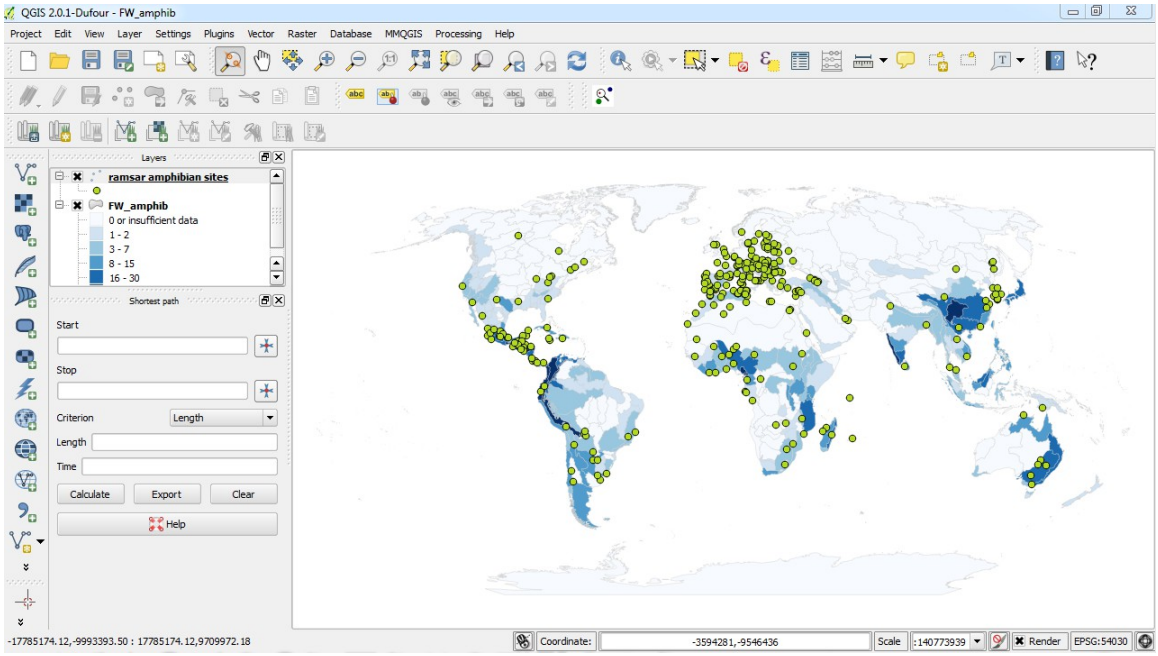


	Site name	Designation date	Latitude	Longitude	Total site area (ha)
1	Réserve Intégrale du Lac Tonga	11/4/1983	36.88333333	8.516666667	2700
2	Marais de la Macta	2/2/2001	35.68333333	-0.166666667	44500
3	Oasis de Tamantit et Sid Ahmed Timmi	2/2/2001	27.75	0.25	95700
4	Petit Loango	30-12-1986	-2.3	9.616666667	480000
5	Setté Cama	30-12-1986	-2.716666667	10.08333333	220000
6	Ain Elshakika	5/4/2000	32.76666667	21.35	33

At the next dialog box, accept the default of WGS84 for the Coordinate Reference System. Notice that the code in the box at the bottom of the screen, which defines the projection, includes the term “longlat.” This tells QGIS that there is no specific projection for the CSV file, just latitude and longitude values – it will then convert to the Robinson projection used for the project:

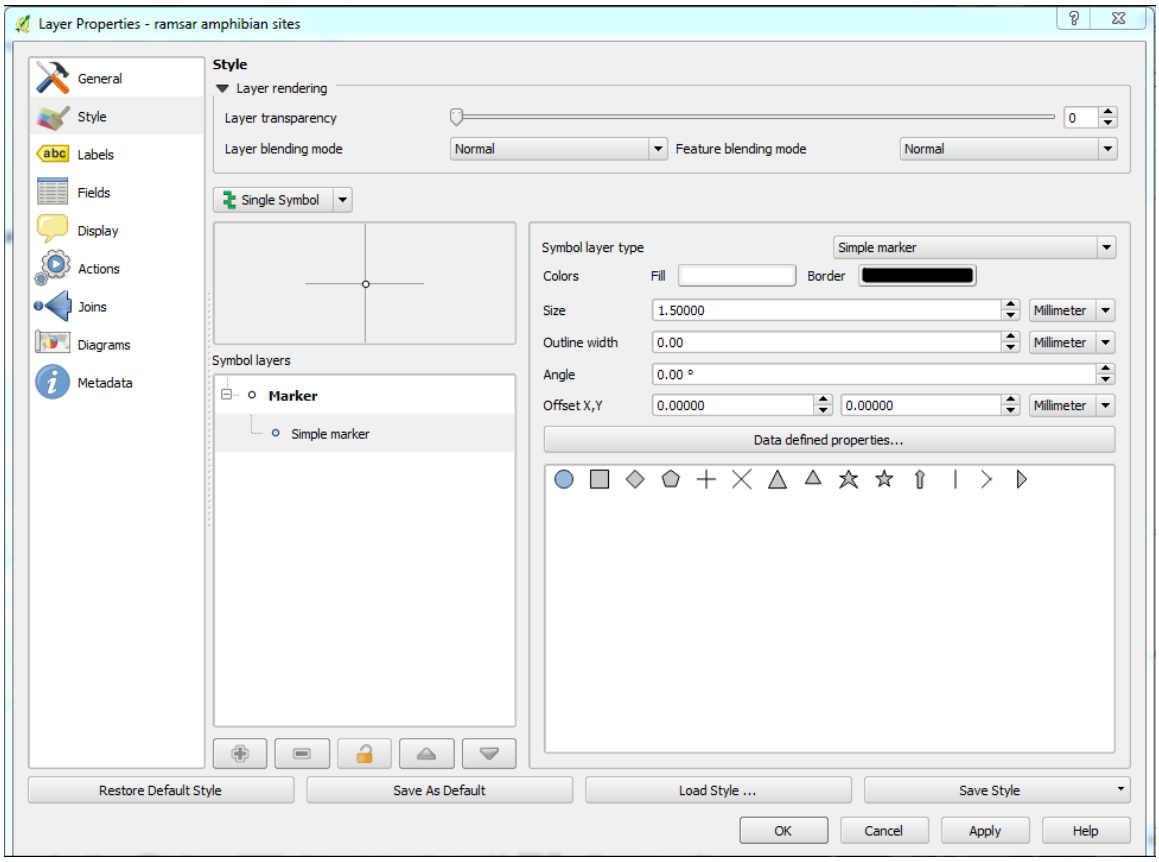


The points should appear on the map:

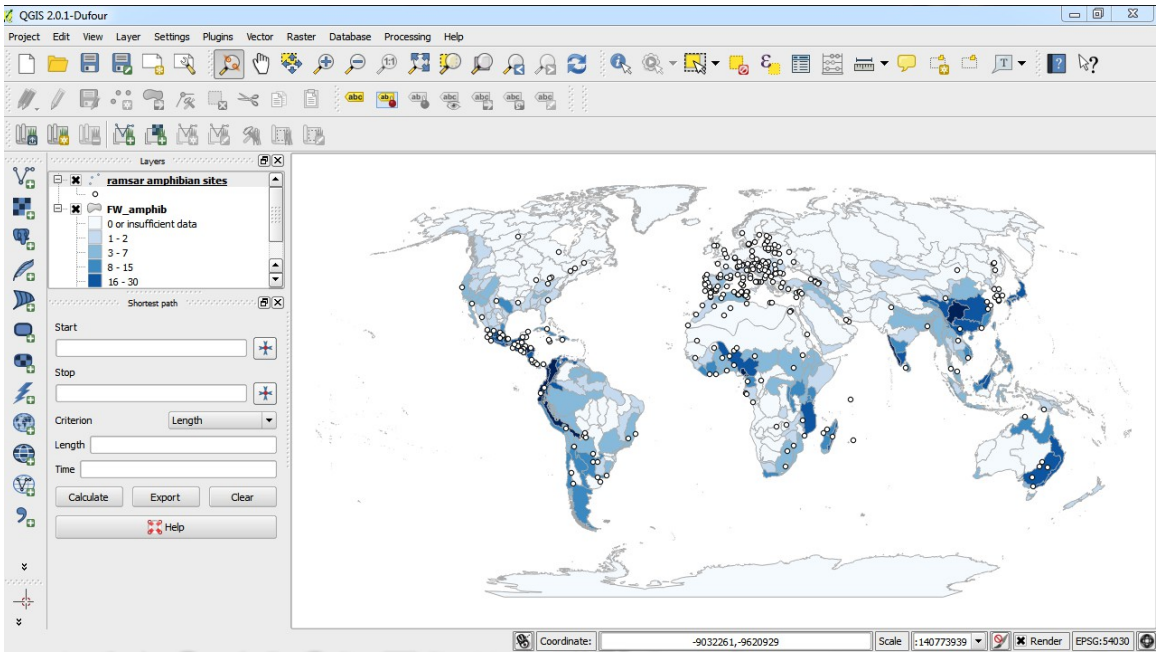


Notice that the entry in the Layers panel has an icon showing that this layer consists of points, rather than polygons.

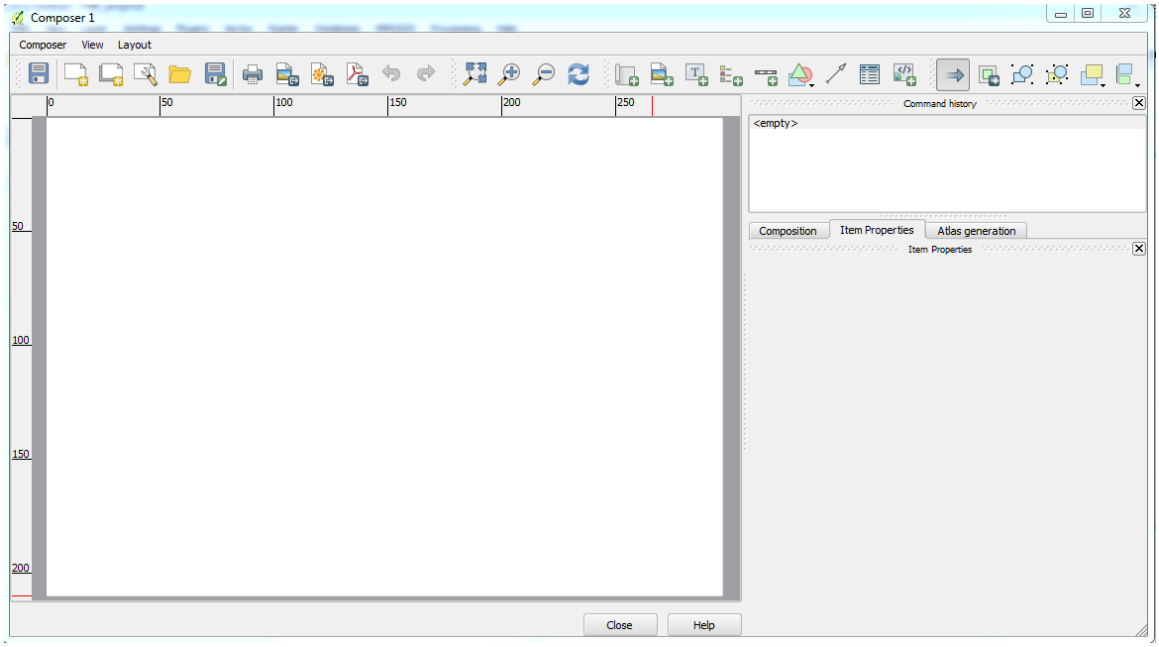
We can style the points much as we did the polygon layer, this time using the **Single Symbol** option.



The map should now look like this:

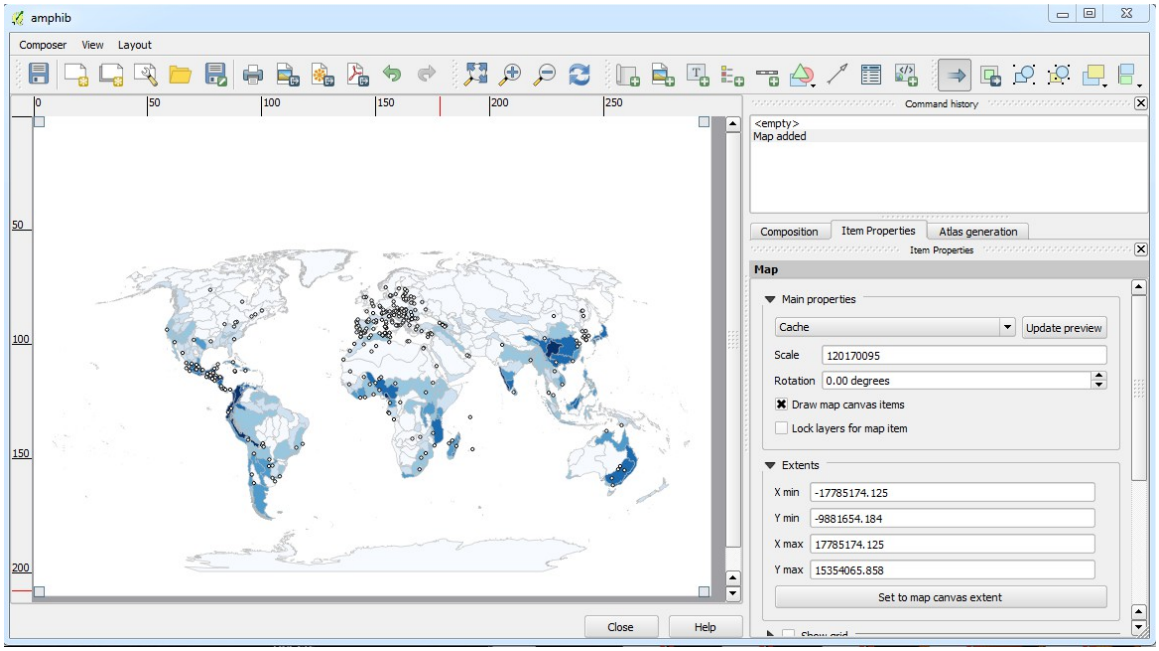


Now we'll export the finished map as a vector PDF, which could be edited further by a graphic designer. **Select Project>New Print Composer**, and add a title at the first dialog box. The following window will open:



To add your map, click the **Add New Map** icon: 

Then click, hold, and use your cursor to draw a rectangle in the main panel. Release the mouse button and the map should appear:



You can add other elements, including a legend, using the options along the toolbar at the top, and customize elements such as typeface and font size, as explained in the QGIS manual.

To export the map as a PDF, click the **Export as PDF** icon: 

You can also export as SVG, another vector graphic format, or as various types of raster image (JPG, PNG etc).