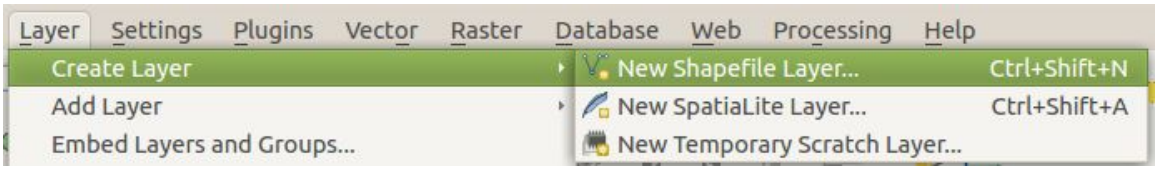
Pontus Olofsson, Christopher E. Holden, Eric L. Bullock
Department of Earth & Environment, Boston University

S2. Methods: Classification

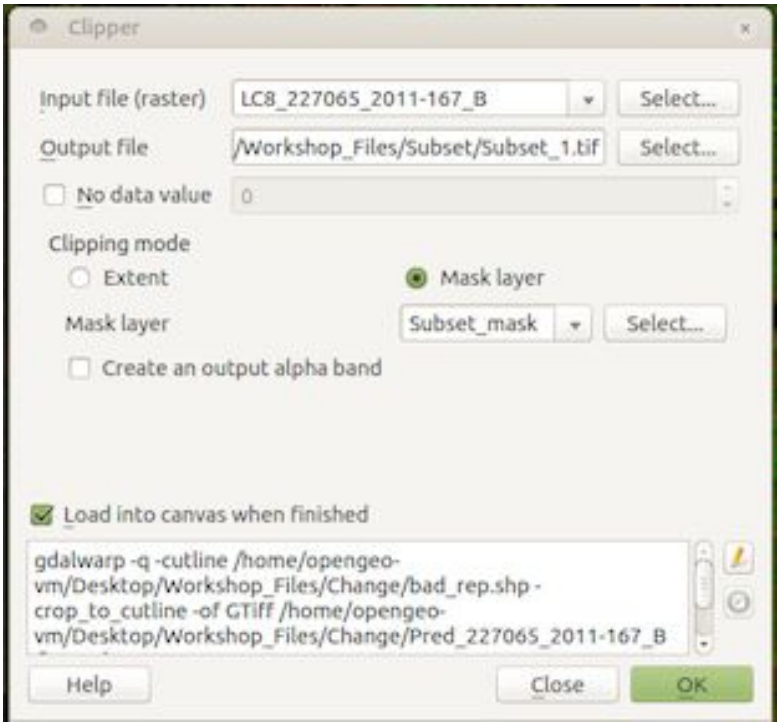
S2.1 Subset an image using *Mask layer*

1. To create your shapefile, go to *Layer > Create Layer > New Shapefile Layer.* Under *Type* click *Polygon.* Click *OK* and save the shapefile somewhere appropriate. (See screenshot below.)



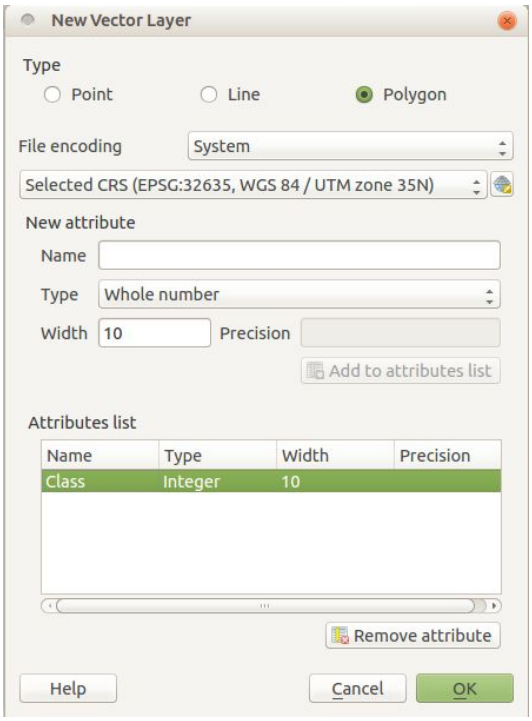
2. To start drawing your subset mask, first right click on the shapefile in the *Layers* panel and click *Toggle Editing.*
3. To start drawing, click the *Add Feature* symbol. It should be on the top of the screen, and looks like this:

4. Click on all corners of your intended boundary. You should see a red polygon being built where you are clicking. Once you have created a polygon, right click to finish drawing the feature. Click *OK.*
5. To save the subset mask, right click on the shapefile in the *Layers* panel and click *Toggle Editing*. Make sure to save your shapefile.
6. To extract the subset, go to *Raster > Extraction > Clipper.* For *Input file (raster)* put the Landsat image you want to subset. For *Output file* click *Select…* and navigate to an appropriate location and come up with a name to save your output. Under *Clipping mode* select *Mask layer* and select the shapefile you just created. Click *OK.* (See screenshot below.)



S2.2 Create your ROIs

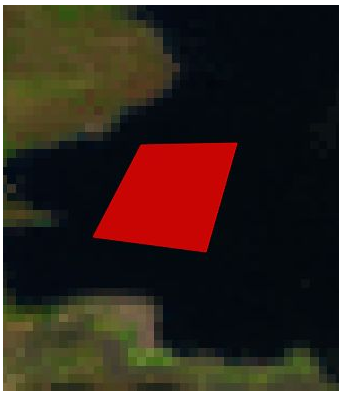
1. Now that you have a subset to easily work with, use what you used in the previous exercise to apply an appropriate 3-band stretch on the image.
2. To collect ROIs in QGIS you need to create a shapefile with features corresponding to your different ROIs. As you did before, go to *Layer > Create Layer > New Shapefile Layer.* Under *Type* click *Polygon*. Now add an attribute to your Polygon with the *Name* 'Class', and *Type* 'Whole Number'. You can delete the other attribute in the *Attribute List* by clicking on the attribute with the name 'id' and clicking *Remove attribute*. Your box should be filled out like the screenshot to the right. Click *OK* to create the shapefile. It will ask you where to save the ROI. Save it to your *Classification* folder.



3. Start drawing your ROIs, right clicking to save them. When doing so, pay attention to the notes below.
4. When you finish drawing an ROI, it will ask you to fill in a value for 'Class'. Each class in

your map will have a corresponding number. If you have 5 classes, then your classes will be labeled 1-5. Make sure to keep track of which number corresponds to which class. Below is an example of an ROI for a water class.

5. After you enter the number for the class, click *OK.* Repeat this multiple times for each class.
6. When you are done creating ROIs, right click on your shapefile in the *Layers* panel and click *Toggle Editing*. Make sure to save your ROI shapefile.



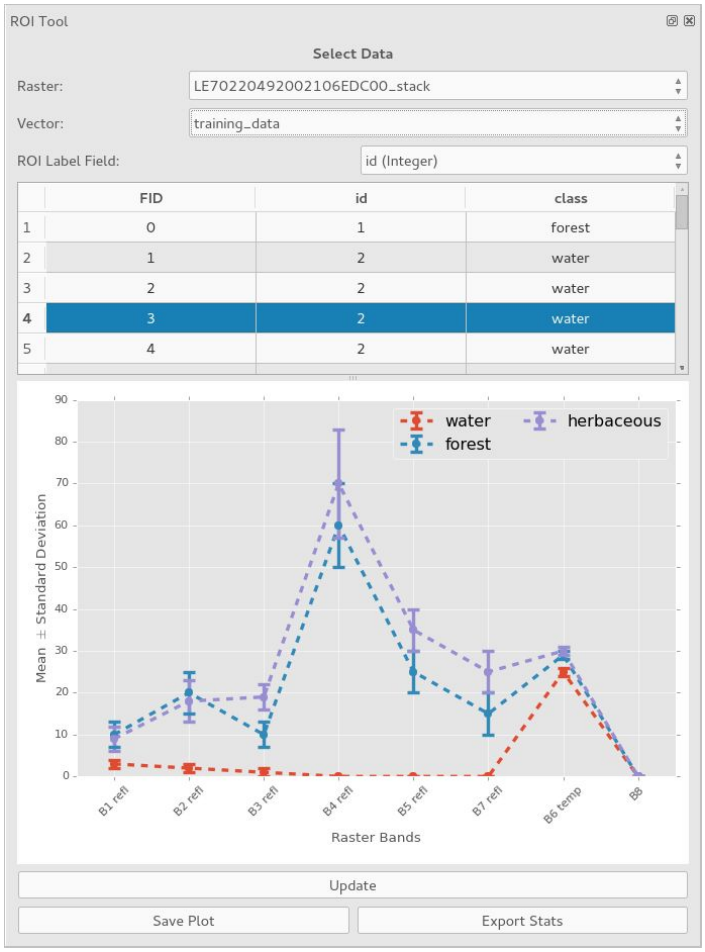
NOTE 1: You can have more than one polygon that represents the same class. In fact, you should typically have many polygons for each class! Many small ROIs are better than a few large ones (a few hundred pixels tops for each ROI). This can help account for the spectral variations in each land cover class. If you provide just one example of forest, it will be hard to recognize the variety of spectral signatures associated with that class. If you have multiple sub-classes under the same class (such as different tree species all being considered forest), your ROIs should reflect the full breath of possible spectral signatures.

NOTE 2: You will find that you will need to adjust your training data a number of times (maybe many times!) in order to produce a good quality classification. The actual classification part is just pressing a button, what reflects the accuracy of your results is the quality of your training data. Make sure to take your time when collecting ROIs.

NOTE 3: The amount of ROIs for each class should be proportional to the total area of that class in the image. If you have more forest than water, then you should have more ROIs for forest than you have for water.

S2.3 Analyze your ROIs [TODO]

1. To visualize the spectral signatures of your training data ROIs, load the "ROI Explorer" plugin within QGIS. Open the "Plugins" menu and click "Manage and Install Plugins". Under the "Installed" tab, enable the "ROI Explorer" plugin. If this plugin is not currently installed, you may need to install it from the "Available" tab.
2. The user interface of the "ROI Explorer" plugin should appear docked in the right side of QGIS. If this user interface did not appear, right click on any QGIS toolbar and show the user interface by clicking the "ROI Explorer" dropdown item.
3. The "ROI Explorer" user interface tracks all raster and vector data opened within QGIS. Select from the "Raster" drop down box the raster image you want to use for your classification. Using the "Vector" drop down box, select the vector layer containing your training data regions of interest. Use the "ROI Label Field" drop down box to select the field (column) within your training data vector file that contains the training data labels. This "ROI Label Field" selects the field that is used to aggregate features into class labels.



4. Now that your raster and vector files are selected, you may view each training data feature within the included attribute table. Selecting features from this attribute

table will select them for visualization within the plot. You may select more than one feature by using the Control key and clicking additional features, by using the Shift key to select a range of features, or by using the Control + A shortcut to select all features.

5. Click the "Update" button below the plot to calculate the mean and standard deviation of all raster bands for each unique class label. The plot will update and display the mean as a point on the plot with the errorbars above and below the point representing plus or minus one standard deviation.
6. Continue to analyze your ROIs for separability. Highlighting ROIs from different classes in the attribute table provides an understanding of the across class variability.
7. You might also want to analyze the within-class variability of your ROIs. You might create another field within your vector layer that provides additional class information, such as "degraded", "growing", or "mature" labels for examples of "forest', or simply provide a unique identifier for each example of "forest". If you select this additional field within the "ROI Label Field" drop down box, then you could analyze the differences among these example forest sub-classes by highlighting only features labeled "forest" and by clicking "Update".
8. You can export the currently displayed plot to a variety of image formats (PNG, JPEG, EPS, etc.) using the "Save Plot" button. A dialog will appear allowing you to specify the filename and file format. Click "Save" to save your plot. Additionally, the computed statistics for the currently plotted classes may be exported as a CSV file using the "Export Stats" button.

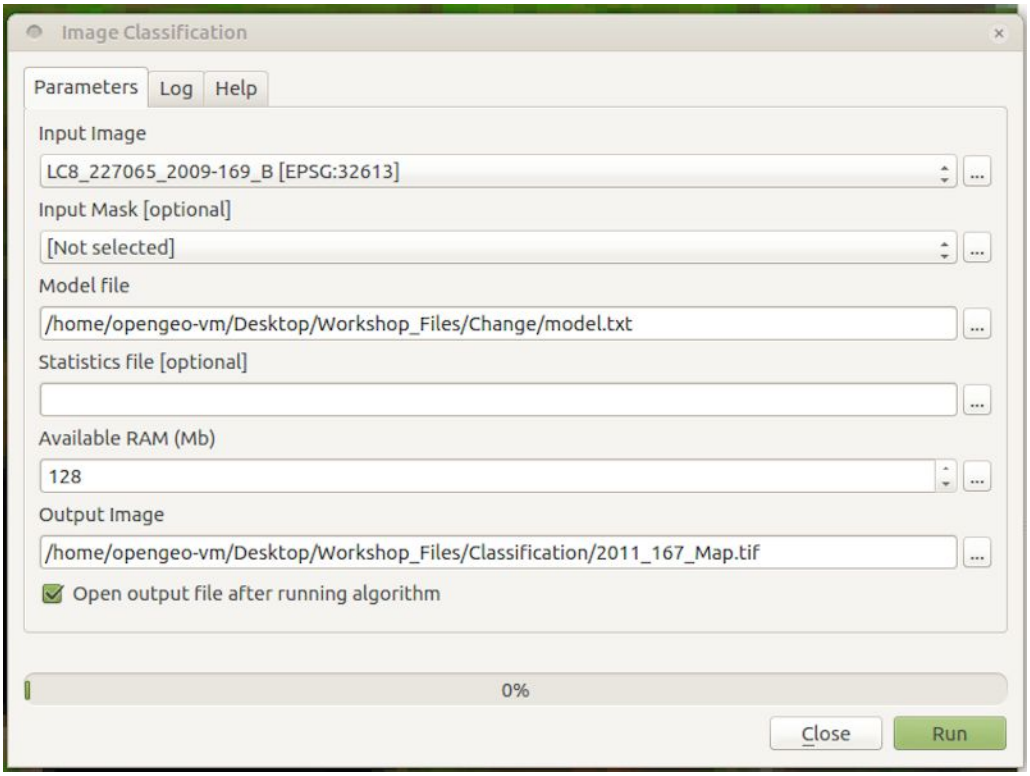
S2.4 Train your classifier

1. To find the tool needed, go to *Processing > Toolbox.*
2. In the toolbox that pops up on the right side of the screen go to *Orfeo Toolbox > Learning > Train > ImagesClassifier(rf).*
3. Then For 'Input Image List' select your Landsat image. For *Input vector Data List* select your ROI shapefile. Make sure the 'Name of the discrimination field' is the name of the attribute you used to differentiate the classes (it should be 'Class'). For *Output Model*, click *Save to file…* and save it as a .txt file in an appropriate location.
4. For now, leave the other fields as they are. These are different parameters for what goes into the Random Forest classification. If you'd like, you can play around with the different parameters and see how they affect the end results. When doing so, you will need to retrain the classifier with new output models. To train the classifier click *Run.*

2.5 Classify your map

1. To find the tool needed for classification, return to the *Processing Toolbox* and go to *Learning > Image Classification.*
2. For *Input Image* select your Landsat image. For *Output Image* select *Save to file…* and

select an appropriate location and filename to save the image. Leave the rest of the fields as they are. Your *Image Classification* box should look something like the image below. Select *Run.*



S2.6 Mean-Shift Clustering

1. [To do.]