Greg Grube

Remote Sensing Lab 7 and 8

Introduction:

Horicon Marsh National Wildlife Preserve in east central Wisconsin and the areas surrounding it show a wide amount of diversity in land cover types. In this lab, my goal is to perform a classification of the Horicon Marsh and the surrounding area's land cover in order to determine which areas are home to healthy vegetation and which are either built up areas or bare ground. To complete my goal, I am implementing a supervised classification scheme, where the image analyst determines pixel categorization by choosing examples training sites that correspond to each class. Once these training samples are collected, I will be using two supervised classification methods: the parallelepiped classifier, and the maximum likelihood classifier. After performing these algorithms, I will analyze the map produced and the computed statistics to determine which classification is most effective. Finally, I will perform an accuracy assessment of the classifications using random independent test sites in order to determine sources of error and correct them to produce a higher quality map. This will ultimately allow me to make good decisions when discerning land cover classes from each other.

Methods:

In attempting to complete the supervised classification process, the first step was to use training sites to define the regions of interests, in this case, I attempted to find representative samples of five land type classes: water, wetland, forest, agricultural/grassland and urban/bare fields. My main strategy while selecting training sites was to use a divide and conquer method, where I used the ROI tool to label pixels in my image and describe them as one of the five categories. So, if there was a large area of forest that I could spot using a combination of the six band layers or Google Earth, I would label it as forest, and then create a new ROI and continue the process with other forest areas. After labeling a few areas in all 5 category areas, I merged the individual classes into one land type grouping so that I had five final ROI classes that I could work with. While trying to pick out representative samples of each land type, I figured out that the best strategy was to select areas that were distinctive and then begin to move toward the transition areas between land types. By doing this, I could have core examples of water, wetlands, etc. and then move toward less clear areas to refine my results as needed. I also found that it worked better to select smaller areas of pixels I was confident with so I wouldn't have to go back and change them later if I had problems with the algorithms. As mentioned, the next step that I completed was to run the parallelepiped and maximum likelihood classifiers to apply my selected training sites to the entirety of the image.

The parallelepiped classifier was the first that I ran. This method attempts to find the high and low limits of each class and if a pixel's DN value fits within that range, it is assigned to that class. The advantages of using this algorithm is that it is computationally efficient and easy to create, however, it also can create overlap in classes, covariance in two band areas and unclassified pixels in areas outside the parallelepiped boundaries. These disadvantages can be problematic, but I could assign the unclassified or misclassified areas to the classes that I wanted or set the max standard deviation from the mean to higher or lower values to account for errors. In the case of the maximum likelihood classifier, there are no decision boundaries as in parallelepiped; instead, decisions are based on probability and how the spectral bands co-vary with one another. A variance/co-variance matrix helps to evaluate the relationship of the classes, assuming normal distribution. After evaluating variance and covariance, unknown pixels are assigned to groups on probability of them belonging to one class versus another. The advantages to this algorithm are that it is the most accurate if samples have normal distribution and that it takes most variables into account. Disadvantages are that it is non-spatial, computationally intense, and tends to over classify places with high variance/covariance matrix values where one class shows up everywhere. This last issue is the most problematic for a user, as it can make the classification scheme too simple. They can be addressed by going back and creating training sites in the areas that are causing problems, or adjusting the probability threshold number to be more or less selective as needed.

The final step in my methodology was to complete an accuracy assessment of the results of my classification. The accuracy assessment is performed for simple curiosity on how well our classification was done, and to locate errors in our work. It is important because it allows us to fix errors, increase map quality, provide information to the map users and offer an objective means of comparison between observation and reality. During the accuracy assessment process, the classification scheme is compared against randomly selected test points, an important step as we want to eliminate biasing our results. These randomly selected points were placed over the original, non-classified image and then the random points were classified into each of the five categories. Once this process was completed, I compiled an error matrix that showed the results. This error matrix shows your overall accuracy, which is the total amount of points predicted correctly, as well as the Kappa Statistic, which is a number from 0 to 1 showing how much of your accuracy is due to knowledge and how much is due to chance. The error matrix also shows user's and producer's accuracy numbers. The user's accuracy measures how well you avoided errors of commission, which are the amount of pixels committed to an incorrect class. The producer's accuracy measures how well you avoided errors of omission, which is when a pixel is omitted from a class that it belongs in. The errors of commission and user's accuracy amounts can be added together to obtain the total pixels in a class and the same can be done for errors of omission and producer's accuracy amounts.

Results:

The first training sites that I created were helpful in allowing me to use the classifier algorithms, but ultimately I needed to return and refine the sites many times until I got a result I was satisfied with. Looking at the final statistics, each class is broken down into the six spectral bands with a minimum and maximum value, a mean value and a standard deviation. While trying to decide on the training sites I most often used Bands 4 and 5 (nir and swir), so I also found these to be the most useful in describing the results. For the forest class, my results were very distinct for the most part. The forests were usually found in bunches on the map and their DN values were in a relatively small range (83 -120 for Band 4) and had a smaller standard deviation than the other classes. Water was also a relatively distinctive class, with low DN values for band 4 (Mean of 27.35) and was easy to separate from all classes except for some wetland areas. Wetlands had some pixel mixing with water, but overall its spectral signature was distinctive, fitting in the area between water and forest (Mean of 82). The two classes that dominated the map, agriculture/grassland and urban/bare ground, were also not surprisingly the two that were confused the most and had the most overlap. Agriculture had a mean value of 130 in band 4, which makes sense given its high vegetation amounts. It often overlapped with urban and bare ground areas however because their DN value ranges were both wide and the standard deviation amounts were larger. That the urban and bare ground values had a wide variation is not surprising given the range of features from buildings to bare agricultural fields and the amount of pixel mixing that occurred in those areas. The DN value of the mean in band 4 (94) expresses the variation well, and also shows this class caused me the most problems, which I solved by choosing very characteristic example training sites that limited mixing as much as possible.

After looking at the statistics and completing the classification process using each classifier, I was left with the decision of evaluating which algorithm was most effective in displaying my classes. The first few times I ran both of the classifiers, there was a lot of confusion between the classes, with the parallelepiped having a lot of unclassified pixels and the maximum likelihood classifier tending to over classify areas. To help and clarify my classification problems, I began to go back and find the areas where the trouble was occurring and try to classify those areas very carefully by zooming in to the pixel level and defining the transition areas myself wherever possible to separate the two classes. In the case of parallelepiped, I increased the maximum likelihood classifier slightly to eliminate the unclassified areas, but not too much as to oversimplify the map. For maximum likelihood, I also experimented with the probability threshold. In both cases, I accounted for the variations that occurred by frequently referencing Google Earth to make sure that the pixels were accurately describing an area. In the end, I decided that my best choice for a map was the maximum likelihood classification because it took into account more variables and delineated areas more effectively and accurately than the parallelepiped algorithm, which even with refining still left many areas unclassified. Once I chose this map, I also completed the majority/minority analysis process to get rid of many of the stray pixels and make the image smoother.

After creating the map, I completed my accuracy assessment for both the original training sites and for the randomly generated test sites. Overall, the accuracy for the original training sites was very high at 98.6%. This telling you the accuracy of my ROI values, that is, of the ROI sites that were labeled on the image, how many were correctly identified. For example, 98.8% of the water pixels were correctly identified as water, and 98.41% of pixels labeled as water actually were. The user's and producer's accuracies in this assessment were all very high, with none of the areas below 97%. For the random sample of my maximum likelihood classifier, the results were not as good. The overall accuracy of this accuracy assessment was only 68.3%. As I only tested 60 pixels for the random samples, my results are quite varied. The producer's accuracies and user's accuracies as a whole show that there were problems in both categories, and the amount of pixels misclassified as agriculture/grassland shows there were major issues with that category being mistaken for other land types. The classes that I did the best in were those where I had more sample points, such as urban/bare ground where I got a 63.16% producer's accuracy, but 100% for a user's accuracy. The classes I didn't do well in were those that had a small amount of sample pixels, such as water, where I predicted 2 pixels of water when there were no water pixels actually sampled or in the case of wetlands, where my user's accuracy was only 40% (2/5 Pixels) and my producer's accuracy was even worse at 28.57% (2/7 Pixels). The implications of these numbers will be talked about in more detail in the discussion.

ROI: Agriculture/Grassland [Chartreuse] 168 points

Basic Stats	Min	Max	Mean	Stdev
Band 1	57	71	62.70833	3.221102
Band 2	24	35	27.69643	2.609882
Band 3	18	34	22.56548	3.791955
Band 4	81	165	130.8155	15.53965
Band 5	57	113	84.13095	12.15053
Band 6	16	46	28.29167	5.489625

ROI: Forests [Green3] 118 points

Basic Stats	Min	Max	Mean	Stdev
Band 1	54	61	57.27966	1.146422
Band 2	22	25	23.18644	0.569137
Band 3	17	20	18.01695	0.739404
Band 4	83	120	107.7627	7.800589
Band 5	57	81	65.12712	4.259823
Band 6	17	25	19.36441	1.412189

ROI: Urban/Bare Ground [Red1] 198 points

Basic Stats	Min	Max	Mean	Stdev
Band 1	86	255	129.6515	41.20932
Band 2	43	190	70.76768	26.27146
Band 3	52	239	88.96465	32.79708

Band 4	63	215	93.69697	26.19197
Band 5	84	245	141.7222	27.25884
Band 6	51	155	80.07576	15.28575

ROI: Water [Blue] 625 points

Basic Stats	Min	Max	Mean	Stdev
Band 1	52	78	60.5072	5.52068
Band 2	21	37	26.5632	4.251273
Band 3	16	40	22.08	5.656004
Band 4	18	58	27.3536	8.704999
Band 5	3	54	14.064	15.61637
Band 6	2	25	7.3184	6.590304

[Sienna]

Basic Stats	Min	Max	Mean	Stdev
Band 1	55	78	60.18349	2.910317
Band 2	22	35	25.43486	1.914504
Band 3	18	37	24.00551	2.94027
Band 4	46	109	82.05872	11.78204
Band 5	16	90	60.63853	10.99196
Band 6	9	37	23.07706	4.986888

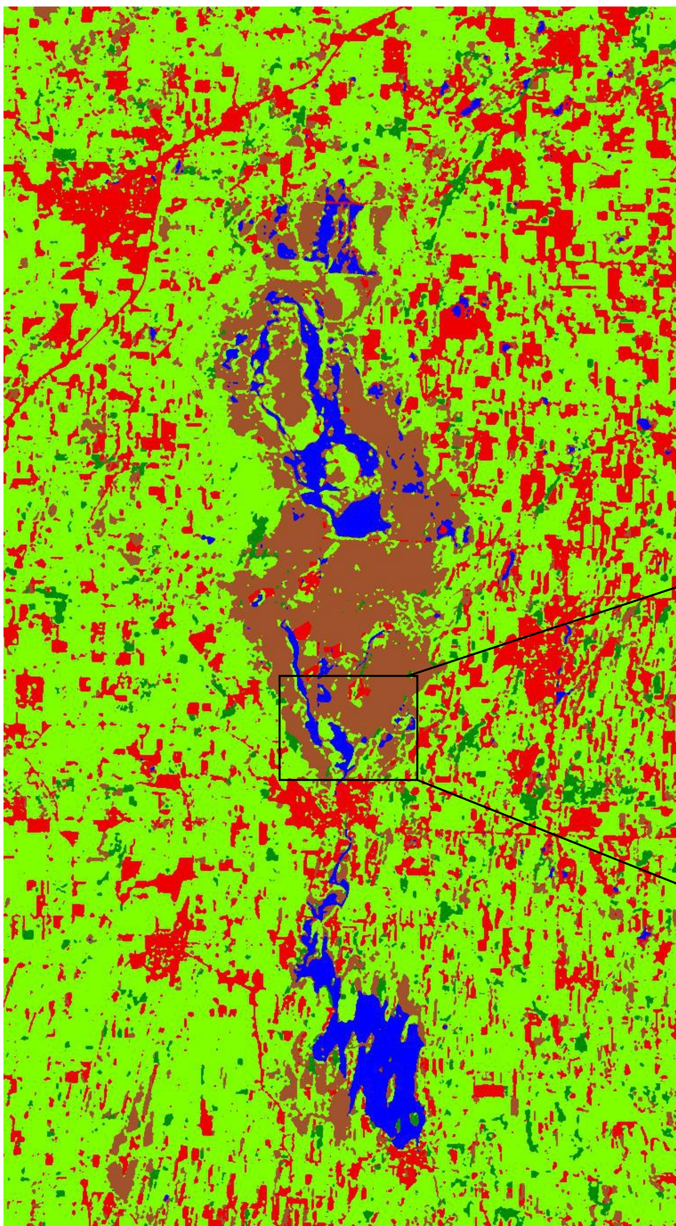

Figure 1: My final map of the Horicon Marsh area produced using the maximum likelihood supervised classification procedure. In the map, water is shown as blue, wetland as brown, forests in dark green, agriculture/grassland as light green, and urban/bare ground as red.

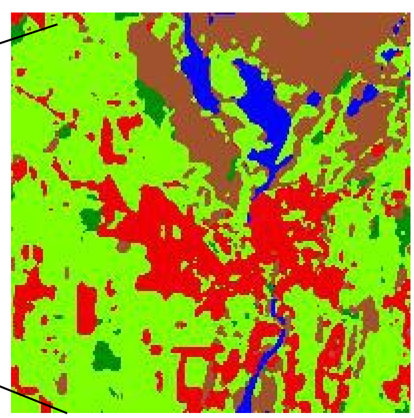

Figure 2: A zoom of the map showing detail of the area near the city of Horicon.

Confusion Matrix: Cross Validation of Original Training Sites

Overall Accuracy	= (1631/1654)	0.986094
Kappa Coefficient	0.9807	

	Ground Truth	(Pixels)			
Class	Forests	Water	Agriculture	Urban	Wetland
Unclassified	0	0	0	0	0
Forests [Green]	115	0	0	0	0
Water [Blue]	0	618	0	0	10
Agriculture [Chartreuse]	3	0	167	0	0
Urban [Red1]	0	0	0	198	2
Wetland [Sienna]	0	7	1	0	533
Total	118	625	168	198	545

	Ground Truth	(Pixels)
Class	Total	
Unclassified	0	
Forests [Green]	115	
Water [Blue]	628	
Agriculture [Chartreuse]	170	
Urban [Red1]	200	
Wetland [Sienna]	541	
Total	1654	

	Ground Truth	(Percent)			
Class	Forests	Water	Agriculture	Urban	Wetland
Unclassified	0	0	0	0	0
Forests [Green]	97.46	0	0	0	0
Water [Blue]	0	98.88	0	0	1.83
Agriculture [Chartreuse]	2.54	0	99.4	0	0
Urban [Red1]	0	0	0	100	0.37
Wetland [Sienna]	0	1.12	0.6	0	97.8
Total	100	100	100	100	100

	Ground Truth	(Percent)
Class	Total	
Unclassified	0	
Forests [Gree	6.95	
Water [Blue]	37.97	
Agriculture [Chartreuse]	10.28	
Urban [Red1]	12.09	
Wetland [Sienna]	32.71	
Total	100	

Class	Commission	Omission	Commission	Omission
	(Percent)	(Percent)	(Pixels)	(Pixels)
Forests [Green]	0	2.54	0/115	3/118
Water [Blue]	1.59	1.12	10/628	7/625
Agriculture [Chartreuse]	1.76	0.6	3/170	1/168
Urban [Red1]	1	0	2/200	0/198
Wetland [Sienna]	1.48	2.2	8/541	12/545

Class	Prod. Acc.	User Acc.	Prod. Acc	User Acc.
	(Percent)	(Percent)	(Pixels)	(Pixels)
Forests [Green]	97.46	100	115/118	115/115
Water [Blue]	98.88	98.41	618/625	618/628

Agriculture [Chartreuse]	99.4	98.24	167/168	167/170
Urban [Red1]	100	99	198/198	198/200
Wetland [Sienna]	97.8	98.52	533/545	533/541

Confusion Matrix: Random Sample for my final max likelihood

Overall Accuracy	= (41/60)	68.3333%
Kappa Coefficient	= 0.5116	

	Ground Truth (Pixels)				
Class	Forest	Water	Agriculture	Urban	Wetland
Forests [Green]	4	0	0	0	0
Water [Blue]	0	0	0	0	0
Agriculture [Chartreuse]	3	1	23	7	5
Urban [Red1]	0	0	0	12	0
Wetland [Sienna]	0	1	2	0	2
Total	7	2	25	19	7

	Ground Truth (Pixels)
Class	Total
Forests [Green]	4
Water [Blue]	0
Agriculture [Chartreuse]	39
Urban [Red1]	12
Wetland [Sienna]	5
Total	60

	Ground Truth (Percent)				
Class	Forest	Water	Agriculture	Urban	Wetland
Forests [Green]	57.14	0.00	0.00	0.00	0.00
Water [Blue]	0.00	0.00	0.00	0.00	0.00
Agriculture [Chartreuse]	42.86	50.00	92.00	36.84	71.43
Urban [Red1]	0.00	0.00	0.00	63.16	0.00
Wetland [Sienna]	0.00	50.00	8.00	0.00	28.57
Total	100.00	100.00	100.00	100.00	100.00

	Ground Truth (Percent)
Class	Total
Forests [Green]	6.67
Water [Blue]	0.00
Agriculture [Chartreuse]	65.00
Urban [Red1]	20.00
Wetland [Sienna]	8.33
Total	100.00

Class	Commission (Percent)	Omission (Percent)	Commission (Pixels)	Omission (Pixels)
Forests [Green]	0.00	42.86	0/4	3/7
Water [Blue]	0.00	100.00	0/0	2/2
Agriculture [Chartreuse]	41.03	8.00	16/39	2/25
Urban [Red1]	0.00	36.84	0/12	7/19
Wetland [Sienna]	60.00	71.43	3/5	5/7

Class	Prod. Acc.	User Acc.	Prod. Acc	User Acc.

	(Percent)	(Percent)	(Pixels)	(Pixels)
Forests [Green]	57.14	100.00	4/7	4/4
Water [Blue]	0.00	0.00	0/2	0/0
Agriculture [Chartreuse]	92.00	58.97	23/25	23/39
Urban [Red1]	63.16	100.00	12/19	12/12
Wetland [Sienna]	28.57	40.00	2/7	2/5

Discussion:

When attempting to classify the image into five groups, there were a few classes that I had trouble discerning. Two classes that I confused were the wetland and water classes. The problem of classifying these groups was mostly due to the large transition area between the two classes being difficult to define. Specifically, there were areas that had distinct pixel colors that fit neither the water nor wetland and yet in Google Earth looked to be water. The variety of wetland vegetation was also hard to discern, as some were heavily vegetated and others were highly mixed with water. In the parallelepiped algorithm in particular there was a lot of unclassified pixels between these two classes and they were difficult to fit into one class easily. I ultimately decided to look specifically at the unclassified areas and assign the pixels to one area or another manually and I had to raise the max likelihood parameter in the parallelepiped classifier slightly to 3.1 or 3.2 to get a better result. Then, when I completed the maximum likelihood algorithm, the transition areas were better defined than before.

The other classes most often confused were the agriculture/grassland and the urban/bare ground classes, because of the large amount of overlap in the DN values in these regions. The agriculture vs. bare ground decision was most often the hardest to decide. In some cases the fields would show up completely bare when using a 4,3,2 band image and in other cases the fields would show up bright with vegetation, making classification straight forward. However, in many cases the agricultural fields would be in a transition area between bare ground and healthy vegetation, sometimes even within the same field. In that case, I was forced to zoom in to the pixel level and attempt to determine which pixels were more agriculture or more bare ground. Compounding the difficulty even further was the fact that using Google Earth wasn't all that helpful in this case because the dates between the images were not the same and with farm fields there is always rotation of crops and differences throughout the year. In the end, I classified carefully in these areas and manipulated the probability threshold numbers until I got an image that was acceptable.

After completing both the unsupervised and supervised classification, I found that the supervised classification worked better overall. The main reason it works well is the amount of control that the user has in picking of the pixels. I was able to go and choose pixels in places where I knew fit into a certain class and were the best examples of that land use type. Similarly, I was able to avoid the transition areas that had mixed pixels until I really wanted to narrow down a class. It became easier to go back and narrow these fields down in supervised classification by adding or removing pixels, whereas in unsupervised you would have to go back and create all new classes and start labeling again. I do think, however that the unsupervised classification was easier to complete. There was only the k-means classifier to use and my only job was to assign the number of classes that you chose into groups. The process did not require as much knowledge of the site ahead of time.

The varying results of my two accuracy assessments make sense when looking at how they were created. For the cross validation of the original ROI sites, the high percentage correct overall and the high percentages of user's and producer's accuracy for each class make sense because these are the sites that I specifically picked out as representative samples of the classes. The amounts that were picked correctly should be high because I picked them specifically for that reason. It is important that this number is high to show that you have ROIs that you can trust as accurate. For the random sample accuracy assessment, the fact that this number is relatively low is not a good sign, but it is understandable given the circumstances. First, the amount of sample points taken (60) is a very small amount and given the randomness of them, many of the points ended up being in transition areas that were difficult to judge. Another issue is the fact that I was comparing the pixels in the original image to my classified image in which I had used a majority/minority filter on. As a result, it was not easy to know if the specific pixels I was looking at in my unbiased accuracy assessment were smoothed out in my final classification map. The last issue that affected my final class percentages was the amount of total pixels from each class that actually ended up being sampled. In the case of water, there were no random samples in those areas even though I claimed there were two. So, both my producer's and user's accuracies ended up as zero, even though that does not mean all of my water points for the entire map are incorrectly labeled.