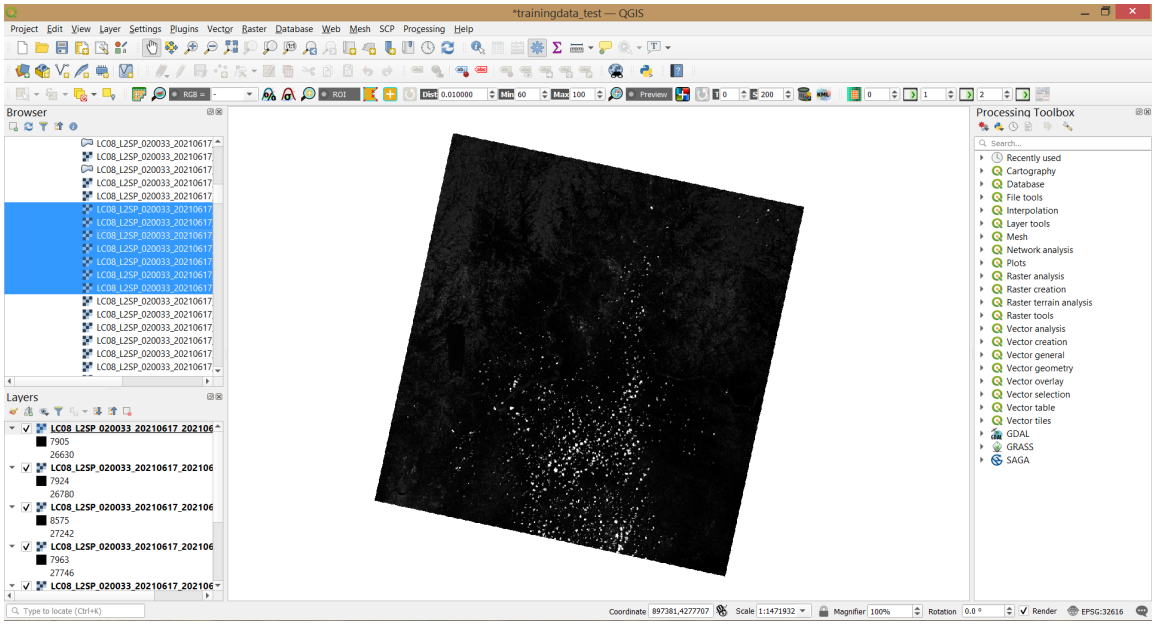
Advanced GIS
Computer Science in Modern Biology Workshop 2021
Instructor: Gwendolyn Lloyd (lloydg@miamioh.edu)
TA(s): Tereza Jezkova (jezkovt@miamioh.edu)

How to Create Training Data in QGIS: Day 1

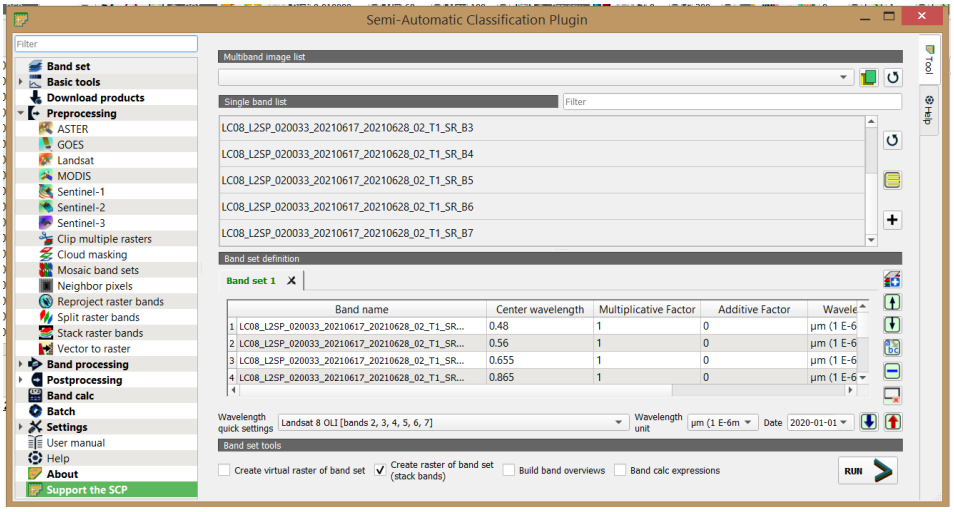
Part 1: Importing Raster Data (Preprocessing)
In this part we will learn how to import raster data and what manipulation we need to do to make it usable.

1. Raster data is saved as TIFF files, which are image files. As we learned before, different bands contain different information. You'll notice that there are multiple images for each raster. Today we will be working with the one that starts LC08….
2. Open a new QGIS file, from the Browser window, navigate to the place you saved the bands, and drag and drop them into the layers window (this is most likely under the home window.



3. We will then use the extra plugin you installed. If you haven't done so, now is the time! Under the SCP window (if it isn't there you will need to turn on the plugin) select Band Set.

4. Add all the individual bands currently open and select Landsat 8 in the Wavelength Quick Settings. Check create a raster of band sets and click run. This may take some time to run.



5. We now have an image!! Is it not what you expect? We've been accustomed to seeing satellite imagery but the raw imagery is much different from what we may expect. We can change the colors within Symbology, by selecting multiband colors. Each satellite has a different combination of bands and colors. Landsat 8, which we will use for this class, has the combination below.

Landsat 8-9 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS)

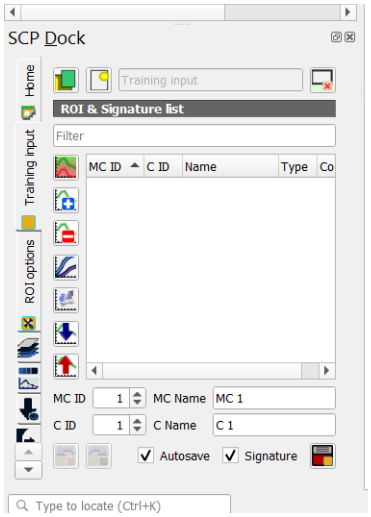
Bands	Wavelength (micrometers)	Resolution (meters)
Band 1 - Coastal aerosol	0.43-0.45	30
Band 2 - Blue	0.45-0.51	30
Band 3 - Green	0.53-0.59	30
Band 4 - Red	0.64-0.67	30
Band 5 - Near Infrared (NIR)	0.85-0.88	30
Band 6 - SWIR 1	1.57-1.65	30
Band 7 - SWIR 2	2.11-2.29	30
Band 8 - Panchromatic	0.50-0.68	15
Band 9 - Cirrus	1.36-1.38	30
Band 10 - Thermal Infrared (TIRS) 1	10.6-11.19	100
Band 11 - Thermal Infrared (TIRS) 2	11.50-12.51	100

6. We are used to Natural Color (red=red, blue=blue, green=green). Change the bands to reflect this and click apply.
7. There are lots of different band combinations used. Play around with a few to see how colors change and different land cover types are emphasized. https://gisgeography.com/landsat-8-bands-combinations/

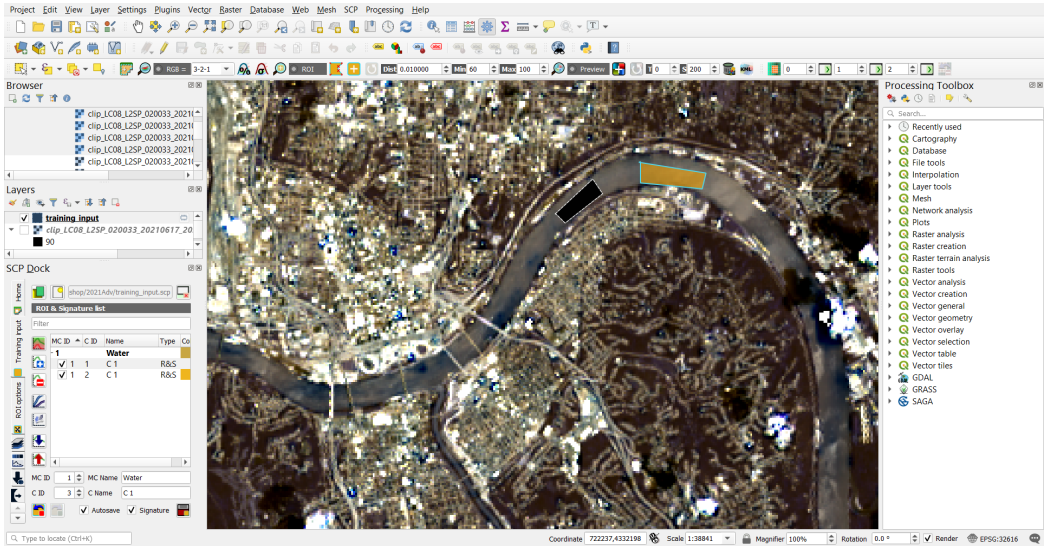
Part 2: Classification

Once we have our data as a single raster, we can classify our data, or define different categories within it.

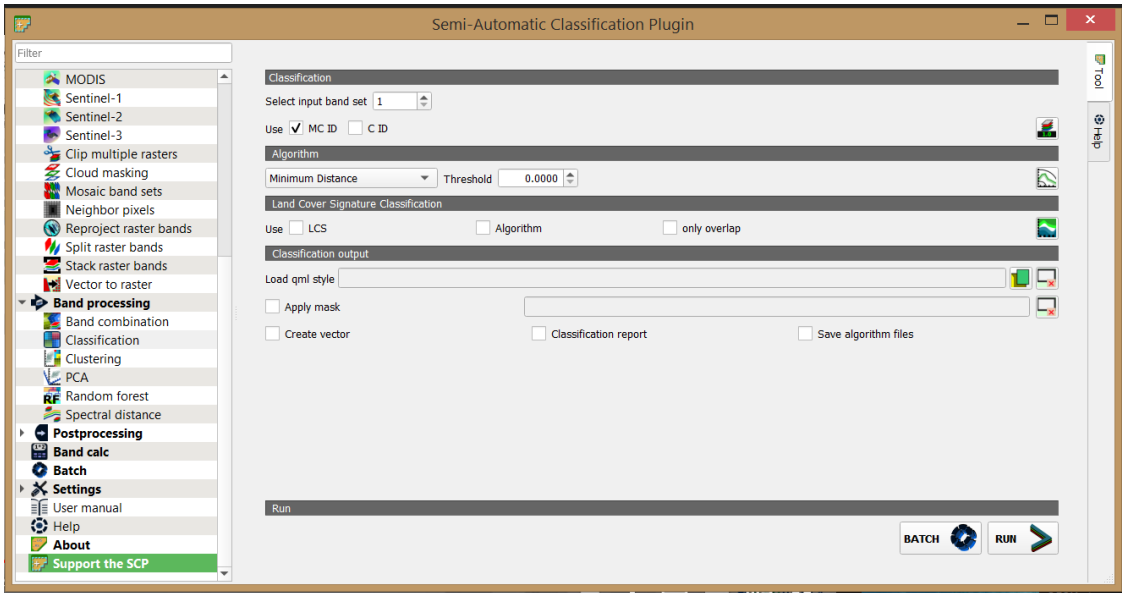
1. Open the SCP dock if it isn't currently open (View → Panels → SCP Dock), and select the training input tab.



2. Selecting the second icon in this tab, create a new file, called training_input.
3. Now we can start classifying our data! In the MC Name box add a type of land cover - lets's start with Water. For C Name you can add more detail like pond or river (this is optional!)
4. Once we have a category we can create an ROI polygon using the tool on the bottom toolbar. To do so you will click the vertices and right click on the last one when you are done.
5. Save your polygon with the criteria in the SCP Dock, and repeat the process 2-3 times for each land cover type.



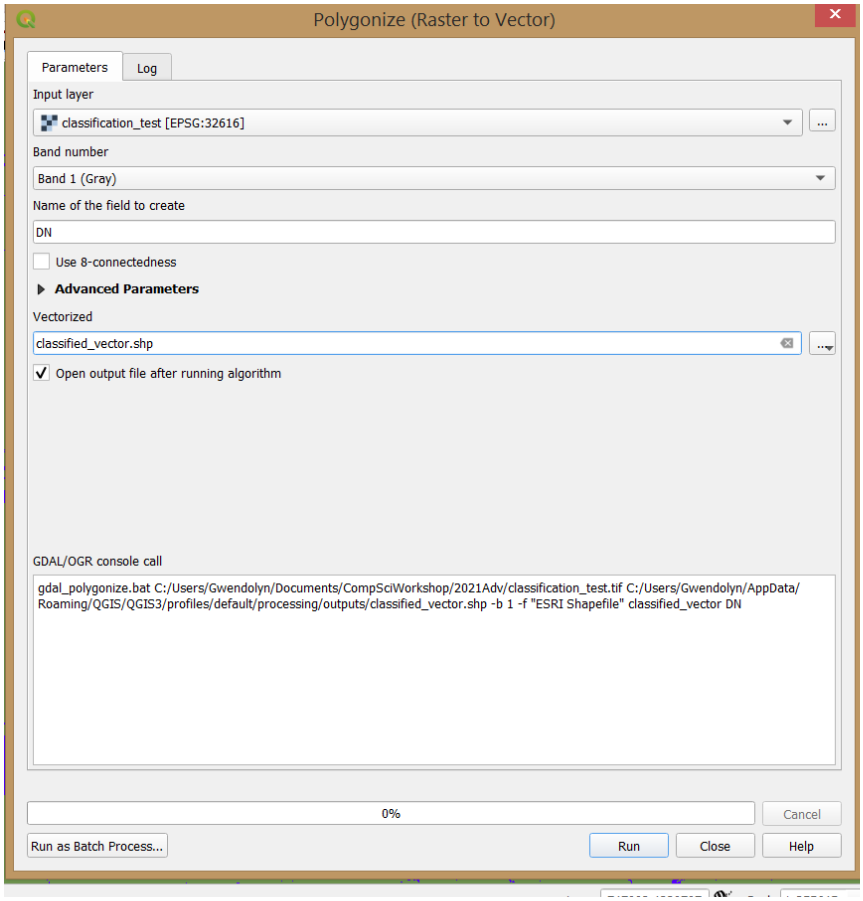
6. Repeat this process for different land cover types - ex.Urban, Suburban, Forest, Agricultural. Be sure to change the MC ID as well as the MC Name for each type of land cover.
7. Once we have the training data we can finally classify our data! Open the SCP window → Band Processing → Classification. Make sure to use the MC ID for classification. Click Run.



Part 3: Using our data

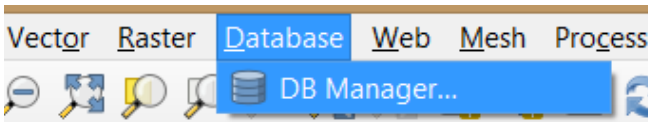
Okay so we finished our classification...now what?

1. To do spatial statistical analysis on our raster we can transform it into a vector shapefile (hint: Raster Conversion!)



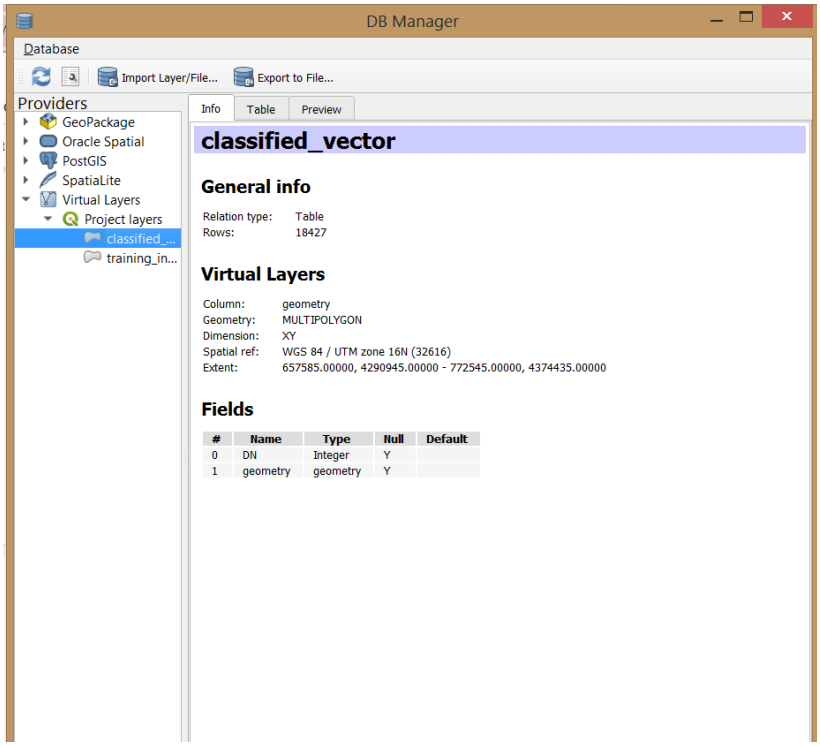
   From here if you were in the intro class this should look familiar :)

2. In our new vector file we can create a new column for shape area in the attribute table, and then get information about the total area covered by different land cover classes using the DB manager
3. This method will involve creating a database and using the database manager (Database → DB Manager)

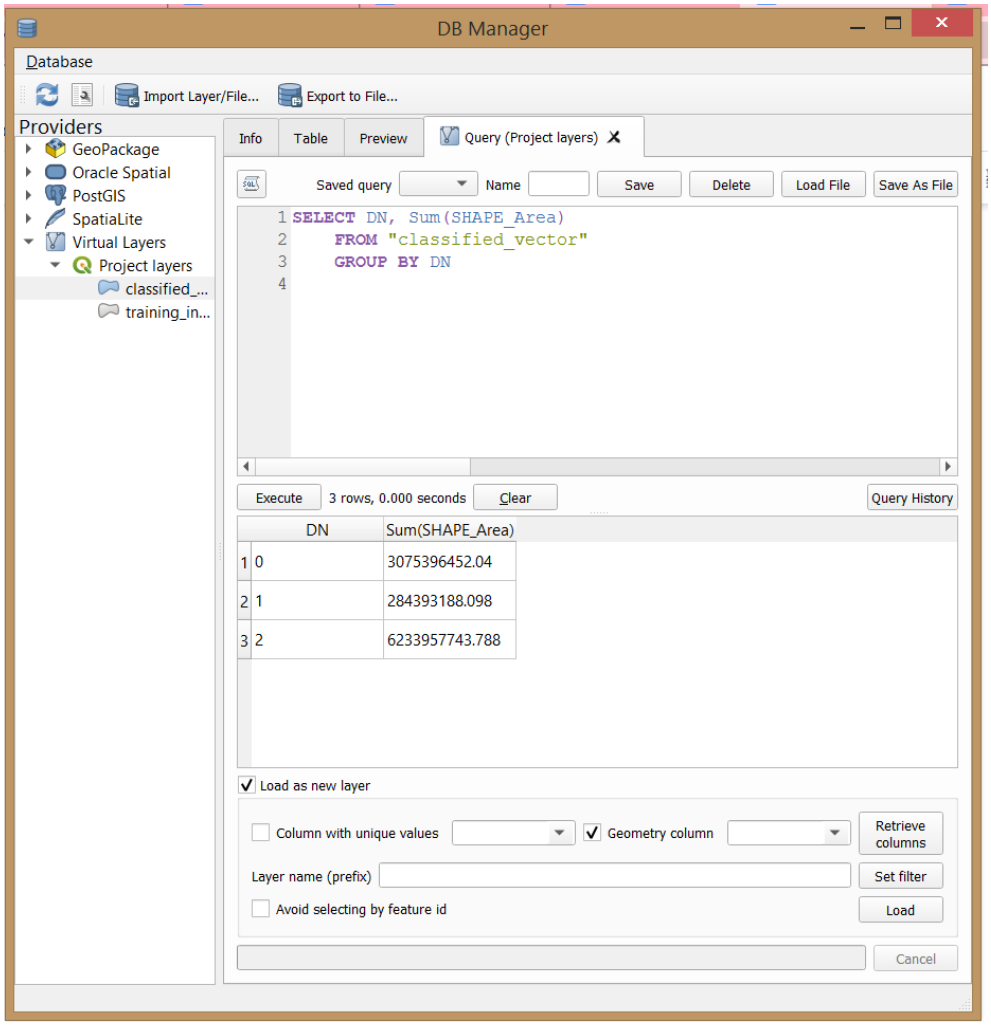


4. Then under the virtual layers database we will then select the layer that we want to get information for. A virtual layer creates a temporary file that does the calculations from another shapefile.

   Next, click on the second button from the left to open an SQL Query window.



1. Once open, you should type the code as written below. The second row, FROM “ ”, is the layer which you are getting the calculations from, so that should change. Otherwise what it means is that it is selecting the two variables, DN and SHAPE_Area from the created data, and taking the sum of the area and grouping it by the categorical variable Class.
   SELECT DN, Sum(SHAPE_Area)
   FROM "*LAYER OF INTEREST HERE*"
   GROUP BY DN
   After you have written your code click execute.
2. Your screen should look similar to the image below, with a table containing the total area covered by each land cover class. These values are in meters squared, due to the original data.
   To open this table as an attribute table under the layer menu, check the box that states ‘Load as new layer’ and then click load in the window.



DN are the values from the MC IDs for your different types of land cover.

How does your classification compare to the official classification?