Project Summary

The project was initially intended to analyze the area within the zip code of 95678 in Roseville, California for pervious and impervious surfaces using a supervised classification. By the end of the project however, and after a range of approaches to completing the analysis were taken, the final product would not be a reliable summary of land cover types in order to conclusively determine the ratio of pervious and impervious surfaces.

Purpose

The purpose of this project was to complete a supervised classification on multispectral imagery to explore the process. Investigating the ratio of pervious and impervious surfaces was borne out of the current interest in the drought facing California in general and in particular those communities relying on water from Folsom Lake. Due to the drought, Roseville is currently asking residents to reduce water use by 20%, and asking businesses to reduce outdoor irrigation by 30%. Originally, a particular interest was in the number of large spaces with lawns for recreational use. Within the area of the 95678 zip code there are two golf courses of the six in Roseville, one public and one private. The project was however broadened to merely evaluate pervious and impervious land covers. Analysis for thirstier outdoor landscaping specifically will have to take place in the future.

Methods

Study Area and Imagery

The study area of this project was the portion of Roseville, California that falls within the 95678 zip code. It is centrally located within the city and includes low density residential areas, commercial areas and open spaces.

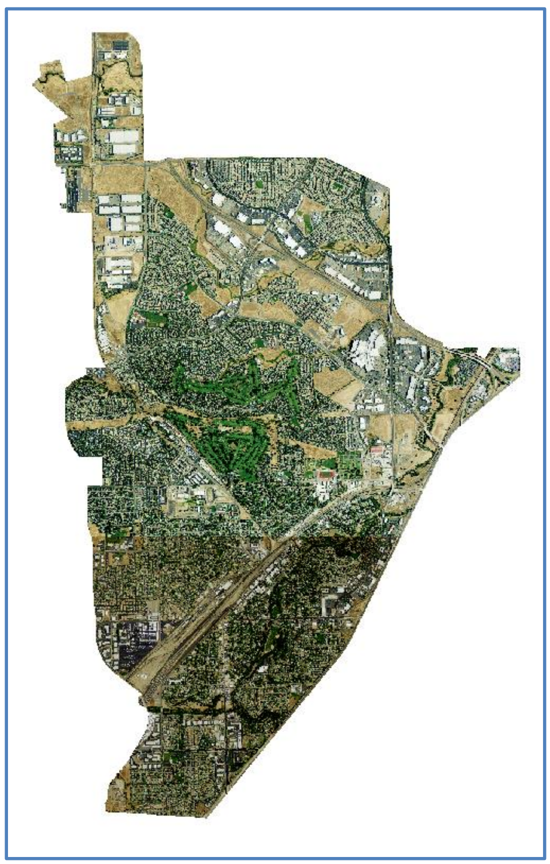


Figure 1: Roseville, CA 95678

Two different image types were obtained over the course of the project for the sake of comparing the outcomes. The final product however was based on aerial imagery obtained from the Cal-Atlas website (http://atlas.ca.gov/imagerySearch.html). It was an orthorectified digital aerial image produced by the USGS at a resolution of 1 meter referred to as "2009 Combined NIR NAIP". The image contained 4 bands including the Near Infrared band.

The image was clipped using the polygon for the 95678 zip code to reduce the project imagery to the area of particular interest which resulted in an area of approximately 80634 acres.

Exploratory Unsupervised Classification

An initial unsupervised classification was run using 18 classes. The results could be more or less

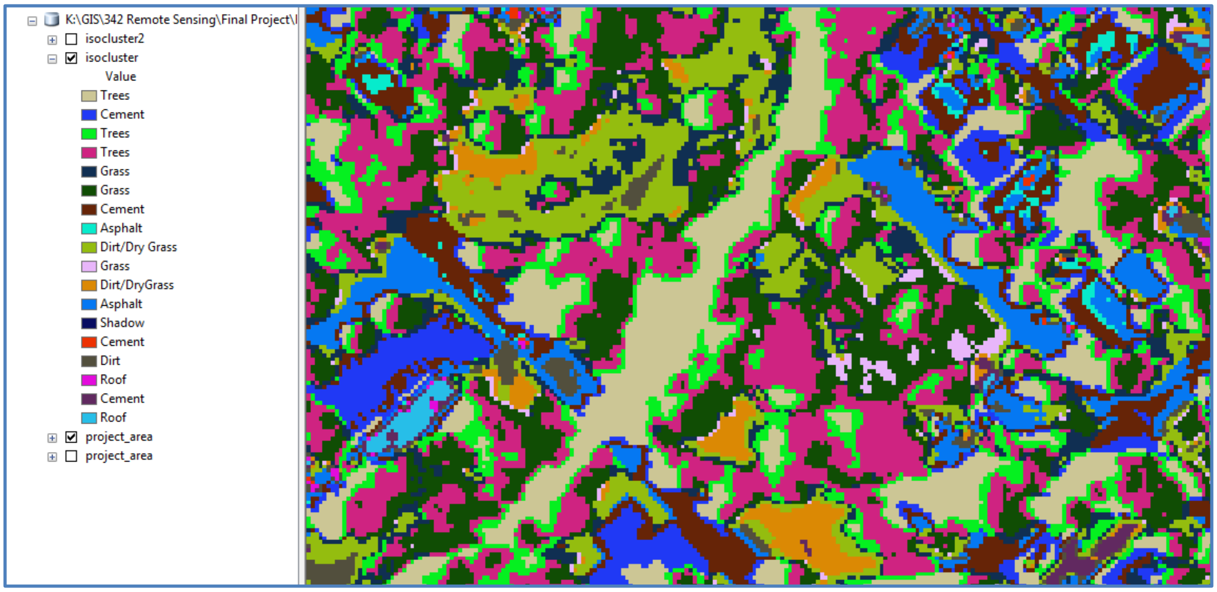


Figure 2: Resulting unsupervised classification using 18 classes

categorized into 7 land cover types. These were Trees (x3), Grass (x3), Cement (x4), Asphalt (x2), Dirt and Dry Grass (x3), Roofs (x2) and Shadows (x1). There is a creek that runs through the project area; however the water was categorized as cement. The unsupervised classification did not differentiate between what would be a road and a water body.

Combination of cover types was noticed between things that would fall in the category of "urban". Examples of this kind of cover type confusion would be between hard surfaces for roads and sidewalks and roofs of building both residential and industrial. For the purposes of this project this confusion would not be too much of a problem since the goal is determining pervious and impervious surfaces. One particular type of confusion that would cause some difficulty in trying to quantify the amount of pervious and impervious surfaces would be the confusion between water and shadows. Water in the form of a creek that runs through the project area was classified the same as shadows cast by trees.

An additional unsupervised classification was run using 100 classes to explore the outcome of further refinement of pixel clusters. The various features within the project area are more discernible. The issue of cover type confusion still exists however, but between more spectral classes since there are more classes representing a given land cover type.

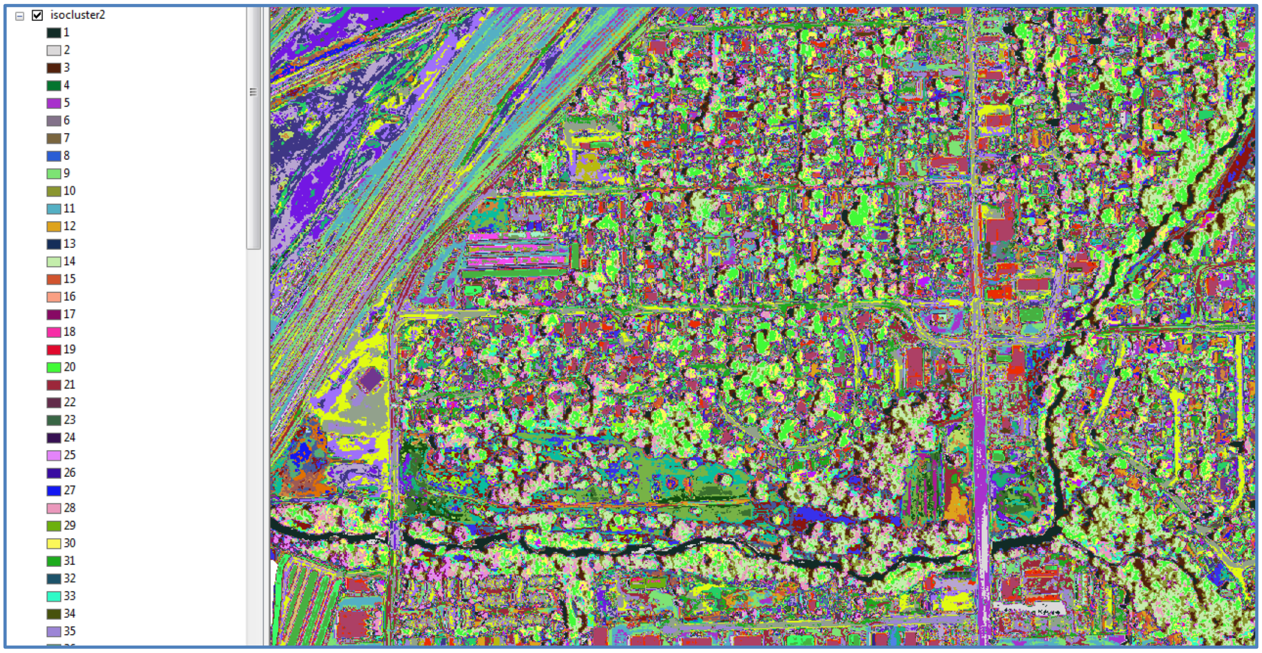


Figure 3: Unsupervised Classification with 100 classes showing more detailed features

Supervised Classification

Based on land cover types identified on the USGS' Land Cover Institute website (http://landcover.usgs.gov/classes.php), training samples were identified for the land cover types present in the project area. Approximately 25 training samples were identified per each land cover type in order to have an appropriate statistical sample for each land cover. The land cover types for the project area were identified to be as follows:

- Grassland
- Urban/Recreational Grasses
- Low Density Residential
- Commercial/Industrial/Transportation
- Trees
- Water

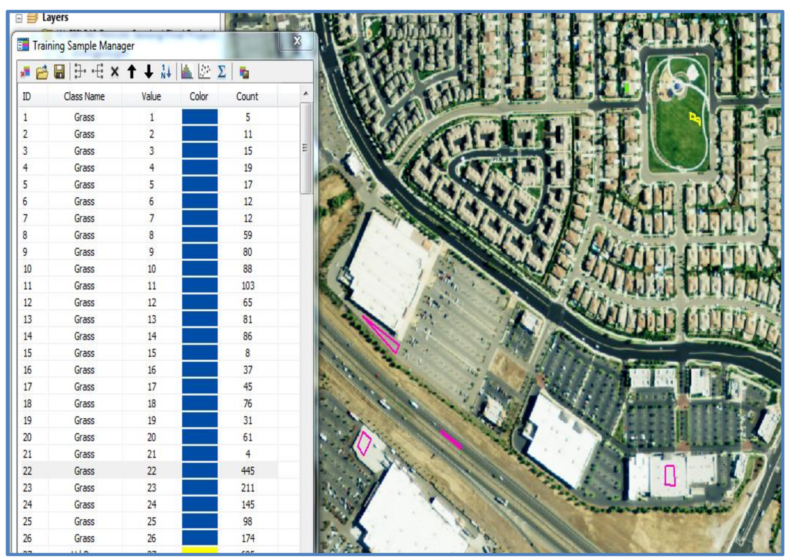
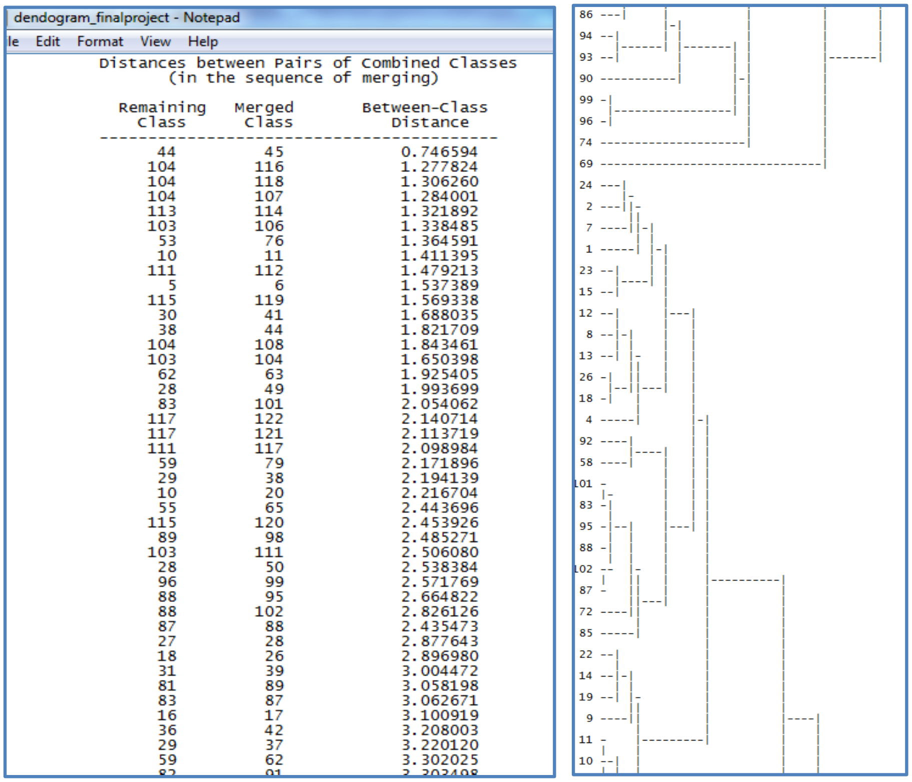


Figure 4: Training Sites for Supervised Classification

As training sites were collected, the statistics and histograms for each sample were reviewed to ensure that the samples were of good quality. When 122 training samples were collected the sites were exported as a signature file to be used in the Maximum Likelihood Classification process. Additionally a Dendogram was created in order to analyze the quality of the signature file.



dendogram_finalproject - Notepad

Distances between Pairs of Combined Classes
(in the sequence of merging)

Remaining Class	Merged Class	Between-Class Distance
44	45	0.746594
104	116	1.277824
104	118	1.306260
104	107	1.284001
113	114	1.321892
103	106	1.338485
53	76	1.364591
10	11	1.411395
111	112	1.479213
5	6	1.537389
115	119	1.569338
30	41	1.688035
38	44	1.821709
104	108	1.843461
103	104	1.650398
62	63	1.925405
28	49	1.993699
83	101	2.054062
117	122	2.140714
117	121	2.113719
111	117	2.098984
59	79	2.171896
29	38	2.194139
10	20	2.216704
55	65	2.443696
115	120	2.453926
89	98	2.485271
103	111	2.506080
28	50	2.538384
96	99	2.571769
88	95	2.664822
88	102	2.826126
87	88	2.435473
27	28	2.877643
18	26	2.896980
31	39	3.004472
81	89	3.058198
83	87	3.062671
16	17	3.100919
36	42	3.208003
29	37	3.220120
59	62	3.302025

Figure 5: Dendogram

The resulting output was provided by the Maximum Likelihood Classification;

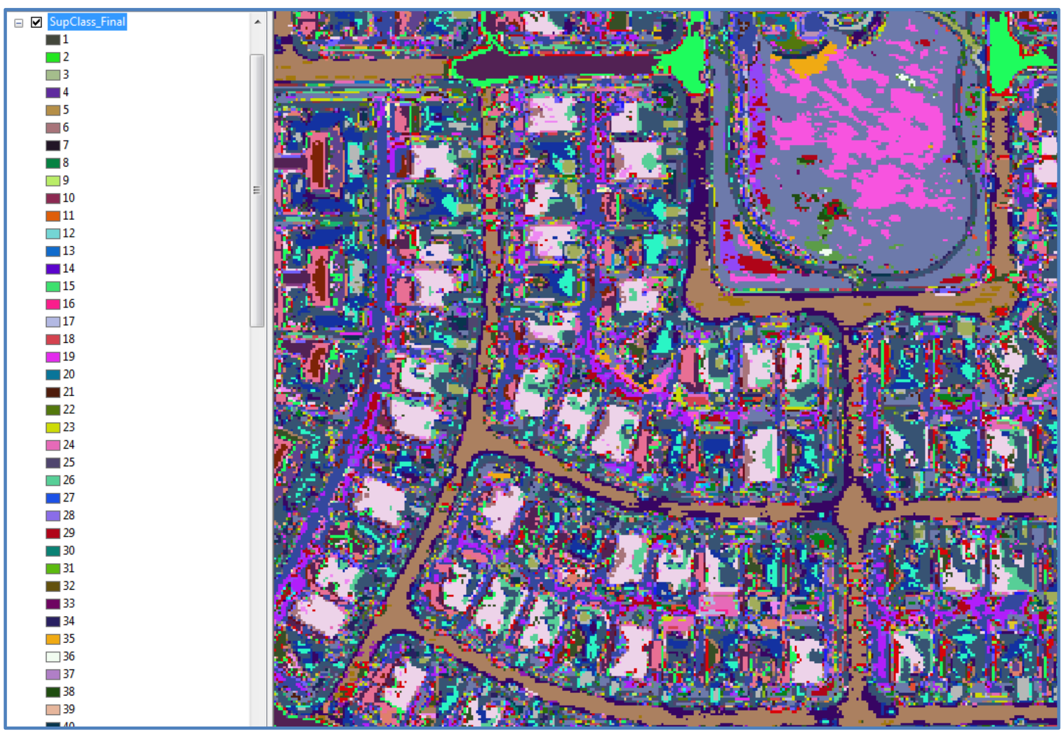


Figure 6: Maximum Likelihood Classification

Using the Reclassify tool in ArcMap, the 122 spectral signatures were combined and organized into information classes.

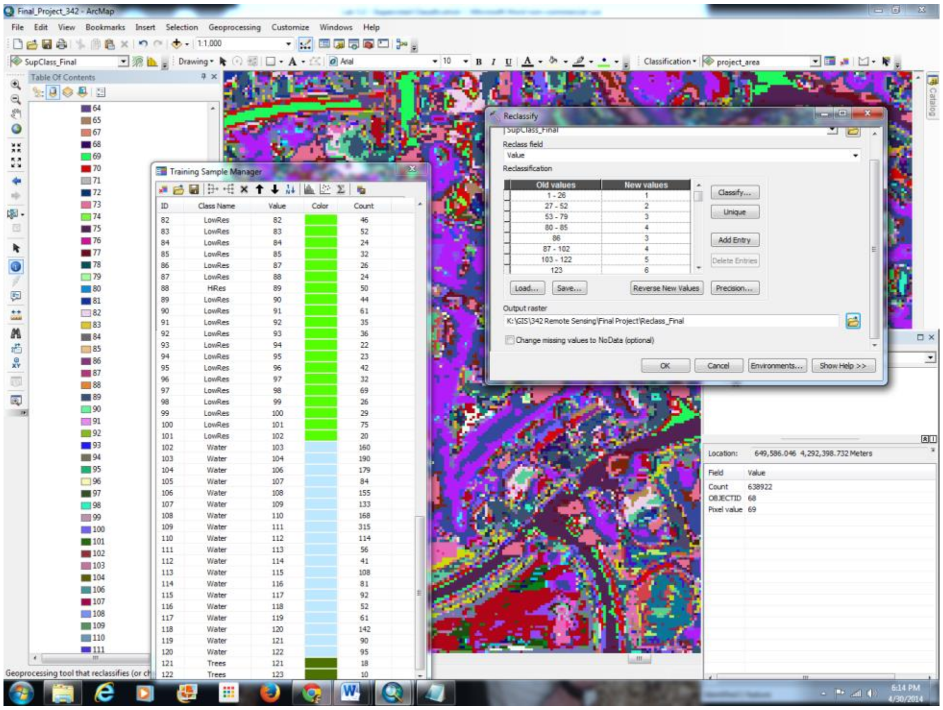


Figure 7: Reclassify Tool

The result of the reclassification then produced a final image that had been assigned colors to match the land cover type. (Figure 8)

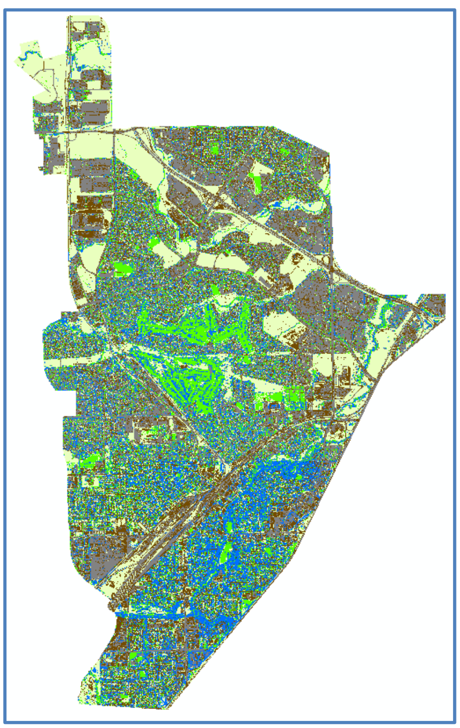


Figure 8: Final result of Supervised Classification

Difficulties

Difficulties were experienced in collecting training sites particularly in areas with a lot of trees due to shadows on the treetops. This made collecting a quality spectral signature representing trees a challenge. In the end, in spite of monitoring the statistics for each training site, the analysis completed by the Maximum Likelihood Classification returned a significant amount of “water” probably due to spectral signatures that too closely resembled another land cover type. The areas identified as “water” however, were also areas prone to more shadows due to the presence of mature trees. Potentially the analysis of this area would be more accurate in leaf-off imagery, due to the reduction of shadows in the image.
Additionally, difficulties were experienced in selecting the right imagery for this project. With the “2009 Combined NIR NAIP” the resolution of the imagery seemed to present some challenges for the software. A single area of the same

cover type would return as different spectral classes (e.g., the train yard, see Figure 3). However, attempting the unsupervised classification using LandSat imagery proved problematic as well.

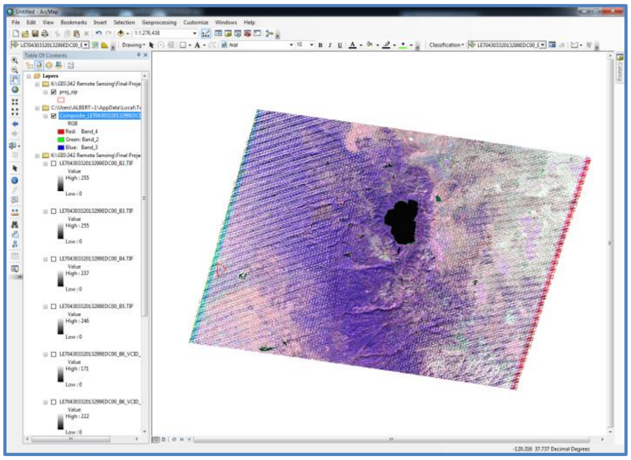


Figure 10: Landsat image used for abandonned unsupervised classification

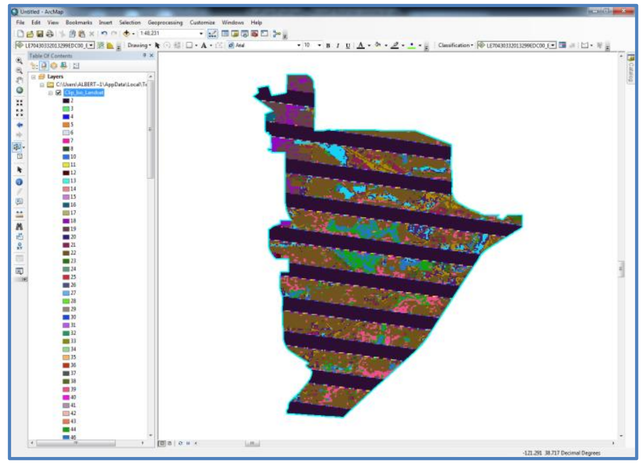


Figure 9: Iso Cluster Unsupervised Classification performed on Landsat imagery

Conclusions

The imagery being used is important. Aspects of when the imagery was captured, like cloud cover and time of year, are important as well. These can influence the kind of analysis the computer is able to do. The software being used to conduct an image classification of a suburban area can also influence the accuracy of the end result. Potentially something like eCognition, an object based classification software, would work better for an area like Roseville with the relatively dense features of low-density residential, mixed with natural features like creeks and recreational spaces. Particularly for an area with mature deciduous trees as one would find in residential areas, object based classification would be better suited and would overcome the issue of shade obscuring the shapes of trees from the remote sensor.

In the end, this classification would not be a reliable summary of pervious versus impervious surfaces due to the confusion between water and shadows. The shadows could be obscuring impermeable surfaces or they could be obscuring landscaping or another permeable surface. Again, the low result for deciduous trees is concerning because a simple review of the natural color image indicates a greater presence of this land cover type than just 5%.

Land Cover	Total	Percent of Total
Grassland	713524	22.30%
Recreational Grasses	412945	12.90%
Commercial/Transportation	800727	25.02%
Low Density Residential	570762	17.84%
Water	514536	16.08%
Trees	187615	5.86%

Table 1: Results from Supervised Classification

What did you learn from conducting this project?

The project was an excellent experience in humility. While I had been taught the process for many of the components of a full image classification, I discovered I had only been given a fraction of all there is to know when endeavoring to classify an image. It seems that to be a proficient at image classification one would to take a whole series of classes on each discreet step.

I learned that there are definite pros and cons to using one kind of software over another, depending on the area of interest for the project. While I did not work with eCognition at all for this project, from what I have seen of its capabilities, it would be far better suited to a project like a residential area. ArcMap seemed to struggle to differentiate surfaces like the dirt and sand of the rail yard from the roof tops of houses. Using an object based classification software, presumably an analyst could identify combinations of shape-specific attributes as well as reflectance levels of a roof material.

Though, I would never tout my results as very accurate, I am nonetheless proud of the resulting classification. Though I felt tremendously unsure of myself throughout the project, in the end, it isn’t completely inaccurate and I now have a better understanding of the process of image classification.