Jason Heutmaker

# Image Classification

*Introduction*

During this lab we learned about the processes of image classification where classification is the process of sorting pixels into a set number of classes or categories based on the pixel values. In geospatials, there are two different main ways that we can classify pixels when using an electronic source: supervised and unsupervised classification. When using supervised classification we collect the spectral signatures that are obtained form training samples in order to classify the image and the computer then uses the users specified areas in order to classify all of the pixels. These training samples are used to collect the spectral signatures from the areas chosen by the user that have been designated as a certain class (water or agriculture). This lab goes over using both the supervised and the unsupervised classification of a Landsat 8 scene taken from San Francisco Bay area. We performed a classification of the same area from two different time frames: 2015 and 1987 in order to compare the two subsetted and radiometrically calibrated images. While doing the lab we learned how to conduct supervised and unsupervised classification procedures on ENVI as well as how to analyze the classification results including labeling and color coding.

*Methods*

After setting up a general workspace on the D: drive we jumped right into the procedure discussing unsupervised classification. To begin, we opened the “BayArea2015.dat” file in ENVI and in the toolbox window selected the Classification Workflow. Once opened the File Section panel appears and the BayArea2015.dat file was chosen, and we then opened the Classification Panel. In this panel we selected No Training Data, selected five as the number of classes to define, and began our classification. Once the classification was complete the classified image loads in the view and the Cleanup panel appears and we allow it to run the cleanup procedure. After this is done we unchecked the Export Classification Vector, checked the Export Classification Image, and checked the box next to "Export the Classification Statistics”. After this was done our image had been completely classified, all that remained was to label and then change the color of the classes. With the classified image in the view we compared the classes to the Landsat image to determine the land cover type of each of the five classes. Next we proceeded to use the tool Edit ENVI Header in order to change the names and colors associated with each class. Under Class Name we updated the name to the land cover type and then selected an appropriate color for the class for all five classes. Once this was done we saved it as a TIFF file and were done with the unsupervised image. We then proceeded to check on the statistics that we had acquired by opening the .txt file we received from exporting the classification statistics. Once this was done we pasted the classification statistics into Excel, changed the color of each cell to match the color of the classification output, and calculated the area in Hectares for each of the classes. We then ran the 2015 file with a supervised classification which is almost exactly the same process. The only difference is that instead we select Training Data at the very beginning and draw three polygons to select an area as a training area for each class (each class gets three training area polygons). Once all five classes have three polygons for training areas you run the process as described above including the statistics. When we had finished with the 2015 data we decided to run an unsupervised classification on the 1987 data and collected the same statistics as well. Once we had acquired

both a TIFF image and the statistics we compared the statistics to each other to find the percent change between 1987 and 2015.

*Results and Discussion*

When we compared the supervised classification to the unsupervised classification we found that the unsupervised classification was extremely better at identifying similar areas than the supervised. It's our opinion that in order to make the supervised become better at choosing areas then the unsupervised, then significantly more training areas than three will be required. The most frequent class size for Unsupervised was water with 37.6% and for Supervised it was Urban with 25.7%. It was interesting to note that each classification type had its own setbacks with being able to identify certain areas better than others. For instance, the Unsupervised had a more difficult time with determining which areas was forest and which were shrubs while the Supervised had issues with classifying water and urban areas. On the other hand, the Unsupervised had no problem with identifying which areas were water and the Supervised had no problems with identifying which parts were forest and shrub land. With both of these classifications there are pros and cons: the Unsupervised is faster and less of a hassle, but it will never be completely accurate while the Supervised, while it is much slower to do and would take many more training sites than we provided, can potentially become significantly more accurate. In the end it comes down to how much time you are willing to put into the classifications and how accurate you need to be. If a rough idea is what is most important, than Unsupervised is the way to go. If accuracy is what is needed, then Supervised should be the way to go, but may often times be overlooked due to how long it would take to setup all of the necessary training points to get an accurate representation. When comparing the Unsupervised Classification between 2015 and 1987 we come across a difference in the percent between the two: water increase by 0.3%, forest increase by 1.59%, shrub increases by 9.27, urban decreases by 8.48%, and mixed vegetation decreases by 2.71%. Interestingly enough the urban decreased the most and the shrub increases the most. This could be due to the strict zoning laws that were enacted between those times.

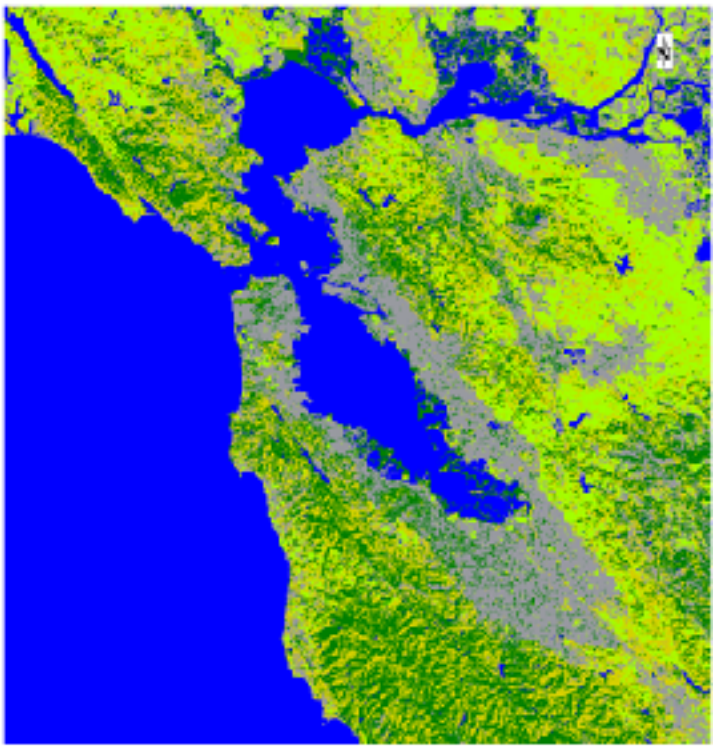

**Figure 1**. Unsupervised Classification of San Francisco Bay Area 2015. This image shows a distinct difference between the urban areas, water boundaries, as well as forest and scrubland.

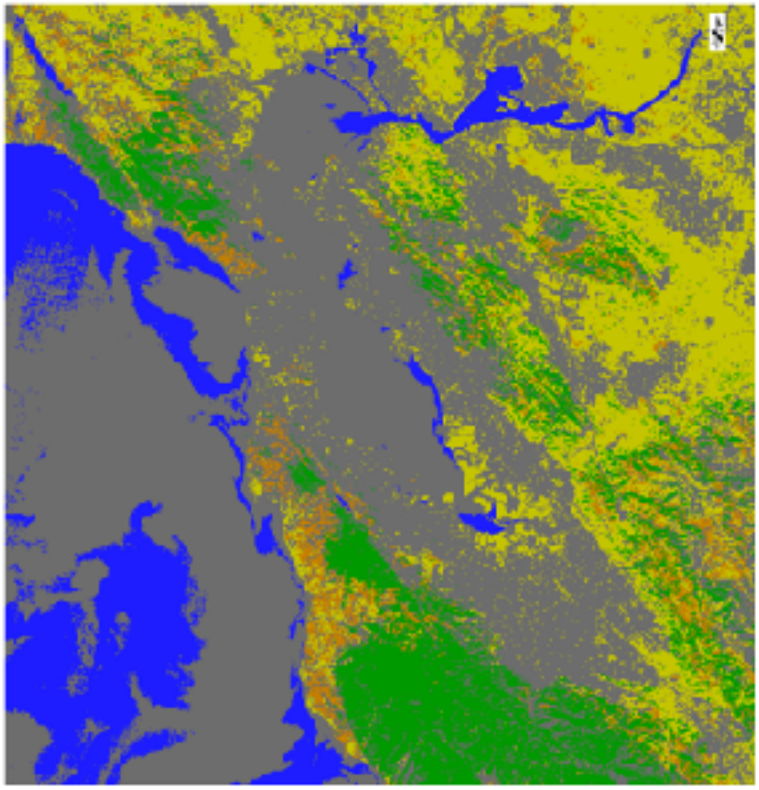

**Figure 2**. Supervised Classification of San Francisco Bay Area 2015. This image shows a lack of distinct difference between the urban areas, water boundaries, as well as forest and scrubland.

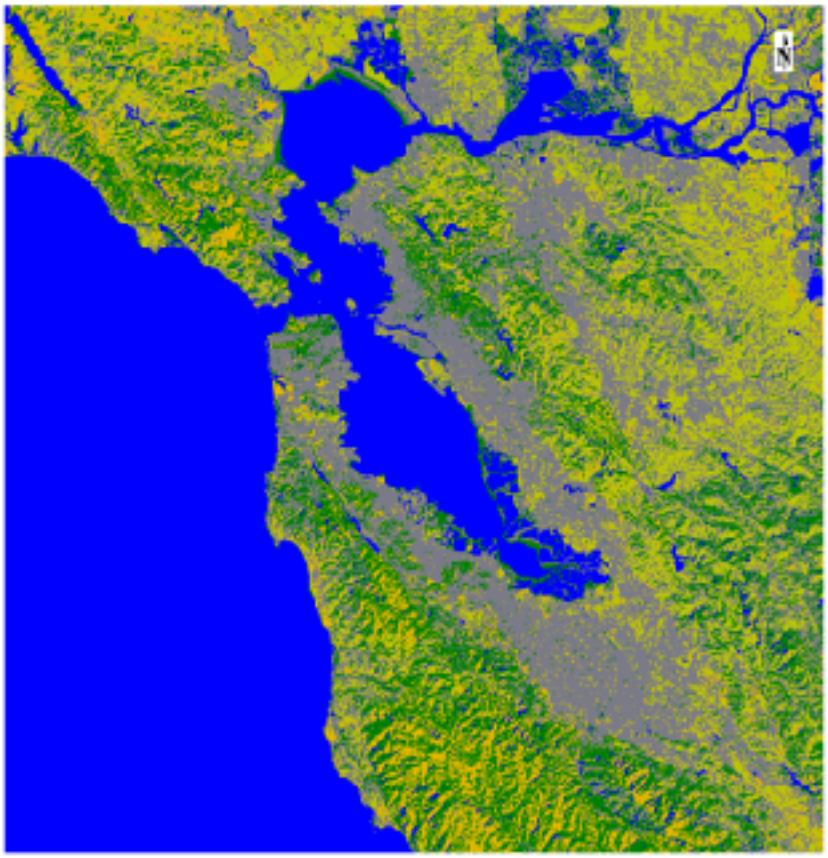

**Figure 3**. Unsupervised Classification of San Francisco Bay Area 1987. This image shows a distinct difference between the urban areas, water boundaries, as well as forest and scrubland. However, it does appear to mistake some areas for water when they should not be.

| CLASS NAME | AREA (Sq. Meters) | PERCENT | Area (Hectares) |
|---|---|---|---|
| Unsupstats | | | |
| Water | 5911596000 | 37.598531 | 591159.6 |
| Forest | 2412899100 | 15.346357 | 241289.91 |
| Shrub/Scrubland | 2637808200 | 16.776808 | 263780.82 |
| Desert/Rock/Urban | 2705034600 | 17.204377 | 270503.46 |
| Mixed Vegetation/Agriculture | 2055606300 | 13.073927 | 205560.63 |
| | | Total Area | 1572294.42 |

**Table 1**. Unsupervised Classification of San Francisco Bay Area 2015 statistics. This table shows the percent coverage of each land class as well as the total hectares that they make up.

| CLASS NAME | AREA (Sq. Meters) | PERCENT | Area (Hectares) |
|---|---|---|---|
| Supstats | | | |
| Forest | 1699854300 | 10.811798 | 169985.43 |
| Water | 1767343500 | 11.240538 | 176734.35 |
| Shrub/Scrubland | 754781400 | 4.800509 | 75478.14 |
| Urban/Rock | 8285523700 | 52.697056 | 828552.87 |
| Mixed Vegetation | 3215436300 | 20.450599 | 321543.63 |
| | | Total Area | 1572294.42 |

**Table 2**. Supervised Classification of San Francisco Bay Area 2015 statistics. This table shows the percent coverage of each land class as well as the total hectares that they make up, although it is not very accurate .

| CLASS NAME | AREA (Sq. Meters) | PERCENT | Area (Hectares) |
|---|---|---|---|
| Unsup 1987 | | | |
| Water | 5860161000 | 37.271397 | 586016.1 |
| Forest | 2161664100 | 13.748469 | 216166.41 |
| Shrub/Scrubland | 1178996400 | 7.498573 | 117899.64 |
| Desert/Rock/Urban | 4038930900 | 25.688134 | 403893.09 |
| Mixed Vegetation/Agriculture | 2483191800 | 15.793428 | 248319.18 |
| | | Total Area | 1572294.42 |

**Table 3**. Unsupervised Classification of San Francisco Bay Area 1987 statistics. This table shows the percent coverage of each land class as well as the total hectares that they make up. The water regions are most likely to high.

| Class | 1987 | 2015 | Percent Change |
|---|---|---|---|
| Unsup (2015 vs 1987) | | | |
| Water | 37.271397 | 37.598531 | 0.32713400000000 |
| Forest | 13.748469 | 15.346357 | 1.597888 |
| Shrub/Scrubland | 7.498573 | 16.776808 | 9.278235 |
| Desert/Rock/Urban | 25.688134 | 17.204377 | -8.483757 |
| Mixed Vegetation/Agriculture | 15.793428 | 13.073927 | -2.719501 |

**Table 4**. Unsupervised Classification of San Francisco Bay Area 2015 statistics vs 1987 statistics. This table shows the percent coverage of each land class as the percent change that occurred between the two time periods.