Unsupervised Classification

Classification = simplification, mapping

The early promise of satellite imagery: (1970s-80s)

A. Rapid map updating

B. Automated mapping of ‘Land Cover’

- avoid manual digitizing … by classifying <u>multispectral</u>

band data

Manual digitizing (yawn …)

e.g. BC VRI (vegetation resource inventory)

BC TRIM data layers 7027 x 1:20,000 tiles

NTS 1:50,000 example

All federal NTS map Sheets (13,370) created from Air photos

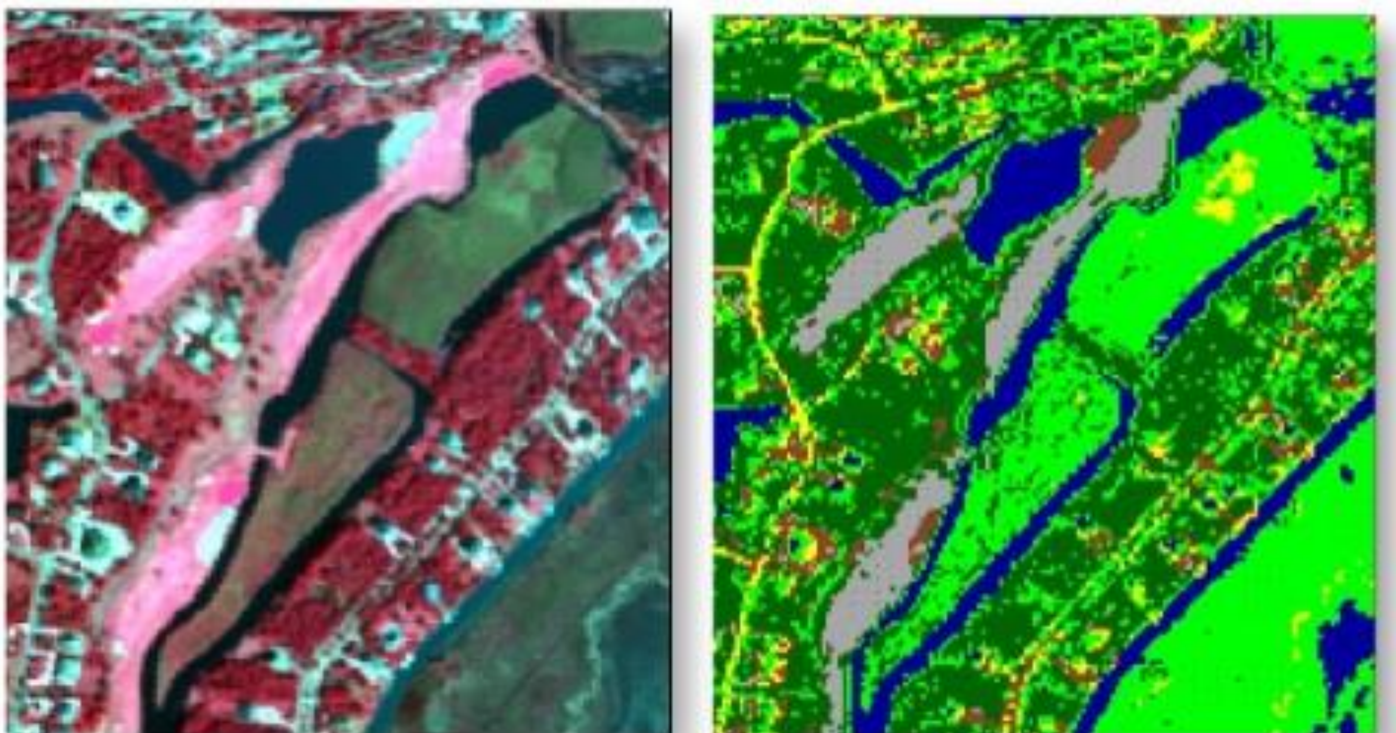
Human interpretation / classification relies on attributes such as:
Shape, pattern, texture, shadows, size, association, **<u>tone</u>**, **<u>colour</u>**

Algorithms mostly use Digital Number (DN) = digital version of tone/colour

Remote Sensing Classification

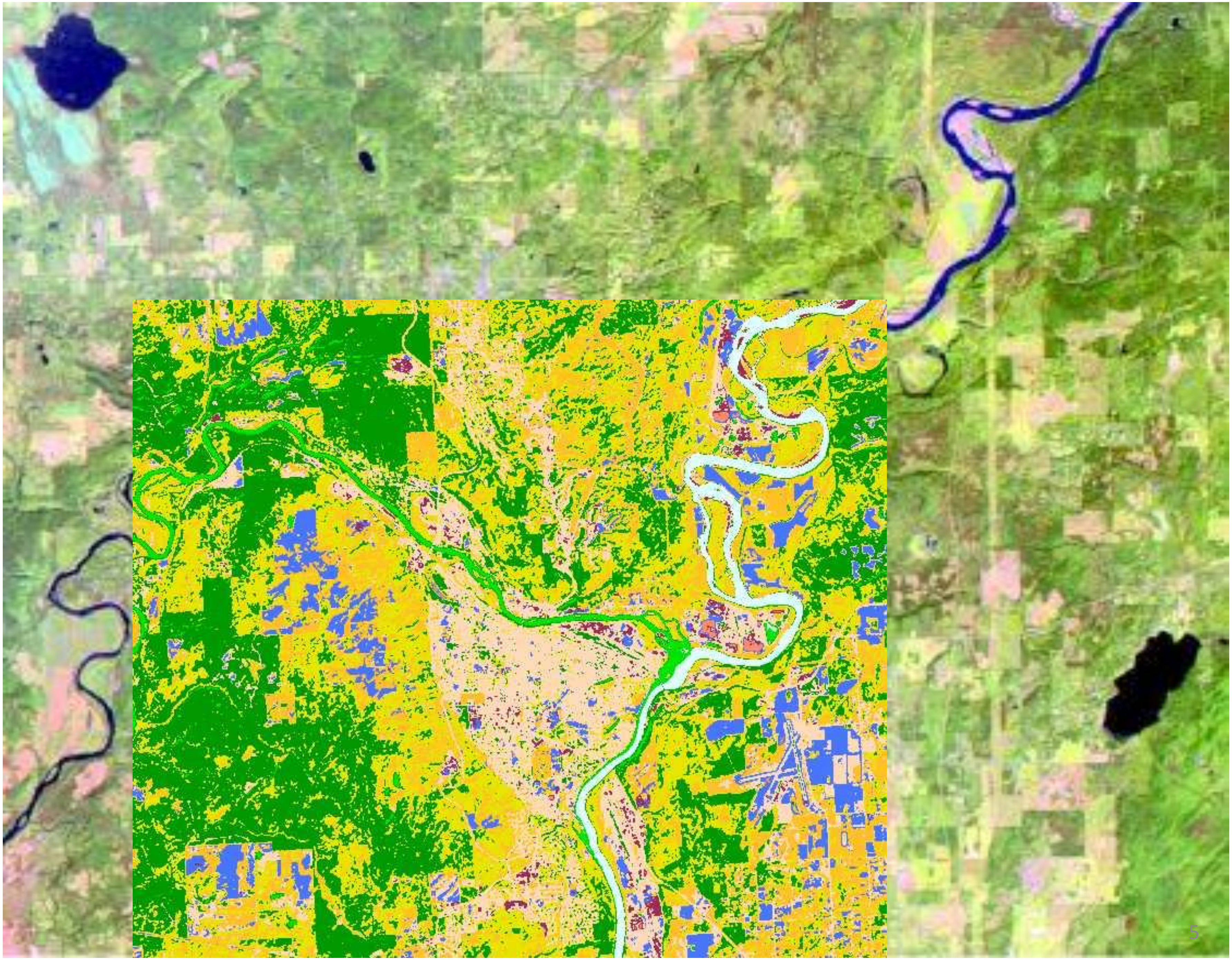
- Automated grouping of similar pixels using <u>multispectral</u> DNs
- Software developed following 1972 (Landsat 1)
- Digital alternative to manual mapping of Land Cover

Information Classes Derived from an ISODATA Unsupervised Classification Using 10 Iterations and 10 Mean Vectors of an Area Near North Inlet, SC



a. Color composite of HyMap data

c. Classification map derived from



Land Use v Land Cover (LULC) e.g. parks

Sugarbowl-Grizzly Den

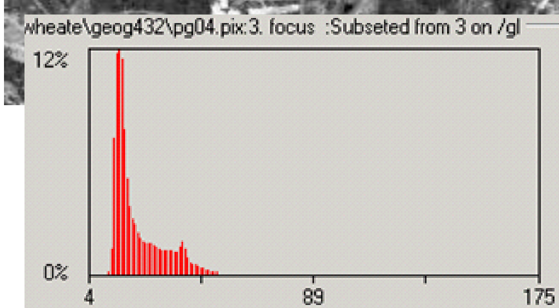
Bowron Lakes

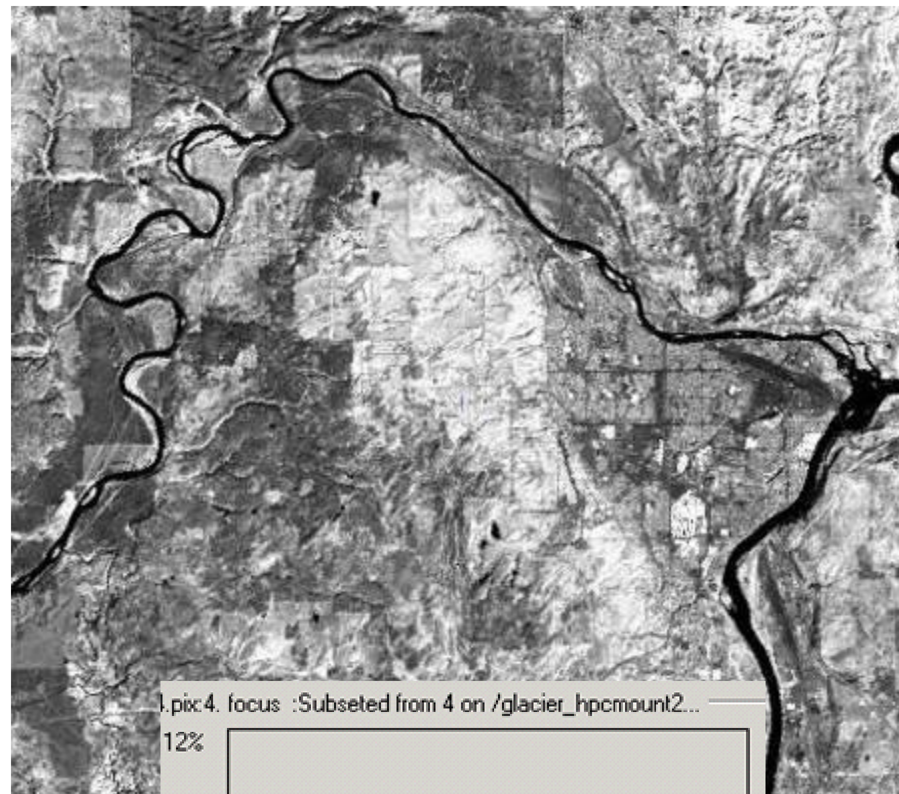
Mt. Egmont / Taranaki, NZ

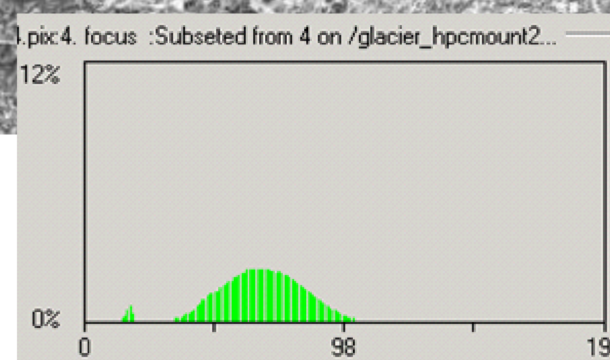
Can we use just one band to classify ?

One image band could be treated as a monochrome air photo (interpretation)

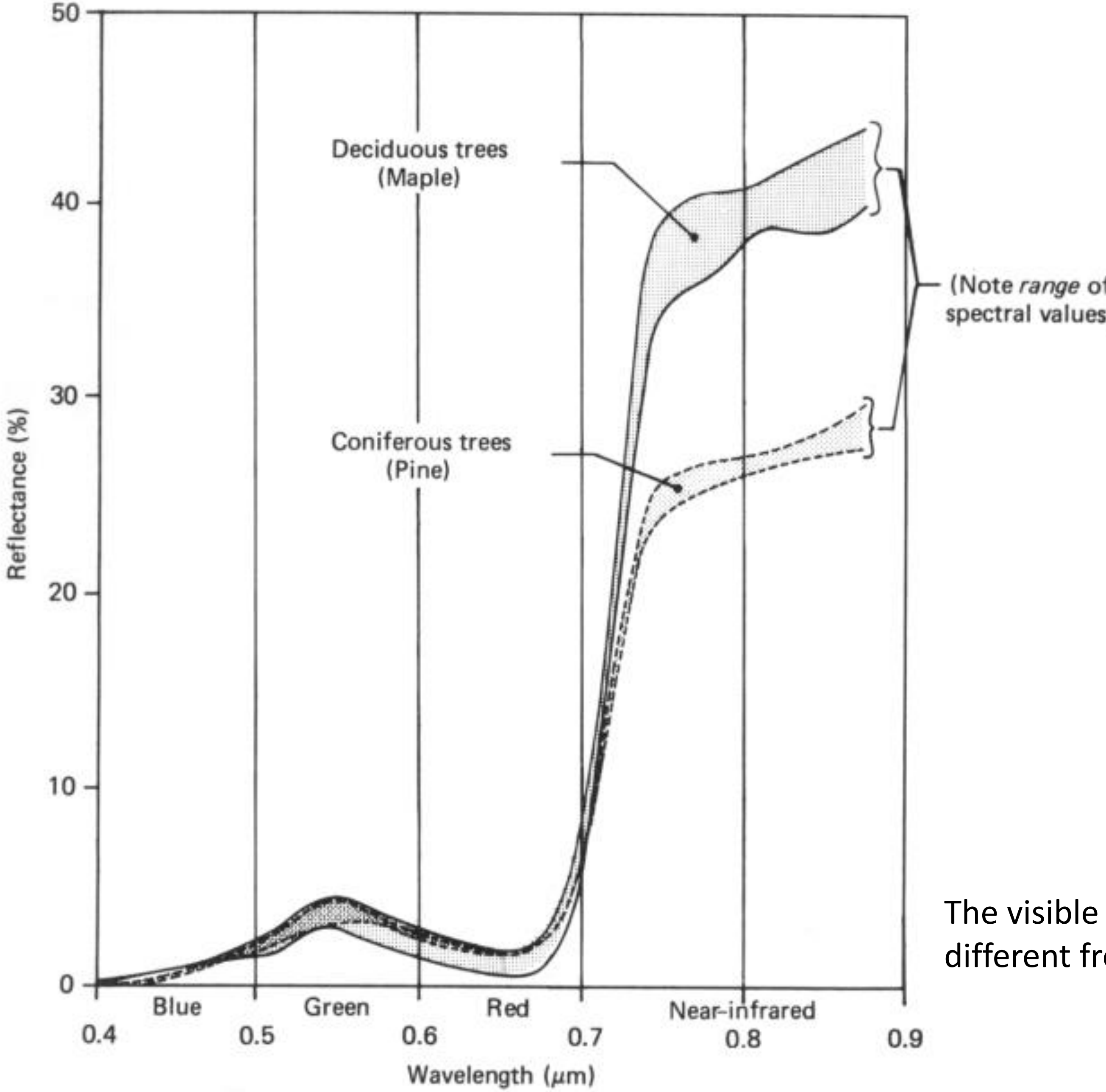
Digital Numbers from one band alone are rarely enough – features are not unique







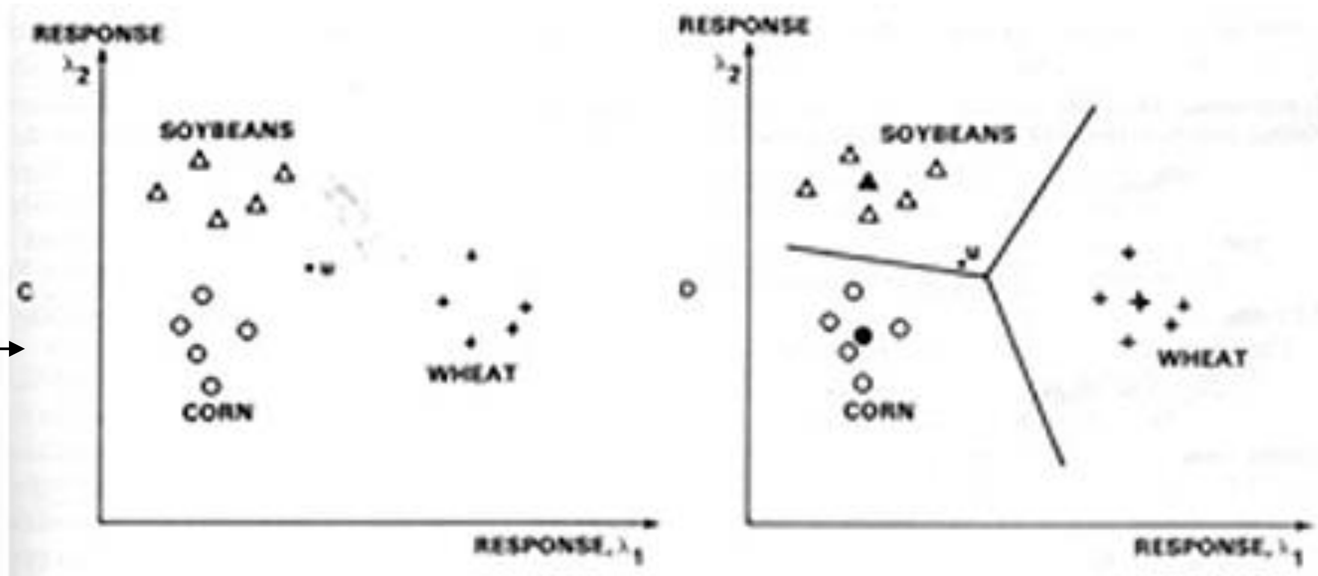
Band 3 Band 4



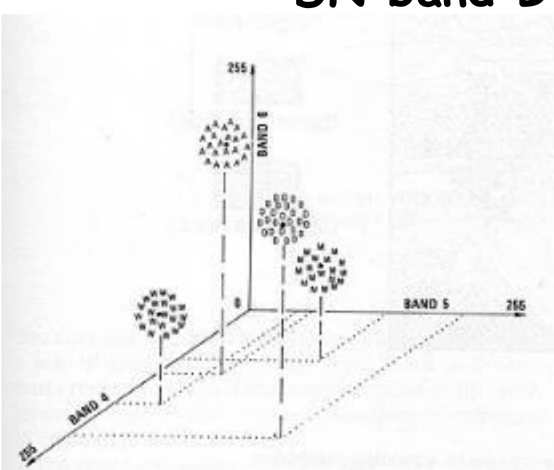
The visible bands are not too different from each other

The role of multispectral sensing in classification

DN
Band A



DN band B



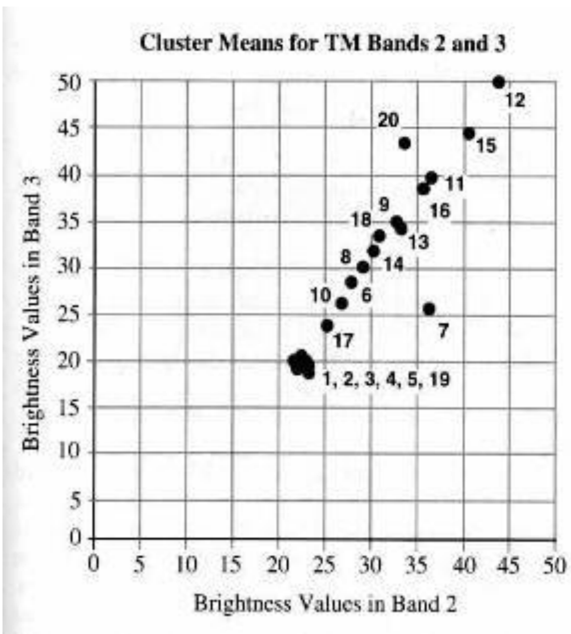
DNs in Band A are similar for Corn and Wheat
DNs in Band B are similar for Corn and Soybeans

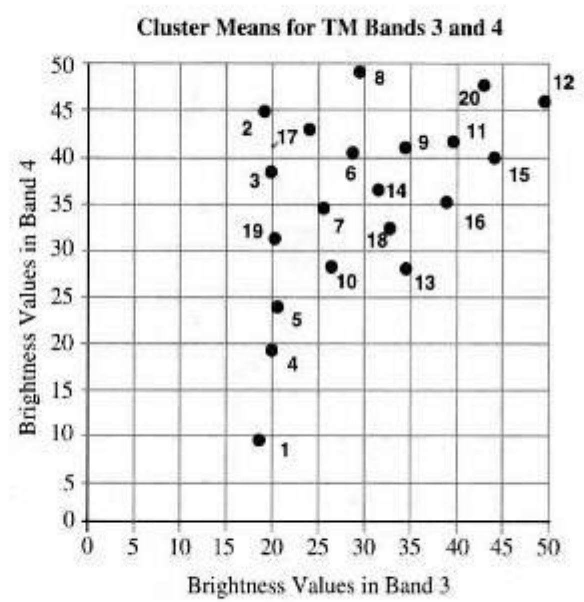
… but if we use both Bands A and B, then all 3 differ

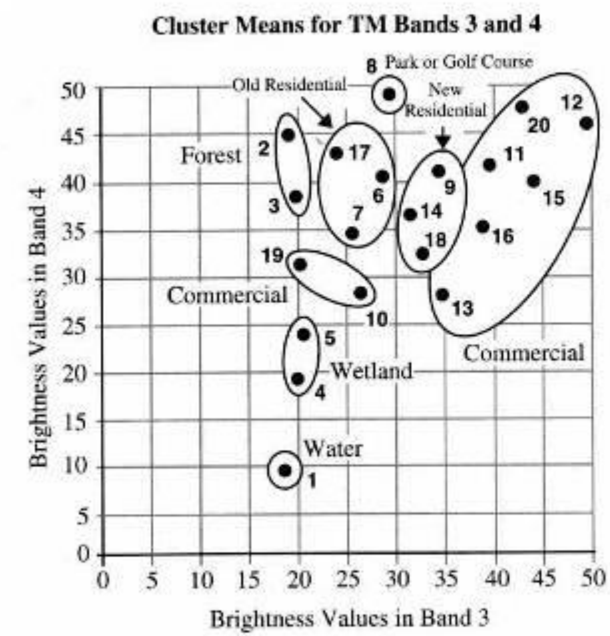
… Algorithms are '**per pixel**' classifiers

Unsupervised classification = 'clustering'

Example of Visible bands only (2,3) versus Visible and Near-IR







Two bands are shown for simplicity

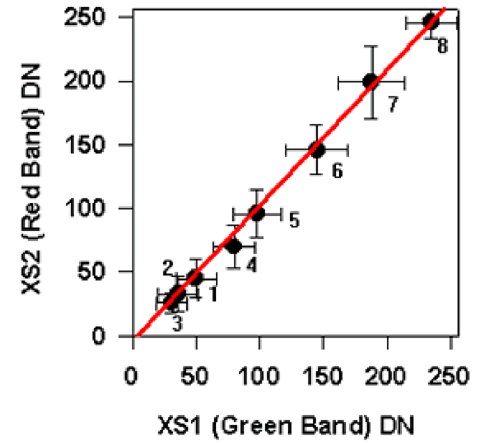
Input bands selected – minimum 3 bands;

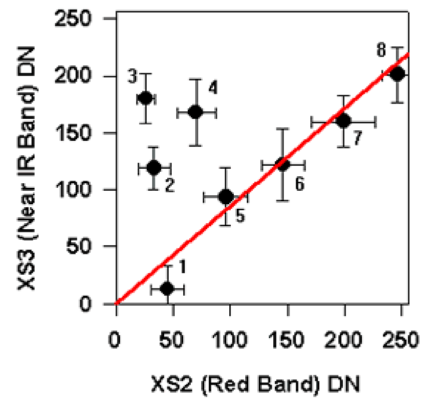
Note: you can only display 3 bands, but you can input **more than** 3 (no limit)

band correlation coefficients and scatterplots

Example: PG Landsat TM data (r values between bands)

	TM1	2	3	4	5	6
TM1						
TM2	.97					
TM3	.96	.96				
TM4	.07	.16	.11			
TM5	.66	.72	.76	.46		
TM6	.77	.77	.81	.14	.80	
TM7	.83	.86	.90	.25	.93	.86





The Visible bands are highly correlated (similar) .. (r = .96 to .97)

.. so also are TM bands 5 and 7 (r = .93)

band 4 (near-IR) is not very correlated with Visible or MIR (nor thermal)

Note: these values will vary for different environments e.g. urban, desert, forested

Classification: Band / Channel Selection

How to choose which ones to use:

1. Low correlation e.g. TM 3-4-5 or 2-4-7 (Visible-NIR-MIR)

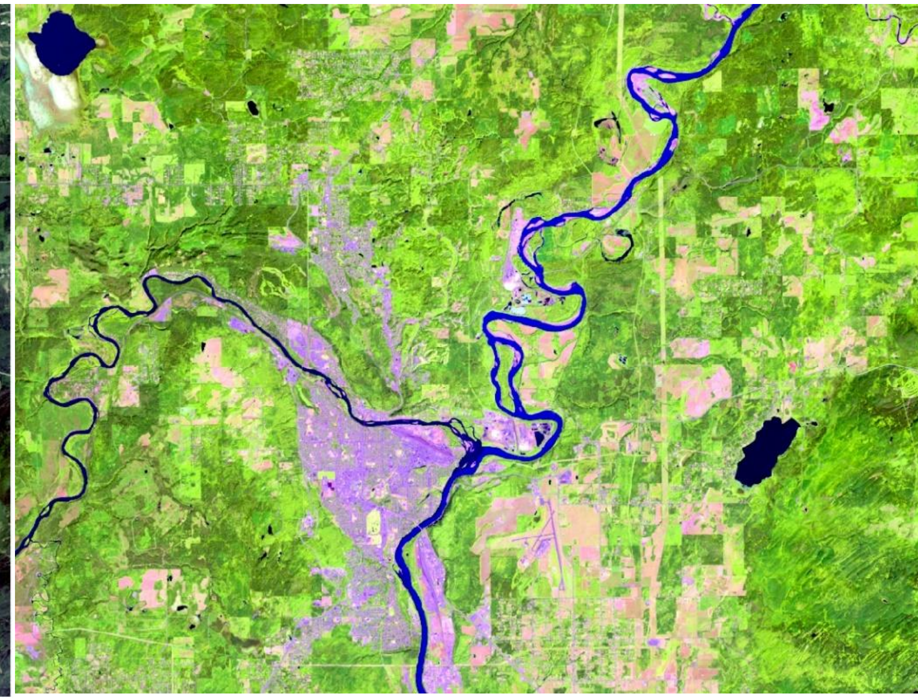
2. Past experience, visual examination, logical thinking

3. Channels that separate the features we want to identify
   (based on DNs / spectral curves / histograms)

4. Or simply just use them all ? … (except the thermal band)
- often this can confuse the classifier and not find clusters

Band / channel selection
Landsat 5 Thematic Mapper (TM) : 1-7
Landsat 8/9 Operational Land Imager (OLI) 1-9

Visible bands Red-Green-Blue



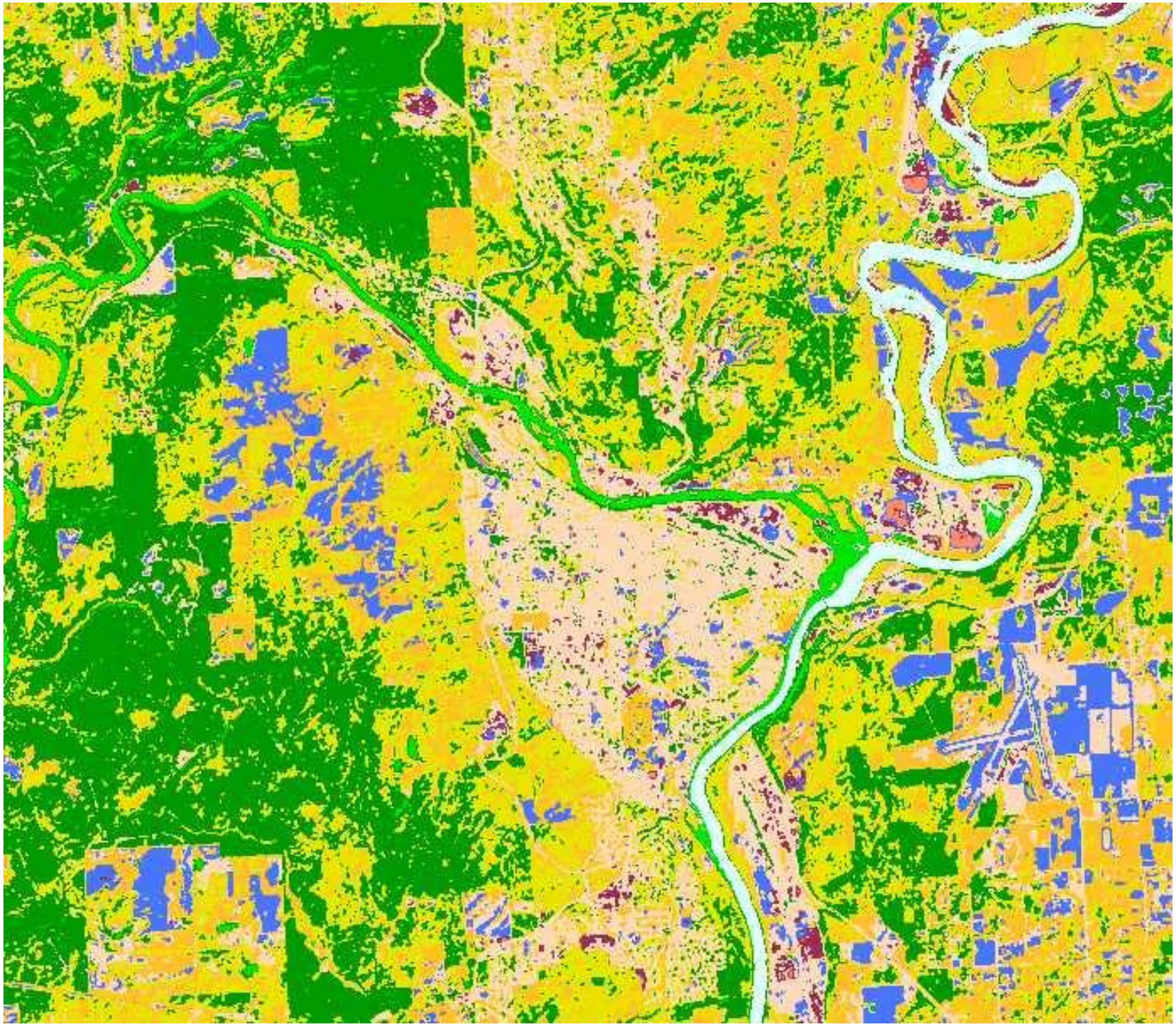
SWIR – NIR- Red

You would NOT select 3 visible bands to classify

The visible bands are similar – and thus the composite is low in contrast


Knob and kettle
Eskers
drumlinoids
Glacial Lake Fraser

Unsupervised result – 10 classes (clusters)



This is a new channel in your .pix file
- It's not a band

Colours are random

Note: urban classification is NOT often easy!

Unsupervised classification

Characteristics

-user initially needs little 'a priori' knowledge of area

-The software clusters pixels by natural DN groupings
  -based on similarity and contrast ~ 'natural breaks'

Steps

- determine how many classes / clusters

- determine which input bands / channels to use

- run classifier : K-means or Isodata

- Rerun with more clusters if needed

- assign names to classes (merge classes if needed)

Unsupervised classification –algorithms and iterations

PCI .. Fuzzy K-means is less common in GIS software

1. K-means minimises within cluster range of DNs ('K' = clustering)

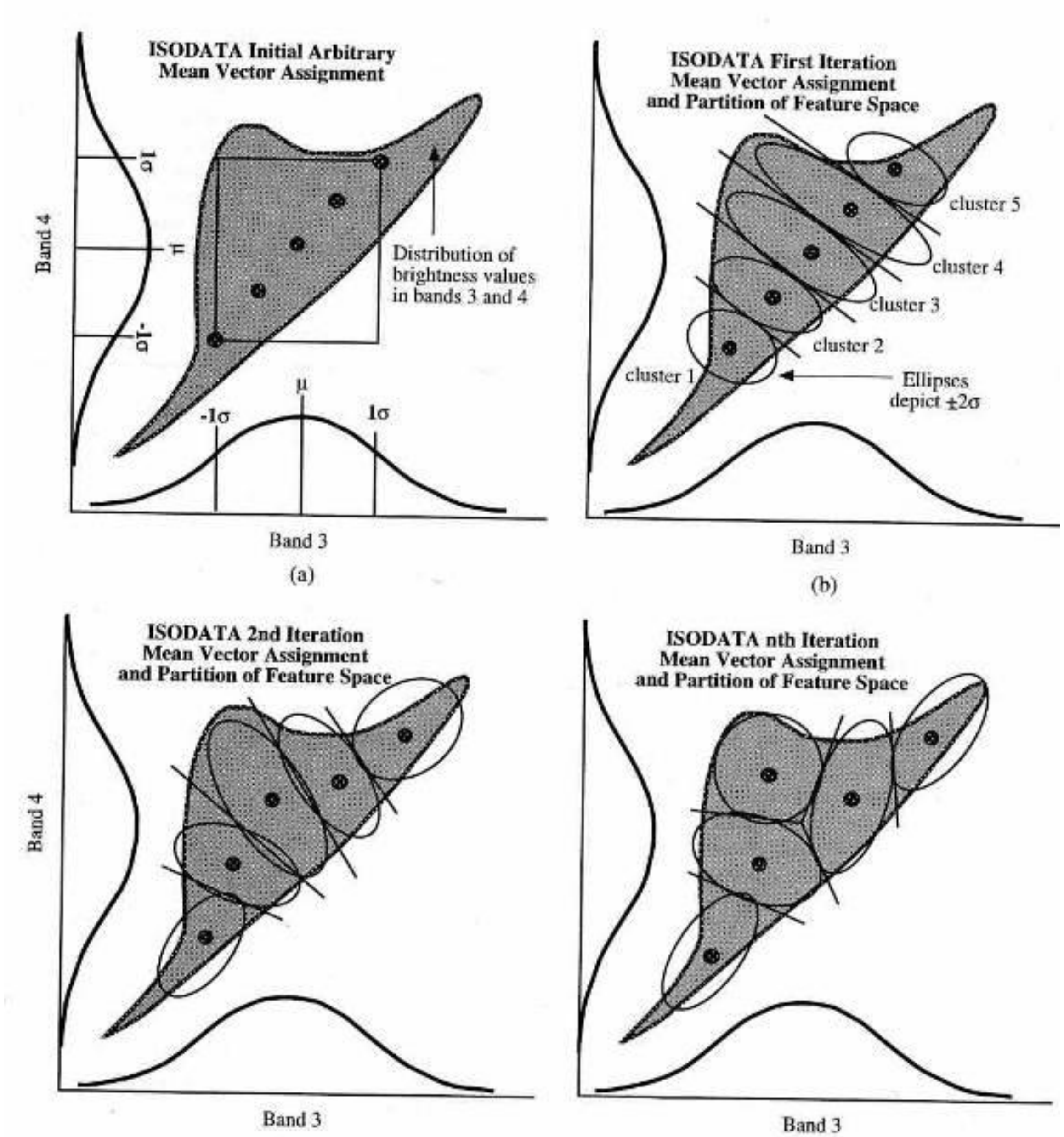
2. Fuzzy K-means enables mixed membership, based on distribution of the clusters

3. Isodata (Iterative Self-Organizing Data Analysis Technique Algorithm) can also merge or split clusters, so the number of clusters is more flexible

The ISODATA algorithm is similar to the k-means algorithm with the difference that **the ISODATA algorithm allows for different number of clusters while the k-means assumes that the number of clusters is known** *a priori*. The objective of the k-means algorithm is to minimize the within cluster variability.

Unsupervised – how it works …. YIKES! (do we need to know this?)

- ❑ Algorithm starts with statistical seed points
- ❑ Assigns each pixel to the closest seed
- ❑ Calculates group mean …. in 'n-dimensional' space
- ❑ Re-assigns pixels to the closest group mean
- ❑ Re-calculates group mean
- ❑ Iterates (10 ?) until relatively little change and fixes groupings



classification report

1 iteration

Note:

clusters with 0 pixels

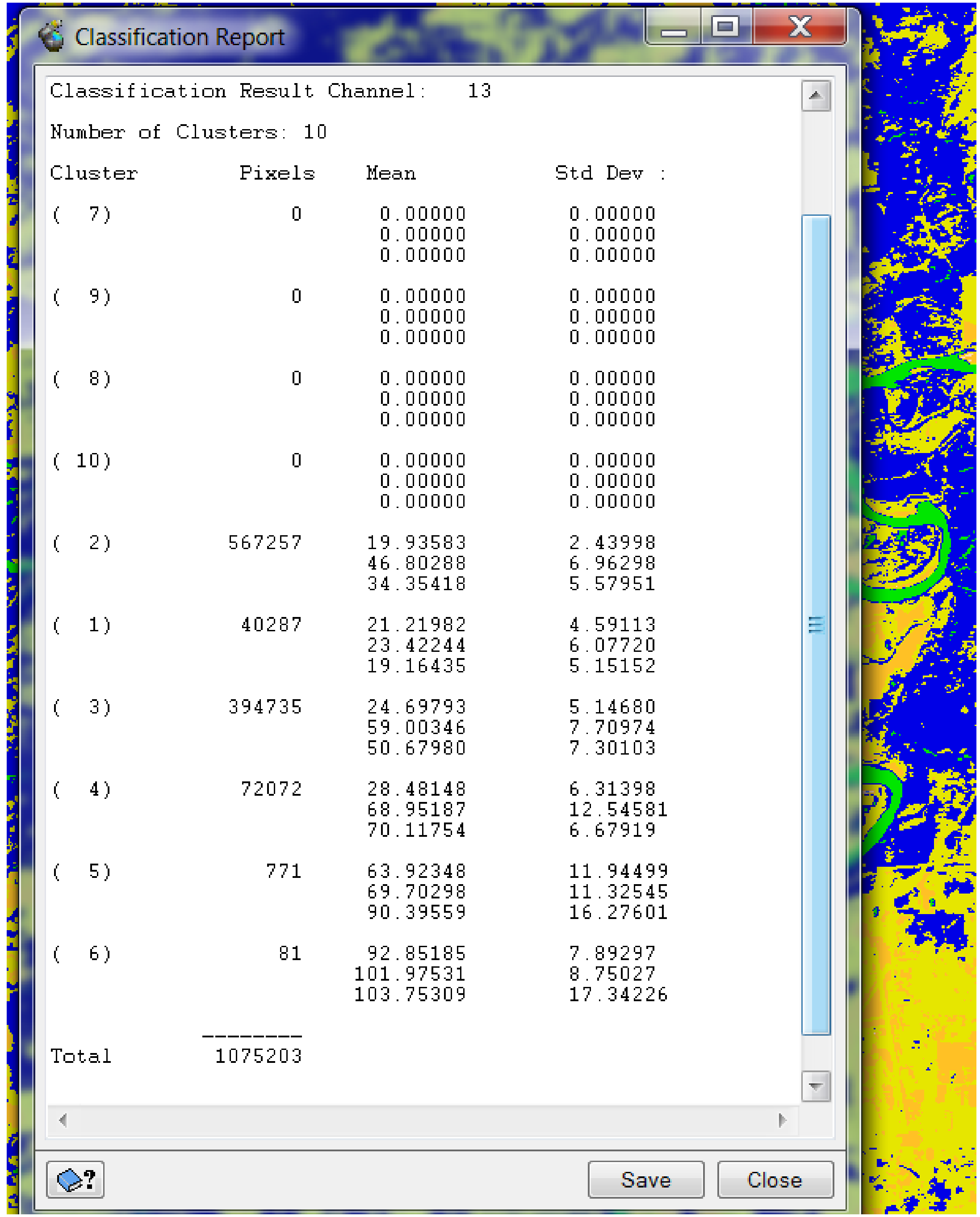
DN values for bands

3,4,5 averages

Final step ..

Assigning names to clusters

(and merge some)


Classification Report

Classification Result Channel: 13

Number of Clusters: 10

Cluster	Pixels	Mean	Std Dev :
(7)	0	0.00000 0.00000 0.00000	0.00000 0.00000 0.00000
(9)	0	0.00000 0.00000 0.00000	0.00000 0.00000 0.00000
(8)	0	0.00000 0.00000 0.00000	0.00000 0.00000 0.00000
(10)	0	0.00000 0.00000 0.00000	0.00000 0.00000 0.00000
(2)	567257	19.93583 46.80288 34.35418	2.43998 6.96298 5.57951
(1)	40287	21.21982 23.42244 19.16435	4.59113 6.07720 5.15152
(3)	394735	24.69793 59.00346 50.67980	5.14680 7.70974 7.30103
(4)	72072	28.48148 68.95187 70.11754	6.31398 12.54581 6.67919
(5)	771	63.92348 69.70298 90.39559	11.94499 11.32545 16.27601
(6)	81	92.85185 101.97531 103.75309	7.89297 8.75027 17.34226
Total	1075203		

After 16 iterations and 16 classes/clusters

Classification Report

Classification Algorithm: Fuzzy K-Means Unsupervised
Classification Input Channels: 3,4,5
Classification Result Channel: 9

Number of Clusters: 16

Cluster	Pixels	Mean	Std Dev :
(2)	296774	23.24140	8.24662
		44.77742	8.91783
		32.44915	10.04080
(3)	292356	24.48324	7.14404
		67.65602	10.67916
		49.51679	9.53926
(4)	155525	24.75149	5.03961
		107.39487	18.20386
		74.22362	13.35878
(1)	135750	42.07941	13.08230
		26.82458	8.12628
		16.47926	11.00162
(5)	151100	42.87475	9.25817
		60.36603	13.13133
		89.47187	18.67191
(7)	86198	84.79987	12.60066
		59.46275	9.37685
		20.03181	14.03484
(8)	85354	122.33620	10.20878
		85.34046	8.68392
		19.40815	16.75611
(9)	79592	151.05591	9.22842
		105.45887	8.13428
		20.59924	20.61743
(10)	60789	175.72850	8.52307
		125.70449	8.00244
		25.36989	25.79001
(11)	55539	201.23238	7.64972
		142.52280	7.16980
		17.82207	16.81002
(12)	54187	225.29511	7.08180
		159.06710	6.45671
		16.18565	13.71707
(13)	56164	247.23974	6.06449
		172.45732	4.56608
		13.11189	7.50873
(14)	113965	254.84619	0.99330
		185.52277	4.41832
		12.44097	4.33174
(15)	51887	254.93781	0.76476
		202.95095	5.59141
		14.38227	7.67388
(16)	33140	254.99879	0.08035
		225.13265	8.29124
		13.28431	3.35810

Fuzzy classification – each pixel has potential membership in more than one cluster

Merging and adding classes

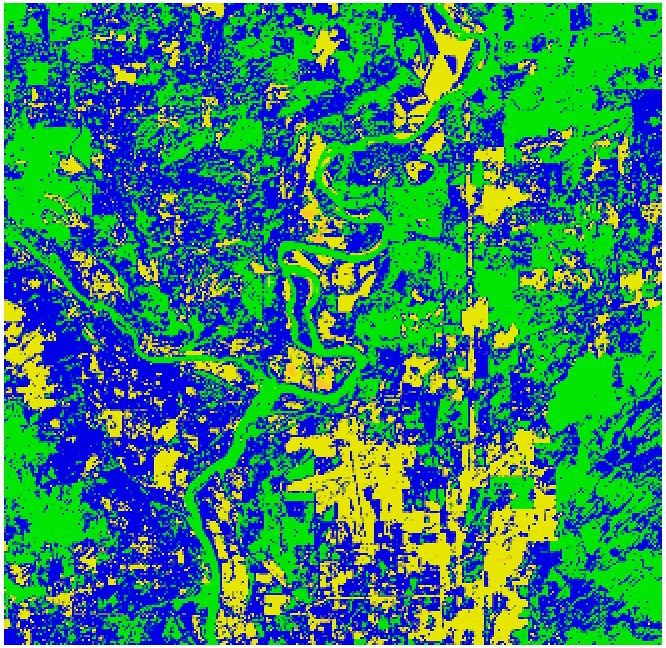
Merging – if clusters are not really separate features; Clusters are merged if they overlap spatially or are similar spectrally. *(visually examine image)*

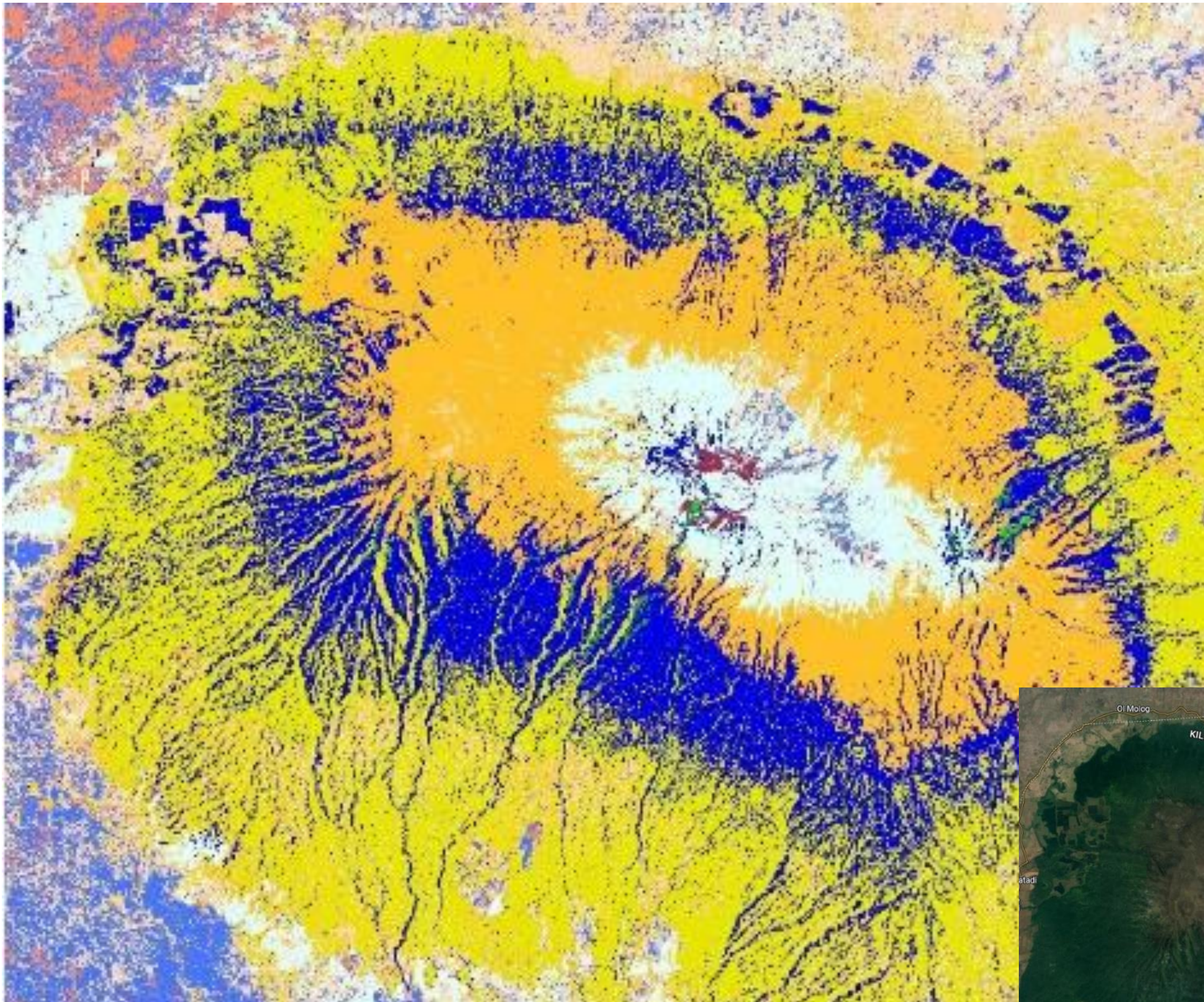
Splitting / adding

If one cluster covers too much area – run again with more clusters

Can also generate many clusters, and then group merge later …

One ploy is to make many clusters (e.g. 50-100 and plan to merge)

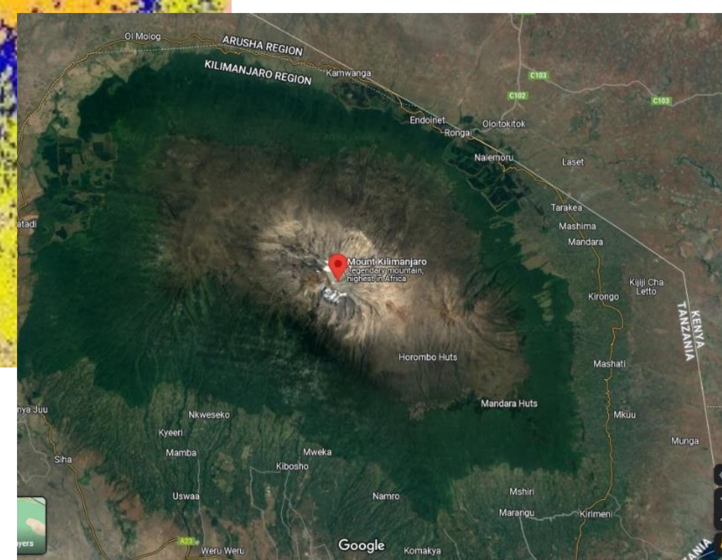




Classification **ALWAYS** produces a 'salt and pepper' effect with isolated pixels

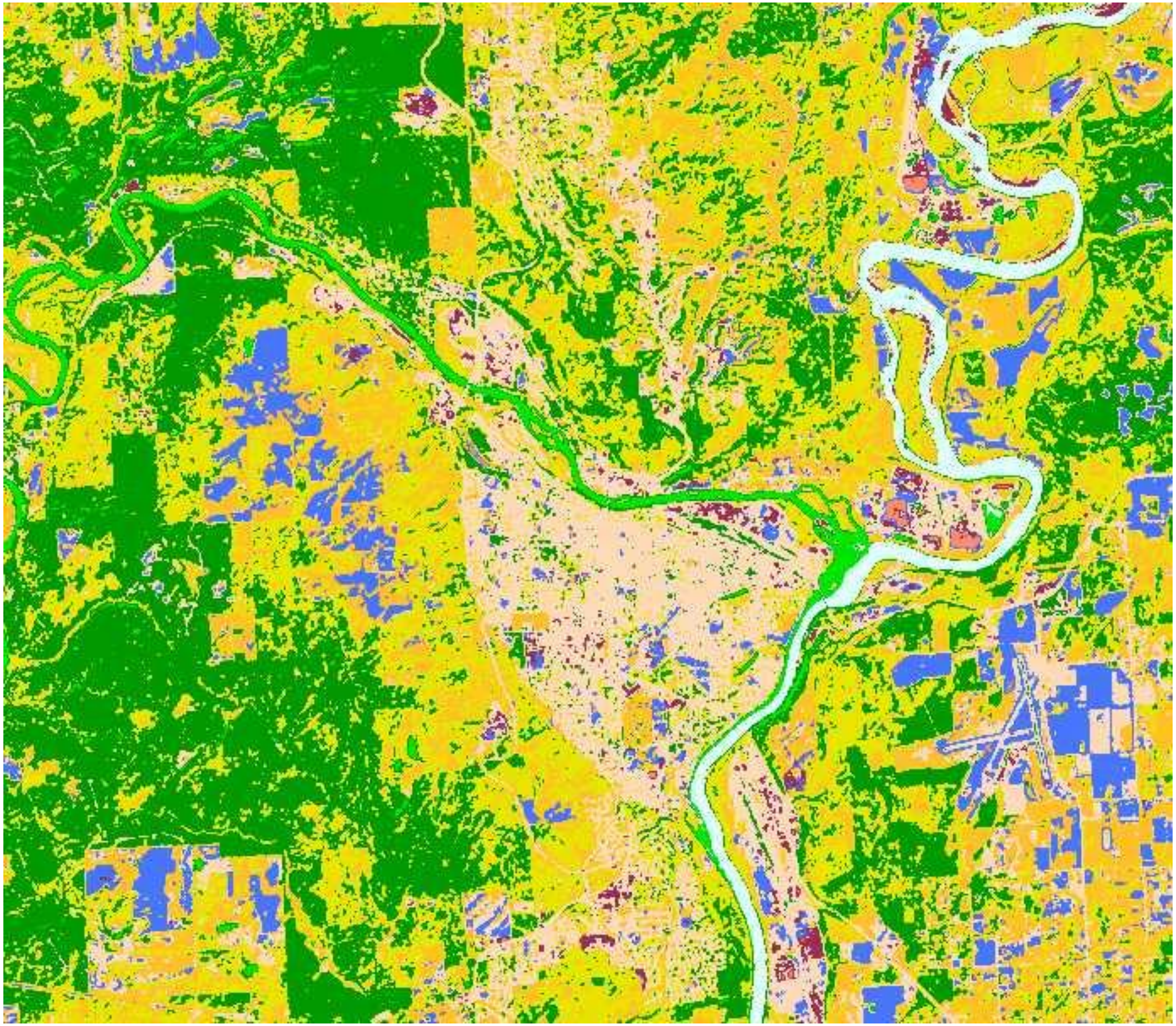
Due to :

- fine local DN variations
- **'per-pixel'** classifiers



Mt. Kilimanjaro

smooth with Filters or Sieve tool



Modal Filters

1 1 1

1 2* 2

1 1 1

replace centre pixel in 3x3 window with mode value (= 1 here)

SIEVE

Merges isolated pixels into adjacent class

Minimum cluster = ?

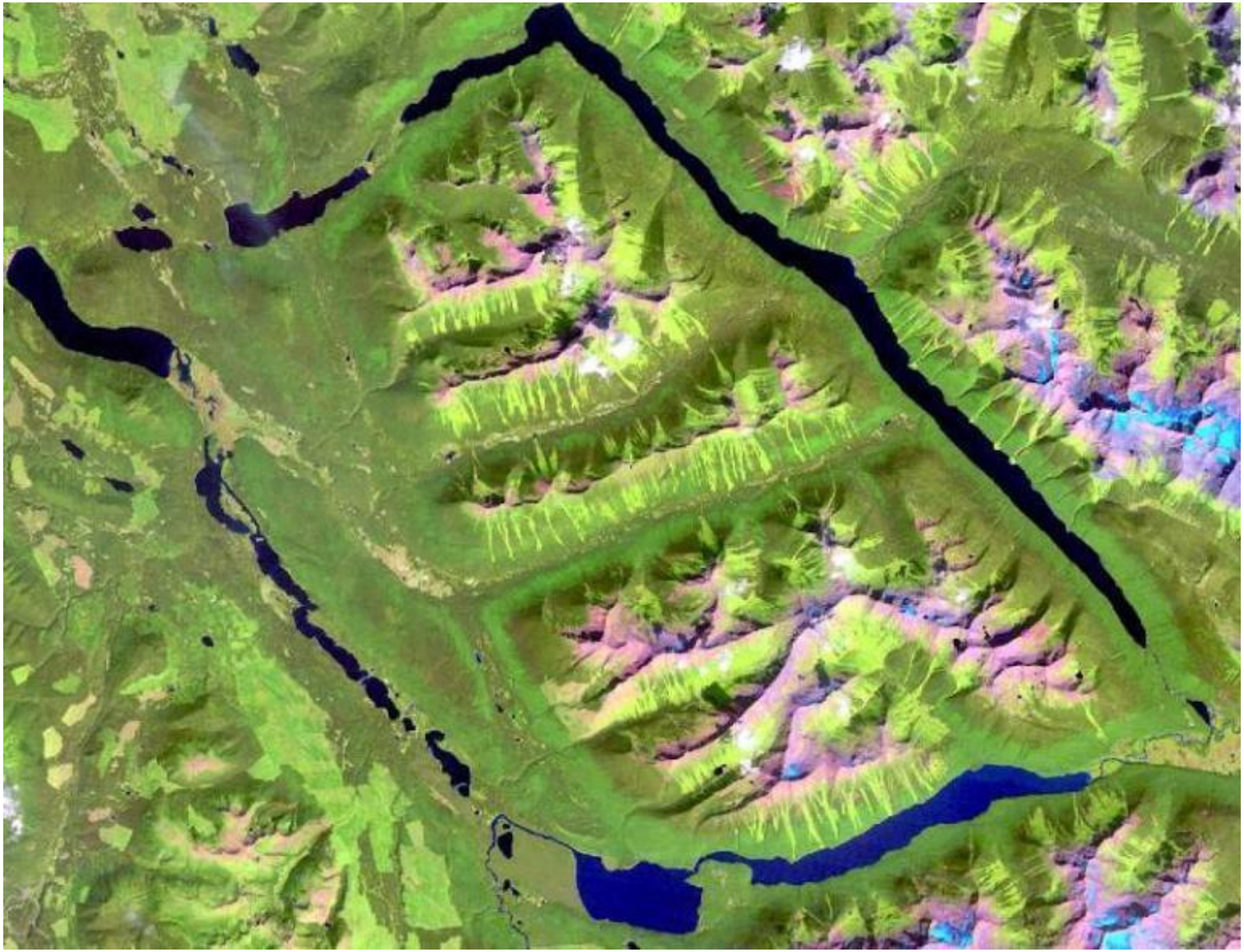
GIS polygon:

1 ha ~ 11.1 pixels

Or use 2 ha or 5 ha ?

Challenges in classification – why it doesn't always beat digitising

There are many spatial variations in reflectance (a range of DNs for a feature)
e.g. stand purity, understory, age/maturity, density, disease, sun angle, **topography**



Classes/clusters: water, bare rock, glaciers, deciduous, coniferous, shadow?, cutblocks, planted..

URBAN / HUMAN – mosaic of smaller features inside a 30 metre pixel

- amount of grass, types of material, roofing colour, weathering, sun angle (building shape)

Overall summary on classification

It is always complex – the classes and contrasts
There are many causes of spatial variations in reflectance
Most (natural) features are continuous, not discrete

Using only DNs (per-pixel classifiers):
Any land cover types have a range of values
Conversely, different cover types can appear similar

Further complications for all images:

a. moisture (recent events) – especially in SWIR/thermal
b. edge (mixed) pixels
c. sun angle (illumination) – usually mid-morning

Textbook classification goal: ~ 85% accuracy
But even manual digitizing may not do any better