Lauren Slawsky

Land Cover and Land Use in the Charlotte University Area

For this study, a 2009 Quickbird image of the university area was used to create an image of single land use types through three separate supervised classifier processes. After collecting GPS locations of pure land use and land cover in the field, the original list was combined down to eight classes: Asphalt and Gravel, Concrete and Light Top Buildings, Bare Soil, Coniferous Trees, Deciduous Trees, Lawn Grass, Natural Grass, and Water Bodies. Regions of interest (ROIs) were then created using ENVI software to train the program for which pixels belonged in what class based on digital number (DN) values in all four bands of the image. To classify the entire the supervised classifiers of minimum distance, maximum likelihood, and parallelepiped were utilized. The first classifies based on an unknown pixel's distance from the mean DN value of the closest class. The second classifies an unknown pixel using its DN values and places it in the class it has the highest probability of belonging to. The third uses a box method, where the shape that classifies an unknown pixel is based on the highest and lowest DN values of the pixels in that class. Using these resources, the main goal was to compare these images of training regions of interest, to pixels of reference, and find which of the three supervised classifiers had the highest accuracy for correctly classifying entire university area image.

Project Goals

Goals of this project include: to successfully identify and accurately record GPS coordinates for regions of pure land cover and land use. To utilize the ENVI software with the 2009 Quickbird image to create good regions of interest of the desired land cover types for use in creating three images of supervised classifiers. And to use ENVI generated confusion error matrices to analyze the accuracy of our software training regions of interest compared to those of reference in order to determine which of the three classifiers works the best.

Image Data

An October 2009 Quickbird image of the University area including UNC Charlotte campus and its surrounding region was used. Quickbird is a four band multispectral sensor with a 16.6km swath width and a revisit time of three days. This sensor has three bands in the visible part of the spectrum from 450-520nm (blue), 520-600nm (green), and 630-690nm (red), as well as one band for near infrared from 760-900nm providing a spatial resolution of 2.4 meter pixels. Quickbird also has a panchromatic band from 450-900nm that produces images of 61cm pixels. Radiometric resolution of 11-bit data for this 2009 Quickbird image gives pixels possible digital number values from 0 to 2048.

Field Work

The project began by finding pure single land use pixels within the University area training site. These sites were used as training and reference regions in the ENVI software in order to create a simplified map of certain types of land cover over the entire original image. Because data was collected in the field and the land use was already known, th data would be organized using supervised classification, and analyzed with supervised classifiers.

Once area of strictly conifer trees, lawn grass, etc. was found, the GPS coordinates were recorded using a hand-held GPS. The goal was to find two of each land use type for more opportunities to create accurate and pure regions of interest in ENVI. The regions found in the field were entered into the ENVI software as GPS. This way, the exact center of that land use region could be found, making it easier to create more accurate regions of interest. Half of these ROIs were used to train ENVI to classify each pixel as a specific land cover type, the other half were used as reference to check how well the software did at categorizing the pixels.

Supervised Classifiers

In order to check ENVI's accuracy, three types of supervised classifiers were used to analyze the regions of interest. Working with the Minimum Distance classifier first - assuming it is the middle of the three classifiers – to check out how well our ROIs were training the software, it is Important to use pure pixels in the ROIs and check statistics of spectral range for each land use in each band to minimize overlap between materials, and thus reduce confusion within the software. After many hours of tweaking both training and reference ROIs, as well as separate standard deviations for each land use, a decent image was produced. The minimum distance classifier works by calculating the Euclidean Distance between an unknown pixel, and the mean digital number value for the class closest to that pixel. In both minimum distance, and maximum likelihood images, everything will be classified unless standard deviations are used to adjust the class thresholds. In this case, multiple standard deviation values were used for each classifier, so all images had some unclassified pixels.

In a maximum likelihood classifier, unknown pixels are assigned to the class they have the highest probability of belonging to. This is based on a normalized curve of the number of pixels in that class, to the digital number values of those pixels. Most groups found this to be their most accurate classifier. This was the second best to minimum distance, but only by a few tenths of a percent.

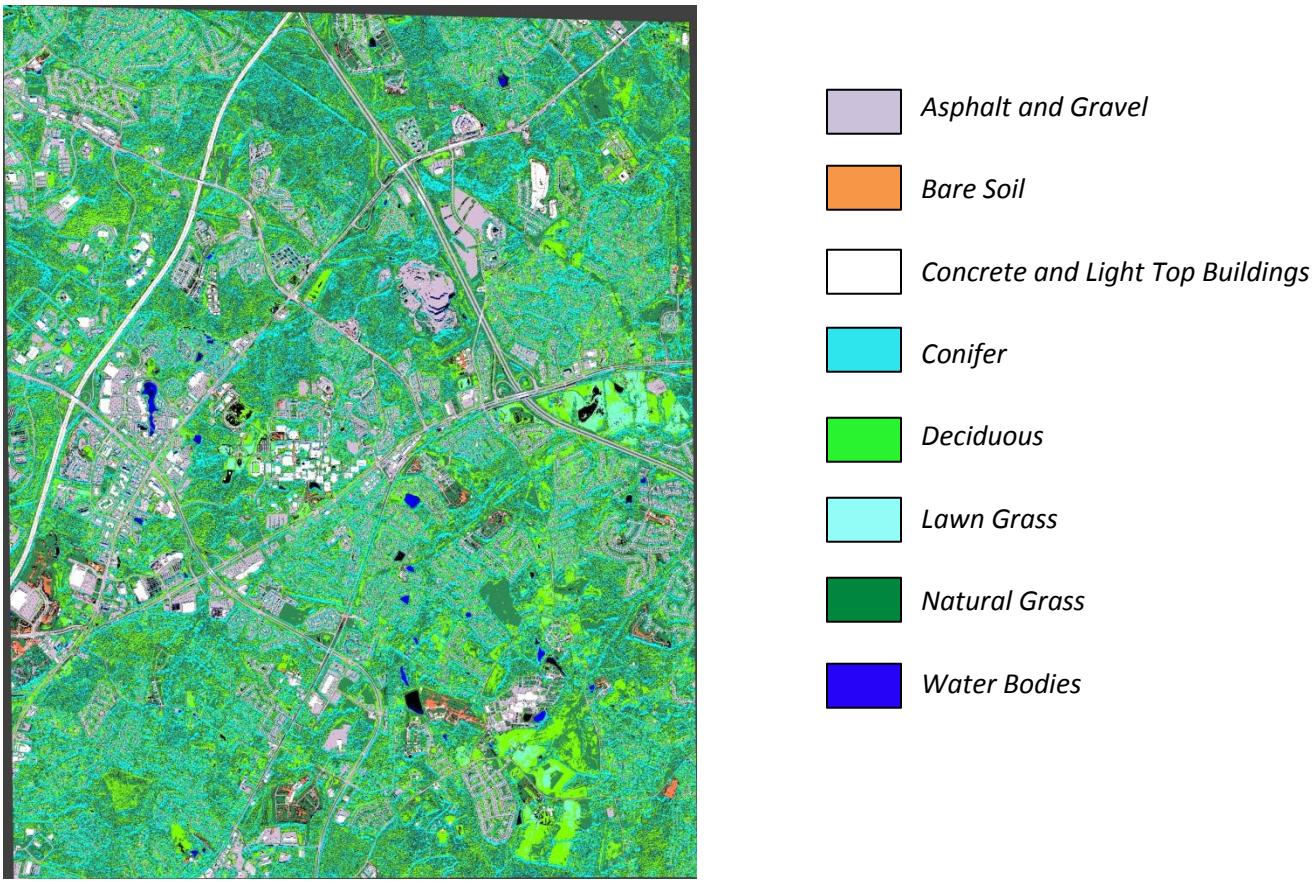
The third type of supervised classifier used was parallelepiped. This classifies an image by defining regions based on the highest and lowest digital number values for each class in each band. Adjusting standard deviation changes the decision boundaries of each class and makes the regions larger; including more pixels as part of that class. Depending on the spectral range of each class, not every unknown pixel is guaranteed to be classified in a parallelepiped classifier image. This was the worst of the three classifiers because there was so much overlap between certain land cover types, such as Deciduous Trees, Lawn Grass, and Natural Grass. With no way to tell which class an unknown pixel was most likely to belong to, this type of supervised classifier resulted in the most error overall.

Classification Error Evaluation Method

In order to evaluate the accuracy of the three supervised classifiers, the training ROIs were compared to reference ROIs using a Confusion Error Matrix. This calculates the overall accuracy of each classifier, for pixel and percentage, as well as the error within each individual class. Errors of Commission are described as a class containing a misclassified pixel that really belongs in a different class. This is calculated with User's Accuracy, which is the percentage from the number of correctly classified pixels divided by the total number of pixels in that class. Errors of Omission describe a pixel that should have been in a class, but was misclassified as something else. This is calculated with Producer's Accuracy, which takes the percentage of the number of correctly classified pixels divided by the total from ROIs made.

Results

Minimum Distance Classifier Overall Accuracy: 97.05%



Percent (%) Error Matrix – Minimum Distance

Class	Asphalt Reference	Bare Soil Reference	Concrete Reference	Conifer Reference	Deciduous Reference	Lawn Grass Reference	Natural Grass Reference	Water Body Reference	Total	User's %
Unclassified	0	3.61	0	0	0	0.78	0	0.93	0.77	
Asphalt/Gravel	**100**	0	0	0	0	0	0	0	27.95	100
Bare Soil	0	**93.57**	0	0	0	0	0	0	16.36	100
Concrete	0	0	**100**	0	0	0	0	0	15.45	100
Conifer	0	0	0	**100**	0	0	0	0	7.44	100
Deciduous	0	0	0	0	**88.78**	0	11.02	0	7.02	87
Lawn Grass	0	0	0	0	11.22	**99.22**	0	0	9.69	92.03
Natural Grass	0	2.81	0	0	0	0	**88.98**	0	7.87	93.75
Water Body	0	0	0	0	0	0	0	**99.07**	7.44	100
Total	100	100	100	100	100	100	100	100	100	
Producer's %	100	93.57	100	100	88.78	99.22	88.98	99.07		

With an overall accuracy of 97.05%, minimum distance was just barely the best supervised classifier. There were three classes without any errors of omission or commission meaning all the training and reference pixels were correctly classified. These classes included Asphalt and Gravel, Concrete and Light Top Buildings, and Conifer Trees. The class that had the most trouble correctly classifying both training and reference pixels was Deciduous Trees with the lowest user's accuracy of 87% and lowest

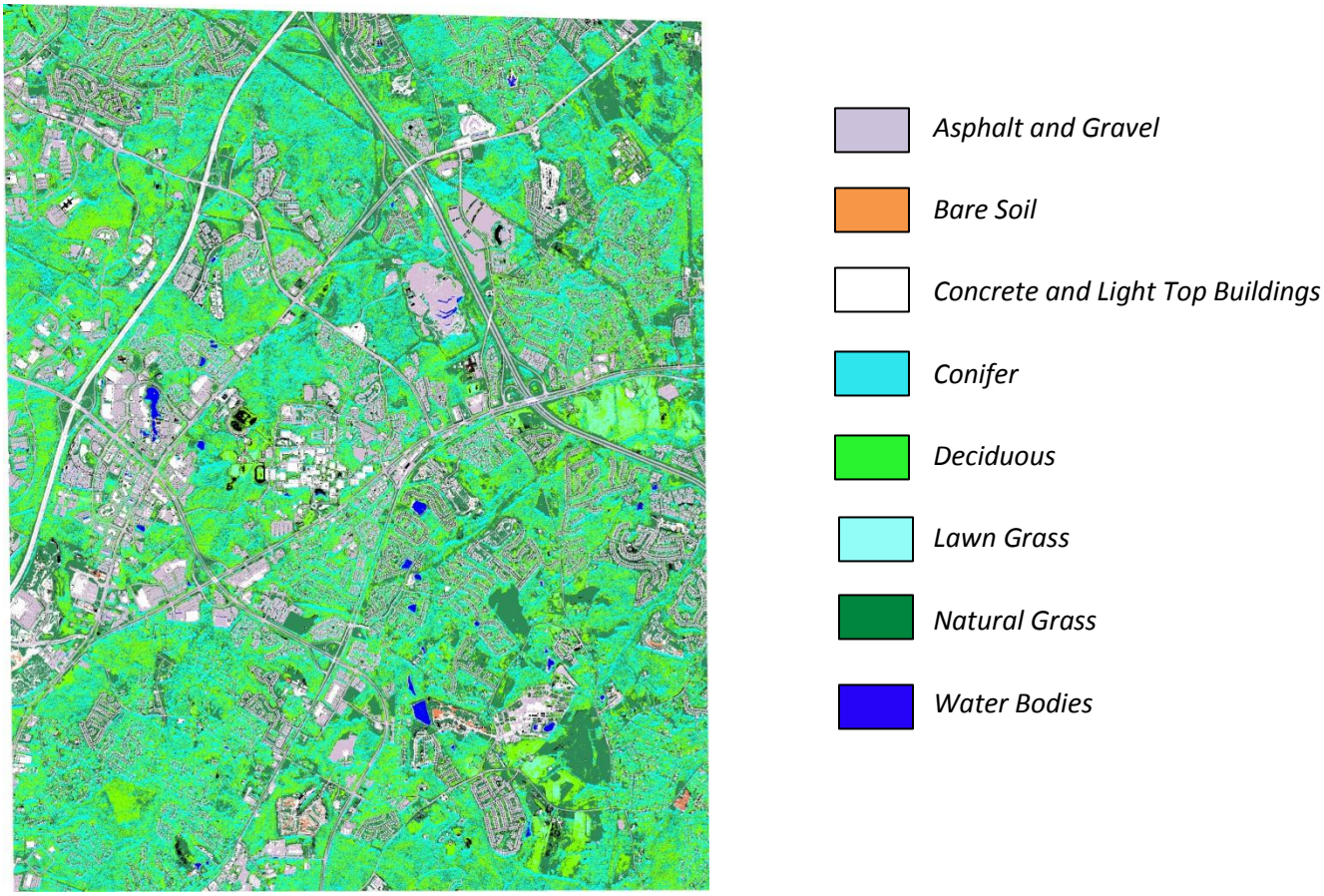
producer's accuracy of 88.78%. There were 11 pixels that should have been classified as Deciduous, but were misclassified as Lawn Grass, and 13 Natural Grass pixels were incorrectly included as Deciduous Trees.

Then, Bare Soil had the greatest number of unclassified pixels at nine, while seven others were incorrectly classified as Natural Grass, giving Bare Soil a producer's accuracy of 93.57%. Natural Grass had the second most errors of omission with a producer's accuracy of 88.98%. User's accuracy was also low at 93.75% because of the Bare Soil pixels that were included but shouldn't have been.

Lawn Grass also had a lower user's accuracy with 92.03% because of the 11 Deciduous Tree pixels incorrectly included in this class. But, with only one pixel misclassified as unclassified, Lawn Grass had a very high producer's accuracy of 99.22%.

Minimum Distance was the only classifier without any errors of commission for Asphalt and Gravel. This classifier was also the only one without any errors of commission, or omission for Concrete and Light Top Buildings.

Maximum Likelihood Classifier Overall Accuracy: 96.91%



Percent (%) Error Matrix – Maximum Likelihood

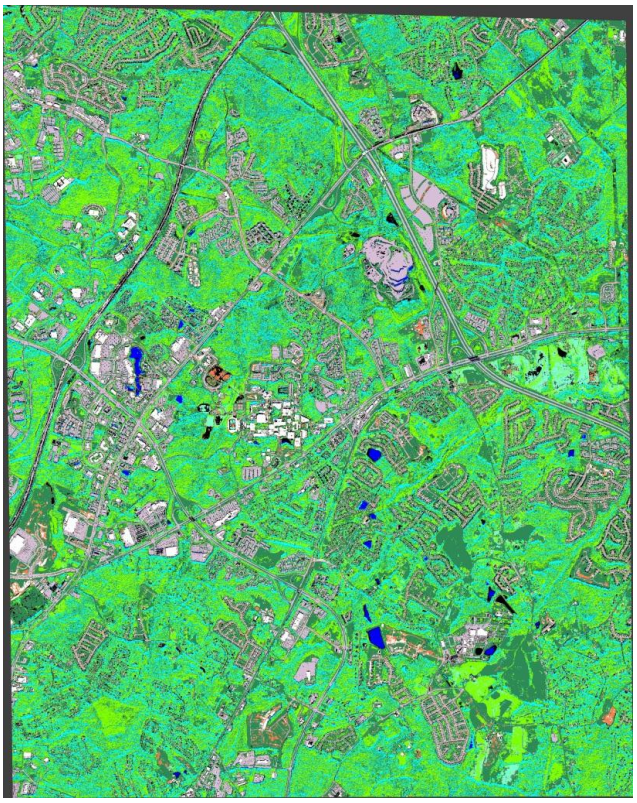
Class	Asphalt Reference	Bare Soil Reference	Concrete Reference	Conifer Reference	Deciduous Reference	Lawn Grass Reference	Natural Grass Reference	Water Body Reference	Total	User's %
Unclassified	0	6.02	0.91	0	0	0	0	0	1.19	
Asphalt/Gravel	**99.5**	0	0	0	0	0	0	0	27.81	100
Bare Soil	0	**88.76**	0	0	0	0	0	0	15.52	100
Concrete	0.05	5.22	**99.09**	0	0	0	0	0	16.36	93.56
Conifer	0	0	0	**100**	0	0	0	0	7.44	100
Deciduous	0	0	0	0	**100**	8.59	0.86	0	7.72	89.09
Lawn Grass	0	0	0	0	0	**91.41**	0	0	8.22	100
Natural Grass	0	0	0	0	0	0	**99.15**	0	8.22	100
Water Body	0	0	0	0	0	0	0	**100**	7.52	100
Total	100	100	100	100	100	100	100	100	100	
Producer's %	99.5	88.76	99.09	100	100	91.41	99.15	100		

Maximum Likelihood was the second best classifier with an overall accuracy of 96.91%. This time there were two classes without any errors of omission or commission and they were Water Body and then Conifer Trees again. Deciduous Trees had a 100% producer's accuracy, meaning all reference pixels were correctly classified. However, 11 Lawn Grass pixels and one Natural Grass pixels were included as Deciduous, giving that class a user's accuracy of 89.09%; the lowest for the maximum likelihood classifier.

For this classifier, instead of Deciduous Tree pixels being classified as Lawn Grass, 11 Lawn Grass pixels were misclassified as Deciduous. This resulted in a producer's accuracy for Lawn Grass of 91.41%. There was also one Natural Grass pixel that was misclassified as Deciduous, giving Natural Grass one of the better producer's accuracy of 99.15%.

Bare Soil had 100% user's accuracy but, with 28 pixels that were misclassified Bare Soil had the lowest producer's accuracy with 88.76%. Of these pixels, 15 were unclassified, while 13 were misclassified as Concrete. Because of this, Concrete and Light Top Buildings had the second lowest user's accuracy of 93.56% as two Asphalt and Gravel pixels were also incorrectly included in this class.

Parallelepiped Classifier Overall Accuracy: 88.06%



Asphalt and Gravel

Bare Soil

Concrete and Light Top Buildings

Conifer

Deciduous

Lawn Grass

Natural Grass

Water Bodies

Percent (%) Error Matrix – Parallelepiped

Class	Asphalt Reference	Bare Soil Reference	Concrete Reference	Conifer Reference	Deciduous Reference	Lawn Grass Reference	Natural Grass Reference	Water Body Reference	Total	User's %
Unclassified	0	0	25	0	0	0	0	0.93	3.93	
Asphalt/Gravel	**99.75**	0	0	0	0	0	0	0	27.88	100
Bare Soil	0	**73.9**	0	0	0	0	0	0	12.92	100
Concrete	0	22.09	**75**	0	0	0	0	0	15.45	75
Conifer	0	0	0	**99.06**	0	0	0	17.76	8.71	84.68
Deciduous	0	0	0	0	**100**	11.72	11.02	0	8.85	77.78
Lawn Grass	0	0	0	0	0	**88.28**	0	0	7.94	100
Natural Grass	0.25	4.02	0	0.94	0	0	**88.98**	0	8.22	89.74
Water Body	0	0	0	0	0	0	0	**81.31**	6.11	100
Total	100	100	100	100	100	100	100	100	100	
Producer's %	99.75	73.9	75	99.06	100	88.28	88.98	81.31		

Parallelepiped was the classifier found to have the most overall error, with a total accuracy of only 88.06%. While the classification results showed no errors of commission for Bare Soil, it was the class with the lowest producer's accuracy with 73.9%. Fifty five Bare Soil pixels were incorrectly classified as Concrete and another ten were misclassified as Natural Grass. This led the Concrete and Light Top Building class to have the lowest user's accuracy with only 75%. Concrete had second lowest producer's accuracy of 75% because 55 of its pixels went unclassified. After Deciduous Trees, Natural Grass also had one of the lower user's accuracy of 89.74% due to errors of commission from the ten Bare Soil pixels, as well as one Asphalt pixel. The Asphalt and Gravel class had no errors of commission however, that one Natural Grass pixel included in this class lowered the producer's accuracy to 99.75%.

Like Maximum Likelihood, the Deciduous Tree class had a large error of commission at only 77.78% because 15 Lawn Grass pixels were incorrectly included, as well as 13 Natural Grass pixels. But no Deciduous Tree pixels were misclassified so the producer's accuracy for this class was 100%.

Unlike Maximum Likelihood, two of the classes with a lot of error were Conifer Trees and Water. There were 19 water pixels misclassified as Conifer, giving Conifer a low user's accuracy of 84.68%. However, no pixels were misclassified as water, so the Water Body class had 100% user's accuracy. The classes switch when it comes to errors of omission. Only one Conifer pixel was incorrectly classified as Natural Grass, so Conifer Trees had the third highest producer's accuracy with 99.06%. Because those 19 Water pixels were incorrectly classified as Conifer, as well as one pixel going unclassified, the Water class had a producer's accuracy of only 81.31%.

Percent of Reference Pixels Correctly Classified (%)

	Minimum Distance	*Maximum Likelihood*	*Parallelepiped*
Asphalt and Gravel	100	99.5	99.75
Bare Soil	93.57	88.76	73.9
Concrete/Light Top Buildings	100	99.09	75
Conifer Trees	100	100	99.06
Deciduous Trees	88.78	100	100
Lawn Grass	99.22	91.41	88.28
Natural Grass	88.98	99.15	88.28
Water Body	99.07	100	81.31

Best Classifier: MINIMUM DISTANCE

	Land Cover Area(km^2)	*Land Coverage Percent (%)*
Unclassified	8.79	13.631
Asphalt and Gravel	7.845	12.165
Bare Soil	0.412	0.639
Concrete and Light Top Buildings	2.858	4.432
Conifer Trees	13.36	20.713
Deciduous Trees	11.3	17.519
Lawn Grass	1.46	2.259
Natural Grass	18.28	28.341
Water Body	0.194	0.3

Discussion

Sources of error within the confusion error matrices are most readily identified through the spectral range of each training class in each band and how it overlaps in one or more bands with another training class. Much of the error was between Concrete and Bare Soil, as well as between Deciduous Trees, Lawn Grass, and Natural Grass. Some overlap also occurred over Water and Conifer Trees in the Parallelepiped classifier.

When the standard deviation of each class is adjusted for each separate supervised classifier image, there is more of a chance classes will overlap, even if the ROIs are such that little if any overlap occurs. It was hard to avoid overlap in such similar classes like vegetation. There were also very bright Bare Soil pixels that couldn't avoid being classified as Concrete. Also, darker pixels in the Conifer and Natural Grass classes were occasionally misclassified as Water due to its low spectral reflectance. This also led to error in the Water class when it would be unclassified sometimes, as well as dark shadows classified as Water when looking at the image itself.

Spectral Range of the 8 training classes in all four bands

Band 1 – Visible Blue.
Band 2 – Visible Green.
Band 3 – Visible Red.
Band 4 – Near IR

Deciduous

Basic Stats	Min	Max	Mean	Stdev
Band 1	132	155	142.546296	5.746001
Band 2	161	214	187.648148	13.387339
Band 3	77	109	91.703704	9.236518
Band 4	435	609	521.685185	44.649130

Lawn Grass

Basic Stats	Min	Max	Mean	Stdev
Band 1	150	169	158.609375	3.332575
Band 2	211	257	223.500000	7.522146
Band 3	98	131	110.585938	4.357142
Band 4	615	700	657.671875	21.240323

Natural Grass

Basic Stats	Min	Max	Mean	Stdev
Band 1	138	183	159.574627	11.662147
Band 2	183	264	223.820896	21.444052
Band 3	94	179	135.589552	24.473159
Band 4	330	455	386.455224	32.612470

Conifer

Basic Stats	Min	Max	Mean	Stdev
Band 1	120	134	128.067308	3.718922
Band 2	128	158	139.173077	6.425741
Band 3	50	77	58.913462	4.979786
Band 4	113	279	170.461538	32.013323

Asphalt and Gravel

Basic Stats	Min	Max	Mean	Stdev
Band 1	173	270	223.989183	17.451494
Band 2	208	401	308.385817	31.572299
Band 3	114	262	191.450721	25.087908
Band 4	105	338	189.085337	38.290971

Water

Basic Stats	Min	Max	Mean	Stdev
Band 1	132	148	140.739583	4.620879
Band 2	144	191	174.437500	14.281245
Band 3	57	93	79.229167	10.141607
Band 4	37	56	48.760417	3.695644

Concrete and Light Top Buildings

Basic Stats	Min	Max	Mean	Stdev
Band 1	294	1067	682.300866	303.415021
Band 2	497	1663	1077.735931	446.094457
Band 3	365	1295	797.679654	309.901708
Band 4	372	1253	787.132035	268.547475

Bare Soil

Basic Stats	Min	Max	Mean	Stdev
Band 1	173	205	188.600000	6.092256
Band 2	294	353	322.570000	14.138052
Band 3	296	382	335.300000	22.114367
Band 4	359	473	401.390000	19.829370

When directly considering the spectral range within each band of the Deciduous Tree class, five pixels of overlap are seen in Band 1 between this class, and Lawn Grass, as well as 17 pixels of overlap between Natural Grass and Deciduous. Band 2 shows less of an overlap with Lawn Grass, but the overlap with Natural Grass is 31 pixels. This is important because Band 2, the Visible green band, shows a distinct increase in spectral reflectance when it comes to vegetation. However, the most important band to look at is Band 4. This is because vegetation is most reflective in Near Infrared. A 4-3-2 image of the 2009 Quickbird was generated to create the vegetation ROIs, and statistics were constantly checked for overlap, especially between these three classes. In the end, six pixels in the upper DN values of Band 4 for Deciduous overlapped with the lower end of DN values for Lawn Grass. Deciduous pixels were overall darker than Lawn Grass, but still caused noticeable error in all three supervised classifiers. There was also a 20 pixel overlap from the low end of the spectral range for Deciduous in Band 4 with the upper end of the range for Natural Grass, causing error. These errors were magnified when standard deviations were increased, especially seen in the Minimum Distance image and error matrix. The Deciduous class was assigned a standard deviation of 7.3, while Lawn grass was given 6.9. The original overlap between

these classes was not that significant, but altered standard deviation values caused Deciduous to have the greatest errors of both omission and commission for this classifier.

The darker pixels for Conifer Trees and Water also led to confusion. Overlap was most significant in Band 3, the Visible blue band, with 20 pixels of overlap right in the middle of the spectral range of each class. This was beneficial for the Maximum Likelihood classifier, because both classes had the best overall accuracy as pixels were correctly placed in the class they had the highest probability of belonging to. However, with Parallelepiped, because the boxes completely overlapped, the software created a lot of error for these two classes because it couldn't decide in which class to put unknown pixels.

On the other end of the spectral range, the brighter pixels of Bare Soil were being confused with those of Concrete and Light Top Buildings. Bands 3 and 4 for Concrete had a very wide spectral range of around 900 for those ROIs, while the range for Bare Soil was much smaller, generally around 110 pixels total. This caused many Bare Soil pixels to be classified as Concrete because the likelihood of belonging to that class was greater, with only 17 pixels of overlap in Band 3, and 13 of overlap in Band 4.

Another source of error related to spectral variability is found in the single Water class. Because this class was not separated into two classes - one of dark, more blue water, and one of lighter, green water - only the dark pixels were chosen as classifiers. This kept a tight spectral range, but also resulted in a lot of the lighter water being completely unclassified. This error was vaguely noted in the Parallelepiped and Minimum Distance error matrices, but when directly viewing the image, whole plots were unclassified, when they were actually bodies of lighter water.

Conclusion

The classifier with the best calculated results was the Minimum Distance classifier with an overall accuracy of 97.05%. Unclassified pixels covered 13.631% of the image, an area amounting to 8.79km^2. According to data, Natural grass covered the most land area at 28.341% of the region. Deciduous Trees and Conifer Trees covered roughly the same amount; 11.3km^2 and 13.36km^2 respectively. Water and Bare Soil didn't even cover 1% of the image combined, while man made materials such as Asphalt covered 7.845km^2, and Concrete, which covered 2.858km^2.

I think this was the best classifier because the greatest amount of time was spent working with it, perfecting the ROIs, and adjusting standard deviations, resulting in the best overall accuracy.

Best Classifier: MINIMUM DISTANCE

	Land Cover Area(km^2)	*Land Coverage Percent (%)*
Unclassified	8.79	13.631
Asphalt and Gravel	7.845	12.165
Bare Soil	0.412	0.639
Concrete and Light Top Buildings	2.858	4.432
Conifer Trees	13.36	20.713
Deciduous Trees	11.3	17.519
Lawn Grass	1.46	2.259
Natural Grass	18.28	28.341
Water Body	0.194	0.3