Exercise 2:
Improving Your Supervised Classification

Introduction

Digital image classification techniques are used to group pixels with similar values in several image bands into land cover classes. Common approaches are unsupervised, supervised, and object-based. This webinar series will focus on the supervised approach. In supervised classification, the user selects representative samples for each land cover class in the digital image. These sample land cover classes are called "training sites." The image classification software uses the training sites to identify the land cover classes in the entire image. The classification of land cover is based on the spectral signature defined in the training set. The digital image classification software determines each class on what it resembles most in the training set.

In this exercise we will build on Exercise 1 by learning how to create multiple training sites (ROIs) for each class. We will also learn how to improve the classification by analyzing the training sites both visually and through statistics.

Data requirements

For this exercise you will need:

- Landsat_calaveras.tif (from exercise 1)

Objectives:

- Learn how to generate multiple training sites (regions of interest) for each land cover type
- Analyze and refine spectral signatures of training sites by
    - Creating spectral plots
    - Definining spectral thresholds
- Use Maximum Likelihood classifier to classify land cover image

Part 1: Create Multiple Training Sites (Regions of Interest)

The goal of training is to provide examples of the variety of spectral signatures associated with each class in the map. As a reminder, here are some general rules to follow when creating training sites:

- Select as many training sites per class as possible

- Select training sites throughout the entire image, not just one area
- Training site selection must be of spectrally homogenous areas (as much as possible)
- Training sites must be as large as possible

In the first exercise you only created one ROI for each land cover class. One of the important rules of creating supervised land cover classifications is to capture all the spectral variability in each land cover class. If you only choose one ROI for each land cover class, then you are not likely to capture all the spectral variability. To improve your classification, you must create several training sites for each class located in different parts of the image. In this exercise we will create 3 training sites per class.

For this exercise we will use the same 6 land cover classes that we used last week, however because the QGIS Supervised Classification Plugin (SCP) requires you to define land cover classes by a macroclass ID (MC ID) and class ID (C ID), we will define the classes a little differently.

Macro Class Name	Macro Class ID	Class Name	Class ID
Water	1	Water 1	1
		Water 2	2
		Water 3	3
Forest	2	Forest 1	1
		Forest 2	2
		Forest 3	3
Oak	3	Oak 1	1
		Oak 2	2
		Oak 3	3
Burned Area	4	Burned 1	1
		Burned 2	2
		Burned 3	3
Forest Harvest	5	Harvest 1	1
		Harvest 2	2
		Harvest 3	3

| Bare Ground	6	Bare 1	1
		Bare 2	2
		Bare 3	3

- Open QGIS and display the Landsat_calaveras image in the Layers Panel by clicking on the Add Raster Layer icon
- Right click on Landsat_calaveras and select Properties. Make sure Style is selected on the left side.
- Select band 5 for the red band, band 4 for the green band and band 3 for the blue band. Make sure Contrast enhancement is set to Stretch to MinMax. Click Load on the right, then click OK at the bottom.
- Click on the SCP Dock tab to bring it forward.
- Under SCP input, you will see **Input image**. Click on the drop down button and select your clipped image (Landsat_calaveras). If you do not see it appear, then click on the Refresh list button [icon], then look again.

- For this exercise we will create a new Training input file. Under SCP input (below Input image), you will see **Training input**. Click the "**Create a new training input**" icon [icon], navigate to the right folder, and define a name (training2.scp). Click **Save**. The path of the file is displayed in the Training input. You will also notice that if you click on the Layers Panel, a vector layer called "training2" is added.
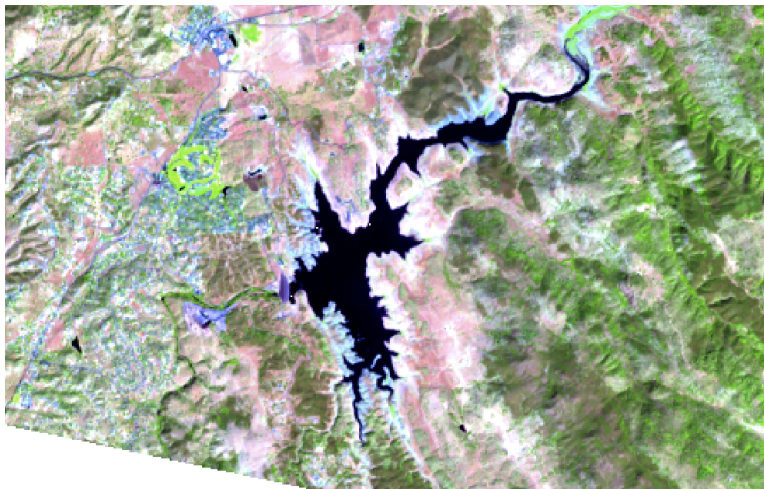
Using the table above we will begin by creating 3 ROIs for the water class.

- We will create the ROIs by using the Activate ROI pointer icon . First zoom into the large lake that is to the left (west) of the burn scar.
- Keep the Dist at 0.01000, the Min 60 and the Max 100. Click on the Activate ROI pointer icon and click in the middle of the lake. You should end up with something like this:
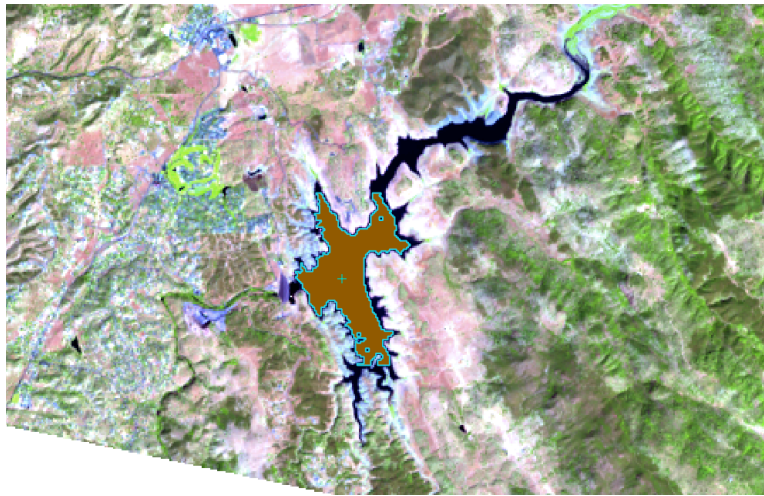

- In the SCP dock on the left, open the **Classification dock**. This is in the main SCP dock and is shown as a table below the **SCP input** tab. At the top of that dock you will see **ROI Signature list** and below that you will see **ROI creation**. In the **ROI creation**, make sure the MC ID = 1 and the C ID = 1 (from the table above).

- Next to MC Info, put Water and next to C Info put Water 1. Now click to save the ROI in the Training input. You may have to scroll down on the ROI creation box to see the save button.

You will then see the signature appear in the ROI Signature list. Don't worry about the Color for now, we will change that later.
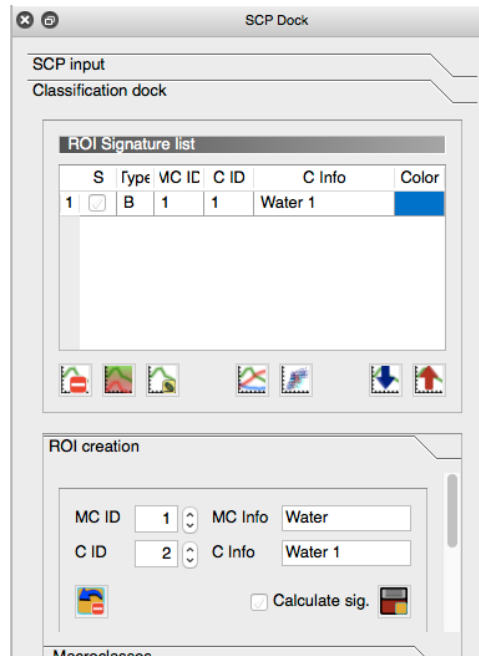

Now we are going to create a second ROI for the water class.

- Zoom out to the entire image using the Zoom full tool at the top [icon] . Or you can do this by right clicking on the Landsat_calaveras file in the **Layers Panel** and clicking on **Zoom to Layer**.
- There are two small lakes to the east of the fire scar. Zoom in to one of those.
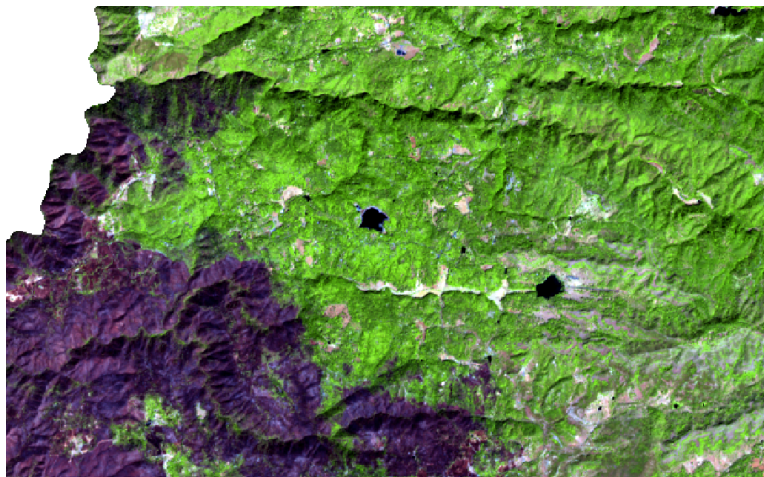
- Click on the **Activate ROI** pointer icon and click in the middle of the lake. An orange polygon will appear. Keep the MC ID as 1, and make sure the C ID is 2, keep the MC Info as Water and change the C Info to Water 2 and save the signature.
- Zoom back out to the full image and zoom in to the partial lake north of the first lake, near the edge of the image.
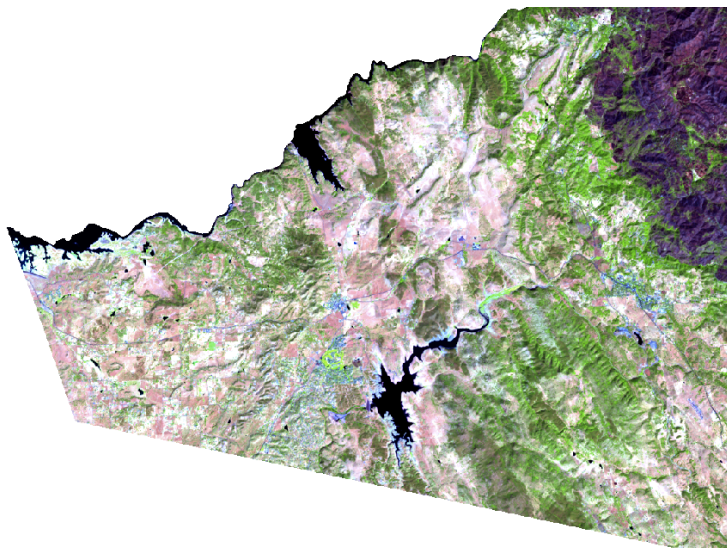

- Click on the **Activate ROI** pointer icon and click in the middle of the lake.
- Make sure the MC ID is still 1 and the C ID is now 3. Leave MC Info as Water, and change the C Info to Water 3. Save the signature. Your ROI signature list should now look like this:

- Next we will do the Forest Macro Class.
- Zoom out to the full image, then zoom into one of the darker green areas to the left (east) of the burn scar. You will see a lot of color variation in the forested area, but we will try and just pick ROIs in the darker green areas.
- Click on the Activate ROI pointer tool, then click in a dark green area.

You will notice that the ROI may be rather small. This is because there is more spectral variation in the forested area than in the water.

- Increase the Dist to 0.04000, then click on the **Activate ROI** pointer tool and click in the dark green area again. The ROI should become much larger.
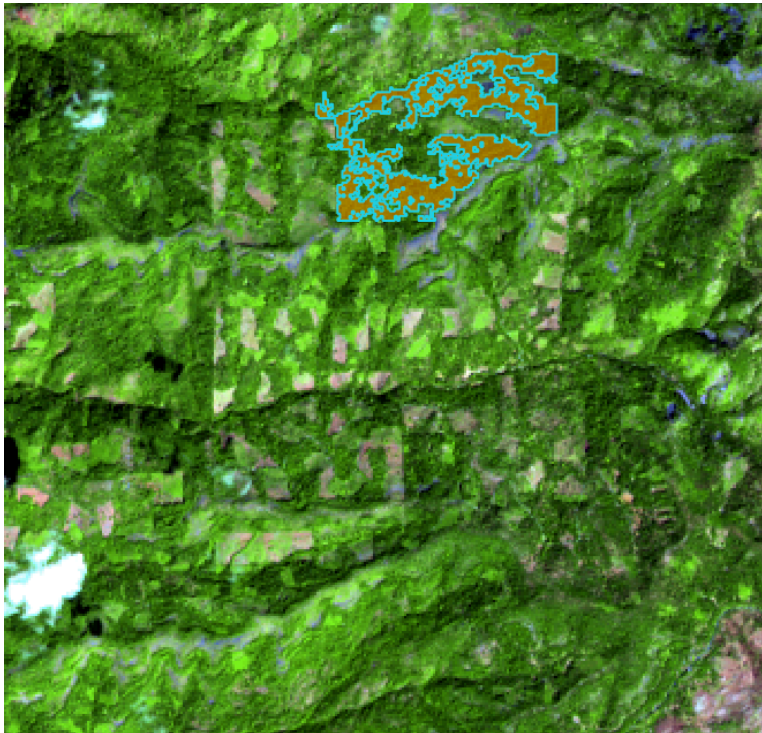

- Increase the MC ID to 2 and decrease the C ID to 1. Change the MC Info to Forest and the C Info to Forest 1.
- Repeat the same process for two additional forest ROIs. Make sure the MC ID remains at 2 increase the C ID to 2 and then 3. The MC Info will remain as Forest, but the next two ROIs will be Forest 2 and Forest 3. Remember to try and pick areas in different parts of the image.

After you finish, the ROI Signature list should look like this:

	S	Type	MC ID	C ID	C Info	Color
1		B	1	1	Water 1	
2		B	1	2	Water 2	
3		B	1	3	Water 3	
4		B	2	1	Forest 1	
5		B	2	2	Forest 2	
6		B	2	3	Forest 3	

The next Macro Class will be Oak. Focus on the light green areas that are around the fire scar. Zoom out to the full image then zoom into one of the green areas just to the left (west) of the fire scar.

- Keep the Dist as 0.04000. Click on the **Activate ROI** pointer tool and click in one of the light green areas. Change the MC ID to 3 and the C ID to 1. Change the MC Info to Oak and the C Info to Oak 1. Save the signature.
- Find two more light green ROIs in different areas, keep the MC ID as 3 and change the C ID to 2 and 3. Keep the MC Info as Oak and change the C Info to Oak 2 and Oak 3. Make sure you save your signatures as you complete them.
- Next we will choose ROIs in the burned area. Follow the same directions as above, but make sure to change the MC ID to 4 and the C ID to 1,2 or 3. Change the MC Info to Burned Area and the C Info to Burn 1,2 or 3.
- The next ROIs will be for Forest Harvest. Zoom into an area somewhere between the far right (east) of the image and the eastern border of the burn scar.
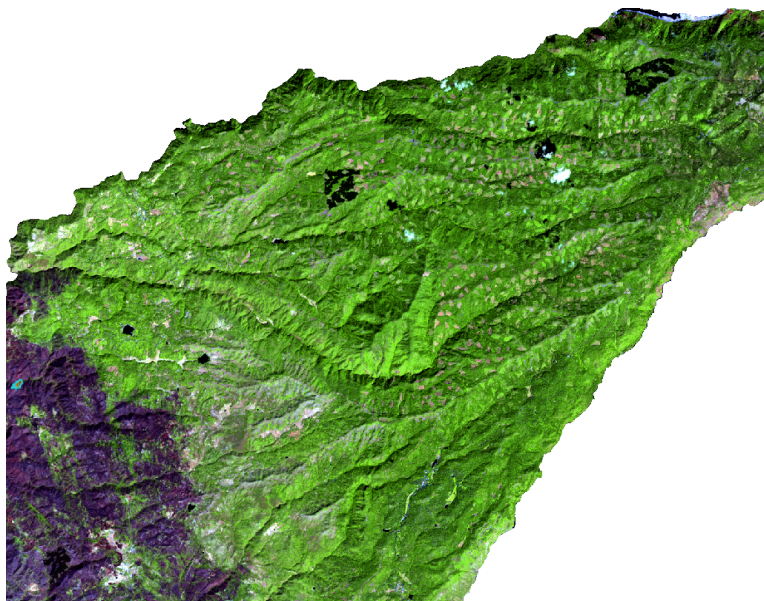
- Choose 3 ROIs that capture the spectral variability of these forest cuts. We are only interested in more current cuts, so do not choose anything that looks like it is regrowing with new vegetation. Change the MC ID to 5, and the C ID to 1,2 or 3. Change the MC Info to Forest harvest and the C Info to Harvest 1,2 or 3.

- The last class is Bare Ground. Zoom into an area to the left (west) of the burn scar.
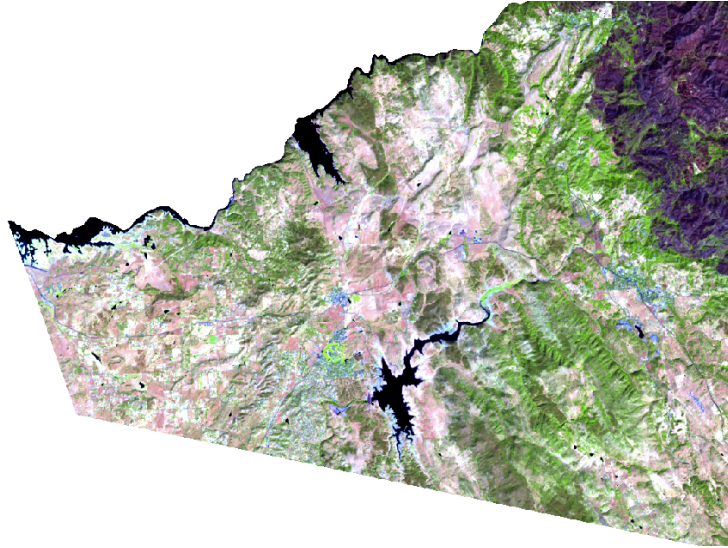
Choose 3 ROIs in the white/pink areas. Make sure you change the MC ID to 6 and the C ID to 1,2 or 3. Change the MC Info to Bare Ground and the C Info to Bare 1,2 or 3.

You should have a total of 18 signatures in your ROI Signature list.

Part 2: Analyzing ROI Signatures

Before we analyze signatures, we can use the SCP plugin to create a preview of your classification to assess your signatures.

- Click on **Macroclasses** to pick the colors for your classification. You can choose any color you want. To change the colors, double clicking in the Color box. For now, make water blue, forest dark green, oak light green, burned area purple, forest harvest orange, and bare ground light yellow.
- Click on **Classification algorithm** and click Use MC ID. Under Land Cover Signature Classification, click **Use LCS**. This means that the ROI signatures you created will be used to create the classification.
- In the SCP toolbar above the image, make sure Preview (located to the right of ROI) is clicked on. T allows you to set the transparency of the preview image and S allows you to set the size of the preview image in pixels. Set S=500.
- Zoom into the region to the left (west) of the burn scar.

- Click the **Activate Classification Preview Pointer** button and then click somewhere near the large lake.

You will notice the blue of the lake and lots of yellow and green pixels indicating bare ground (yellow) and oak (green). You will also notice many black pixels. These are pixels that did not get classified because their signatures did not match any of your ROI signatures or that there is a lot of overlap between classes.

There are several different things you can do at this point. First, we will create a new class to include some of the unclassified pixels.

If you click the Preview image on and off (by clicking the radio button to the left of Preview) you may notice some blue/purple areas north of the lake that did not get classified. These are small urban regions. To include these regions, we need to create a new class and add more signatures.

- Turn off the Preview and zoom into one of the blue/purple regions.
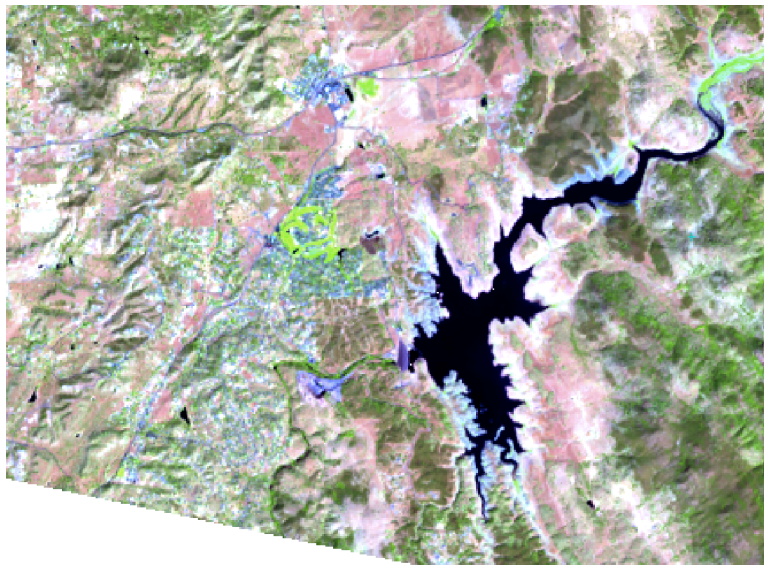

-
- Make sure ROI is clicked on. Increase the Dist to 0.080000 and click on the **Activate ROI** pointer tool. Click in the blue/purple urban area. Go to the ROI creation area in the SCP Dock and increase the MC ID to 7 and change the C ID to 1. Change the MC Info to Urban and the C Info to Urban 1. Save the signature.
- Create two more urban ROIs, remembering to increase the C ID to 2 and 3 and the C Info to Urban 2 and 3.
- Now we will create a new preview classification with the new urban class included. Click on **Macroclasses** in the SCP Dock and change the urban color to a light purple.
- Click on **Classification algorithm** and make sure you are using the MC ID and the LCS.

- Click the Preview on and the first Preview image will appear. You can remove

    this Preview classification by clicking on the trash can icon to the right of the S=500 along the top of the window. This allows you to remove temporary files.
- Pan to where you can see the lake.
- Click on the **Activate classification preview** pointer icon and click somewhere near the lake again. You will see a lot more light purple pixels appear.
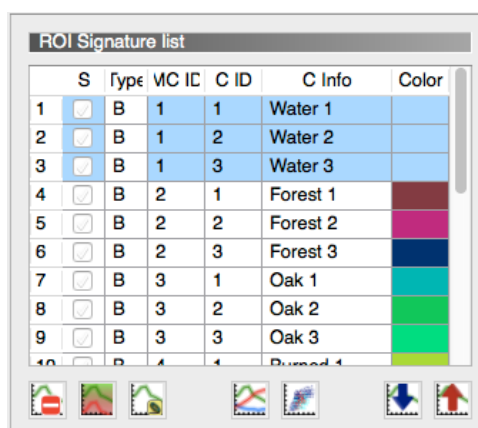
As you can see, the urban areas were classified but a lot of areas that are not urban were also classified as urban. At this point you could determine what those other land cover types are and create a new class using the same method as above. We are going to move on to some different ways to analyze signatures, but you can see how supervised classification can be very time consuming and challenging!

Click the Preview image back on. You will still see several black pixels outside the edge of the lake, and you might even see black pixels where there clearly is water (if you click on and off the Preview image).
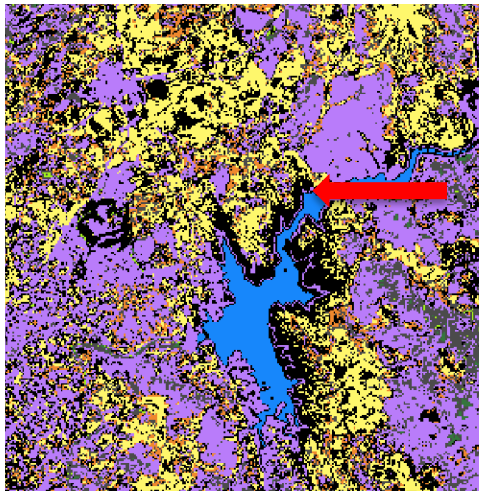

We can adjust the spectral range of the lake signatures to include the portion of the lake that was not classified.

- Highlight the three bare ground signatures in the ROI signature list by left clicking and dragging the mouse down the numbers 1,2, and 3 (the three Water signatures), or you can left click to select signature 1 and then Shift click to select signatures 2 and 3. Once you highlight them the rows should turn blue.

Click on the button "Add highlighted signatures to spectral signature plot" [image], located below the ROI signature list. The spectral signature plot dialog box will appear.
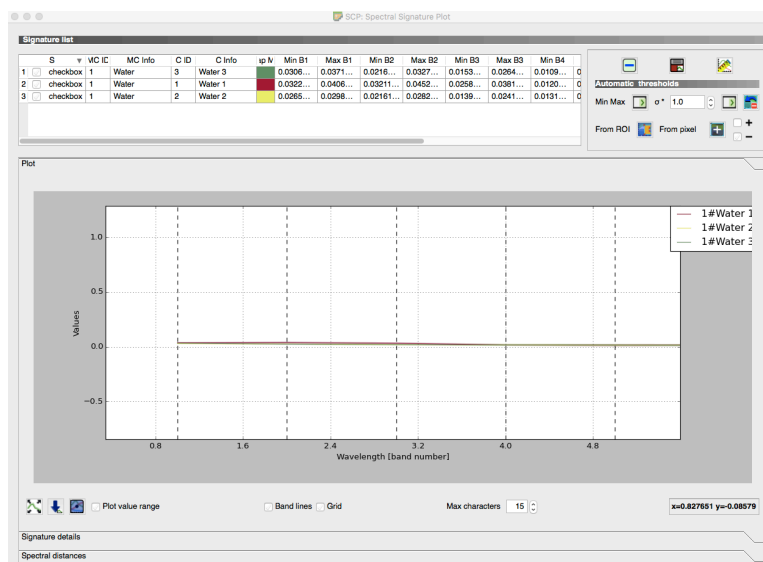

This box includes the Signature list on the top. This list shows the minimum and maximum values for all the bands included in the image. Below the list is the spectral plot of the signatures. The x-axis is the wavelength by band number and the y-axis are the pixel values.

You can display the plot value range, which shows the minimum and maximum values for each band, and the location of the band lines. You can see that the water signatures do not vary much between bands. You can use tools in the Automatic thresholds area (right of the Signature list) to adjust the thresholds of the signatures. You can select individual pixels from the image or you can create a ROI to change the thresholds of the signatures. We will expand the threshold of the water signature using a ROI.

- Go back to the image with the preview classification and make sure the ROI is clicked on. Make sure the Dist is 0.04000. Click on the **Activate ROI** pointer and click in the area on the edge of the lake where the pixels are black. A ROI will appear.
- Go back to the Spectral Signature Plot box. Select Water 1 by clicking on it's number (the whole row should be highlighted).

|	S	▼	MC ID	MC Info	C ID	C Info	ap M	Min B1	Max B1	Min B2	Max B2	Min B3	Max B3	Min B4	
1	✓ checkbox		1	Water	3	Water 3		0.0306...	0.0371...	0.0216...	0.0327...	0.0153...	0.0264...	0.0109...	0
2	✓ checkbox		1	Water	1	Water 1		0.0322...	0.0406...	0.03211...	0.0452...	0.0258...	0.0381...	0.0120...	0
3	✓ checkbox		1	Water	2	Water 2		0.0265...	0.0298...	0.02161...	0.0282...	0.0139...	0.0241...	0.0131...	0

- In the Automatic Thresholds area, make sure the + is clicked on. Click **From ROI**. That will change the minimum and maximum values for that signature using the ROI you just created.
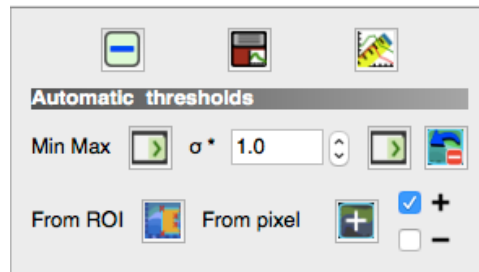
- Go back to the image and delete the preview image by clicking on the trashcan to the right of Preview.
- Click on the Activate classification preview pointer again and click somewhere near the lake.

A new preview classification will appear. You will see that the area that previously had black pixels, is now classified as water.

Now we will work on identifying clouds and cloud shadows. In the area to the right (east) of the burn scar you will see two small white areas and two small black areas above them. Zoom into that region.
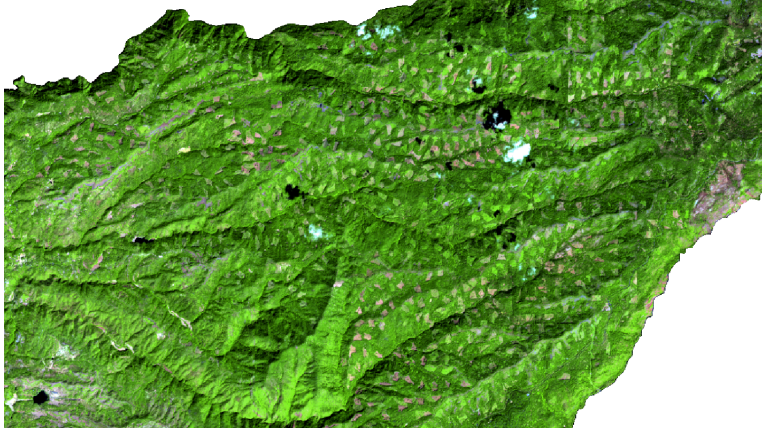
Create a Preview classification in between the two clouds. You will see that most of the clouds and their shadows are not classified (although there is some confusion with urban).

Use the same method as above to create an ROI for the clouds and create a new Macroclass (8) called Cloud. We will only create one ROI for clouds and one for Cloud Shadow.

Since we are creating a new Macroclass, you do not need the Spectral Plot window for now, so you can close it.
- Zoom into the larger cloud and use the **Create a ROI polygon** tool to make an ROI within the cloud area.
- In ROI Creation, in the SCP Dock increase the MC ID to 8 and make the C ID 1. Change MC Info to Cloud and the C Info to Cloud 1. Save the signature.
- Do the same for the cloud shadow but change the MC ID to 9 and make the C ID 1. Change MC Info to Cloud shadow and the C Info to Shadow 1. Save the signature.
- Under Macroclasses, make the Cloud white and the Cloud shadow brown.
- Zoom back out and create a new Preview classification.

You will notice that most of the cloud and cloud shadow pixels for the cloud that you used for the ROI are now classified, but the other prominent cloud may not be. In this case, you could either create an additional ROI signature and add it to the list or you

could use the Spectral Signature Plot tools to increase the threshold of that signature (like we did for water).

Now we will analyze the signatures using the information in the Spectral Signature Plot.

First highlight all the signatures in the ROI Signature list by left clicking and dragging down the entire list of numbers. If you closed the Spectral Signature Plot box then click on the icon to reopen it.

All the ROIs in the signature list and all the Signatures are plotted below the signature list. You will also notice that some of the signatures are highlighted orange and some are not. The signatures that are highlighted orange indicate that they overlap with other signatures. You can scroll down to see which signatures overlap and which do not. To see how signatures overlap, you can expand the column labeled Color (overlap MC_ID-C_ID)

Signature list

	S	▼	MC ID	MC Info	C ID	C Info	Color [overlap MC_ID-C_ID]	Min B1	Max B1	Min B2	Max B2	M
1	checkbox		5	Forest Harvest	2	Harvest 2	7-2;	0.0400...	0.0559...	0.0556...	0.0872...	0.0
2	checkbox		4	Burned	3	Burned 3	7-3;7-2;	0.0414...	0.0669...	0.0390...	0.0728...	0.0
3	checkbox		2	Forest	1	Forest 1	7-3;	0.0178...	0.0426...	0.0218...	0.0550...	0.0
4	checkbox		6	Bare ground	3	Bare 3		0.0788...	0.10119...	0.1088...	0.1453...	0.1
5	checkbox		4	Burned	2	Burned 2	7-3;1-1;	0.0326...	0.0580...	0.0274...	0.0603...	0.0
6	checkbox		7	Urban	3	Urban 3	4-3;2-1;4-2;3-1;1-1;3-3;2-3;3-2;	0.0377...	0.1360...	0.0401...	0.1517...	0.0

In the example above, the C Info signature Harvest 2 (from the Forest Harvest Class) overlaps with signature 7-2, which is Urban 2. Urban 3 overlaps with many other signatures: 4-3 (Burned 3), 2-1 (Forest 1), 4-2 (Burned 2), 3-1 (Oak 1), 1-1 (Water 1), 3-3 (Oak 3), 2-3 (Forest 3), 3-2 (Oak 2). This is not surprising because urban is very heterogeneous and is often confused with many different things.

One of the other things you can do is look at the standard deviations of each of the signatures in the different bands. Your goal is to have a reasonably small standard deviation for each of your signatures in each band. If you have a large standard deviation, it may indicate that your ROI is not homogeneous enough. This is really only important if you used a polygon to define your ROIs, rather than the ROI pointer. With the ROI pointer you are able to specify the maximum spectral distance of your ROI, so the standard deviation should be quite small, unless you specified a large spectral distance.

To look at the standard deviations, click on **Signature details** below the plot.

For each signature, you will see the average pixel value and the standard deviation of all the pixels in that signature for each band. You can also see the total number of pixels in the ROI. If you scroll down until you see the Urban and Cloud signatures, you will see that they have the largest standard deviations. As I mentioned before, urban is not very homogeneous. As you recall, you defined the cloud ROI using a polygon rather than the pointer tool, which can often result in higher standard deviations.

Lastly, you can look at the spectral distances between the signatures, which is useful for assessing the separability between the ROIs. If ROIs are very similar then you may want to consider deleting one, or altering thresholds.

- Make sure all the signatures are checked on in the Signature list. To the right, click on the Calculate spectral distances icon.

The spectral distances tab will open.
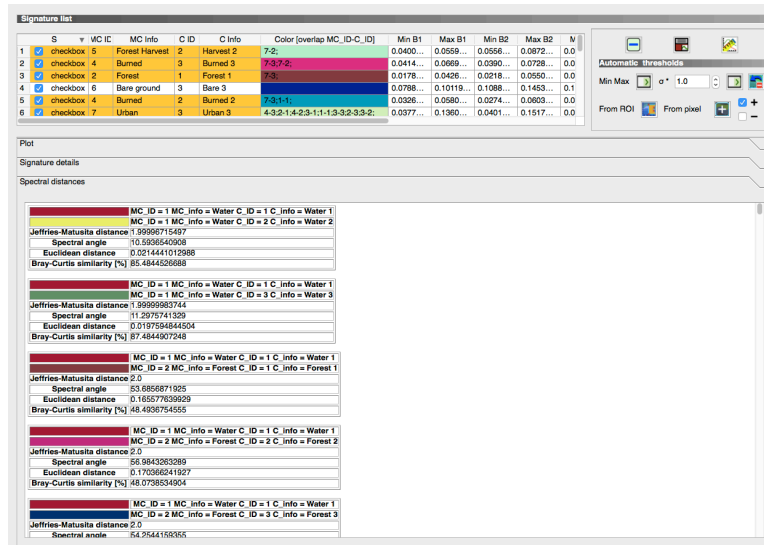
The top two lines of each box show which two signatures are being compared. Below that are four different measures of separability:

- Jeffries-Matusita Distance (0=identical; 2=different); useful for maximum likelihood classifications
- Spectral Angle (0=identical; 90=different); useful for spectral angle mapping classifications
- Euclidean Distance; useful for minimum distance classifications
- Bray-Curtis Similarity; (0=different; 100=identical); useful in general.

Note that values are displayed in red if signatures are particularly similar. Scroll down until you see the first red values. In my example, my first values are the comparison between Water 2 and Water 3.

	MC_ID = 1 MC_info = Water C_ID = 2 C_info = Water 2
	MC_ID = 1 MC_info = Water C_ID = 3 C_info = Water 3
Jeffries-Matusita distance	1.99999809509
Spectral angle	3.7891777015
Euclidean distance	0.00418593583114
Bray-Curtis similarity [%]	97.0912513054

This is not of great concern since these are both Water ROIs. You would expect the ROIs from the same Macro class to be very similar. However, the spectral distances between the forest and the oak ROIs are similar using a couple of the distance measures: spectral angle and Euclidean distance.

	MC_ID = 2 MC_info = Forest C_ID = 1 C_info = Forest 1
	MC_ID = 3 MC_info = Oak C_ID = 1 C_info = Oak 1
Jeffries-Matusita distance	2.0
Spectral angle	8.14361844243
Euclidean distance	0.137606307165
Bray-Curtis similarity [%]	73.2796879748

	MC_ID = 2 MC_info = Forest C_ID = 1 C_info = Forest 1
	MC_ID = 3 MC_info = Oak C_ID = 2 C_info = Oak 2
Jeffries-Matusita distance	1.99999999996
Spectral angle	9.82365198457
Euclidean distance	0.139345037022
Bray-Curtis similarity [%]	72.4096089423

	MC_ID = 2 MC_info = Forest C_ID = 1 C_info = Forest 1
	MC_ID = 3 MC_info = Oak C_ID = 3 C_info = Oak 3
Jeffries-Matusita distance	2.0
Spectral angle	7.4056744874
Euclidean distance	0.152388752788
Bray-Curtis similarity [%]	71.1446568518

Depending on the classification algorithm you are planning to use, you may want to adjust the thresholds of some of these signatures or delete them. In our case, we are going to use the Maximum Likelihood classifier, so we will not change the signatures. At this point you might want to scroll through the entire list to see if there are any comparisons between signatures that are especially similar and decide if you need to make any changes to the signature.

Once you are satisfied with your signatures, you can run the classification.
- Click on the **Classification algorithm** in the SCP Dock. Make sure Use MC ID is selected and select Maximum Likelihood as the Algorithm.
- Click **Algorithm** under Land Cover Signature Classification.

- Click on **Classification output** and click on the **Run** button. Save it as Calaveras_class2.tif. It will take a little while to run the classification algorithm. Once it is finished, the final classification image appears.
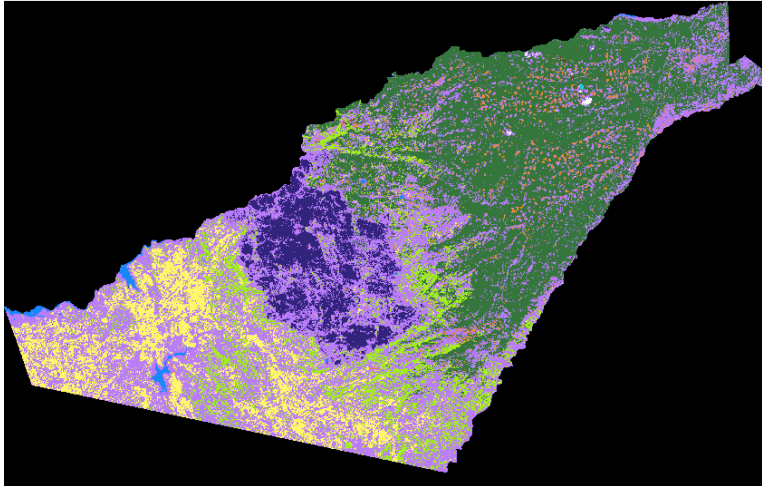
In my final classification, urban gets confused with everything. As mentioned previously, urban is extremely heterogeneous spectrally, so this result is not surprising. The decision of what to do depends on the goals of your project. You could eliminate the urban signatures and reclassify, or you could find some other way to separate out urban from other signatures. It is very difficult to create an accurate land cover map of a diverse area based solely on spectral information. It can be quite useful to incorporate other ancillary information, such as slope, aspect and elevation and create a decision tree. Many people divide their study area into ecoregions, or bioregions, and then classify the imagery, which often improves the land cover map and results in less confusion.

This exercise has demonstrated a few tools that can be used with the Supervised Classification Plugin in QGIS. The SCP has several other tools available and I encourage you to read the Semi-Automatic Classification Plugin Documentation by Luca Congedo for many more ideas.