**Week 4: Spatial data manipulation**

**Overview**

During this exercise, you will use several tools to manipulate both raster and vector layers. These tools will allow you to change and modify the geography and the information of the original layers. You will use these processes to answer two questions. For the second part (B), you will have to download data from the internet.

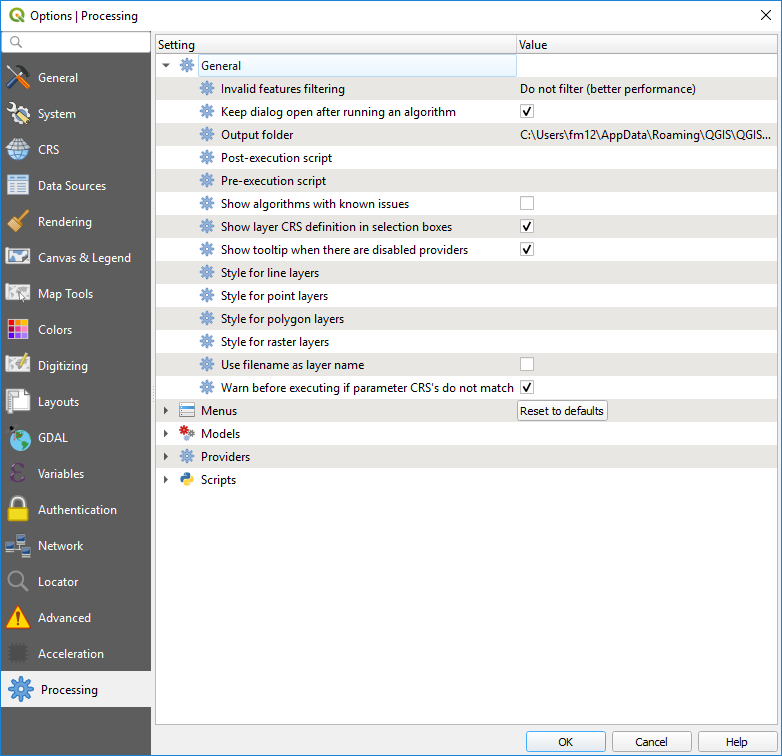
**Download the dataset for this week’s practical**

1. On canvas, go the week 04 page and download the archive zip folder (don’t OPEN it, SAVE it!). Save the zip file in an easy location to find (desktop for example).
2. Extract the archive: right-click on it and select “Extract all”.

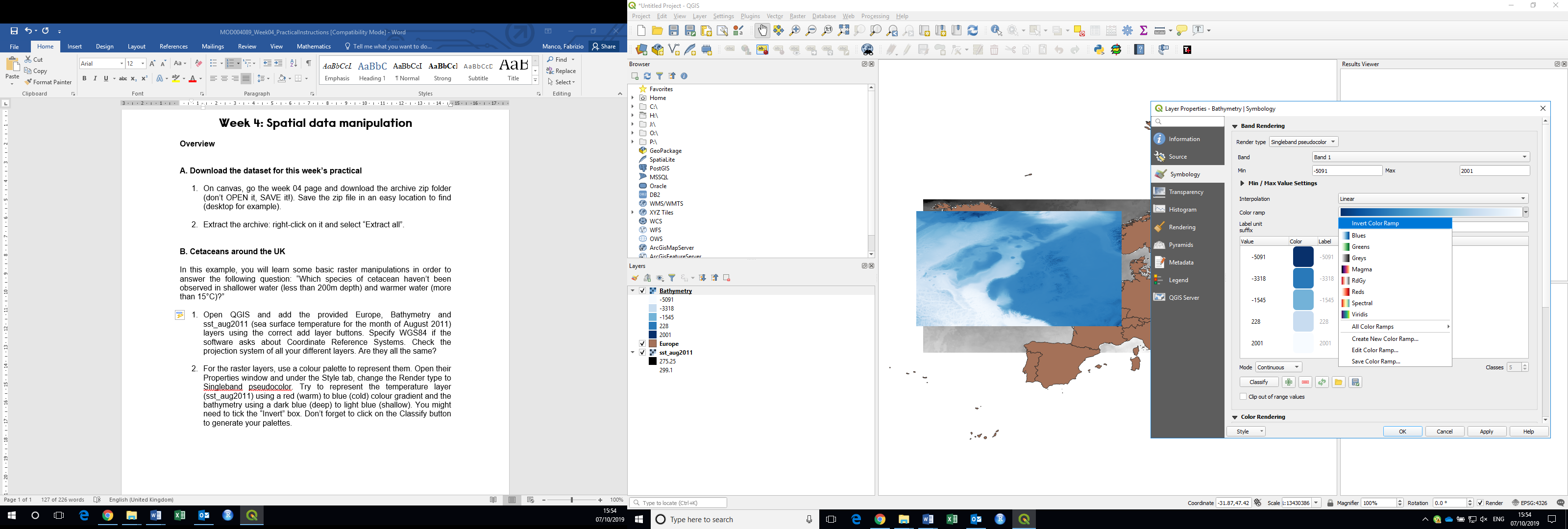
**Part A: Cetaceans around the UK**

In this example, you will learn some basic raster manipulations to answer the following question: **“Which species of cetacean haven’t been observed in shallower water (less than 200m depth) and warmer water (more than 15°C)?”**

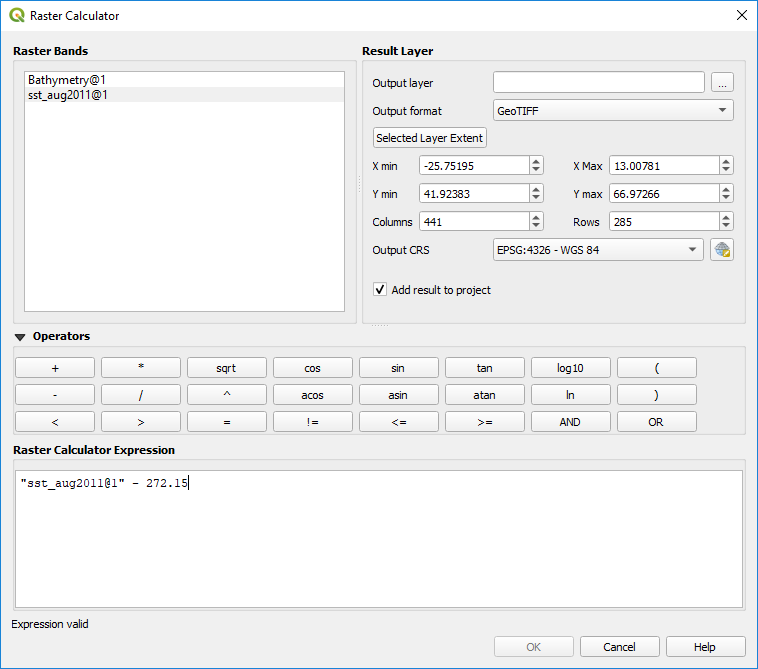
During the transformation of a raster to a polygon (step 6), the result might have some geometry errors which might impact the spatial query (step 7). To ignore the geometry errors, open the Settings > Options window and under the Processing tab, expand the General group of options. Make sure you have “Do not filter (better performance)” for Invalid features filtering.



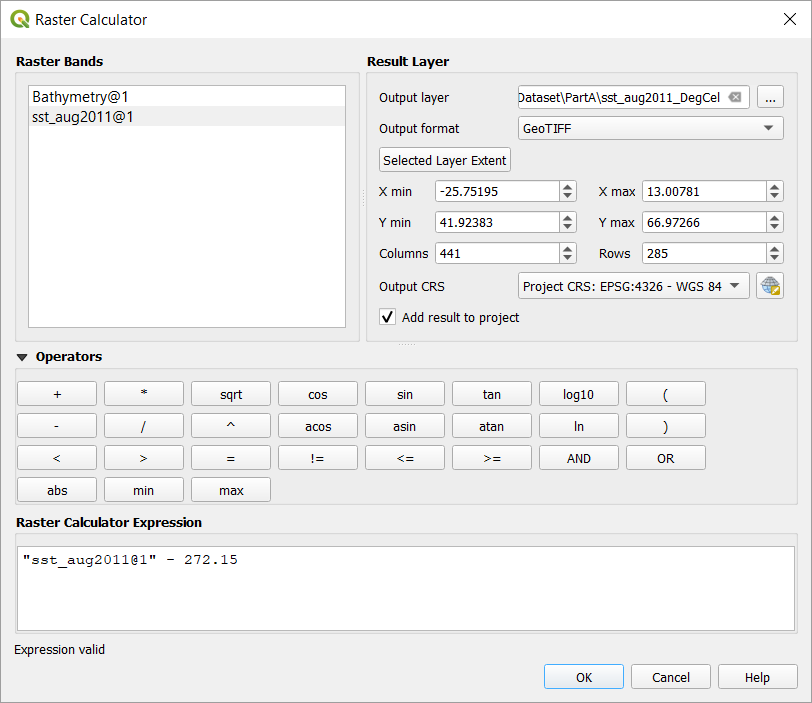
1. Open QGIS and add the provided Europe, Bathymetry and sst\_aug2011 (sea surface temperature for the month of August 2011) layers. Specify WGS84 if the software asks about Coordinate Reference Systems. Check the projection system of all your different layers. Are they all the same?
2. For the raster layers, use a colour palette to better represent them. Open their Properties window and under the Style tab, change the Render type to Singleband pseudocolor. Try to represent the temperature layer (sst\_aug2011) using a red (warm) to blue (cold) colour gradient and the bathymetry using a dark blue (deep) to light blue (shallow). You can find predefined colour ramps by clicking on the little down arrow on the right of the colour ramp button. You might need to select the Invert Color Ramp option to create an intuitive colour ramp.



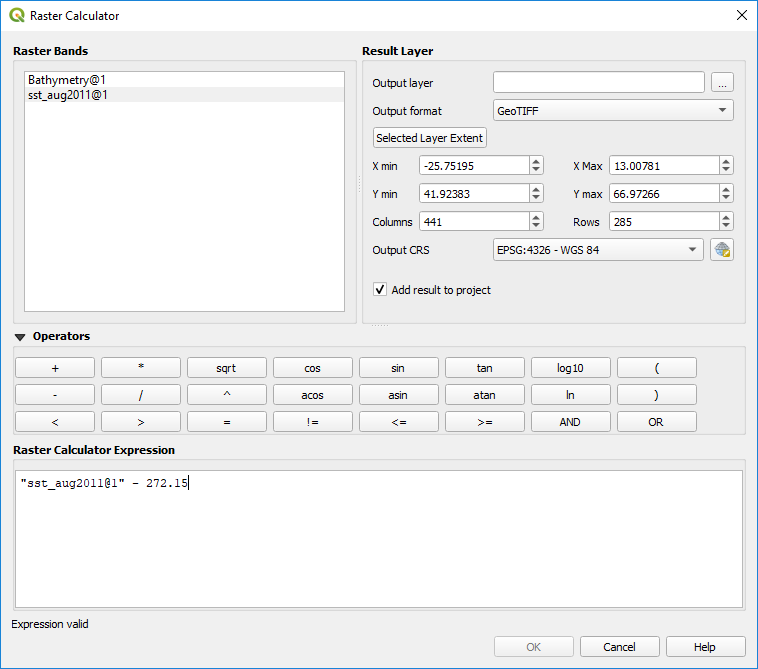
1. Interrogate the values of the sst\_aug2011 layer. What are these values? To convert them in degree Celsius, we can use the raster calculator: in the raster menu, select Raster Calculator. Double-click on sst\_aug2011@1 in the Raster bands list (the “@1” means that we are using the first band of this raster – we will learn about multi bands raster images during the session on remote sensing). In the Raster calculator expression box, write the following expression: "sst\_aug2011@1"-272.15 (the conversion from degrees Kelvin to Celsius).

In the Result layer section, under Output layer, click on the browse button () to specify where to save the new raster file and enter its new name (sst\_aug2011\_DegCel for example). Click on the Selected Layer Extent button to specify the extent of your output layer (the X and Y min and max values should then update).

In the Output CRS, make sure you have EPSG: 4326 - WGS 84 (it should be the Project CRS).

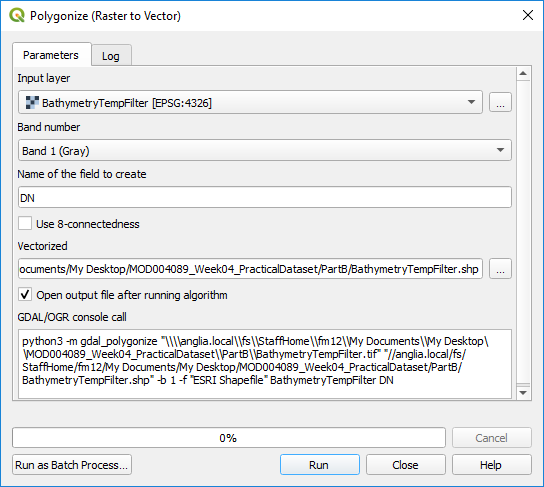


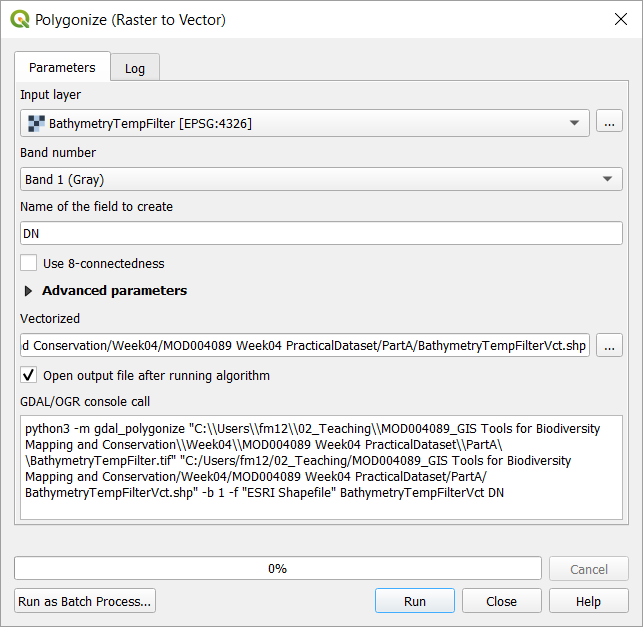
You can now click on Ok to run the calculation. Check the values of the new added layer. Change the style again to get the correct colour palette. You can now remove the old sea surface temperature layer in Kelvin (right-click > Remove).

1. You can add the Cetacean\_Atlas\_Annual\_AllSpecies.csv using the Delimited Text tab in the Data Source Manager menu. Make sure to specify the correct columns for the coordinates (X and Y fields). Under Geometry CRS, search for the EPSG:4326 – WGS 84 (you can use the Coordinate Reference System Selector ).
2. To answer our question ‘**Which species of cetacean haven’t been observed in shallower water (less than 200m depth) and warmer water (more than 15°C)?’**, we need to extract from our bathymetry raster the location that are shallower than 200 metres and warmer than 15°C. We can use the raster calculator again. Try to build the following expression:

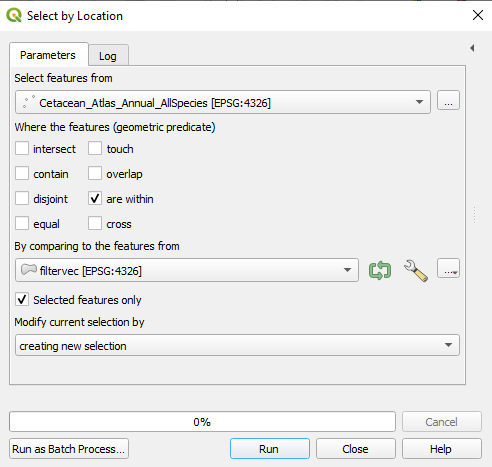
sst\_aug2011\_DegCel@1 >= 15 AND Bathymetry@1 >= -200

Specify an output name for the result (BathymetryTempFilter for example) and for the extent, select the bathymetry layer in the Raster bands list and click on the Current layer extent button. Set the Output CRS as EPSG:4326 – WGS 1984. Run the calculation and explore the result. What are the values in this new layer?

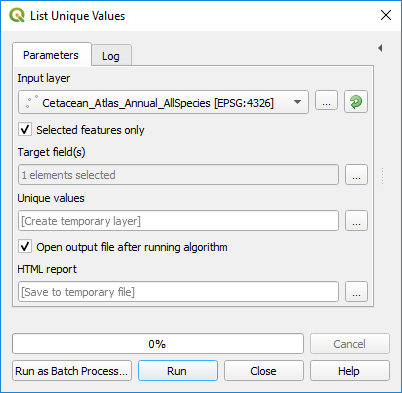
1. To query our cetacean observations, we need to convert this new raster into a polygon. In the Raster menu go to Conversion > Polygonize (Raster to Vector). You can specify the name of the column that will contain the values of the raster or leave it as default (DN). Under Vectorized, click on the  button and select Save to File. Save the new polygon layer as BathymetryTempFilterVct.shp for example (make sure you save your new layer as a shapefile). Explore this new layer (geography and attributes). What are the attributes? What do these values mean?



1. From our filter polygon layer, using any selection tool you learnt in week 2, select the polygons that meet our criteria (DN=1). You can now select your cetacean observation points that are within the selected polygons by doing a selection by location. Go to Vector > Research Tools > Select by location. Create a new selection from the Cetacean\_Atlas\_Annual\_AllSpecies point layer that are within the Vectorized version of the BathymetryTempFilter layer. Tick the Selected features only, so you will only select points within the previously selected polygons. Click on Run.

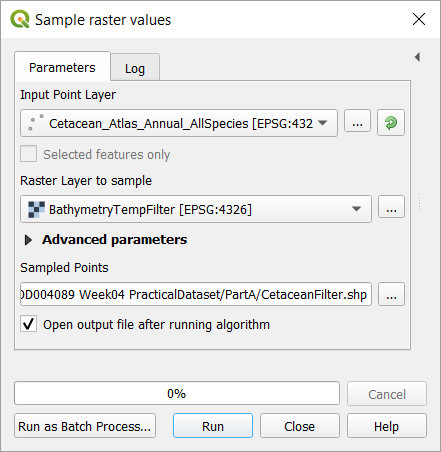


1. To answer the original question, you can now run the List Unique Values tool (Vector > Analysis Tools > List unique values) to compare the list of cetacean species present in the filtered Cetacean\_Atlas\_Annual\_AllSpecies layer (only observations in warmer and shallower waters by ticking the Selected features only box) and all the species present in the unfiltered layer (all the points). The Target field(s) is the column containing the species names.



Note: There are alternative ways of answering this question and one is based on a very useful function: the sample raster values. This tool will sample a raster layer at the location of each point from a point vector layer.

Under the Processing menu, enable the Processing Toolbox. Under the Raster analysis group of tools, run the Sample raster values tool. As Input Point Layer, select the Cetacean\_Atlas\_Annual\_AllSpecies layer and as Raster Layer to sample, select the BathymetryTempFilter layer. Save the Sampled Points as a new CetaceanFilter.shp shapefile. Click on Run.



Check the attribute table of the new point layer. What is the new field (or column) in this layer? What does it contain? How can you now use this layer to answer the question?

You could have also ran that Sample raster value tool on both the bathymetry and then the sea surface temperature rasters to extract the exact values of these two layers for each cetacean point and then use a selection process to find which observation meet the criteria.

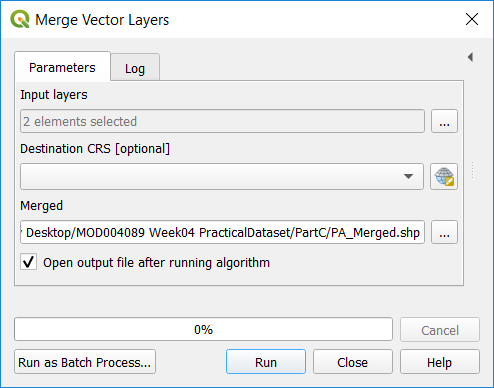
**Part B: Orang-utan and protected areas**

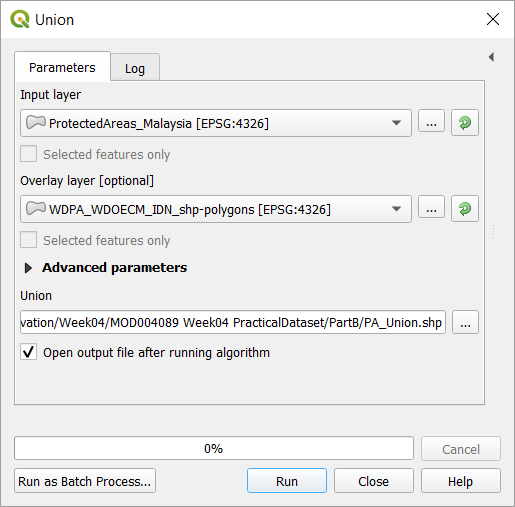
In this example, we are going to compare the estimated distribution area of a species (the Orang-utan) and quantify how it overlaps with protected areas.

1. Download and prepare the data for the species: go to the IUCN redlist website (http://www.iucnredlist.org/) and search for *Pongo pygmaeus*. You can have a look at the distribution through the online map and you can also download the range data as a polygon shapefile. Click on the Download button and select Range data – Polygons (SHP). You will have to register and sign in. Once you manage to download the data, extract the zip file and add the shapefile in an empty QGIS project and you can rename it (right-click > Rename Layer) as Pongo pygmaeus
2. For the protected areas (PA), go to http://www.protectedplanet.net/ and search for the protected areas in Indonesia using the search tool. On the result page, click on the link for the country, then on the Download button and select SHP to get the data as a shapefile. this dataset and select .SHP (shapefile). Save and extract the file.

Do the same for the protected areas of Malaysia (as the species is in both countries on the island of Borneo). If you have any issues downloading this data, please contact me (hannah.white@aru.ac.uk).

In QGIS, add the polygons representing the protected areas from these two countries. Once you extracted all the zip files, you can delete them to clean your file folder.

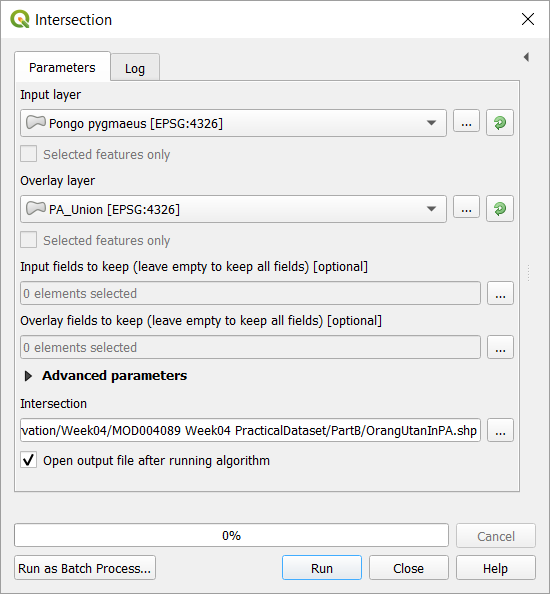
1. To analyse how the Orang-utan distribution overlap with the protected areas polygons, we need to first merge the two countries protected area layers. We are going to use a tool called Union, which allows to merge vector layers even if the fields of the attribute tables are different. In the Vector menu, under Geoprocessing, run the Union tool. As Input and Overlay layers, select the protected area layers from both countries and under Union click on  and select Save to File and save the merged layer as a new PA\_Merged shapefile. You may have to do this multiple times as the protected areas within a country are split into multiple files.



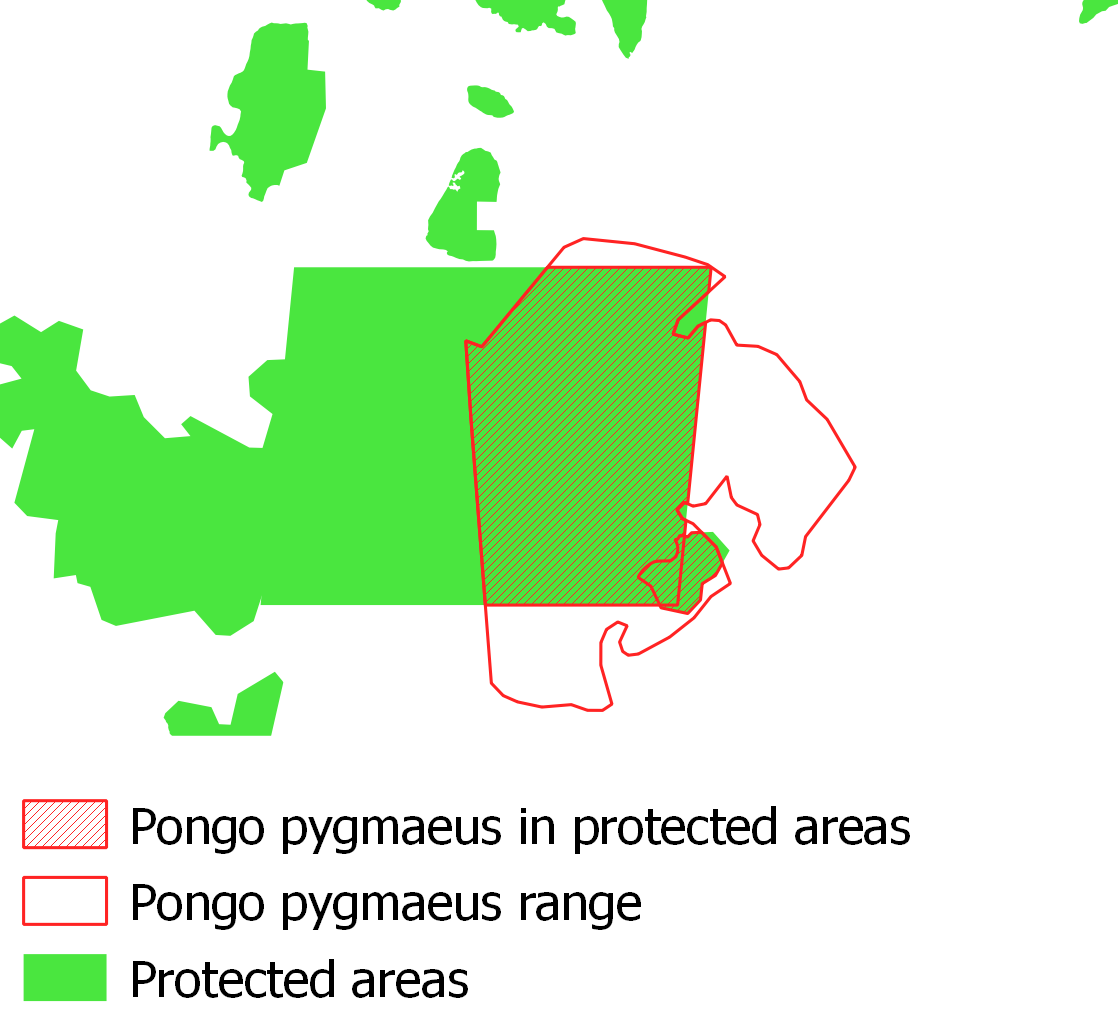
1. Try to change the symbologies of the *Pongo pygmaeus* and PA\_Union layers as on the map below. For the PA\_Union layer, under the Single symbol symbology, click on the Simple fill entry and change the Stroke style to “No pen” to hide the polygon borders. For the orang utan range layer, change the Fill style to “No Brush” to remove the filling of the polygons.

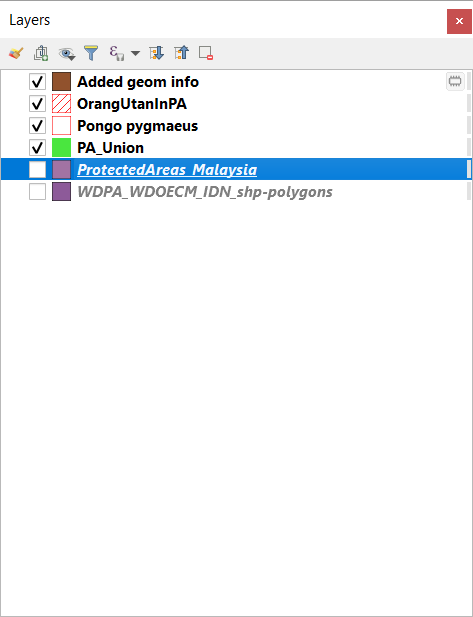


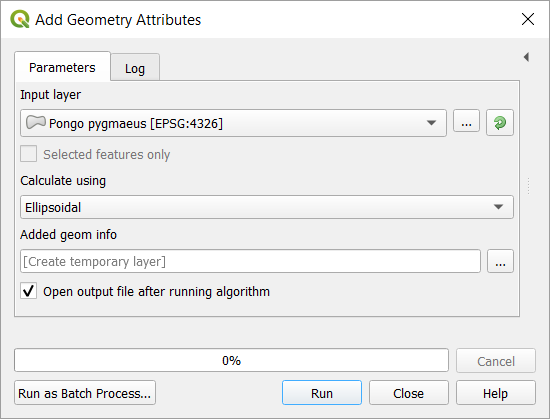
1. To find the area of the Orang-utan polygons that are included into the PA polygons, we need to extract the overlapping area between the two layers (the species and the protected areas), the tool we can use is called intersection. In QGIS, it is in Vector > Geoprocessing Tools > Intersection. The input layer is the species range layer and the overlay layer is the merged PA layer. Save the results as a new shapefile, OrangUtanInPA, for example.



Try to modify the symbology of the intersection layer to recreate a map like this one:



1. Finally, in order to calculate the proportion of the orang utan range that is included within protected areas, you need to calculate the areas of the polygons for these layers. In the Vector > Geometry Tools run the Add Geometry Attributes tool. You can start with the *Pongo pygmaeus* layer. Make sure you set the Calculate using option to Ellipsoidal so the areas will be calculated in metres. You can save the resulting layer as a new file or just ignore that so it will be stored in a temporary layer (indicated by the  symbol in the list of layers.



Check the attribute table of the new Added geom info layer. To calculate the total area of all the polygons, you can use the Vector > Analysis Tools > Basic Statistics for Fields tool you used during week 2. Find the appropriate result you need to answer the question.

Repeat this step for the Intersection layer so you can calculate the proportion of *Pongo pygmaeus* range that is included within protected areas (in percentage).