

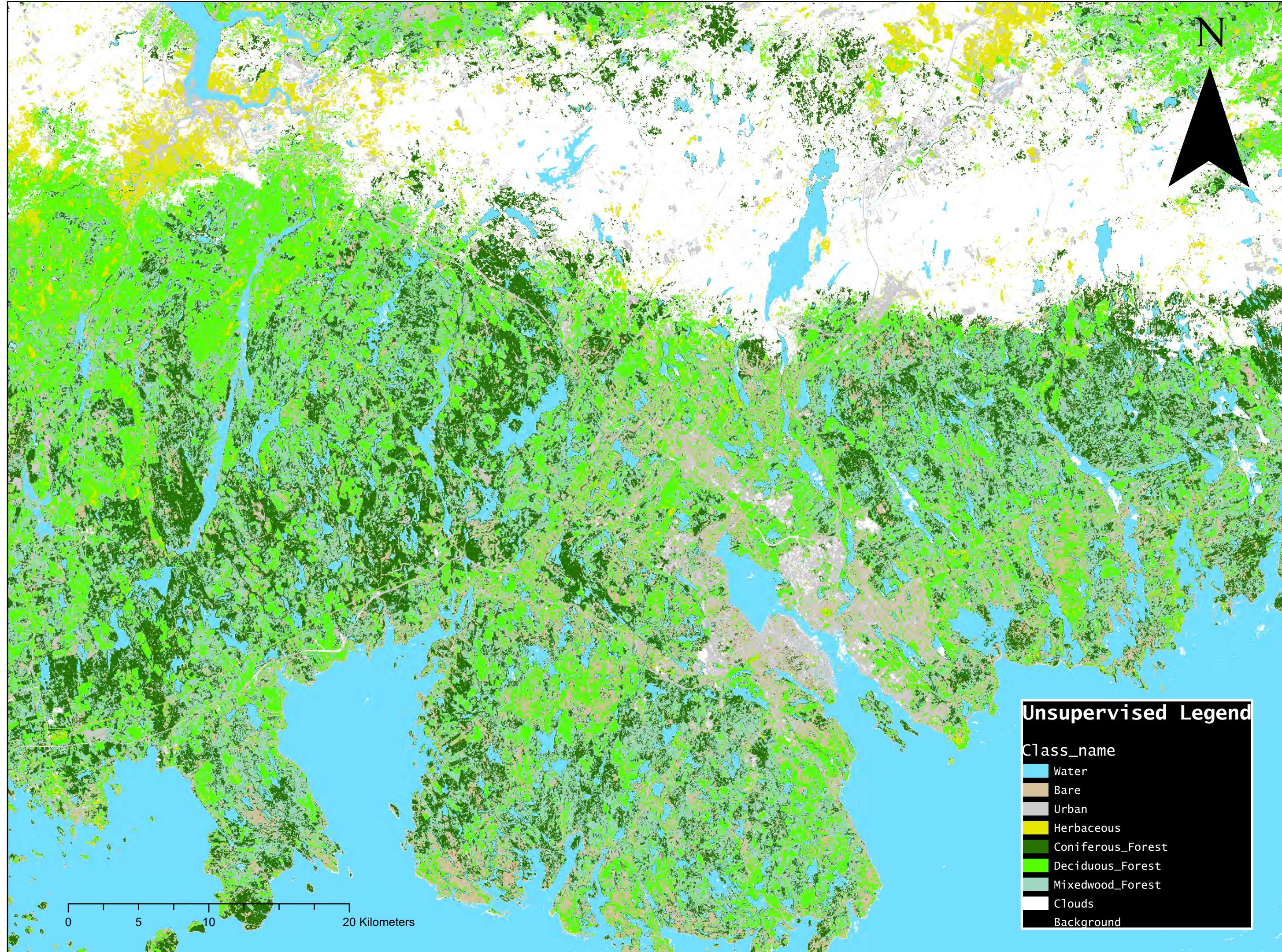
REMS 5001 - Lab 5: Image Classification with ArcGIS Pro

HRM, Nova Scotia - Imagery Provided by Course Instructor

Alex Moss, REMS 5001 - Fundamentals of Remote Sensing & Digital Imaging

Unsupervised, pixel based classification

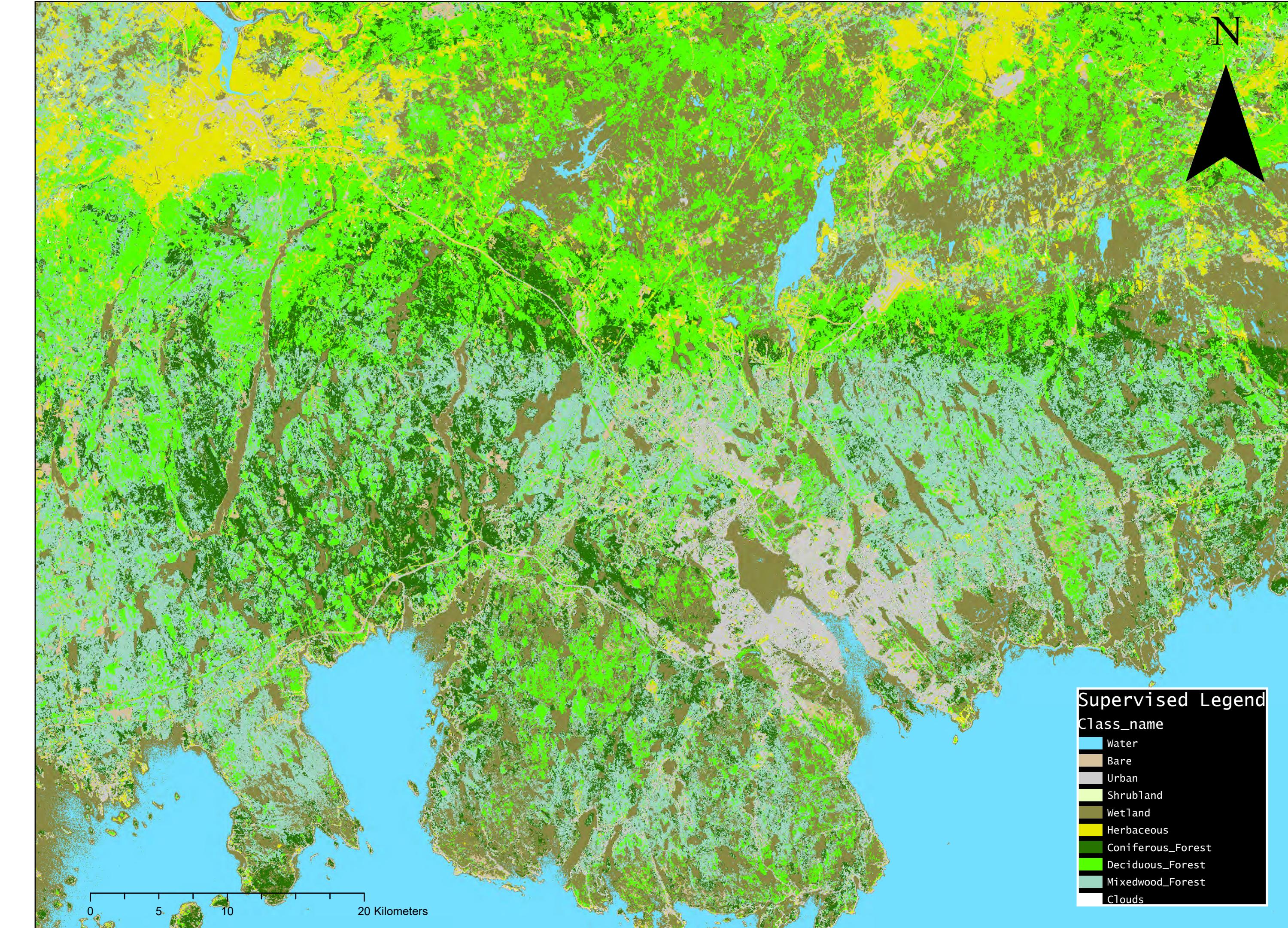
Supervised, SVM, pixel based classification



Cluster % Tables for Unsupervised Classification		
Class	Pixel Count	Cluster %
Water	15320271	37%
Bare	1944957	5%
Developed/Urban	1226712	3%
Herbaceous	987848	2%
Coniferous Forest	3010464	7%
Deciduous Forest	7508917	18%
Mixedwood Forest	5588422	13%
Clouds	5979485	14%
Total	41567076	100%
Kappa	0	0

Classification Matrix for Unsupervised Classification																		
Class	Pixel Count	Cluster %	OBJECTID	ClassValue	Water	Bare	Developed/Urban	Shrubland	Wetland	Herbaceous	Coniferous Forest	Deciduous Forest	Mixedwood Forest	Clouds	Background	Total	U_Accuracy	Kappa
Water	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	85%	0
Bare	0	0	0	0	0	5	0	4	0	0	0	0	0	0	9	0%	0	0
Developed/Urban	11	0	2	14	0	0	0	0	0	0	0	0	0	0	0	25	56%	0
Shrubland	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%	0
Wetland	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%	0
Herbaceous	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%	0
Coniferous Forest	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%	0
Deciduous Forest	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	5	80%	0
Mixedwood Forest	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%	0
Clouds	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%	0
Total	17	13	21	11	8	0	24	13	21	9	0	137	0	0	0	0	0	0
P_Accuracy	100%	0%	67%	0%	0%	0%	17%	38%	25%	56%	0%	0%	37%	0	0	0	0	0.300397839
Kappa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

For the unsupervised, pixel based image classification, most of the default parameters were used, as directed in the lab sheet. A custom schema was created to match the table in the lab sheet. For the max number of classes, instead of setting it to 5-10 times the number of desired classes, it was set to 25. Only 8 of the 10 land cover classes were assigned spectral classes, which is why wetlands and shrubland are missing from the map. For the results, looking at the cluster percentage added into each information class, water had the most with 37% of the total clusters, which is to be expected. The herbaceous class had the least with only 2%. The number that stands out the most is clouds at 14%. This most likely took place due to a lot of bare or urban features appearing almost white because of how much they reflected the light. Clouds should have been the class with the lowest amount of clusters. Turning our attention to the classification matrix, we can see that the overall accuracy is at 37%, which is quite low. This value is obtained by adding up all of the numbers highlighted in a colored box and dividing by the total numbers of instances. The expectation is that the supervised classification will have a better accuracy score. One thing to note about the accuracy score is that it is based on a set of ground truth validation data that is clustered around the Annapolis Valley, which limits how authentic the accuracy data is.



Cluster % Tables for SVM Supervised Classification		
Class	Pixel Count	Cluster %
Water	16578476	26.4%
Bare	642743	1.0%
Developed/Urban	1027070	1.6%
Shrubland	645787	1.0%
Wetland	5062947	8.1%
Herbaceous	23533475	37.5%
Coniferous Forest	2574773	4.1%
Deciduous Forest	4978609	7.9%
Mixedwood Forest	7509466	12.0%
Clouds	265105	0.4%
Total	62818451	100%
Kappa	0	0

Classification Matrix for Supervised Classification, Support Vector Machines																		
Class	Pixel Count	Cluster %	OBJECTID	ClassValue	Water	Bare	Developed/Urban	Shrubland	Wetland	Herbaceous	Coniferous Forest	Deciduous Forest	Mixedwood Forest	Clouds	Total	U_Accuracy	Kappa	
Water	10	0	0	0	0	0	0	0	0	0	0	0	0	0	2	12	83%	0
Bare	0	10	1	0	1	0	0	0	0	0	0	0	0	0	0	11	91%	0
Developed/Urban	0	1	17	0	0	0	0	0	0	0	0	0	0	0	0	18	94%	0
Shrubland	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	1	0%
Wetland	7	0	0	0	0	5	0	0	2	0	0	0	0	0	2	4	20%	0
Herbaceous	0	2	2	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0%
Coniferous Forest	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	4	75%	0
Deciduous Forest	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	21	14%	0
Mixedwood Forest	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	26%	0
Clouds	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%
Total	17	13	21	11	8	0	0	0	0	0	0	0	0	0	24	13	0%	0
P_Accuracy	59%	77%	81%	0%	63%	0%	0%	0	0	0	0	0	0	0	23%	43%	0%	0.42%
Kappa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.351594392

For the supervised, pixel based image classification, a set of training samples were created, along with the created schema from the unsupervised classification, were used. A total of 30 samples were taken for each information class. As for the classifier, all three algorithms were used but support vector machine performed the best so it was used for the poster. Checking the results, in the cluster % table, we can see that the Herbaceous class was clearly being misidentified as it had 37.5% of the clusters. Clouds only had 0.4%, which makes more sense when compared to the previous example. My guess is that the herbaceous training samples, although definitely agriculture plots, were mistaken for the other classes that look similar. When totaling up the accuracy score from the confusion matrix, we get 42% compared to the 37% we got with the unsupervised classification. Although an improvement was expected, only achieving 42% is disappointing. Information classes such as shrubland (even with the file you sent me), herbaceous, and clouds were getting misidentified as other classes. This can be attributed to both the limitation of the validation dataset as well as a lot of classes looking very similar. When comparing the maps, the supervised classification map is visibly classified better, with a more diverse number of colors representing all ten information classes on display.