

Comparing a series of unsupervised classifications ran on a Landsat-5 image of the Salton Sea

Nicole Stotz

Geog 342

Spring 2010

Project Summary

- Compare increasing numbers of unsupervised classes, and the affect on image output using a Landsat-5 image of the Salton Sea.
- Part 1: Ran a series of unsupervised classifications, starting with 6 classes and ending with 30. Compared transformed divergence reports for each classification.
- Part 2: Land cover classification – used the recode tool on a series of unsupervised classifications, comparing how the number of signatures affect a simplified landcover scheme with 5 classes.

Acquiring An Image

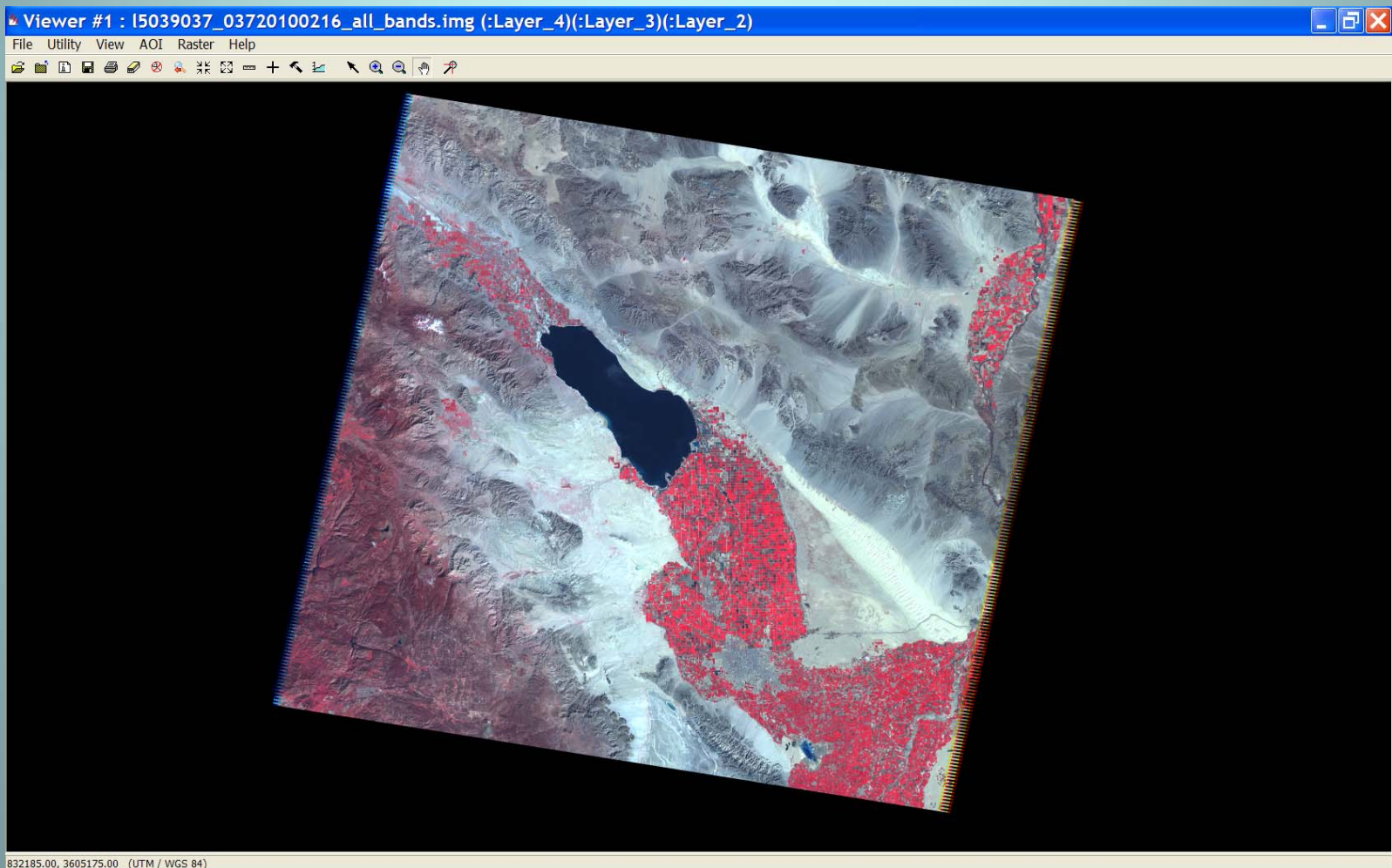
- Study area: Salton Sea/Imperial County, CA
- glavis.usgs.gov for Landsat imagery
- Originally I downloaded a Landsat 7 image – post 2003 images have striping because of damage to the satellite's scan-line corrector. I didn't know how to correct this, so I chose the same view from Landsat 5.



Striping in Landsat 7 image

Salton Sea, Landsat-5 image acquired Feb. 16, 2010

- 28.5-meter data.
- 7 bands, coverage area of 183 x 170 km.
- Image file included individual .tif for each band.
- Band files were converted from .tif to .img files, and then stacked into a single layer using ERDAS.



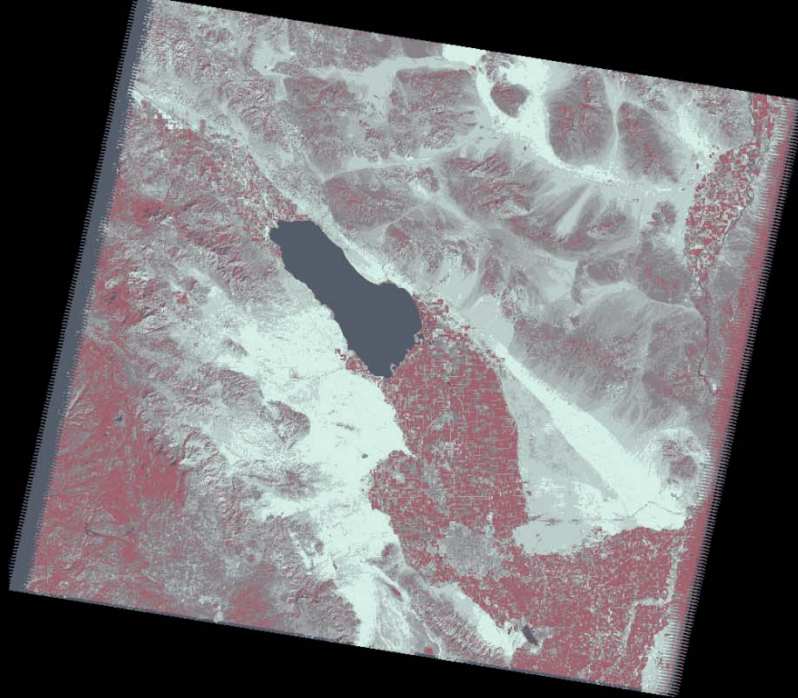
Part 1: Comparing unsupervised classifications

- 6 classes as a starting point – a smaller number wouldn't work with the spectral diversity of the image.
- 30 classes as an end point – a balance of demonstrating a trend, and data overload.
- 14 unsupervised classifications ran, 15 iterations per classification.
- Collected transformed divergence reports and dendrograms for each unsupervised classification.

Output examples....

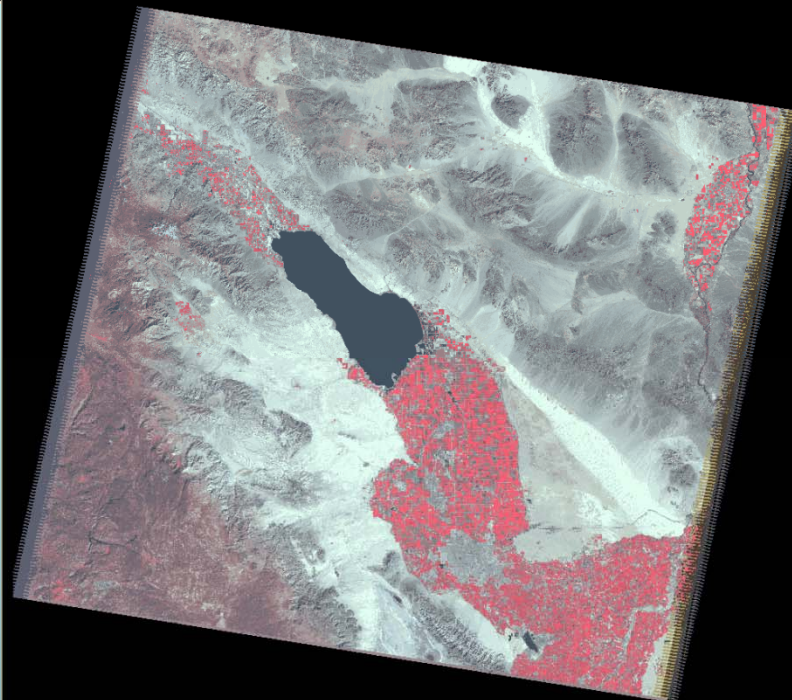
7 classes vs. 30 classes

(Lots of other intermediary classes not shown)



Raster Attribute Editor - 7cl...

Row	Histogram	Color	Red	Green
0	17024338		0	0
1	2807760		0.33	0.36
2	6190561		0.64	0.39
3	7370414		0.53	0.52
4	7424234		0.61	0.62
5	5988959		0.67	0.71
6	6237490		0.74	0.81
7	4263605		0.84	0.94



Raster Attribute Editor - 30classes.1...

Row	Histogram	Color	Red	Green	Blue
0	17024338		0	0	0
1	1029705		0.36	0.38	0.44
2	1529809		0.25	0.32	0.37
3	1101375		0.82	0.37	0.44
4	1110521		1	0.36	0.44
5	865178		0.78	0.47	0.52
6	1932639		0.53	0.44	0.47
7	2429718		0.44	0.41	0.45
8	559836		0.29	0.25	0.2
9	1440481		0.5	0.54	0.55
10	1659617		0.49	0.52	0.54
11	1868564		0.54	0.57	0.59
12	389562		0.58	0.62	0.64
13	1718722		0.56	0.62	0.63
14	1393750		0.6	0.5	0.52
15	1160739		0.64	0.61	0.62
16	1606316		0.6	0.65	0.65
17	1113952		0.52	0.68	0.69
18	1047887		0.67	0.62	0.62
19	1199348		0.65	0.72	0.73
20	1545914		0.65	0.69	0.69
21	244312		0.68	0.63	0.41
22	496443		0.6	0.54	0.56
23	1770522		0.68	0.75	0.75
24	1103367		0.7	0.71	0.69
25	245061		0.73	0.83	0.85
26	2132404		0.72	0.78	0.78
27	1500128		0.75	0.84	0.84
28	1773300		0.76	0.82	0.79
29	2594587		0.81	0.91	0.89
30	1612486		0.89	1	0.95

Transformed divergence tests were ran on every unsupervised classification. What sort of patterns emerged?

-The class confusion threshold is 1800, with 2000 meaning perfect seperability.

-Initial transformed divergence tests ran on unsupervised classifications of 6, 7, and 8 classes revealed a trend in class seperability.

-As more classes are specified for an unsupervised classification, class confusion does increase, but it is a finer grained spectral confusion.

Separability CellArray

Distance Measure: Transformed Divergence
Using Layers: 1 2 3 4 5 6 7
Taken 7 at a time
Best Average Separability: 1910.55
Combination: 1 2 3 4 5 6 7

Signature Name	1	2	3	4	5	6
Class 1 1	0	2000	2000	2000	2000	2000
Class 2 2	2000	0	1832.47	1998.89	2000	2000
Class 3 3	2000	1832.47	0	1606.6	1992.29	1999.95
Class 4 4	2000	1998.89	1606.6	0	1632.13	1995.91
Class 5 5	2000	2000	1992.29	1632.13	0	1600.09
Class 6 6	2000	2000	1999.95	1995.91	1600.09	0

Separability CellArray

Distance Measure: Transformed Divergence
Using Layers: 1 2 3 4 5 6 7
Taken 7 at a time
Best Average Separability: 1921.19
Combination: 1 2 3 4 5 6 7

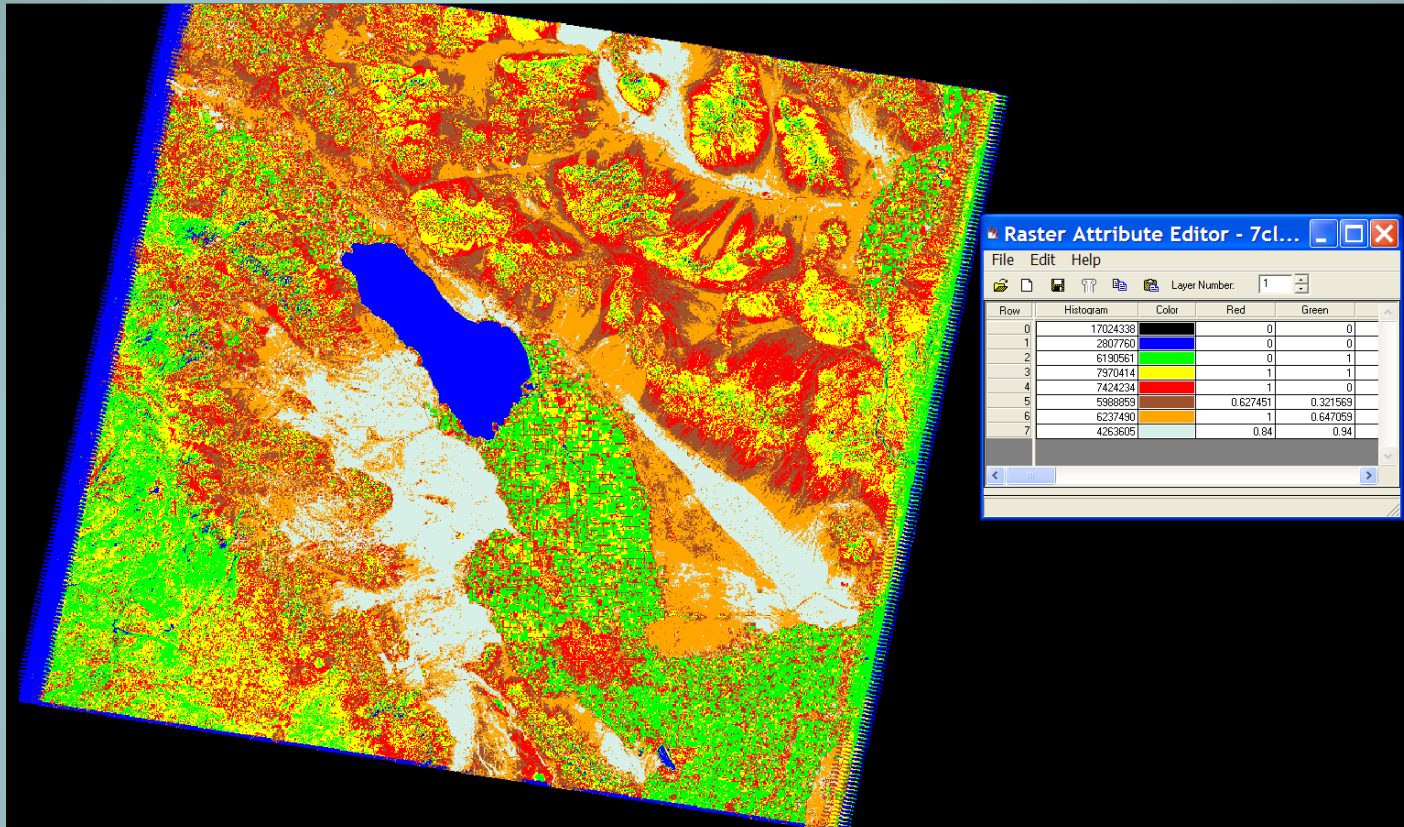
Signature Name	1	2	3	4	5	6	7
Class 1 1	0	2000	2000	2000	2000	2000	2000
Class 2 2	2000	0	1883.07	1998.27	2000	2000	2000
Class 3 3	2000	1883.07	0	1560.87	1990.7	1999.86	2000
Class 4 4	2000	1998.27	1560.87	0	1644.75	1994.4	2000
Class 5 5	2000	2000	1990.7	1644.75	0	1655.44	1997.74
Class 6 6	2000	2000	1999.86	1994.4	1655.44	0	1619.94
Class 7 7	2000	2000	2000	2000	1997.74	1619.94	0

Separability CellArray

Distance Measure: Transformed Divergence
Using Layers: 1 2 3 4 5 6 7
Taken 7 at a time
Best Average Separability: 1939.82
Combination: 1 2 3 4 5 6 7

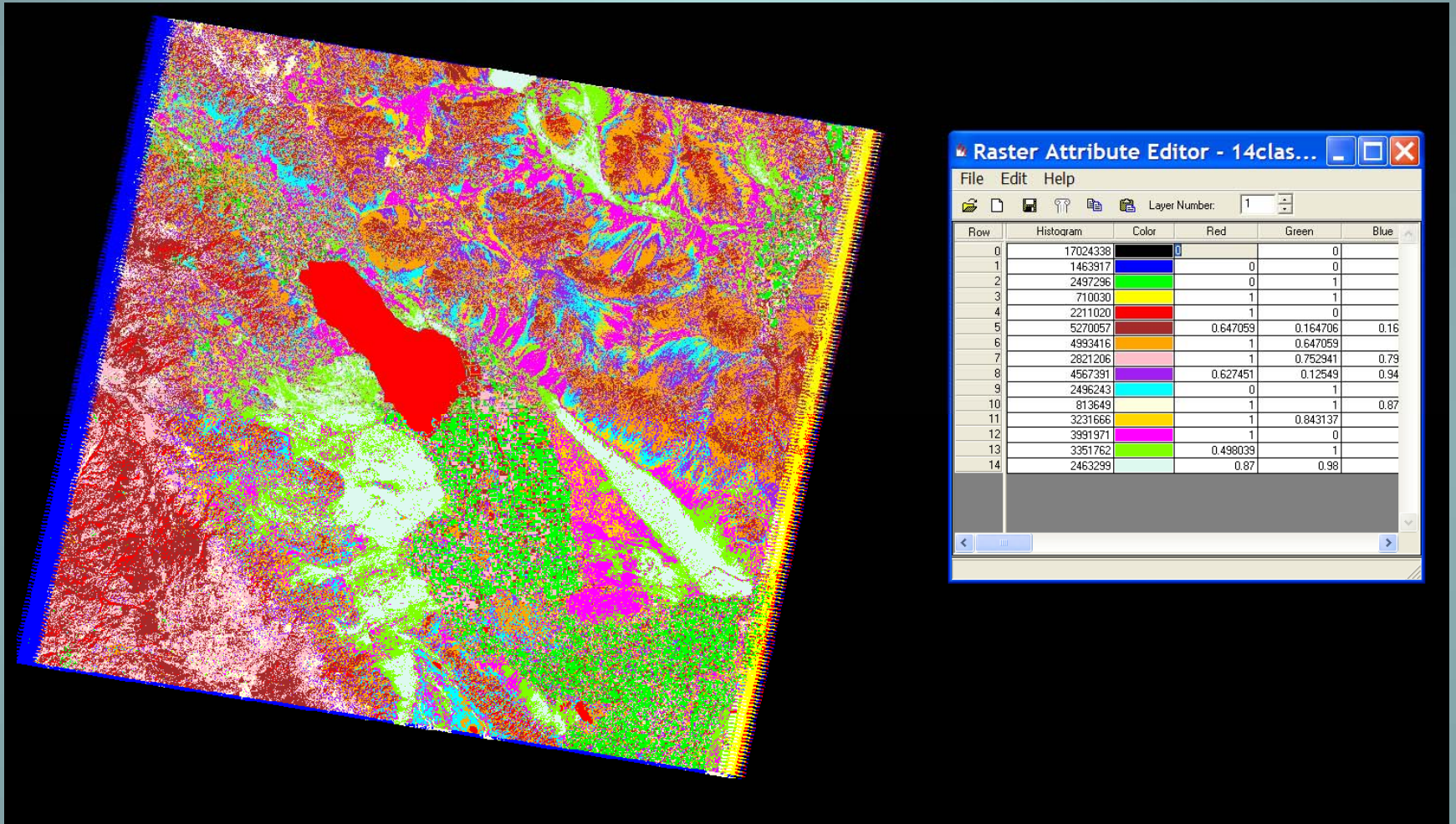
Signature Name	1	2	3	4	5	6	7	8
Class 1 1	0	2000	1999.96	2000	2000	2000	2000	2000
Class 2 2	2000	0	1982.59	1990.75	2000	2000	2000	2000
Class 3 3	1999.96	1982.59	0	1731.79	1999.25	2000	2000	2000
Class 4 4	2000	1990.75	1731.79	0	1663.41	1991.51	1999.91	2000
Class 5 5	2000	2000	1999.25	1663.41	0	1691	1991.86	1999.99
Class 6 6	2000	2000	2000	1991.51	1691	0	1680.73	1997.57
Class 7 7	2000	2000	2000	1999.91	1991.86	1680.73	0	1594.56
Class 8 8	2000	2000	2000	2000	1999.99	1997.57	1594.56	0

7 class unsupervised classification, colored to show class confusion

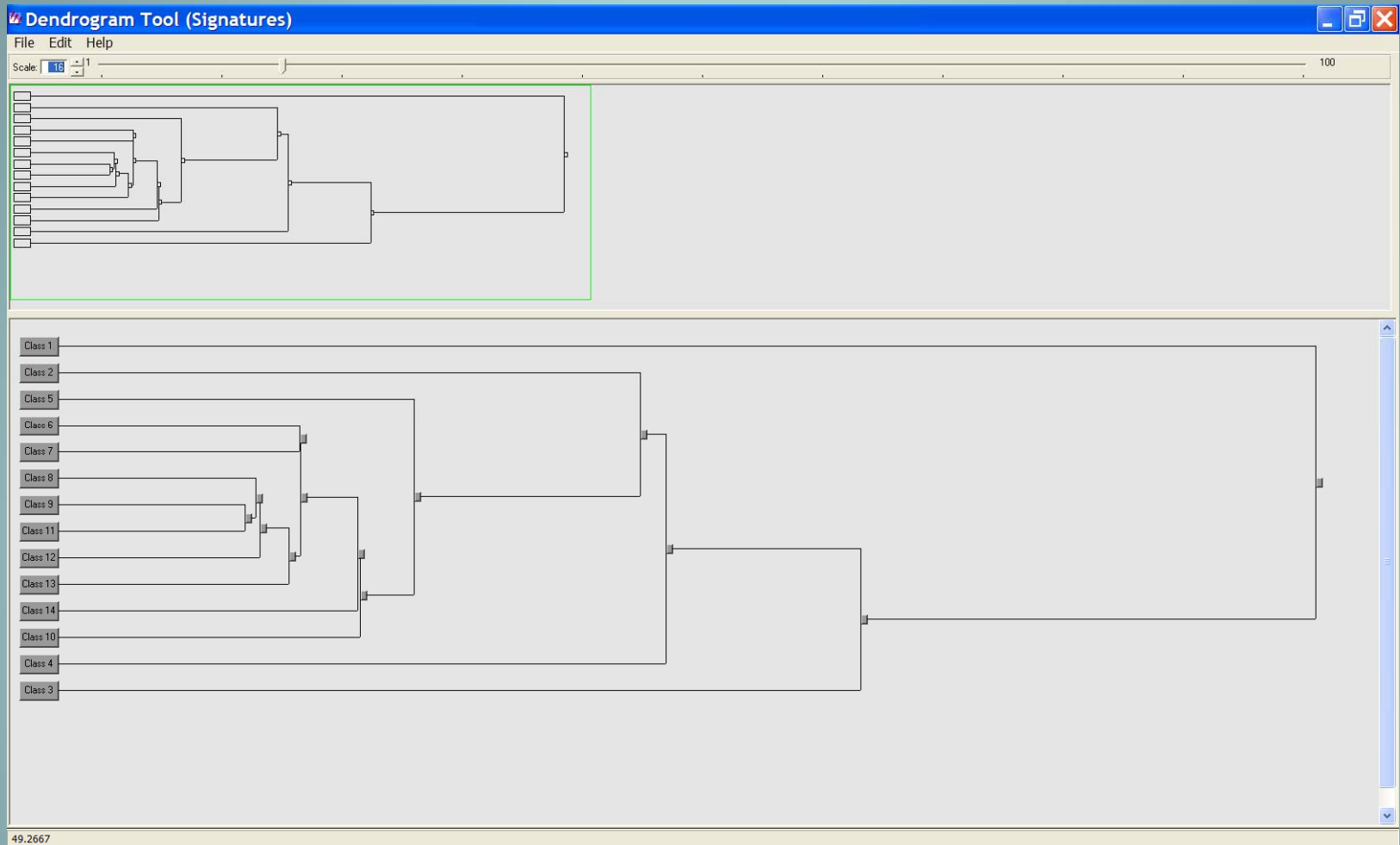


- 4:5(1645), 3:4(1561), 5:6(1655), 6:7(1620) – classes that are spectrally confused.
- Desert areas (dunes, barren, shrub) are difficult to tell from urban areas (yellow, brown, orange, grey).
- Green areas represent both crops and woodland.

14 class unsupervised classification, colored to show class confusion



Crops and woodland are now separate, but there is still much confusion between urban areas and rocky desert hills (orange, pink, blue, purple, etc).



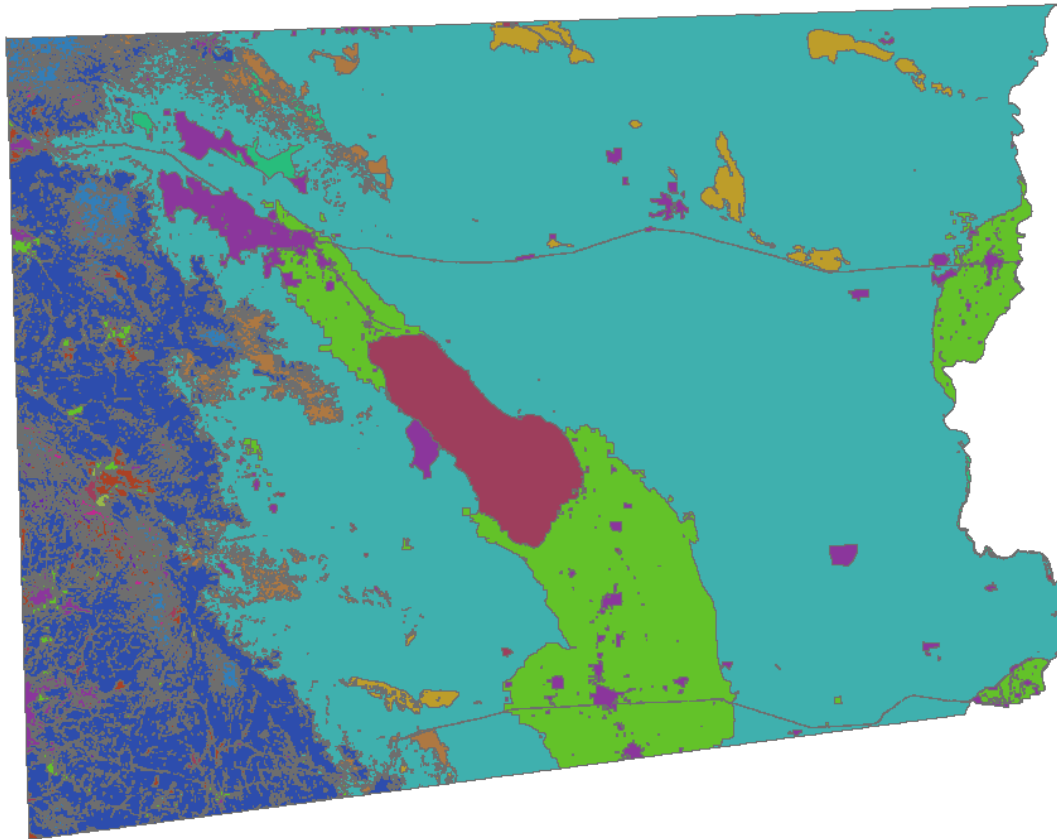
Dendrogram for 14 class unsupervised classification. Classes in the center are spectrally confused desert and urban areas.

Conclusions from Part 1:

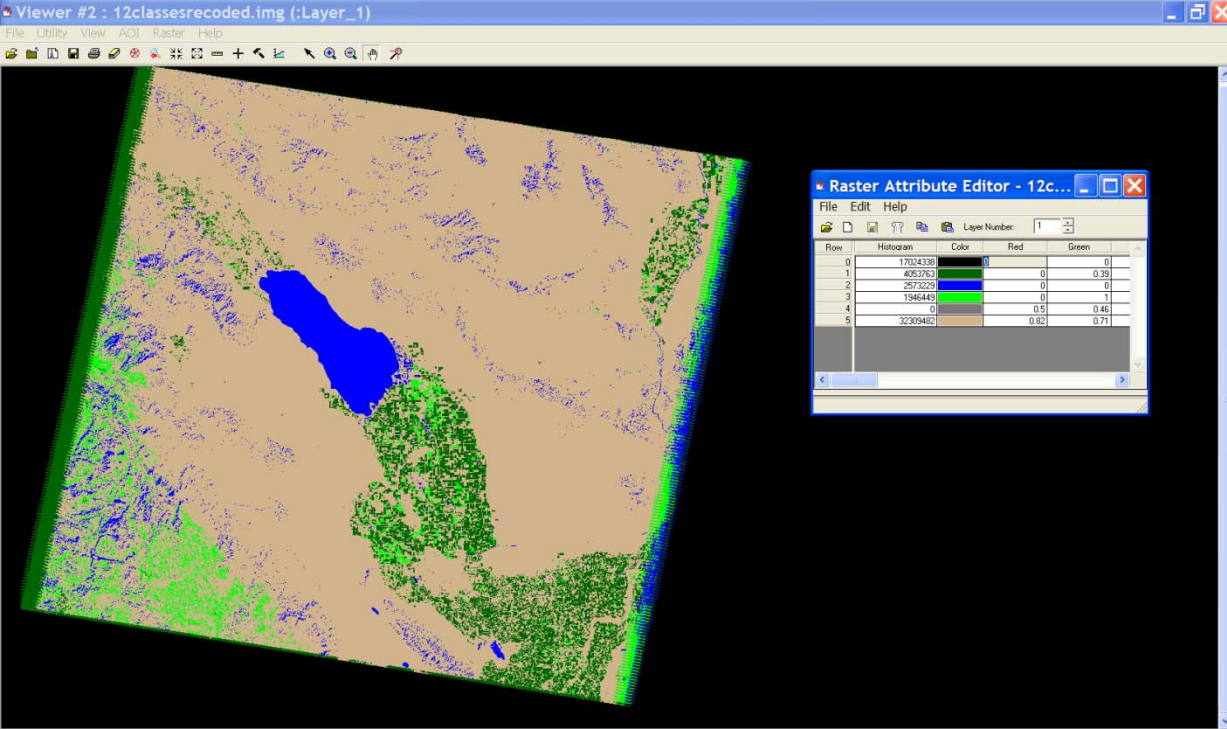
- Increasing classes in an unsupervised classification allows for greater class separability.
- Comparing signature statistics along with transformed divergence tests can help identify trends that may be passed over in a visual inspection.
- Mixed pixel areas lead to the greatest class confusion. Especially noticeable in the belt of irrigated desert cropland between the Salton Sea and Mexicali.
- Did not anticipate that urban areas would share spectral characteristics of desert shrub land and rocky hills.

Part Two: Landcover classification

- Increasing classes in an unsupervised classification affects class seperability: how does this factor into a simple landcover scheme of 5 classes?
- Used the recode tool on unsupervised classifications to create 10 landcover images.
- Planned to increase the signatures per class by 1 for each image: image 1 (6 signatures) , image 2 (12 signatures)...image 10 (60 signatures).
- Started with a 6 class scheme based on a CalFire landcover overlay: cropland, water, urban, woodland/grassland, desert, barren.
- Realized my image wasn't detailed enough to pick out barren areas. Desert became a catch-all category, and I used a 5 class scheme: cropland, water, urban, woodland/grassland, desert.
- Not sure how much this affected my results, but the general trend I saw is still useful.

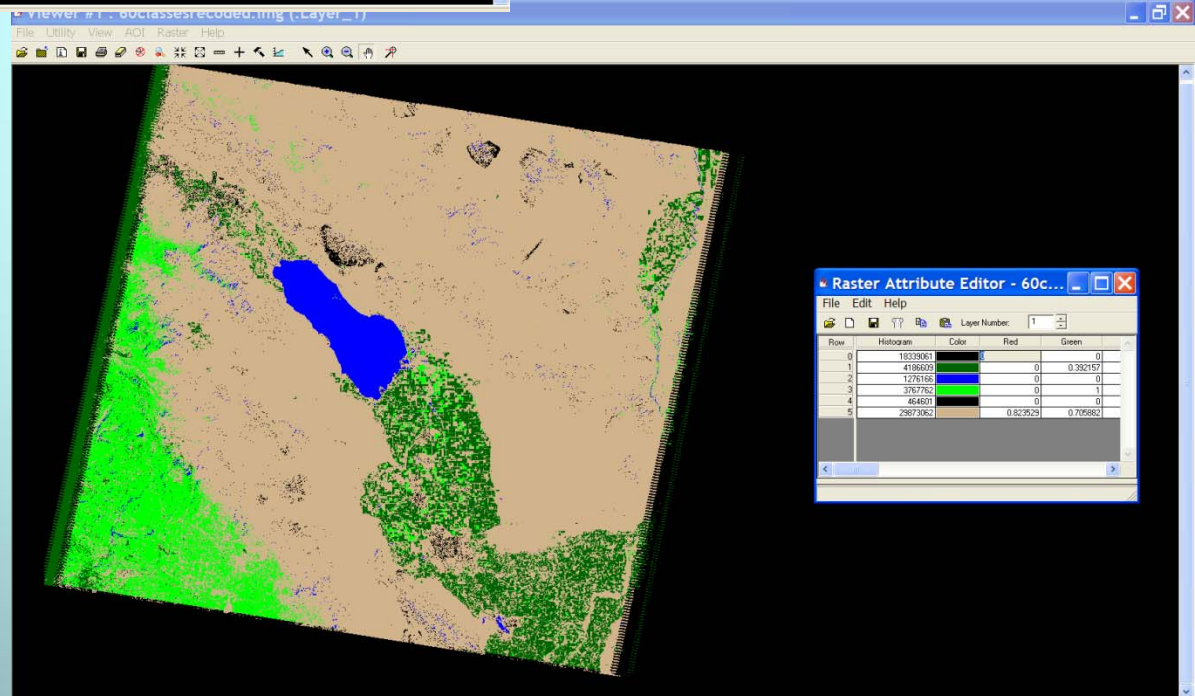


- CalFire landcover overlay, clipped to extent of Landsat scene.
- 13 landcover classes represented.
- The overlay stops at the border. Since the Landsat scene extends only a few miles into Mexico, I assumed that the terrain is close enough to that of the California overlay that I could still use it as for general comparison.



Sample output...

12 classes recoded vs
60 classes recoded



Comparison of 12 vs 60 classes recoded

- Final output for the 12 class recode

1Ag – 1-2

2Water- 3

3Trees/grassland – 6

4Urban- none

5Desert - 4-5, 7-12

*6Other - 1 (not shown)

*6Other = bands only on the edge of the image

- Final output for the 60 class recode

1Ag –6-9, 11

2Water - 2

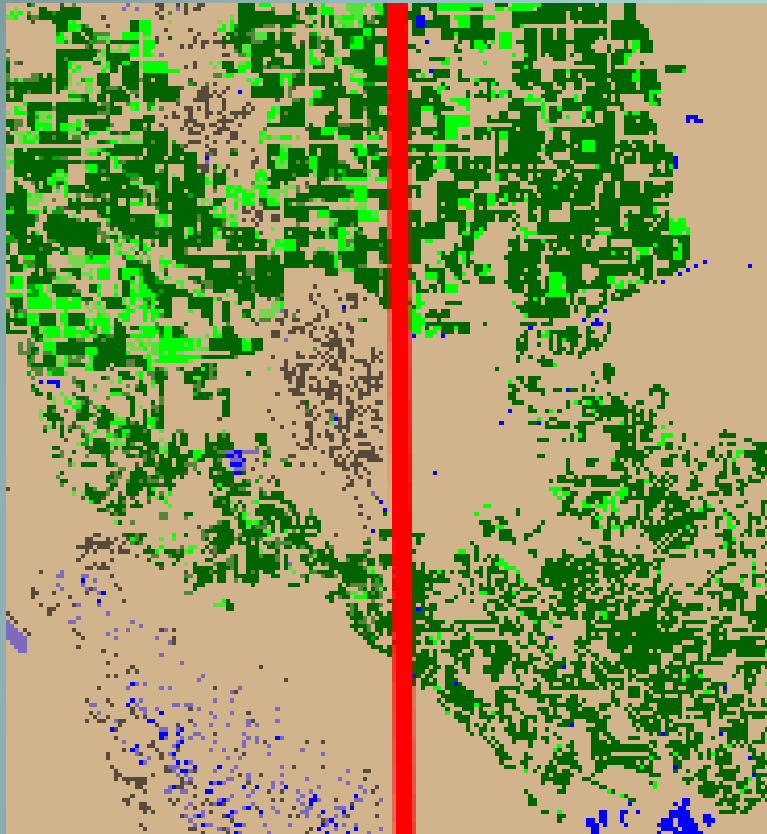
3Trees/grassland- 4, 12, 15, 22

4Urban - 24

5Desert- 13, 14, 16-21, 23, 25-46, 48-60

*6Other - 3, 5, 10, 47 (not shown)

The most signatures ended up in the Desert class, the fewest in the Urban class, in all 10 image recodes.



12 class recode vs 60 class recode

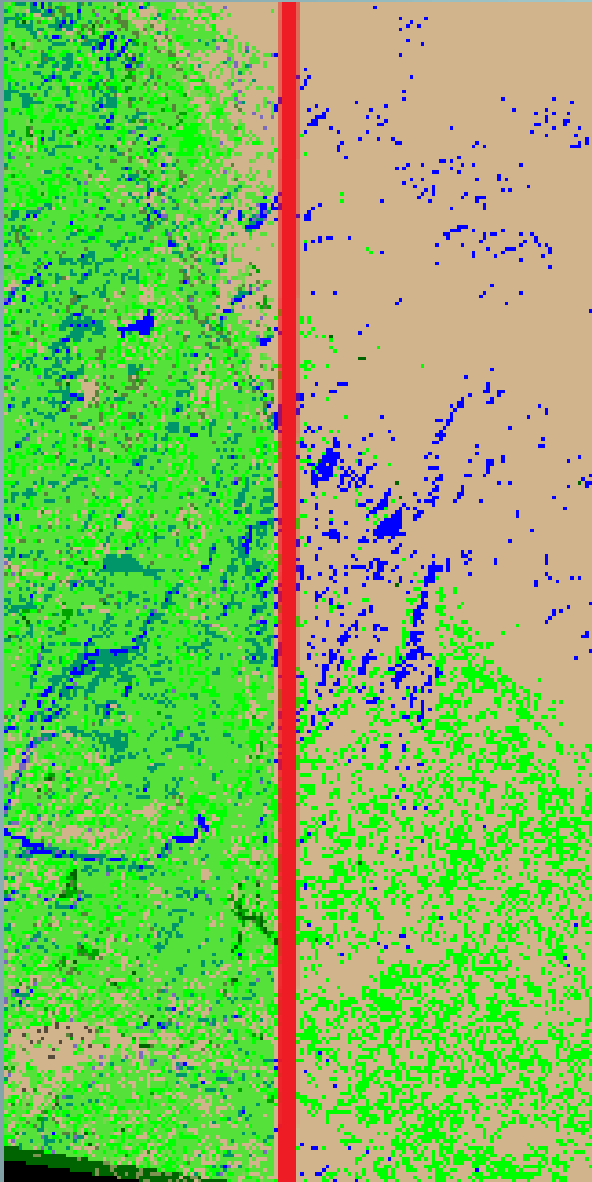
Used the swipe tool in ERDAS to compare the Urban class in my scheme.

Left = 60 class recode

Right = 12 class recode

Black speckled area is about the location of El Centro – it appears as desert in the 12 class recode on the right. The Urban class was the hardest to define.

Even though El Centro did not appear in the 12 class recode image, it's possible to define its boundaries by the missing area in the middle of the cropland.

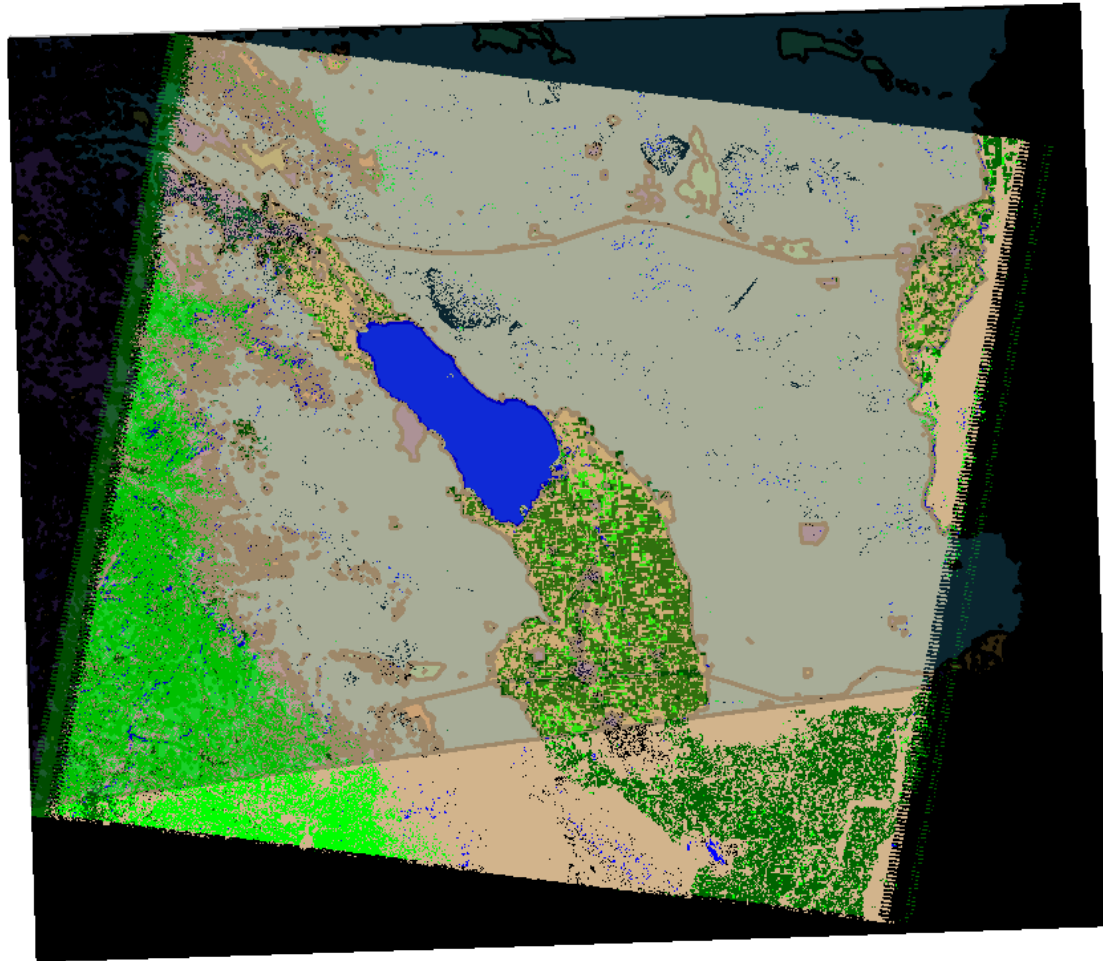


Another comparison...

Left = 60 class recode

Right = 12 class recode

Green area represents Trees/grassland class. This is a range of mountains between the Imperial Valley and San Diego. More signatures assigned to this class in the 60 class recode more accurately represent the actual vegetation.



CalFire 13 class landcover layer as a 75% transparent overlay on the 60 class recode image.

Conclusions from Part 2:

- Signature classification didn't go as planned during the recode. Some classes had more or less signatures assigned than planned for.
- More signatures assigned to a class increases the accuracy of the classification.
- Assigning signatures for the Urban class was difficult, leading to inaccuracies in the landcover images.
- Next time, use higher number of classes in an unsupervised classification if it's a basis for landcover classification.
- A higher resolution image would have made classifications more accurate.
- Using recode on an unsupervised classification is a good way to create a basic landcover classification, but the inaccuracies make it suitable for basic analysis only.