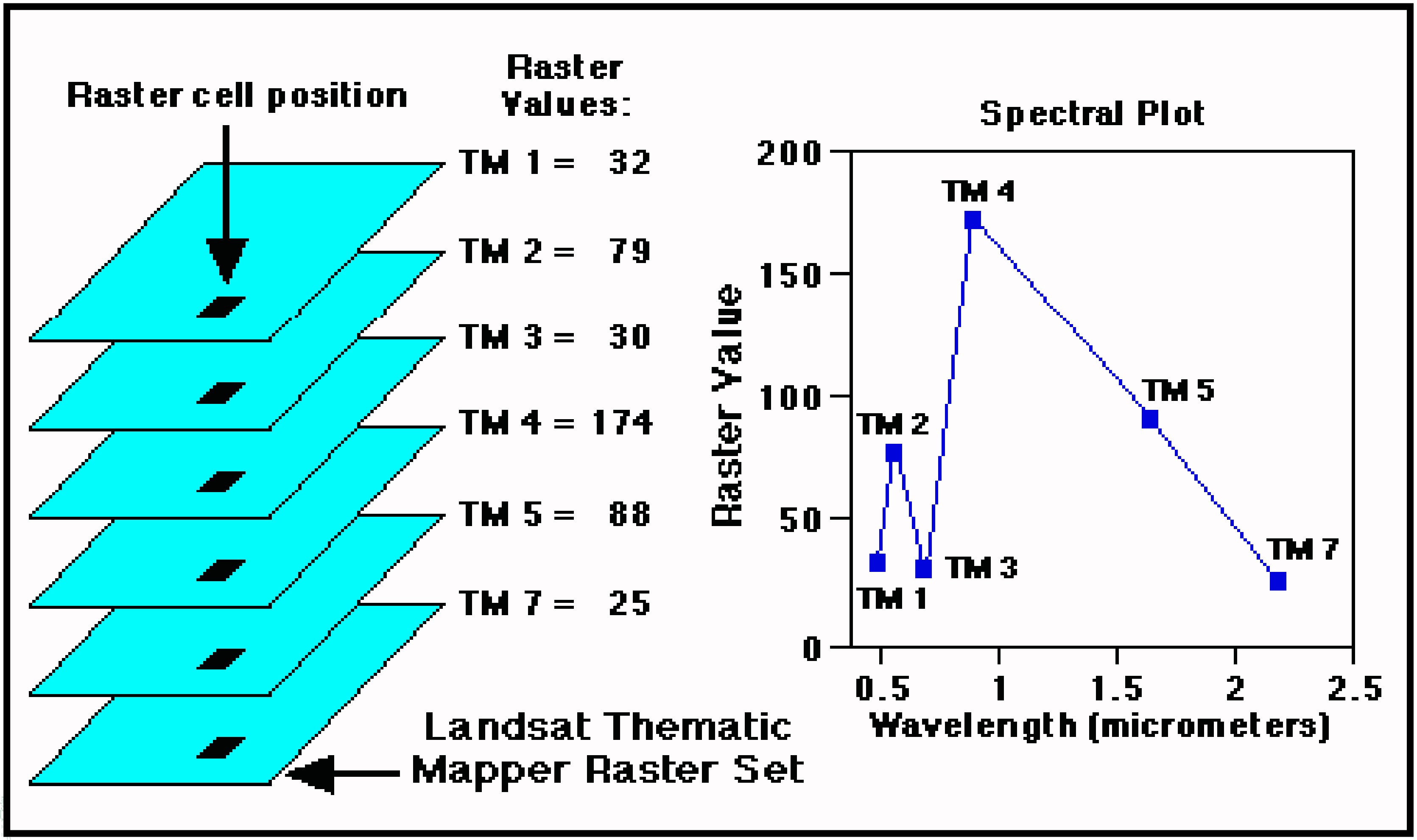
Digital Image Classification

- ❑ Uses the spectral information represented by the digital numbers in one or more spectral bands.

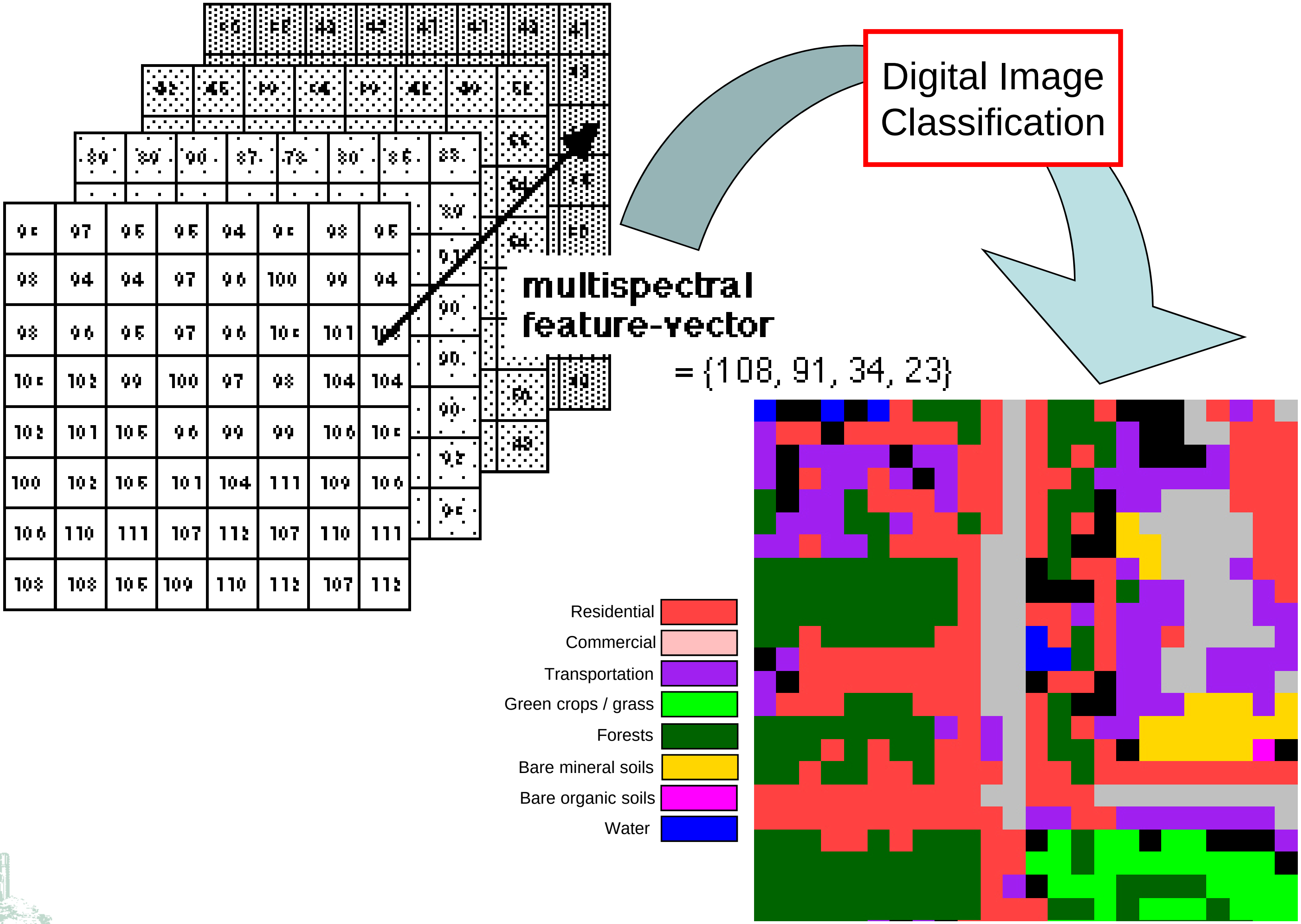
- ❑ Classifies each individual pixel based on this spectral information. Also known as spectral pattern recognition.

  - ➢ ***The objective***: To assign all pixels in the image to particular classes or themes (*e.g.*, water, coniferous forest, deciduous forest, crops, bare soil, etc.)

Digital Image Classification



Digital Image Classification



Digital Image Classification

- The two approaches:
  - **Per-pixel classification**
    - The algorithm categorizes **each input pixel** into a spectral feature class based solely on its ***individual multispectral vector*** (signature). *No* context or neighborhood evaluation is involved.
  - **Object-oriented classification**
    - The input pixels are **grouped into spectral features (objects**) using an *image segmentation algorithm*. These objects are characterized in both the raster and vector domains. The objects are classified using both spectral and spatial cues (metrics).

Digital Image Classification

- The three common **per-pixel** methods are:
  - **Supervised classification**
  - **Unsupervised classification**
  - Rule-based classification

Digital Image Classification

- **Supervised classification**
  - The analyst "supervises" the categorization of a set of specific classes by providing ***training statistics*** that identify each category.
- **Unsupervised classification**
  - The raw spectral data are **grouped first**, based solely on the statistical structure of the data. Then the **analyst must label** each statistical cluster, placing them into the appropriate categories (if possible).
- Rule-based classification
  - Spectrally categorized pixels are classified using ancillary data in a GIS model.

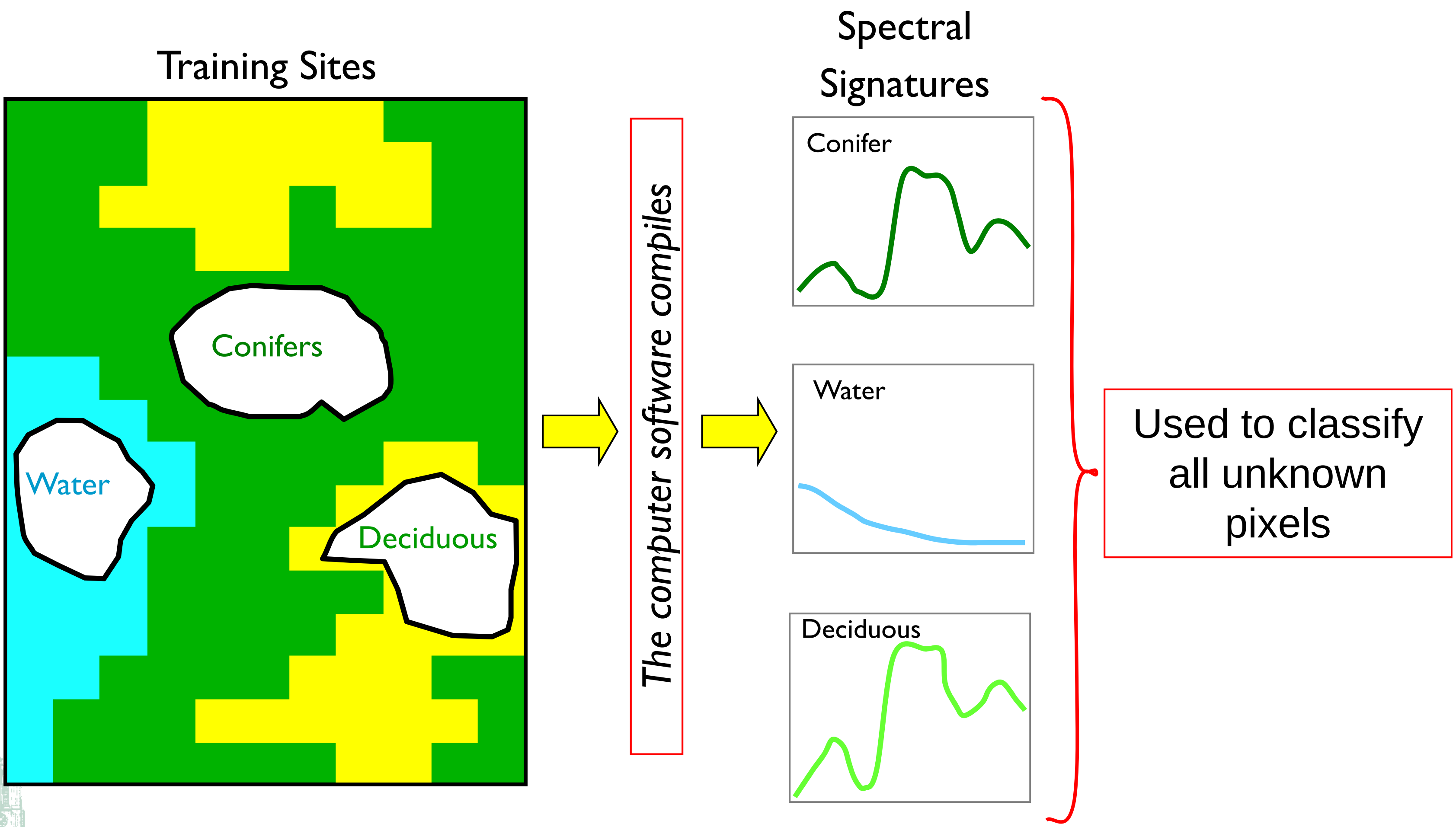
Digital Image Classification

Supervised classification

- In the imagery, the analyst identifies homogeneous, representative examples of the various surface cover types (information classes) of interest.
- These samples are referred to as **training areas**.
- The selection of appropriate training areas is based on the analyst's familiarity with the geographical area and their knowledge of the actual surface cover types present in the image.
- **Good visual interpretation skills are mandatory for success**.

Digital Image Classification

❑ Supervised classification



Digital Image Classification

❑ Supervised classifiers (algorithms)

- Parallelpiped – based on range or variance of class DNs
- Minimum Distance-to-Means – based on mean class DNs
- **Maximum Likelihood** – based on probability of class membership
- **Spectral Angle Mapper** – class membership based on minimum difference from the *n*-dimensional spectral vectors of the classes

Digital Image Classification

❑ Supervised classification process

➢ Step 1: Define training data

✓ Each training site should appear homogenous and representative of the legend class

✓ Delineate several training sites for each legend class

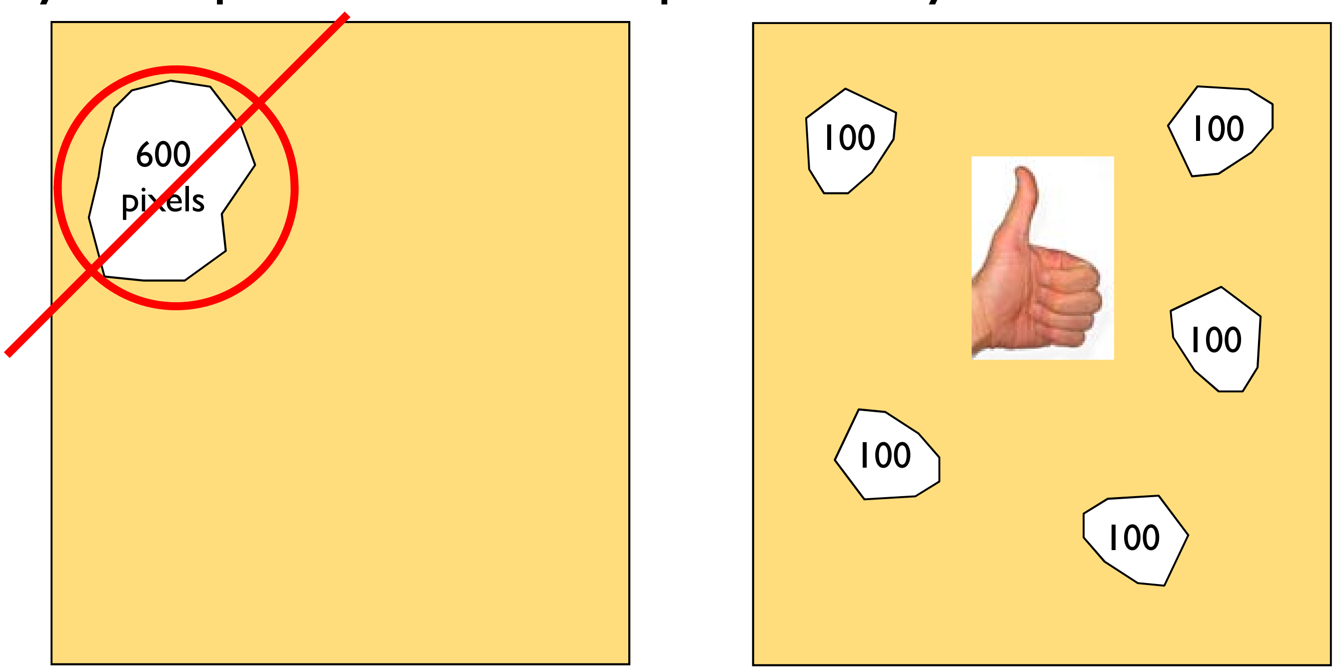
✓ Make each training site at least 20-25 pixels (*i.e.*, > 5 acres for 30 meter pixels)

✓ Each class should be represented by ~ 100 *n* pixels (*where n* = number of spectral bands in the data set)

Digital Image Classification

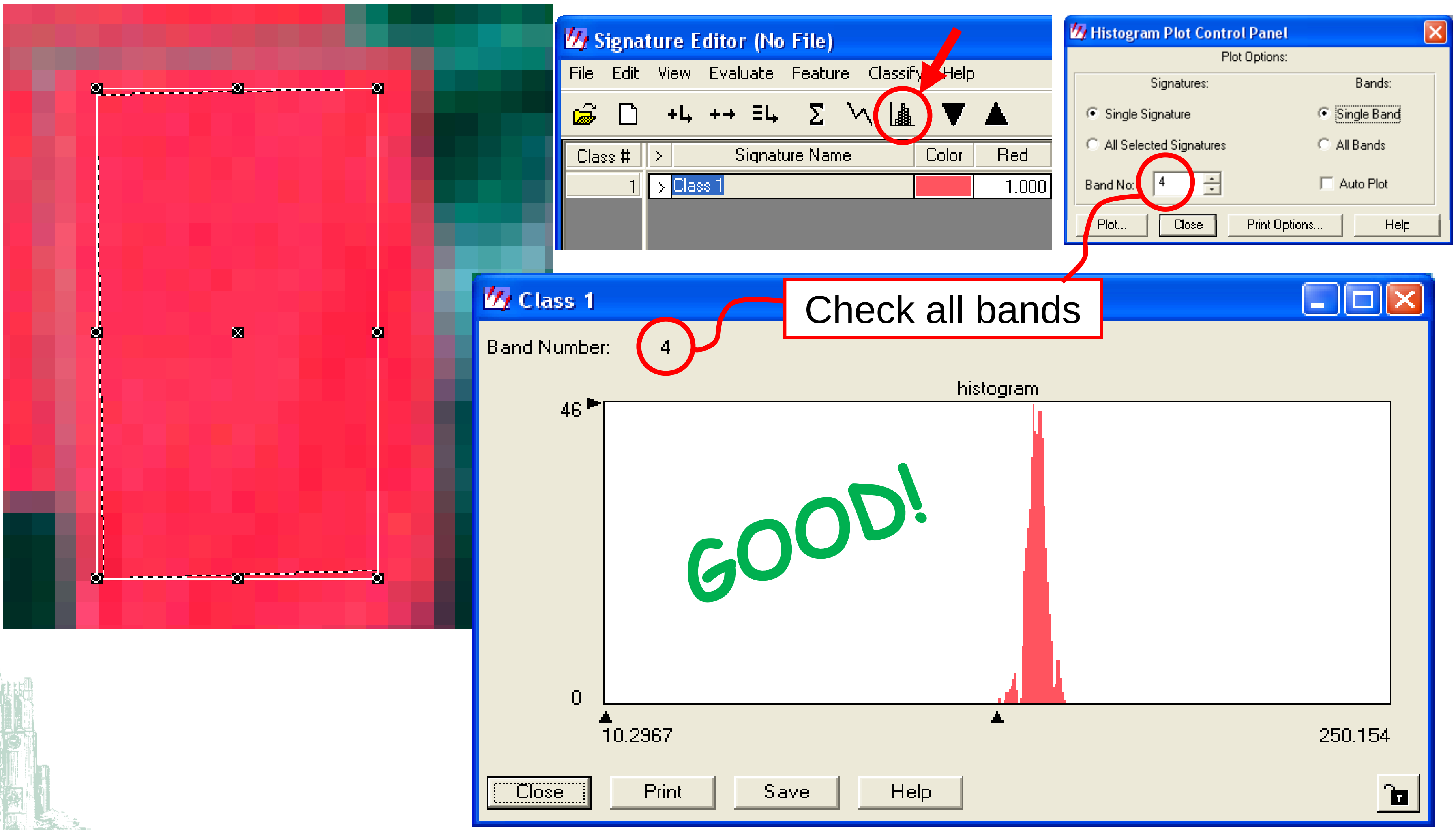
❑ Supervised classification process

- ➢ Training Sites
  - ✓ For 6-band TM & ETM imagery, the total number of training pixels per class should be at least 600
  - ✓ Try to capture the landscape diversity of the class



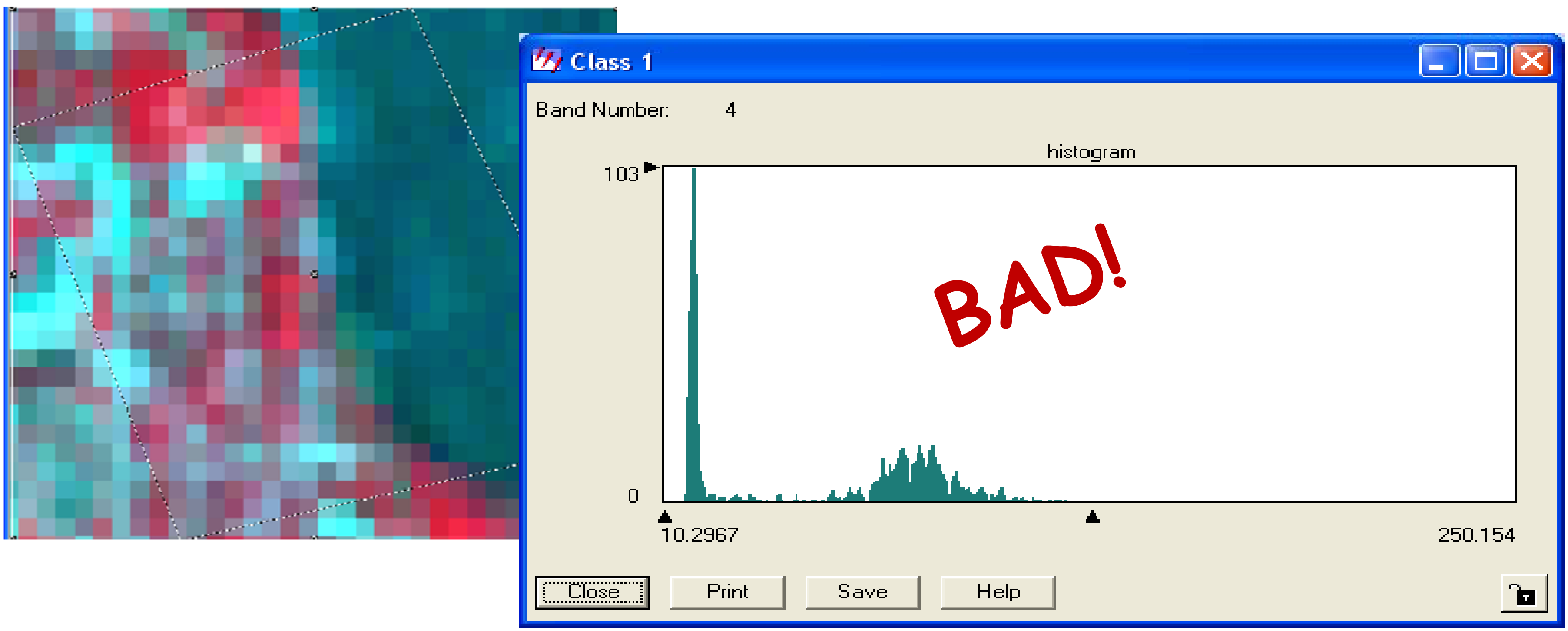
Digital Image Classification

- ❑ Example of a homogeneous training site



Digital Image Classification

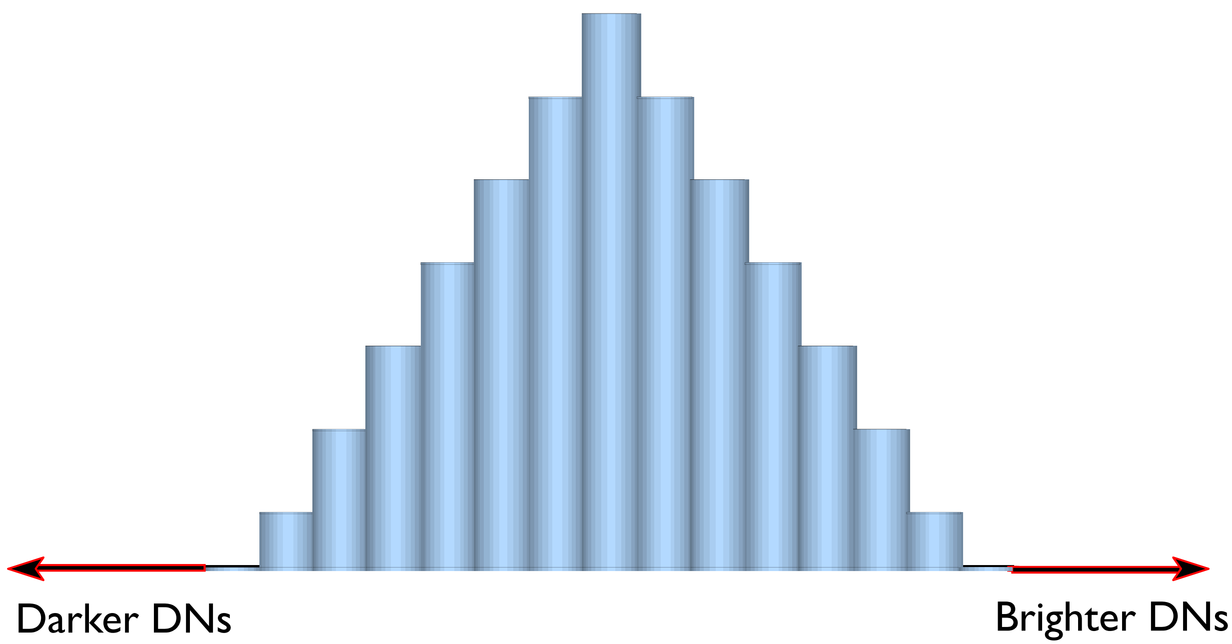
- Example of a heterogeneous training site



Digital Image Classification

❑ Supervised classification process

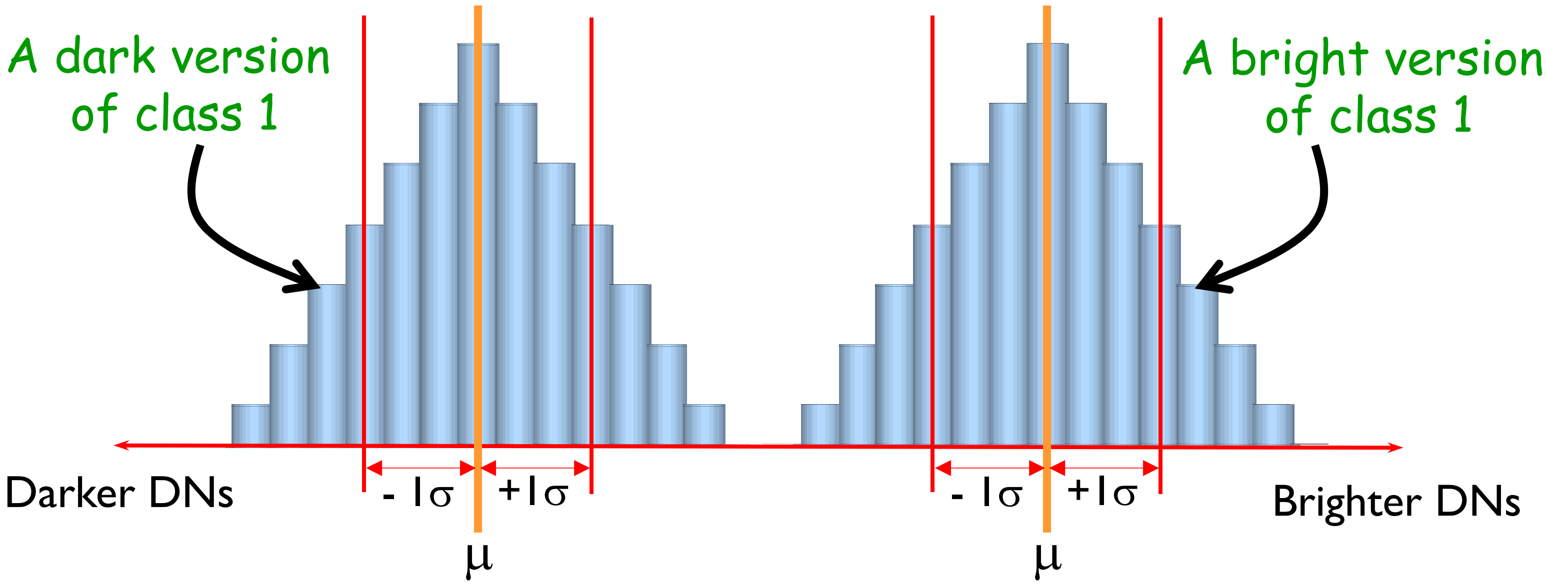
- Training data
  - ✓ Should be ***normally distributed*** ***if*** the maximum likelihood classifier is going to be used



Digital Image Classification

❑ Supervised Classification Process

- Collect several training sites for each class in order to capture the landscape diversity of the class



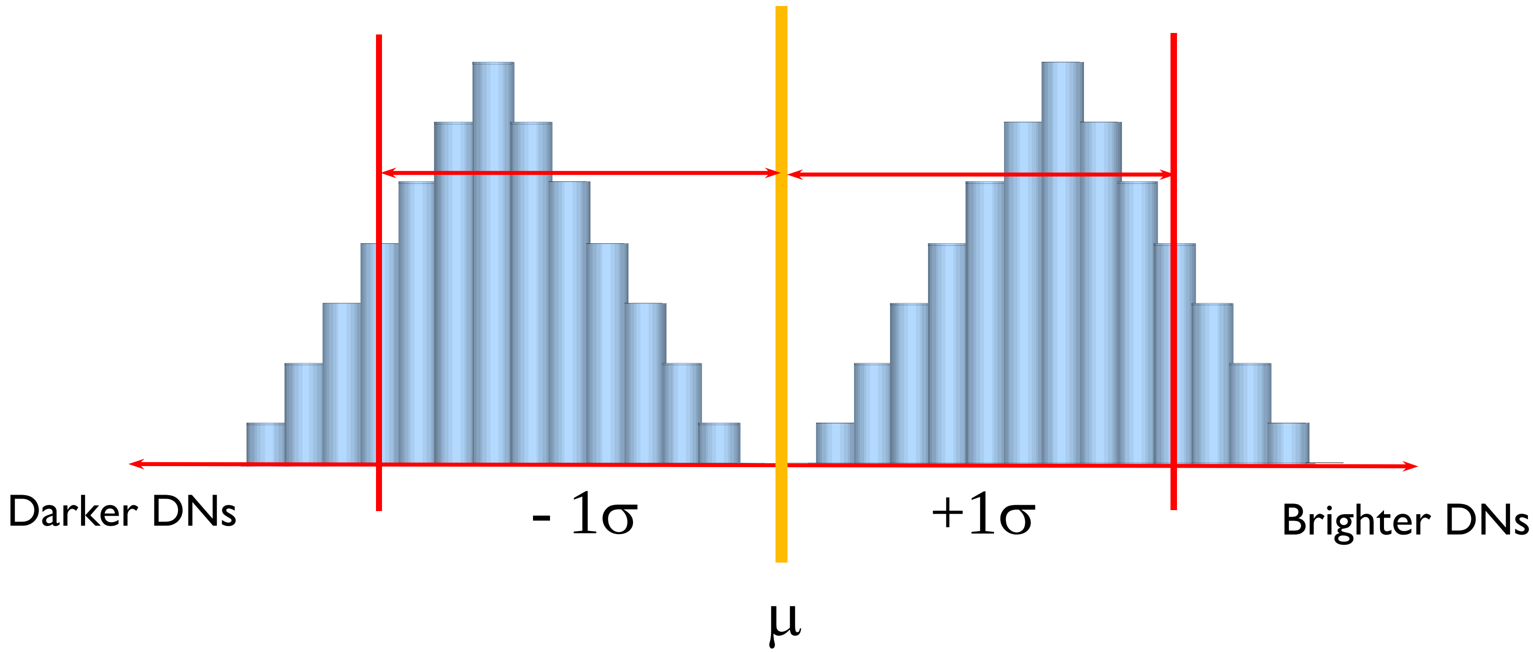
The advantage of individual signatures:

separate mean DN values and **small variances**

Digital Image Classification

❑ Supervised Classification Process

- ➢ For variable classes ***do NOT*** merge signatures

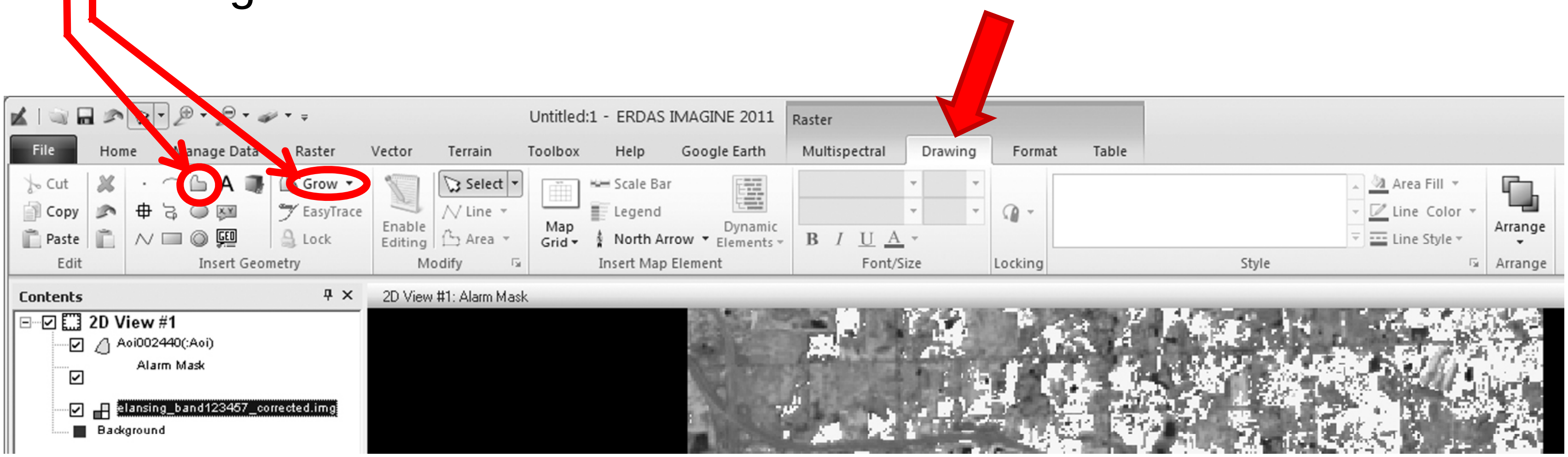


The disadvantage of merged signatures:

shifted mean DN value (**atypical of the class**) and a **large variance**

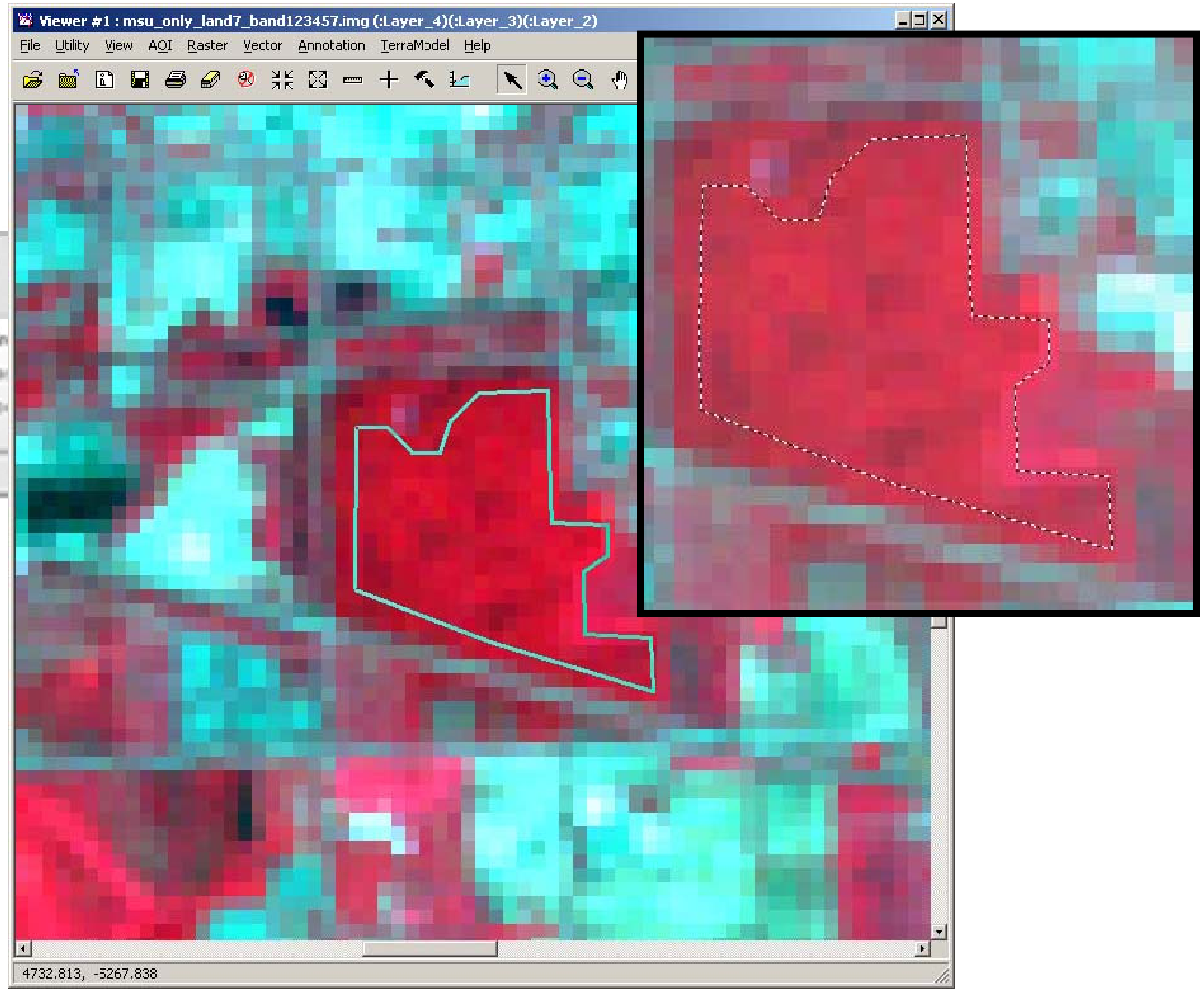
Supervised Classification

- **Training Sites** – Tools in ERDAS Imagine
  - **Drawing tab: Insert Geometry Tools**
    - ✓ Polygon tool
    - ✓ Region Grow tool



Supervised Classification

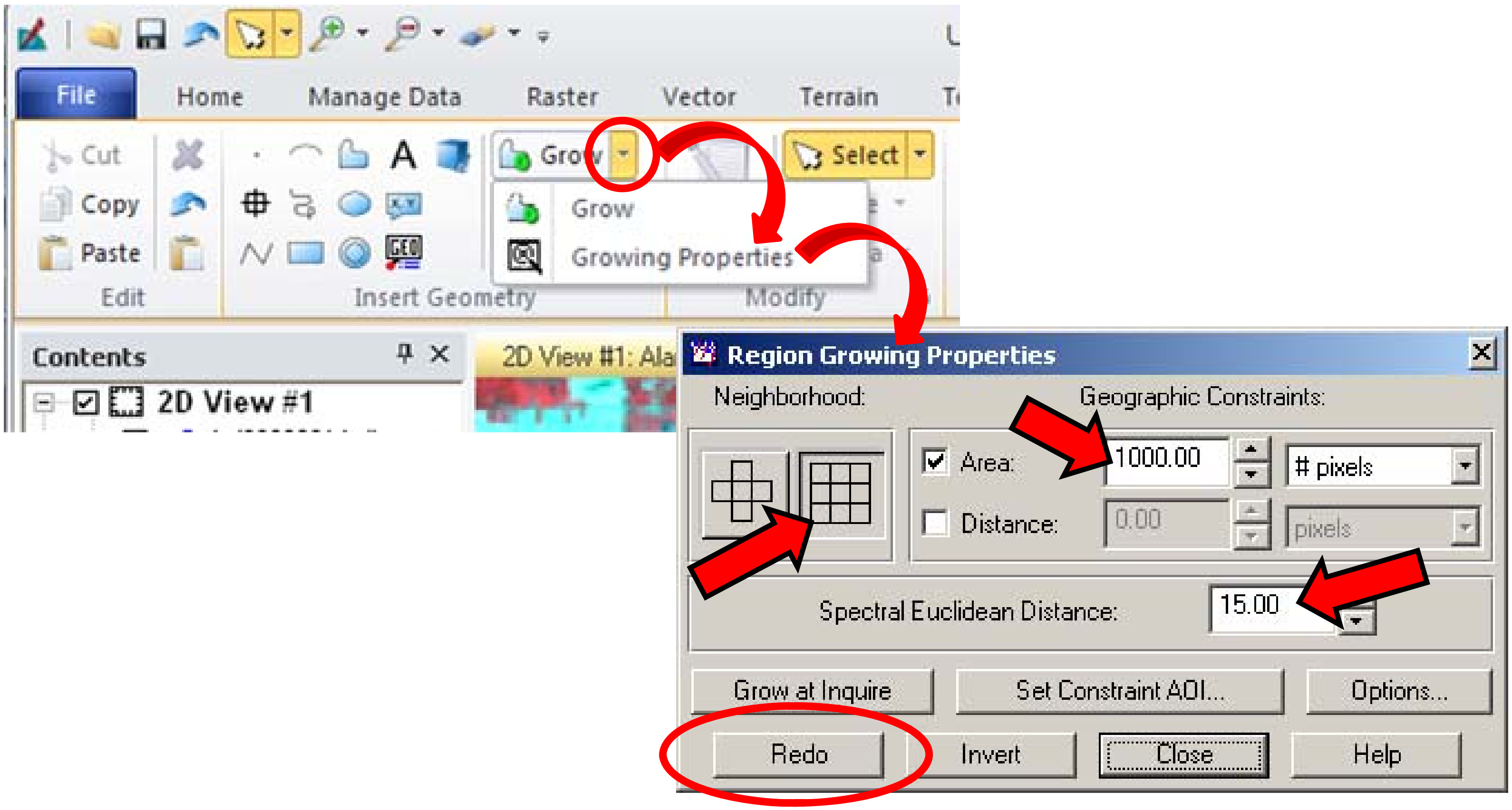
Polygon tool



Supervised Classification

- ❑ **Drawing Tools → Region Grow tool**

  Step 1: Set Seed Properties:



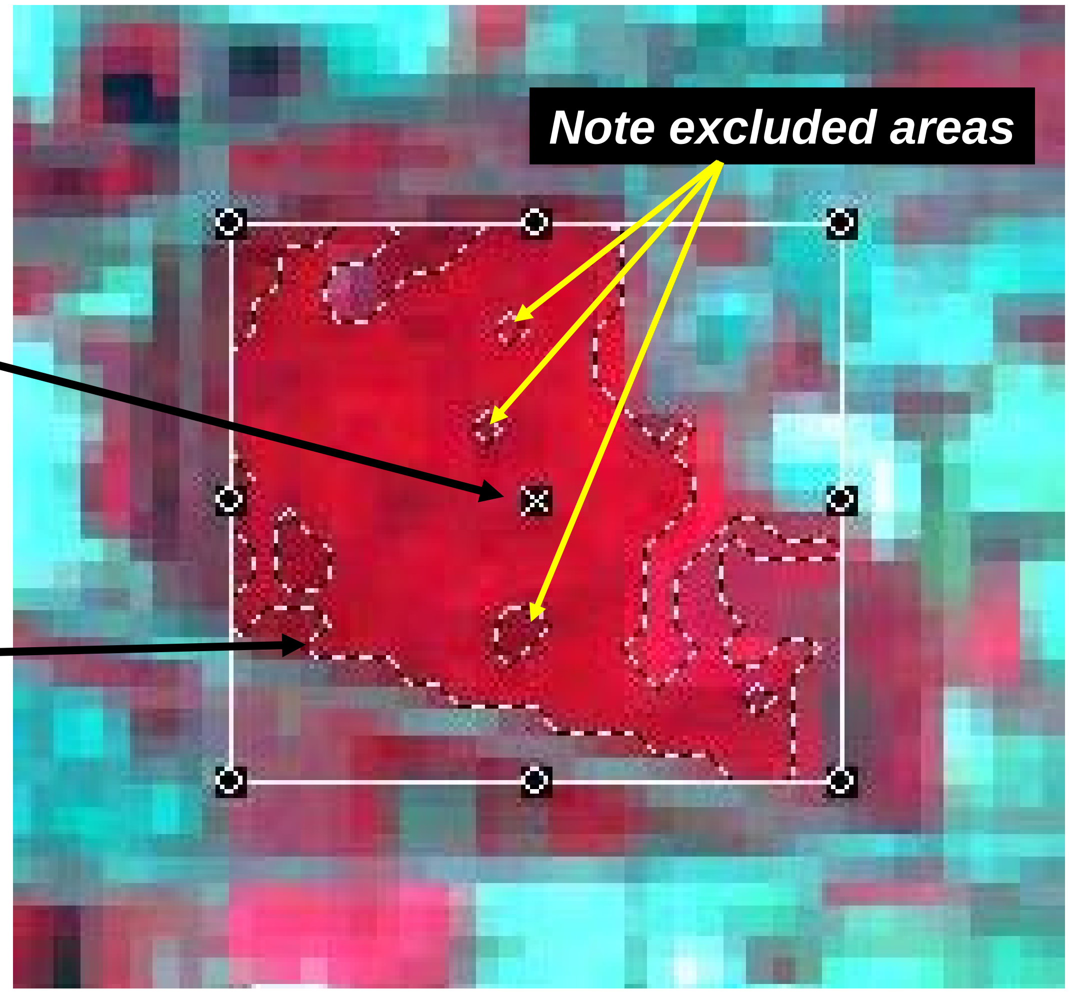
Supervised Classification

❑ Drawing Tools → Region Grow tool

Step 2: Select a seed pixel with the Region Grow Tool:

Selected seed pixel

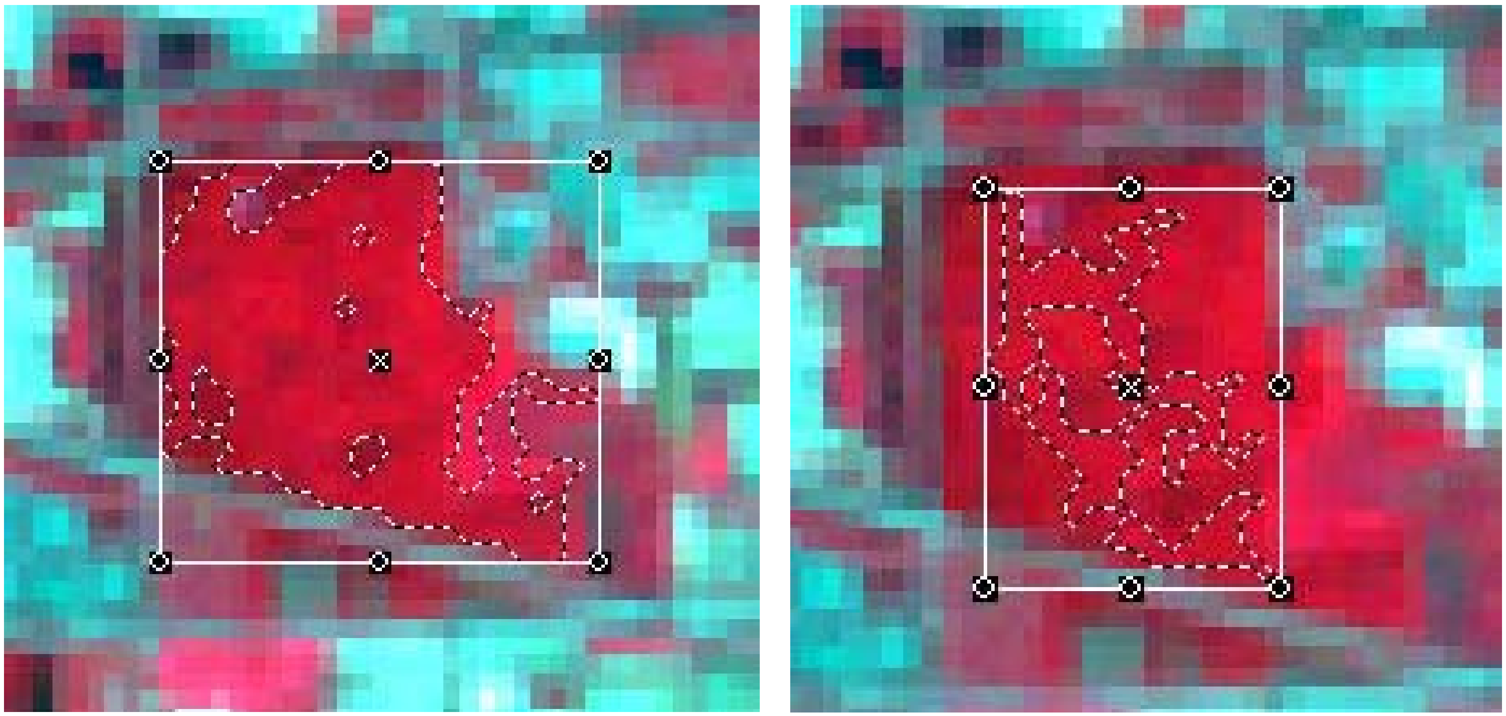
Grown region



Supervised Classification

- **Drawing Tools → Region Grow tool**

  Step 3: Evaluate Spectral Euclidean Distance:



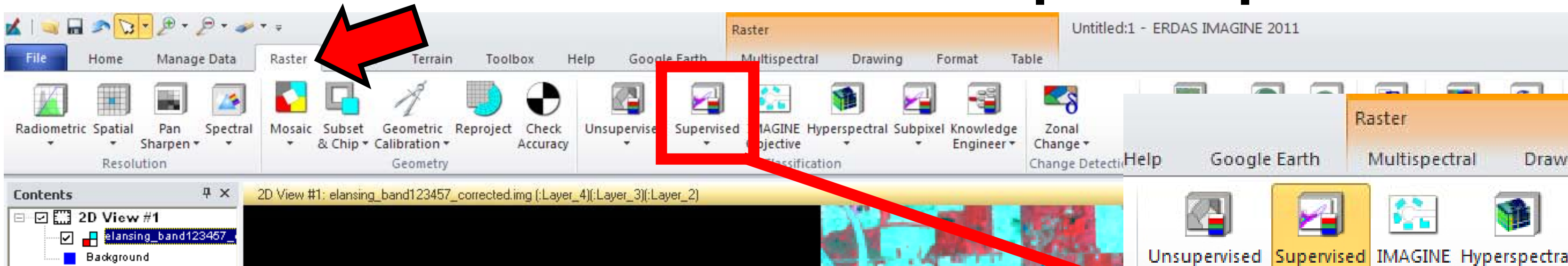
Spectral Euclidean Distance = 15 Spectral Euclidean Distance = 5

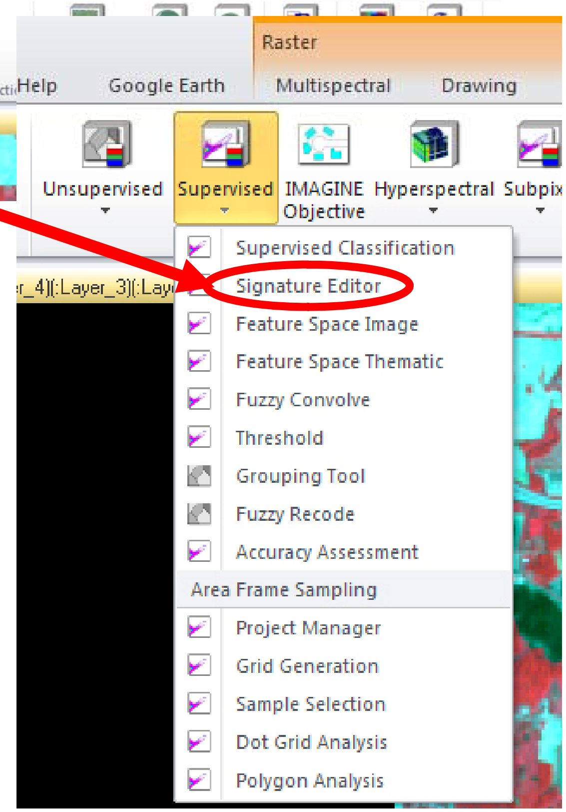
using 8-bit imagery

With 16-bit imagery, SED ~ 900 to 2000 are appropriate

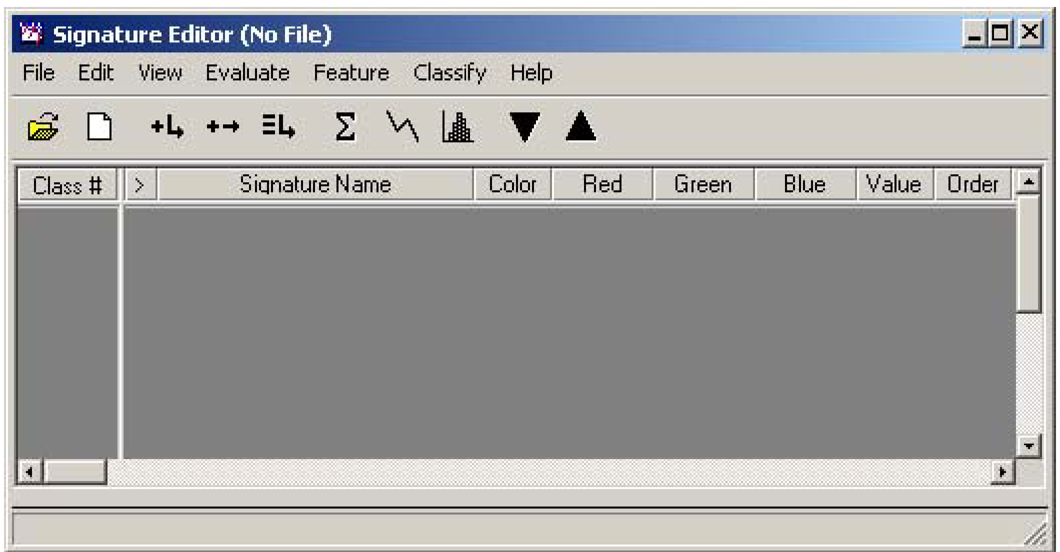
Supervised Classification

- **Training Sites** – Tools in ERDAS Imagine
  - **Raster tab → Classification Group → Supervised**



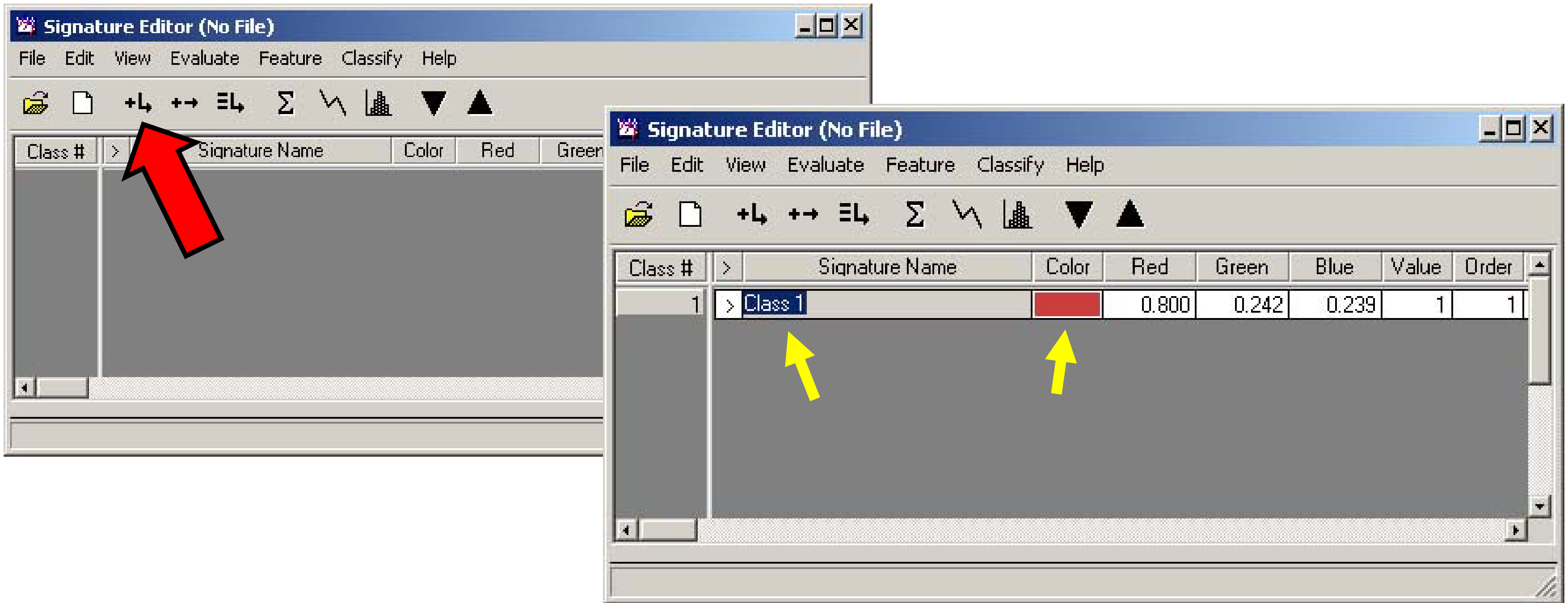


- **Signature Editor**



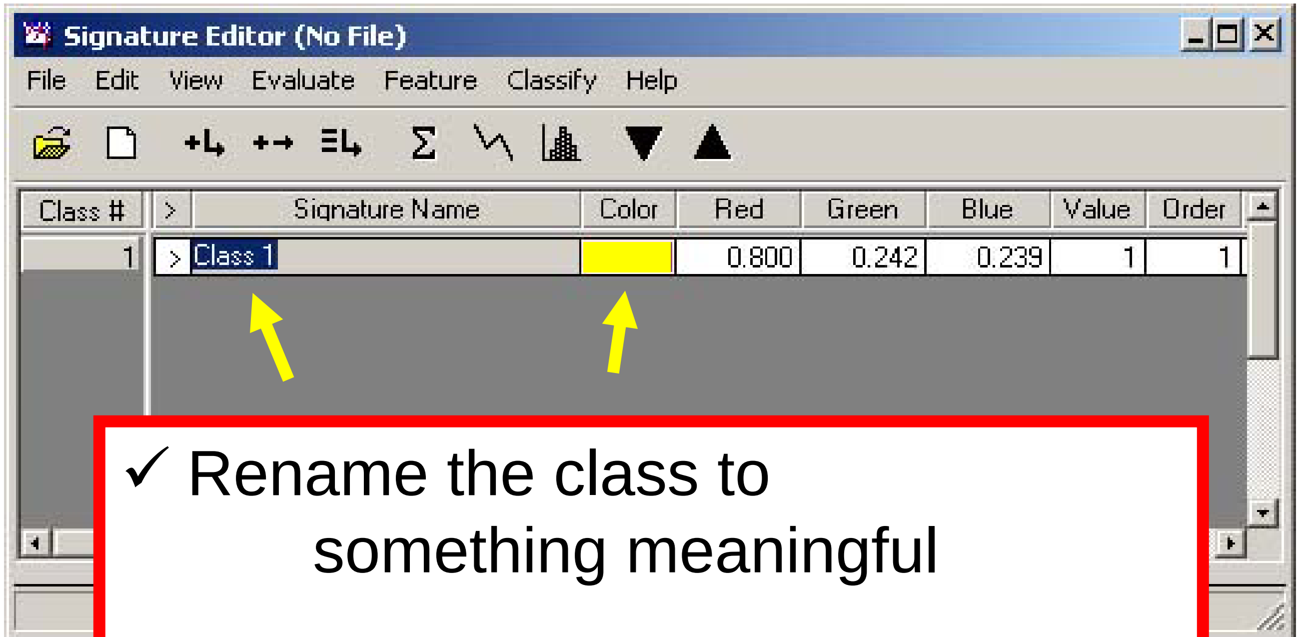
Supervised Classification

- **Training Sites** – Tools in ERDAS Imagine
  - Signature Editor → Create New Signature(s) from AOI(s)



Supervised Classification

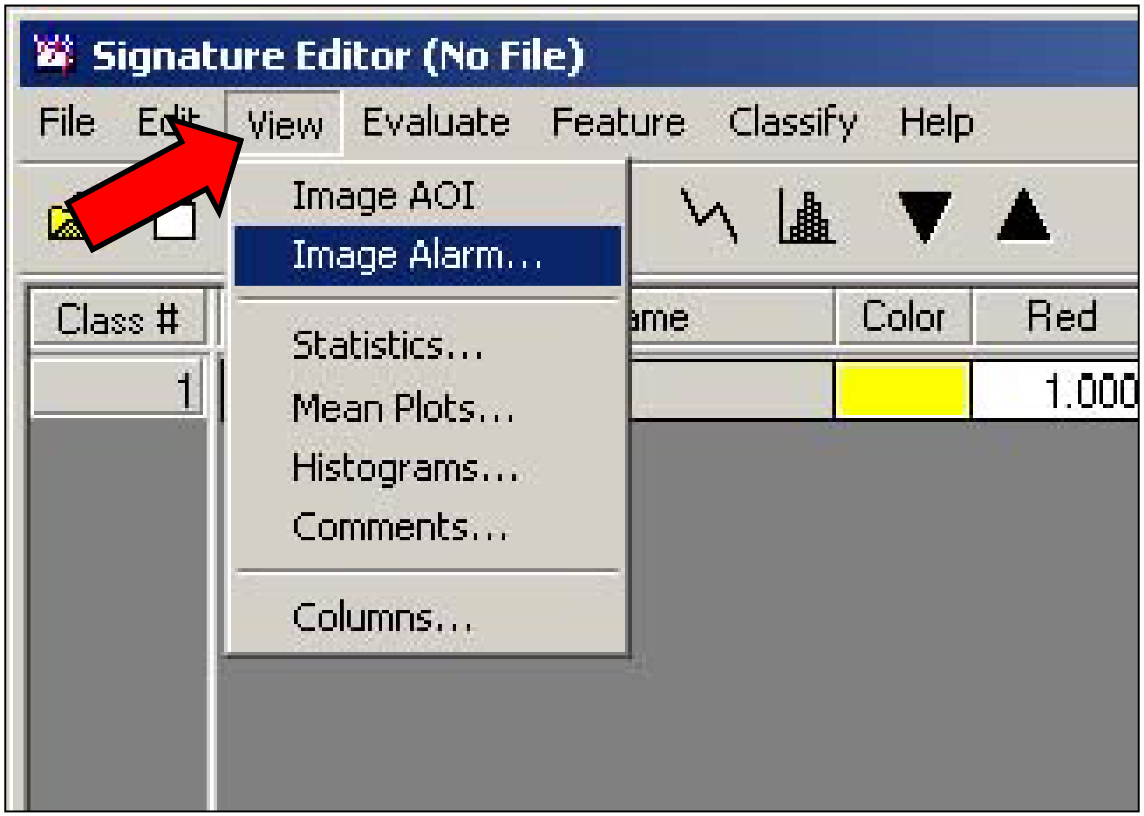
- **Training Sites** – Tools in ERDAS Imagine
  - Signature Editor → **Evaluate each new signature**



- ✓ Rename the class to something meaningful
- ✓ Change color to **Yellow** for Image Alarm evaluation

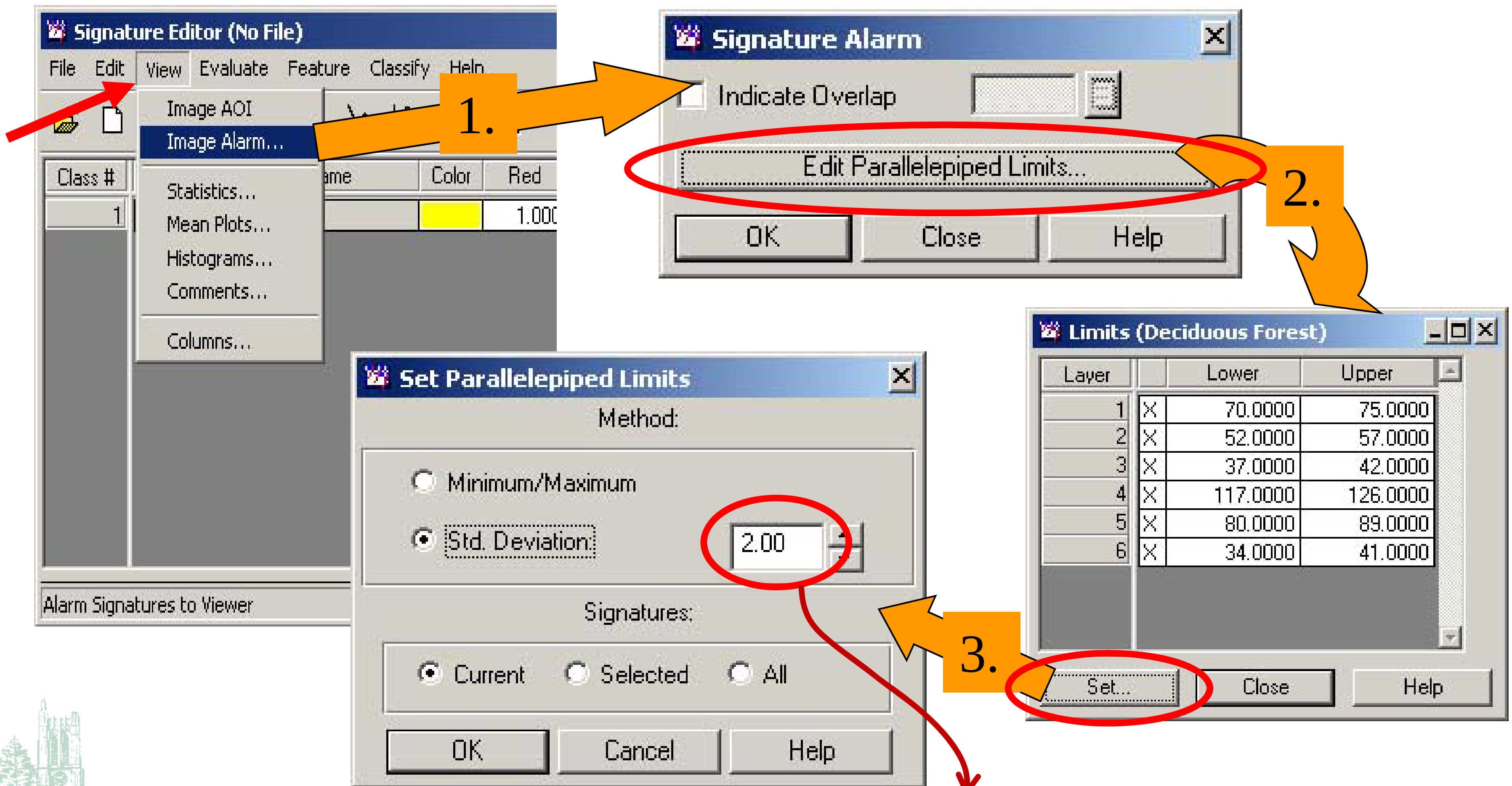
Supervised Classification

- **Training Sites** – Tools in ERDAS Imagine
  - **Signature Editor → View → Image Alarm**



Supervised Classification

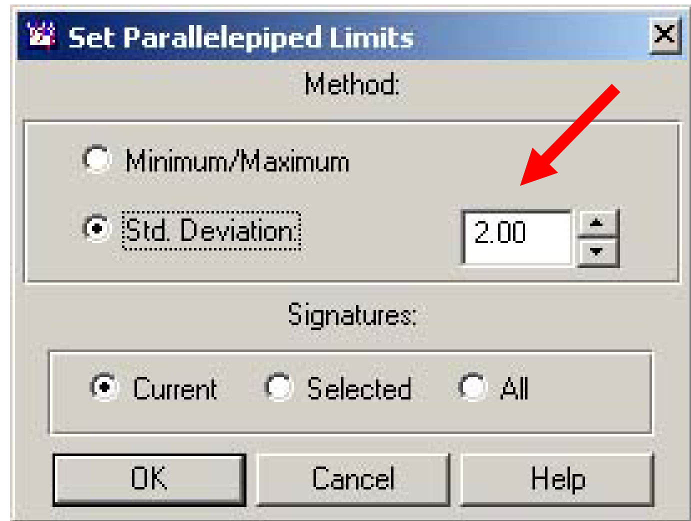
❑ Signature Editor → View → **Image Alarm**



2 SD eliminates the darkest and brightest 2.3% of the data

Supervised Classification

- Signature Editor → View → Image Alarm
  - Set Std. Deviation between:
    - ✓ ***2.00*** (95.4% of the data)
    - ✓ ***3.00*** (99.7% of the data)
      - 1.00 SD = 68.3% of the data

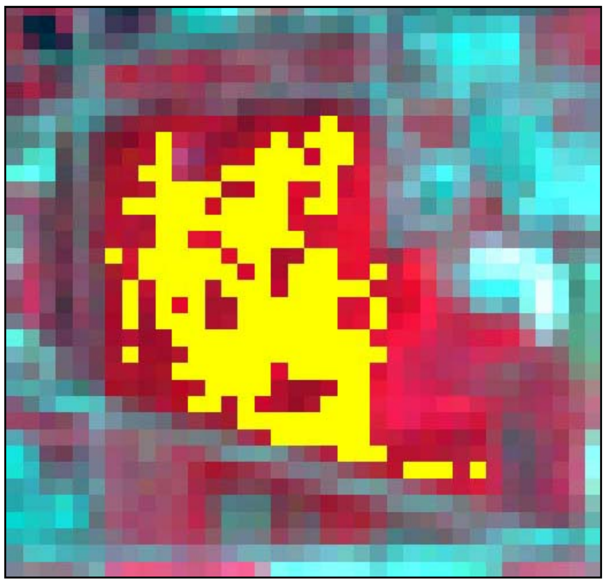


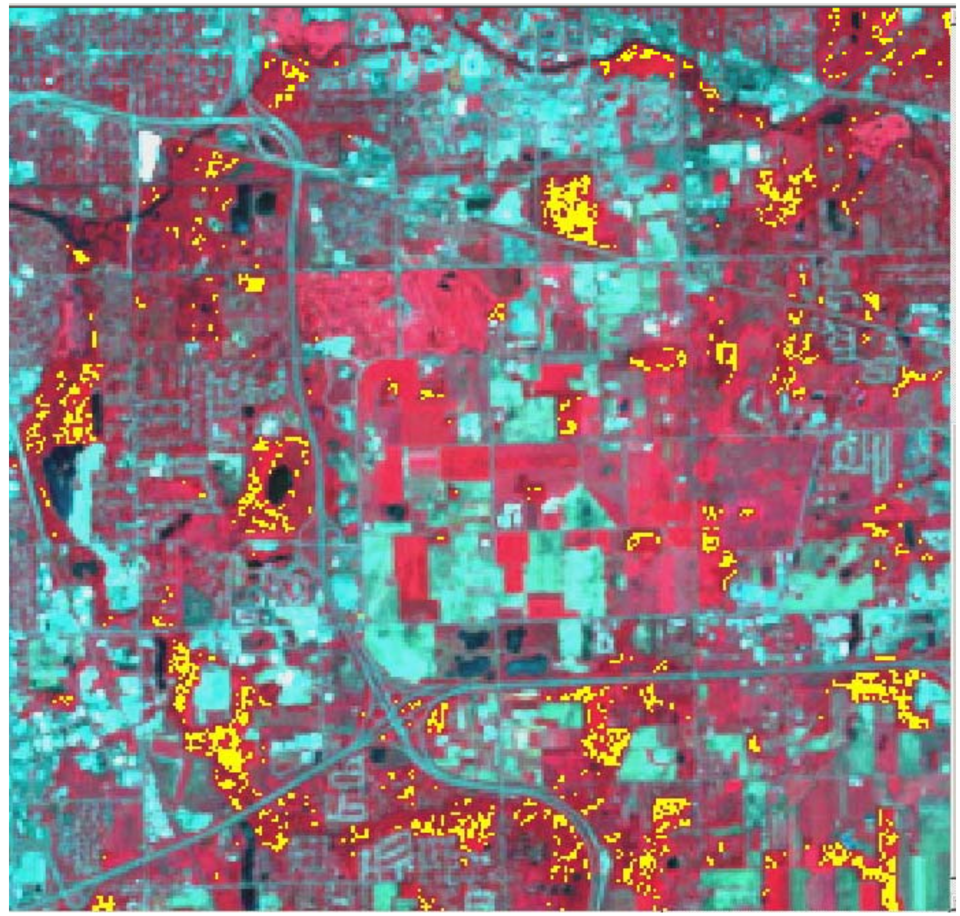
Rationale: Eliminate the most atypical pixel values from consideration.

***Image Alarm** uses the simple Parallelepiped classifier in which each class is defined as an n-dimensional volume whose boundaries in brightness space are set by these limits.*

Supervised Classification

- Visually evaluate Image Alarm for:
  - ***Errors of Omission***
  - ***Errors of Commission***

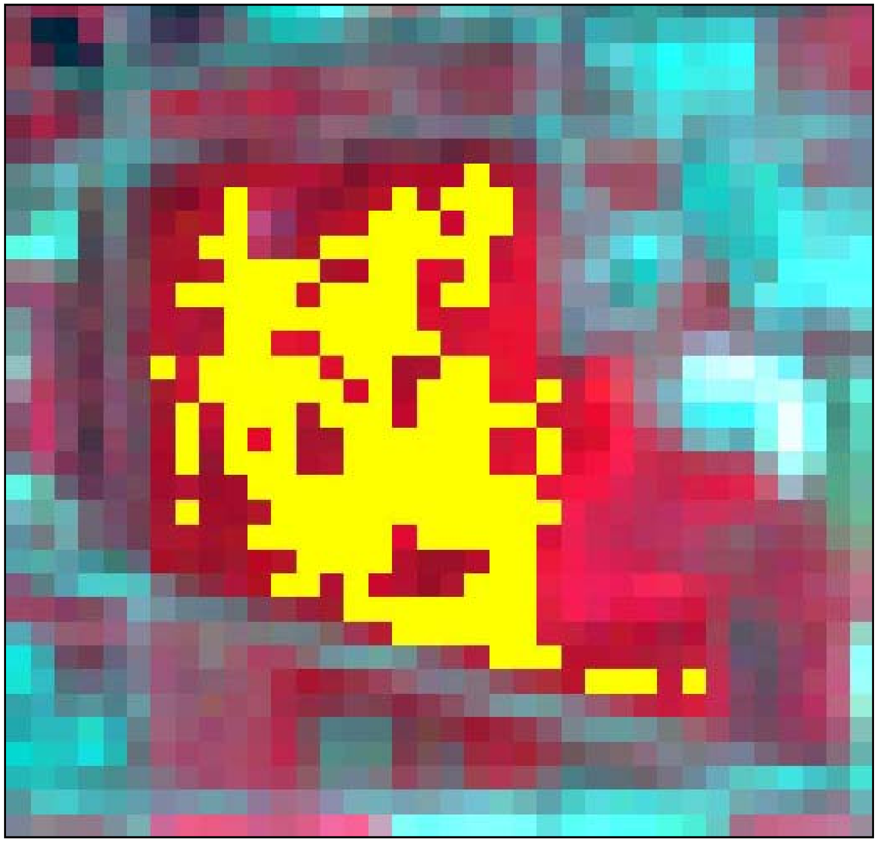




Supervised Classification

❑ *Errors of Omission*

- Indicates the need for **and location of** additional training sites to adequately capture the spectral variability of the class in question.



Supervised Classification

- ***Errors of Commission***
  - Indicates that the training site is not spectrally "pure" regarding the assumed land cover/use class.
  - The analyst must go back to "constrict" the training site (spectrally) by editing the existing training polygon or drawing a new one.
  - *General rule*: "**Do no harm**" (*i.e.*, it is better to leave some pixels unclassified (omission errors) rather than to have large commission errors).

Digital Image Classification

❑ Maximum Likelihood Probability classifier

- Without *A Priori* Probability Information
  - Decide unknown measurement vector X is in class *i* if, and only if, $\mathbf{p_i} > \mathbf{p_j}$ for all *i* and *j* out of 1... m possible classes and

$$p_i = \frac{1}{2}\log_e |V_i| - \left[\frac{1}{2}(X - M_i)^T V_i^{-1}(X - M_i)\right]$$

Where $|V_i|$ is the determinant of the covariance matrix , $(X - M_i)^T$ is the transpose of the vector $(X - M_i)$ and V_i^{-1} is the inverse of the covariance matrix . M_i is the mean measurement vector for class *i*, and V_i is the covariance matrix of class *i* for bands *k* through *l*. Therefore, to assign the measurement vector *X* of an unknown pixel to a class, the maximum likelihood decision rule computes the value p_i for each class and assigns the pixel to the class that has the largest (or maximum) value.

Digital Image Classification

❑ Maximum Likelihood Probability classifier

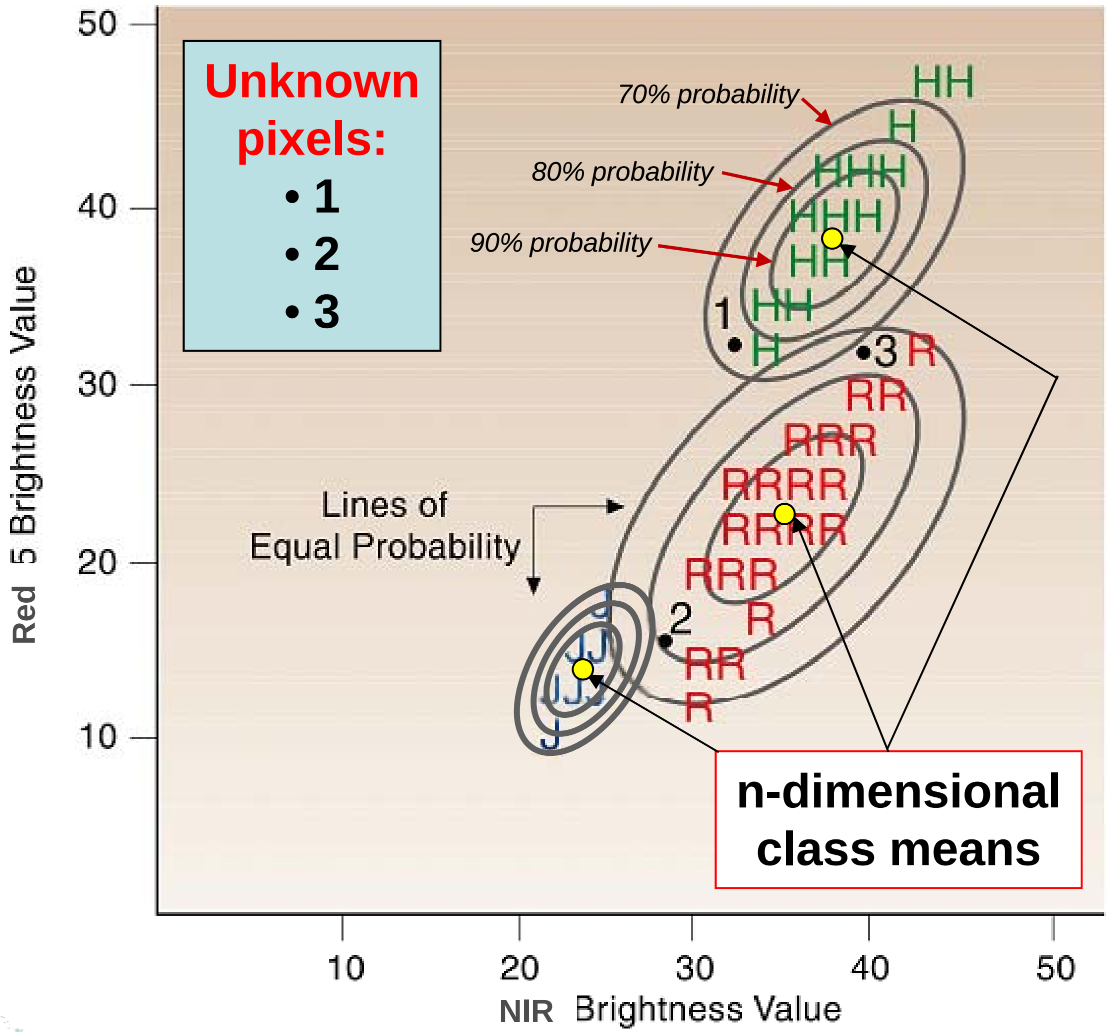
- Without *A Priori* Probability Information
  - The mean vectors M_i and covariance matrix V_i for each class are **estimated from the training data**.

$$p_i = \frac{1}{2}\log_e |V_i| - \left[\frac{1}{2}(X - M_i)^T V_i^{-1}(X - M_i)\right]$$

Where $|V_i|$ is the determinant of the covariance matrix , $(X - M_i)^T$ is the transpose of the vector $(X - M_i)$ and V_i^{-1} is the inverse of the covariance matrix . M_i is the mean measurement vector for class *i*, and V_i is the covariance matrix of class *i* for bands *k* through *l*. Therefore, to assign the measurement vector *X* of an unknown pixel to a class, the maximum likelihood decision rule computes the value p_i for each class and assigns the pixel to the class that has the largest (or maximum) value.

Digital Image Classification

❑ Maximum Likelihood Probability classifier



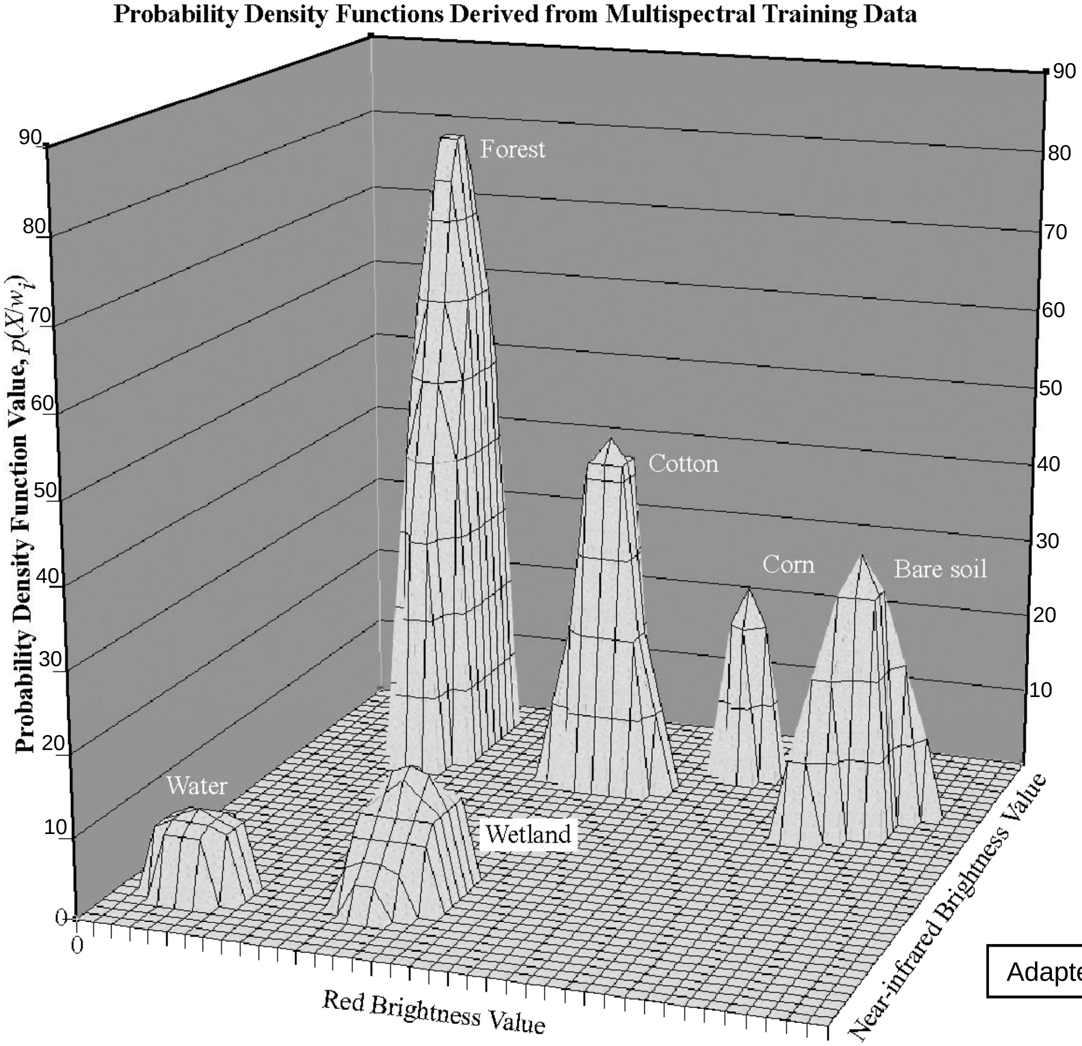
Unknown pixel 1 is assigned to class H because $p_{1H} > p_{1R} >> p_{1J}$

Unknown pixel 2 is assigned to class R because $p_{2R} > p_{2J} >> p_{2H}$

Unknown pixel 3 is assigned to class R because $p_{3R} > p_{3H} >> p_{3J}$

Digital Image Classification

❑ Maximum Likelihood Probability classifier



Adapted from Jensen, 2005

Digital Image Classification

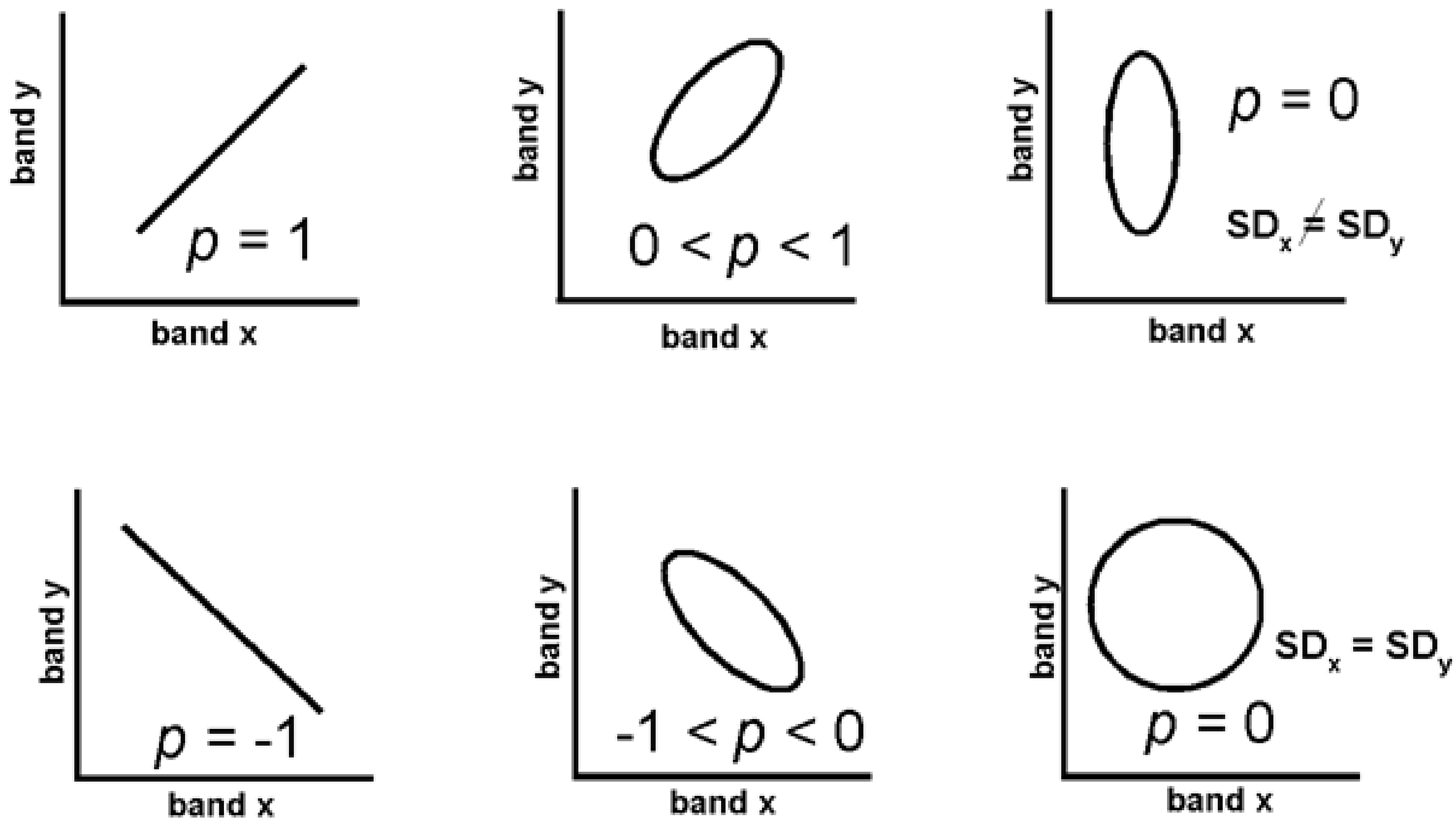
❑ Maximum Likelihood Probability classifier

- ➢ Advantages
  - ✓ Sensitive to both variance and covariance in the training data
- ➢ Disadvantages
  - ✓ Training data must be normally distributed
  - ✓ The per-pixel, rather than pixel neighborhood approach, remains a problem

Digital Image Classification

❑ Maximum Likelihood Probability classifier

➢ Multi-band Variance, Covariance and Correlation



Digital Image Classification

❑ Maximum Likelihood Probability classifier

➢ Variance – Covariance Matrix

$\sigma^2_{11} = 5192.16$	$\sigma^2_{12} = 3866.65$	$\sigma^2_{13} = 2722.83$	$\sigma^2_{14} = 1094.98$
	$\sigma^2_{22} = 3781.67$	$\sigma^2_{23} = 2520.50$	$\sigma^2_{24} = 1462.00$
		$\sigma^2_{33} = 4806.52$	$\sigma^2_{34} = 3652.66$
			$\sigma^2_{44} = 3927.78$

Sum of variances = 17708.13

4-band data set (range 0 – 255)

Mean band 1 = 129.29

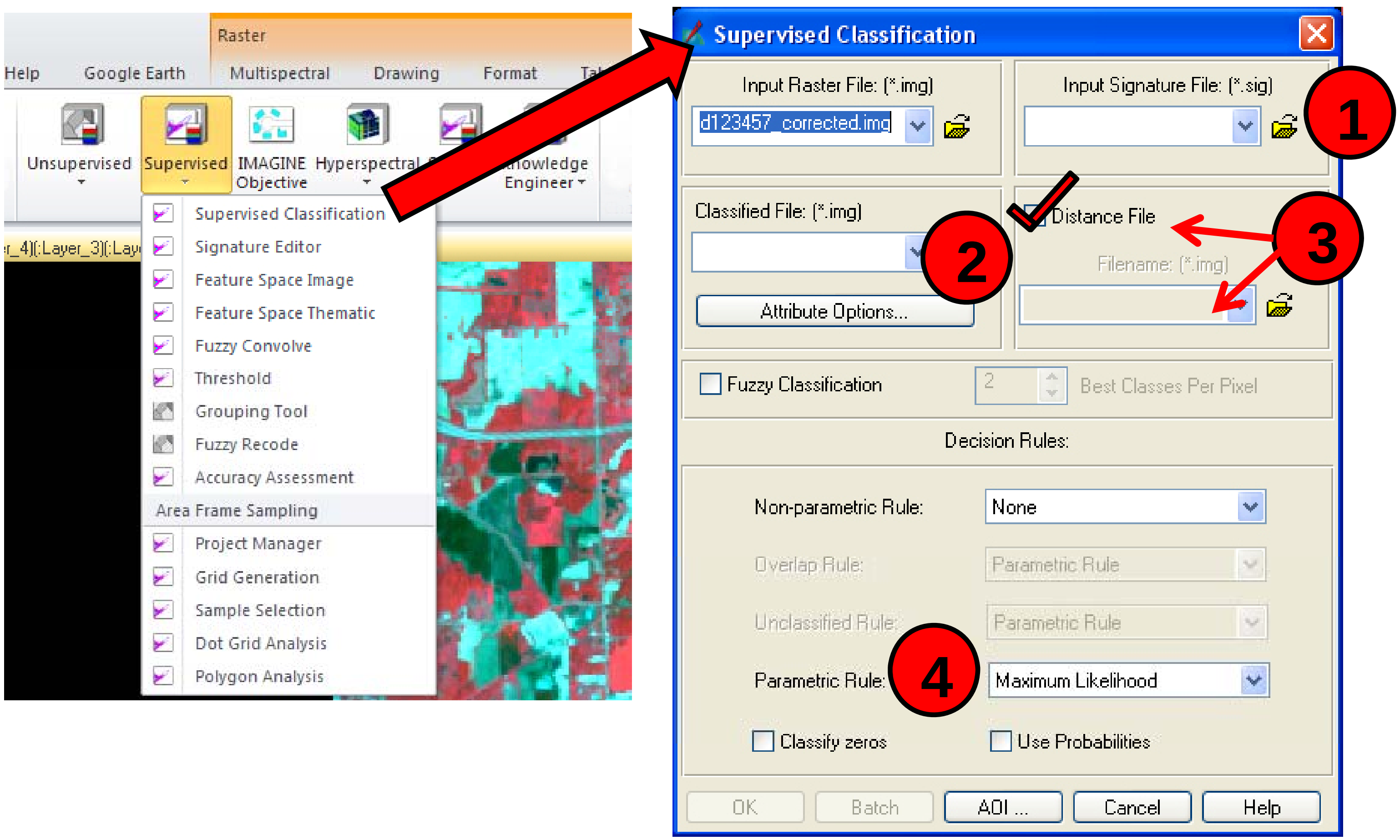
Mean band 2 = 106.63

Mean band 3 = 104.71

Mean band 4 = 121.80

Supervised Classification

- ***Maximum Likelihood Classifier***
  - **Raster tab → Classification Group → Supervised**



Supervised Classification

❑ *Maximum Likelihood Classifier*

- All pixels are classified initially, irrespective of how low the probabilities of class membership are.
- *E.g.*, assume the probabilities of pixel ***i*** belonging to one of eight classes is:
  - $\boldsymbol{p}_1 = 0.0125$
  - $\boldsymbol{p}_2 = 0.0237$
  - $\boldsymbol{p}_3 = 0.0105$
  - $\boldsymbol{p}_4 = 0.0095$
  - $\boldsymbol{p}_5 = \mathbf{0.0275}$
  - $\boldsymbol{p}_6 = 0.0078$
  - $\boldsymbol{p}_7 = 0.0083$
  - $\boldsymbol{p}_8 = 0.0195$

Pixel ***i*** is assigned to Class 5 even though its probability of class membership is only 2.75% (*i.e.*, it's 97.25% likely to NOT be a member of Class 5 (and even less likely to be a member of any other class).

Supervised Classification

❑ *Maximum Likelihood Classifier*

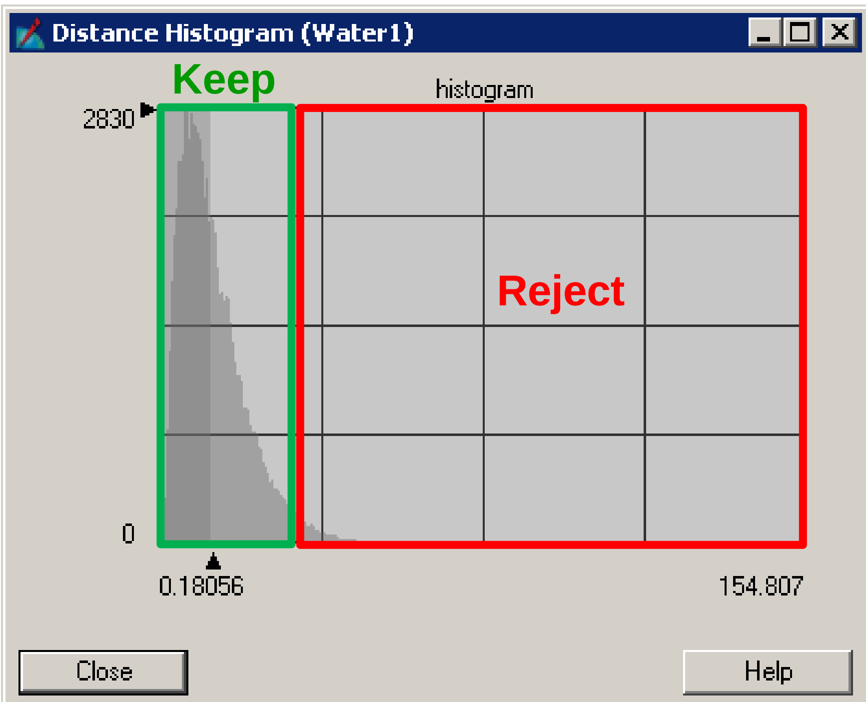
➢ **Threshold tool** (Raster tab → Supervised → Threshold)

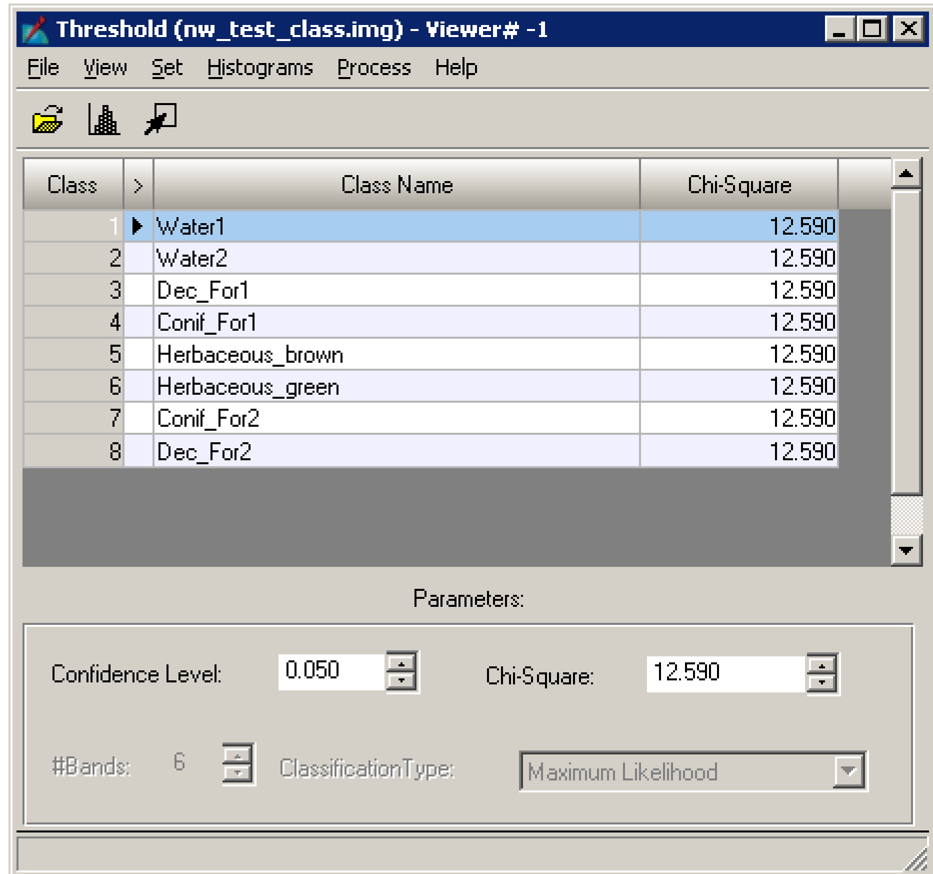
- **Thresholding** is a statistical method used to refine a supervised classification by determining which pixels in the new thematic raster layer **are most likely to be incorrectly classified** (using the **distance file** that was created during the supervised classification). These distances occur in probability space, so larger distances are associated with lower probabilities of class membership.
- The histogram of the distance file usually approximates a ***chi-square distribution***. Normally, there are many pixels with a distance image **value at or near 0**, indicating that these pixels are **likely to be classified correctly**. A histogram is generated for each class in the classified image file.
- Analyst displays, inspects and edits (thresholds) each histogram. The rejected pixels become "unclassified."

Supervised Classification

❑ *Maximum Likelihood Classifier*

- **Threshold tool** (Raster tab → Supervised → Threshold)



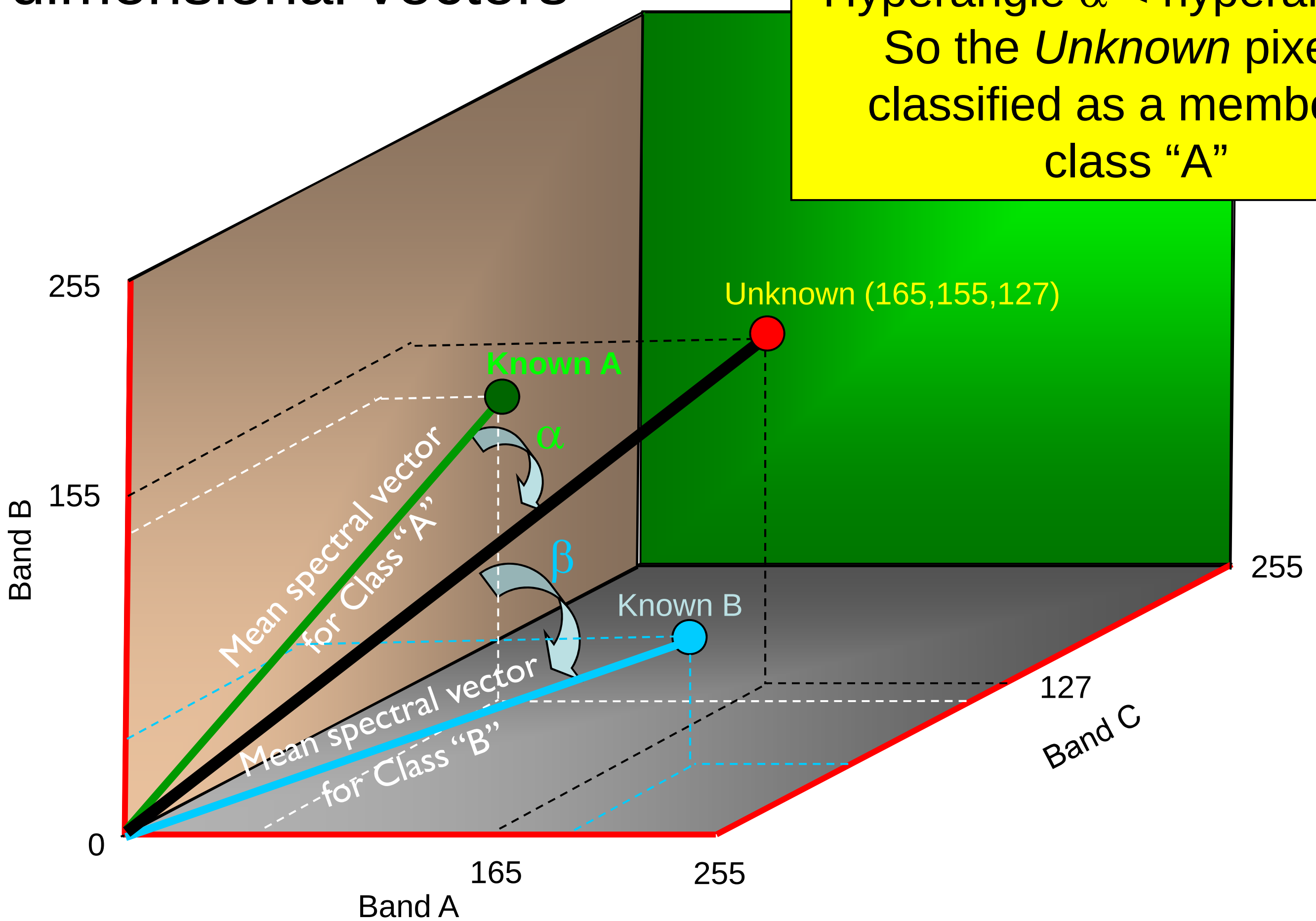


Digital Image Classification

❑ Spectral Angle Mapper

- ➢ n-dimensional vectors

Hyperangle α < hyperangle β, So the *Unknown* pixel is classified as a member of class “A”



Digital Image Classification

Spectral Angle Mapper

$$\alpha = \cos^{-1}\left(\frac{\sum_{i}^{nb} tr}{\left(\sum_{i}^{nb} t_i^2\right)^{1/2} \left(\sum_{i}^{nb} r_i^2\right)^{1/2}} \right)$$

Where *nb* is the number of bands in the image, *t* is the unknown pixel spectrum, and *r* is reference spectrum.

Digital Image Classification

Spectral Angle Mapper

- Given a set of reference signatures collected from an image with *m* bands, the SAM classifier performs the following two steps:
  - *Step 1*: Determine the spectral angle $\theta_{i,r}$, between a pixel *i* in the image and every reference class *r* :

$$\theta_{i,r} = \cos^{-1}\left(\frac{\sum_{m}^{nb} t_{i,m} r_{r,m}}{\sqrt{\sum_{m}^{nb} t_{i,m}^2 \sum_{m}^{nb} r_{r,m}^2}} \right)$$

Digital Image Classification

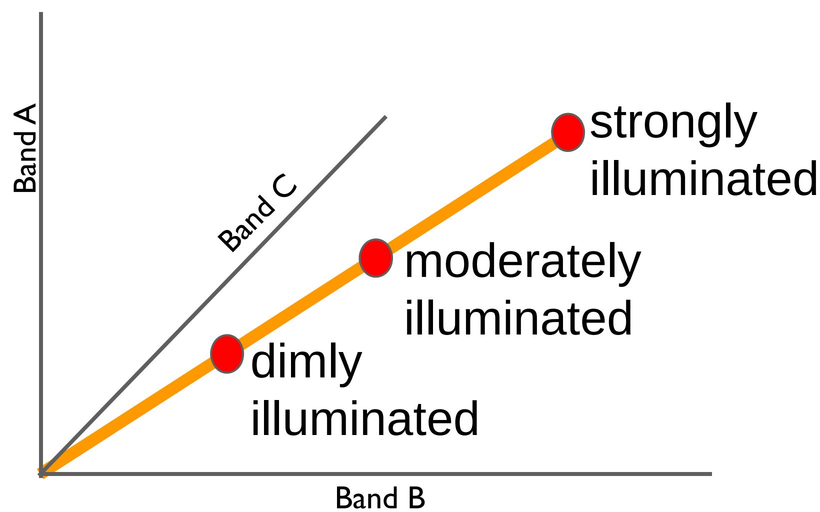
❑ Spectral Angle Mapper

– *Step 2*: Assign each pixel to the reference class *r* that has the smallest spectral angular distance between pixel *i* and reference class *r*. For each pixel $i = 1$ to n, find the reference class *r* such that the $\theta_{i,r}$, is the minimum of all reference classes.

Digital Image Classification

❑ Spectral Angle Mapper Classifier

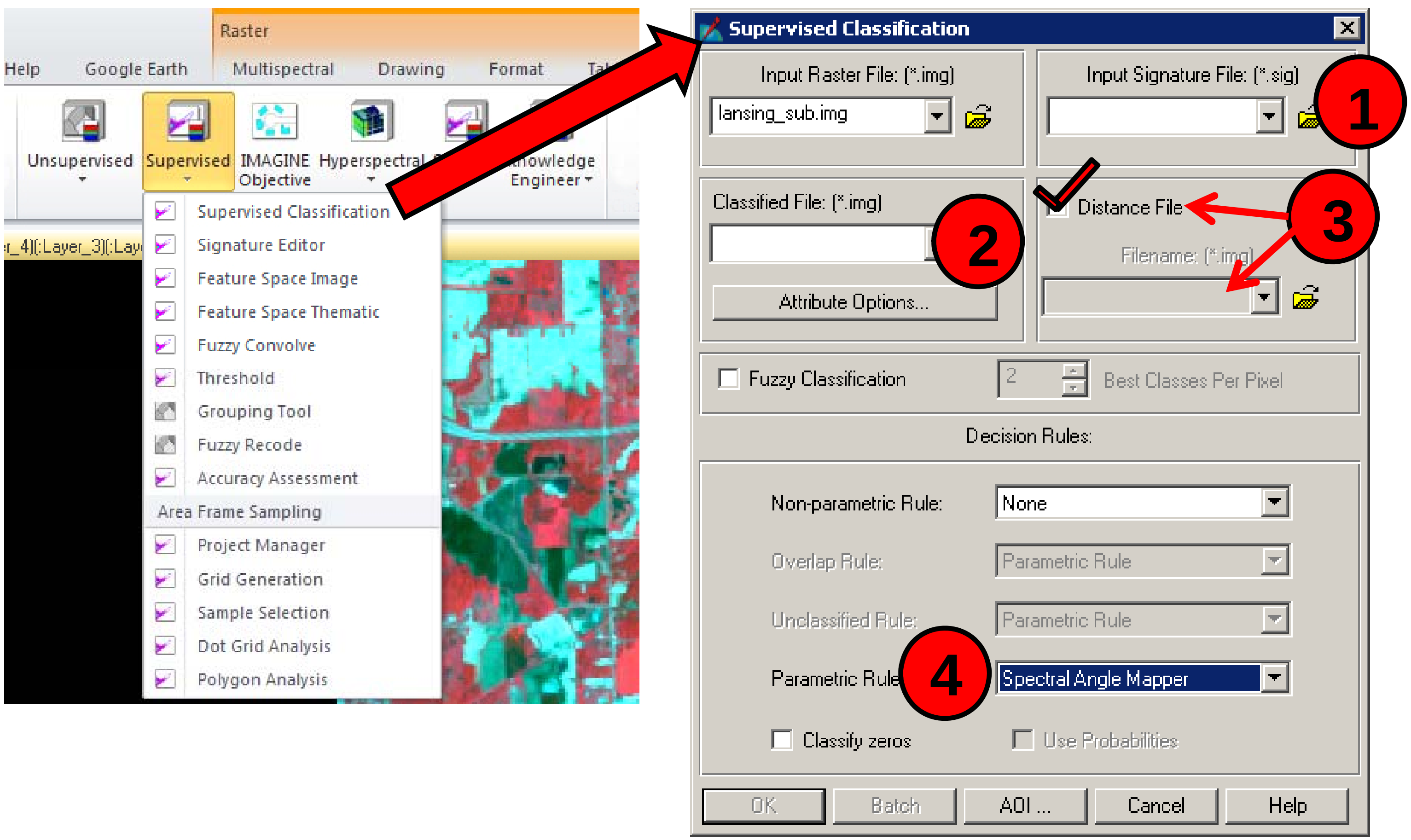
- ➢ Advantages
  - ✓ Often more accurate than maximum likelihood
  - ✓ Less sensitive to illumination differences
- ➢ Disadvantages
  - ✓ Numerous training spectra required to describe the "best" mean vector for each class



Supervised Classification

Spectral Angle Mapper

- Raster tab → Classification Group → Supervised



Supervised Classification

❑ Image Classification

- Classifications can always be improved - perfection is rarely possible, especially using precise LULC categories
- Per pixel classification works best with moderate spatial resolution imagery (GSD >= 20 meters)
- The classification process is iterative – it's rare to have high classification accuracy after the first attempt
- Consider a hybrid classification – part supervised, part unsupervised (especially in areas where your field experience is limited)
- Manual Classification and/or Editing is not cheating, but can be extremely time consuming

Digital Image Classification

- ❑ **Supervised classification**
  - The analyst "supervises” the categorization of a set of specific classes by providing ***training statistics*** that identify each category.
- ❑ **Unsupervised classification**
  - The raw spectral data are **grouped first**, based solely on the statistical structure of the data. Then the **analyst must label** each statistical cluster, placing them in the appropriate categories (if possible).
- ❑ **Rule-based classification**
  - Spectrally categorized pixels are classified using ancillary data in a model.

Unsupervised Classification

- ❑ Requires only minimal initial input from the analyst.
- ❑ However, the analyst has the task of interpreting the many classes that are generated by the unsupervised classification algorithm.
- ❑ Unsupervised classification is also called **clustering**, because it is based on the natural groupings of pixels in image data when they are plotted in feature space.

Unsupervised Classification

- ❑ Clusters are defined with a clustering algorithm.
- ❑ The clustering algorithm has no regard for the contiguity of the pixels that define each cluster unlike the training areas for supervised classification.

  - ➢ The **I**terative **S**elf-**O**rganizing **Dat**a **A**nalysis – **ISODATA** clustering method uses spectral distance and iteratively classifies the pixels.

  - ➢ After each iteration, ISODATA redefines the criteria for each class, and classifies again, gradually “discovering” the spectral distance patterns (*i.e.*, the clusters) in the data.

Unsupervised Classification

- The initial cluster means are evenly distributed across the feature space along a vector that runs from the
  - multidimensional **-1 σ** $(\mu_1-\sigma_1, \mu_2-\sigma_2, \mu_3-\sigma_3, \ldots \mu_n-\sigma_n)$

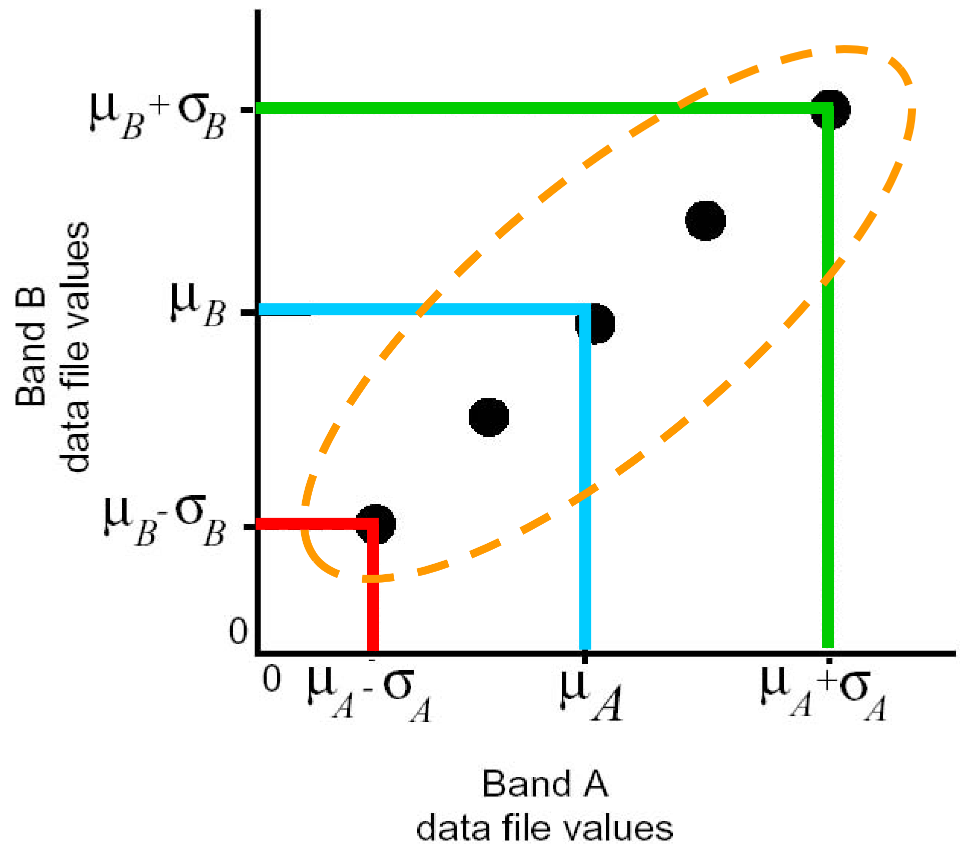
    to the
  - multidimensional **+1 σ** $(\mu_1+\sigma_1, \mu_2+\sigma_2, \mu_3+\sigma_3, \ldots \mu_n+\sigma_n)$

Unsupervised Classification

- ISODATA Initialization

5 arbitrary cluster means in two-dimensional spectral space

User specified number of clusters

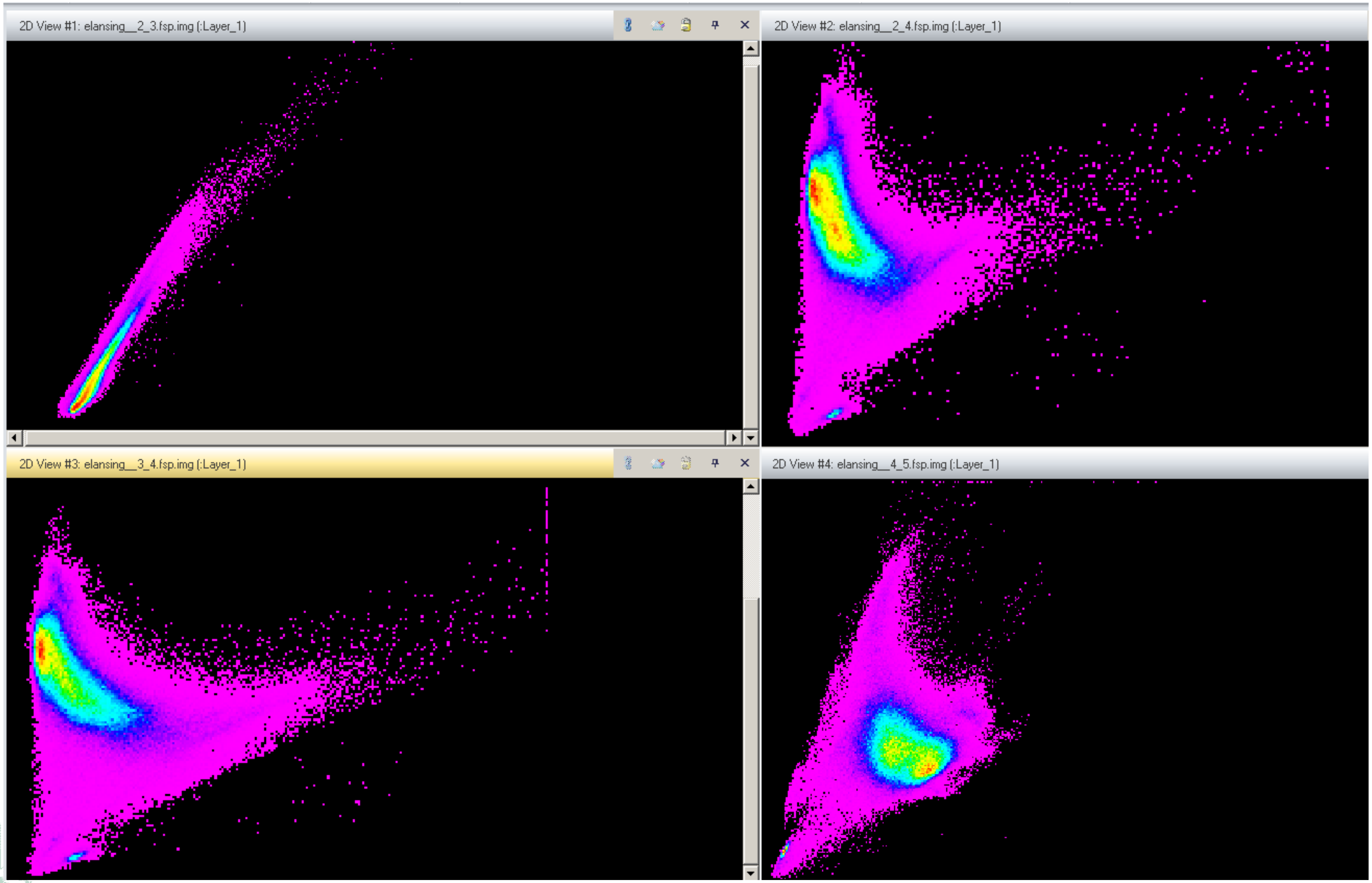


Unsupervised Classification

- ISODATA uses minimum spectral distance to assign a cluster for each candidate pixel.
  - The process begins with a user-specified number of arbitrary cluster means (or the means of existing signatures).
  - With each iteration, the means shift to those of the clusters in the data.
  - Because ISODATA is iterative, it is not biased to the top of the data file, as are the one-pass clustering algorithms.
  - The spectral distance between any candidate pixel and each cluster mean is calculated and the pixel is assigned to the cluster whose mean is the closest.

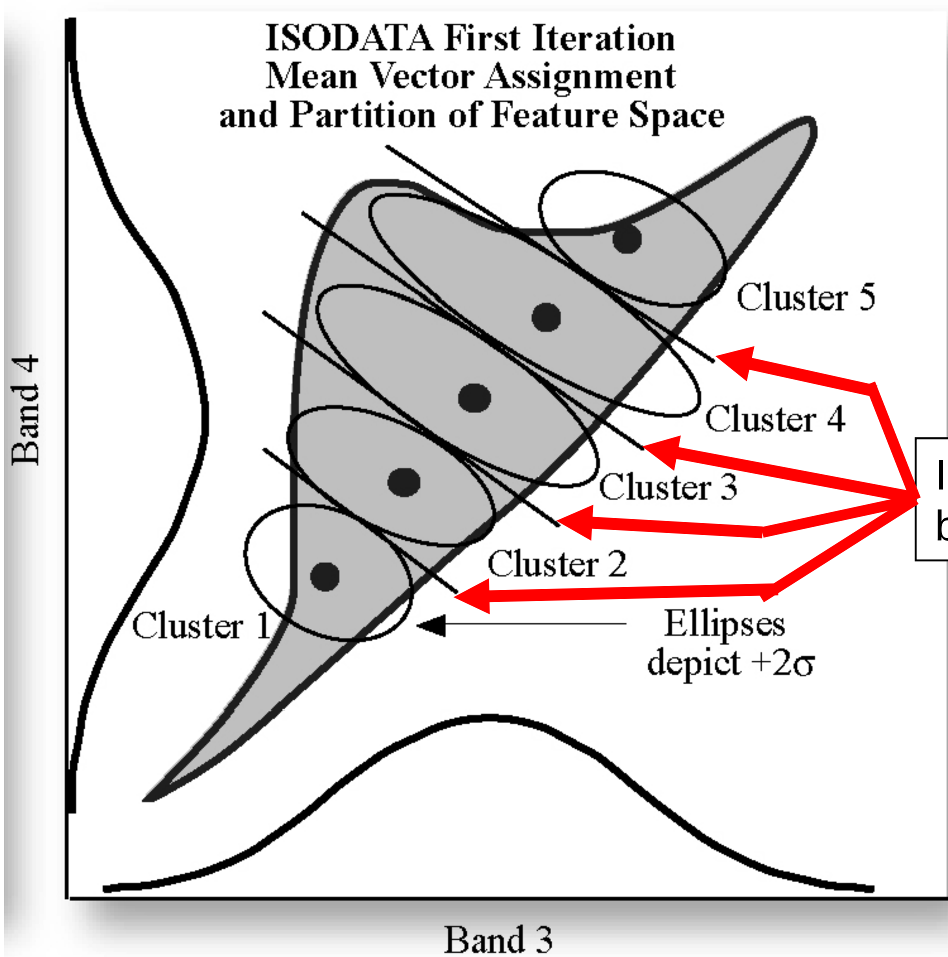
Unsupervised Classification

- Multispectral Feature Space



Unsupervised Classification

ISODATA first pass



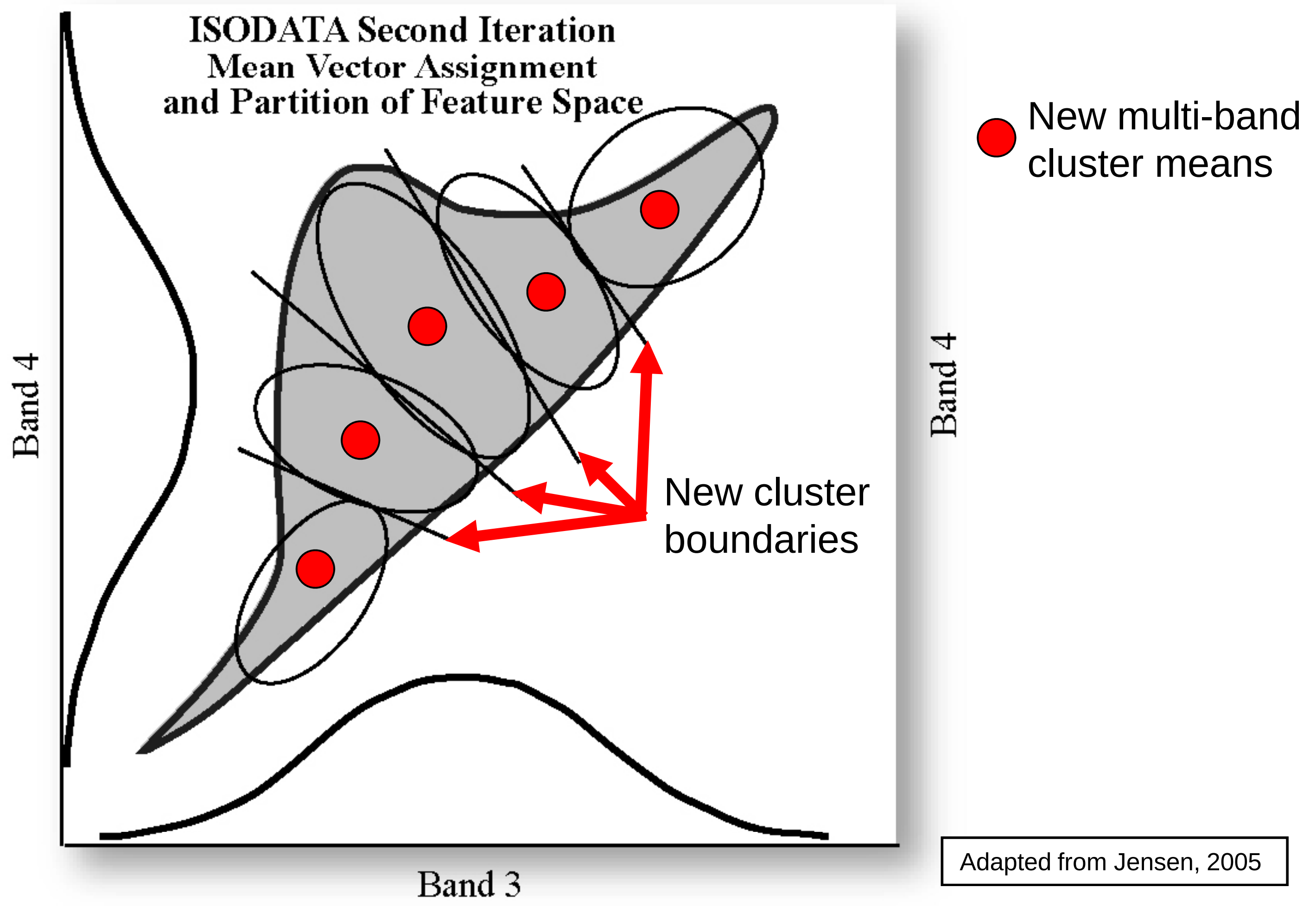
- Initial arbitrary multi-band means

Initial cluster boundaries

Adapted from Jensen, 2005

Unsupervised Classification

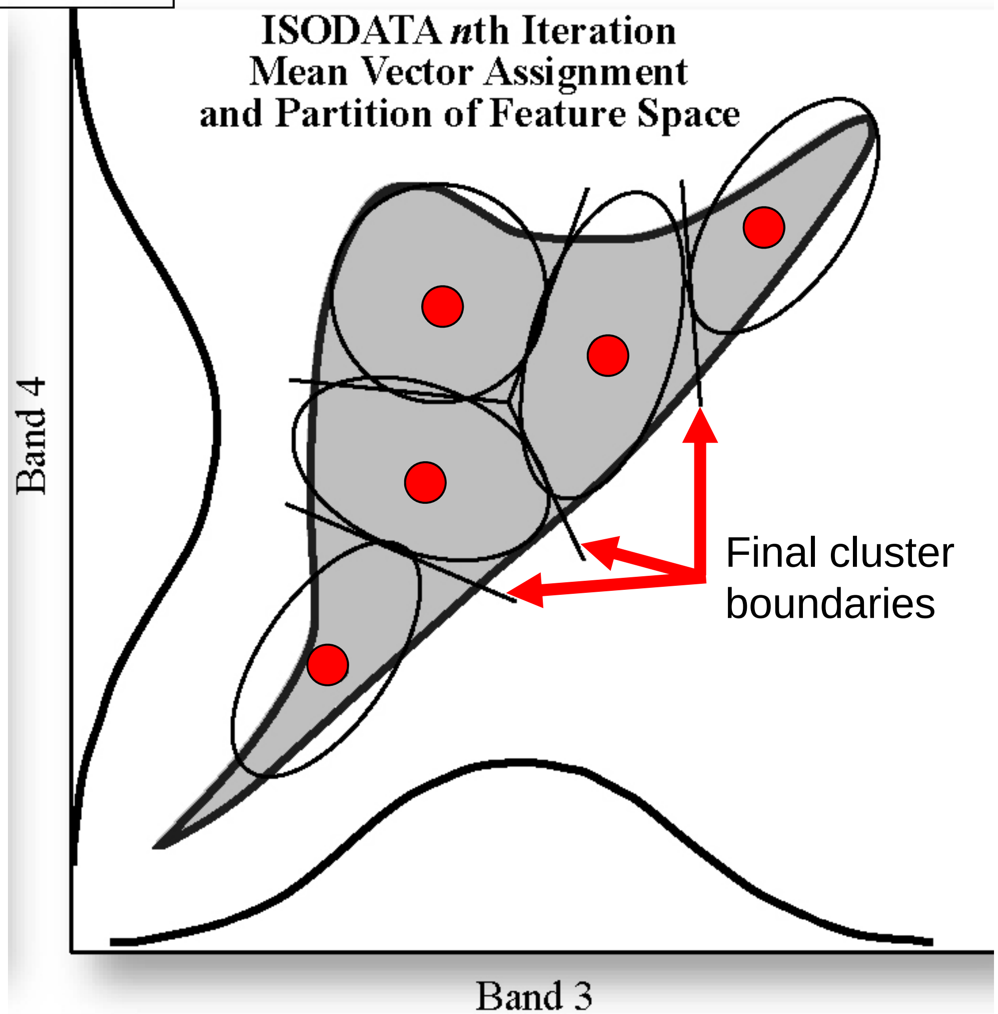
ISODATA second pass



Adapted from Jensen, 2005

Unsupervised Classification

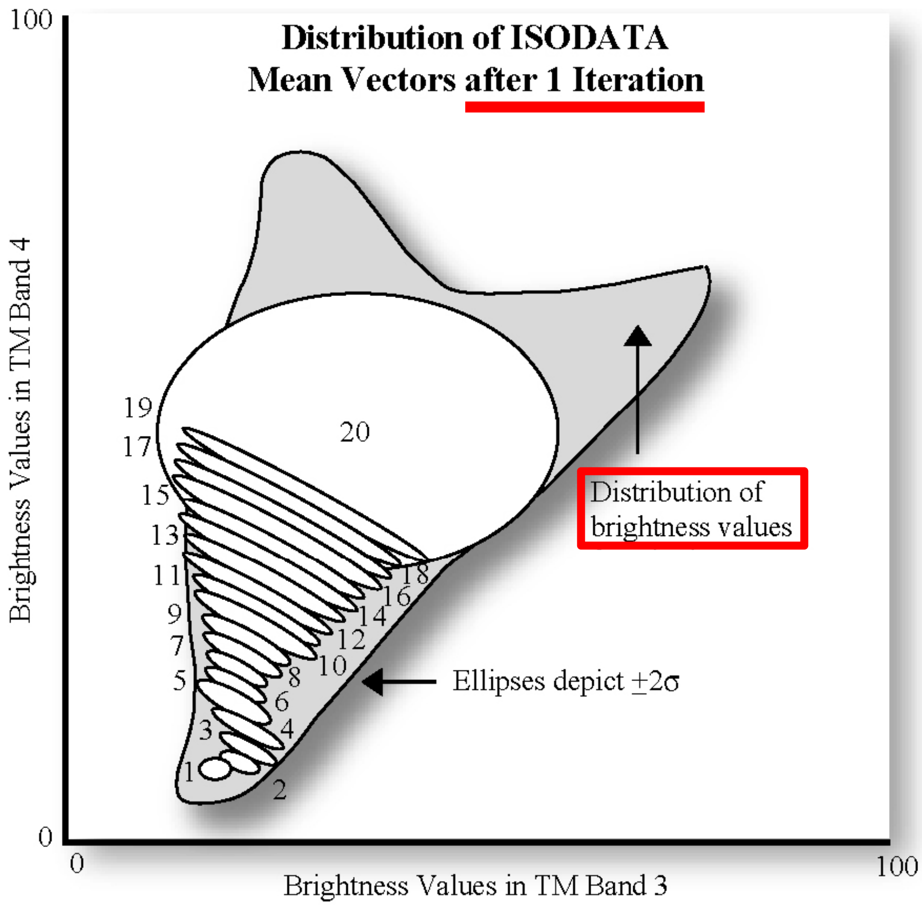
ISODATA n^{th} pass

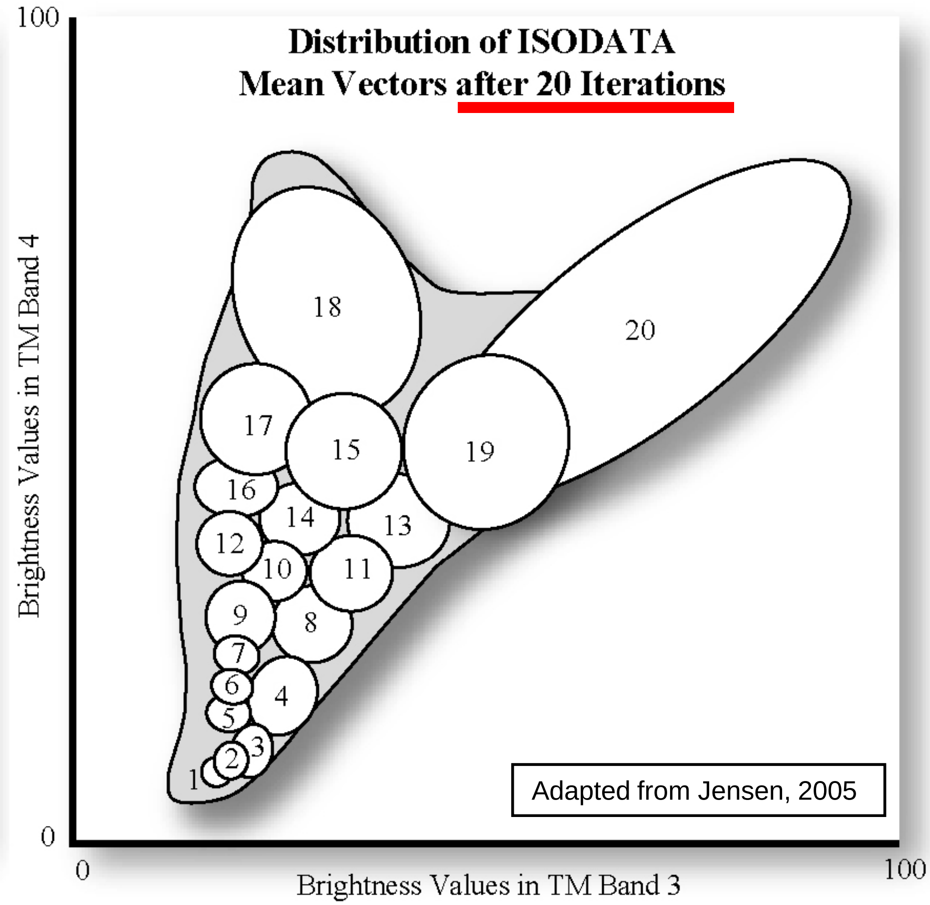


Final multi-band cluster means

Adapted from Jensen, 2005

Unsupervised Classification

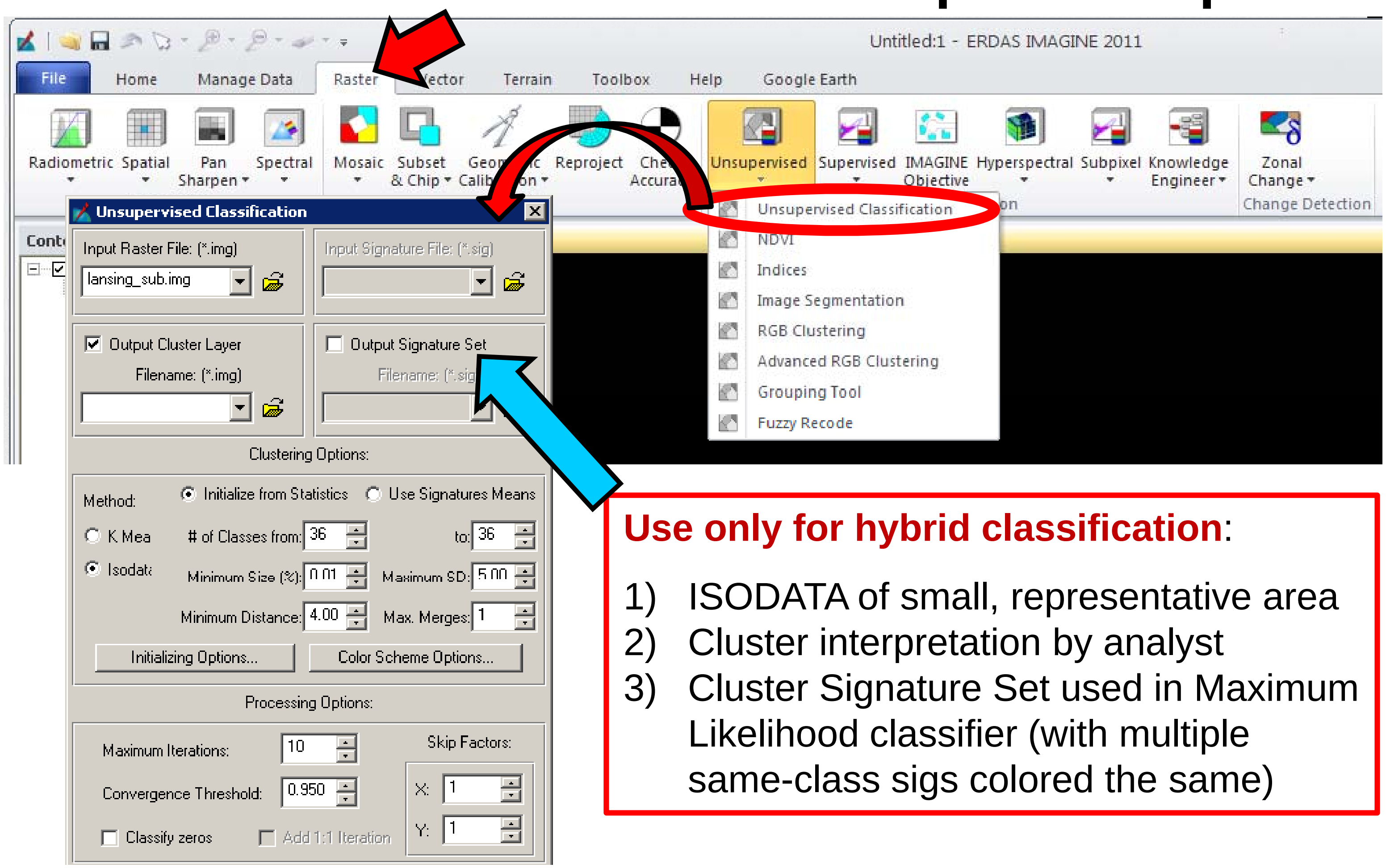




Adapted from Jensen, 2005

Unsupervised Classification

- **ISODATA** in ERDAS *Imagine*
  - **Raster tab → Classification Group → Unsupervised**

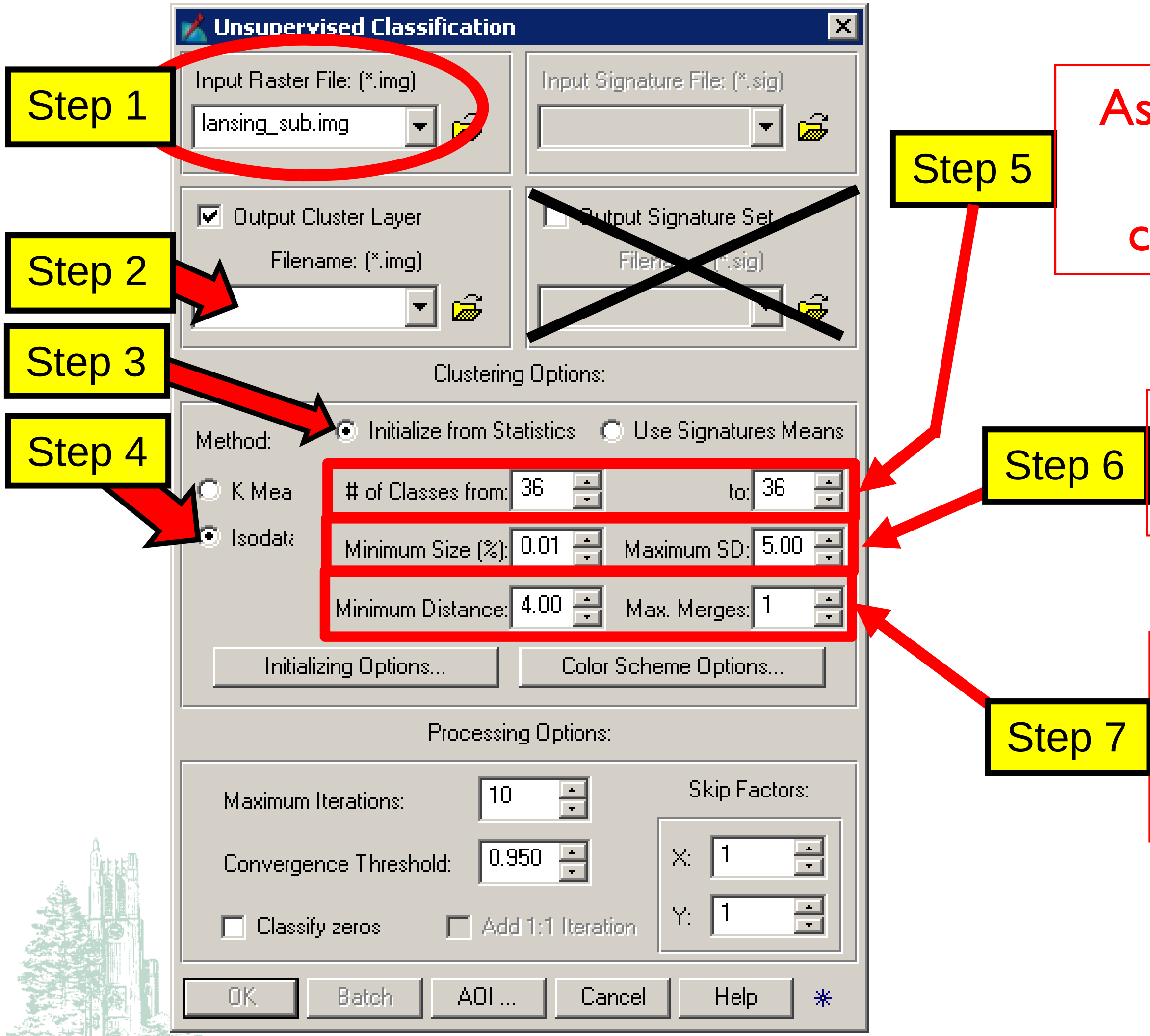


Use only for hybrid classification:

1) ISODATA of small, representative area
2) Cluster interpretation by analyst
3) Cluster Signature Set used in Maximum Likelihood classifier (with multiple same-class sigs colored the same)

Unsupervised Classification

- **ISODATA** in ERDAS *Imagine*



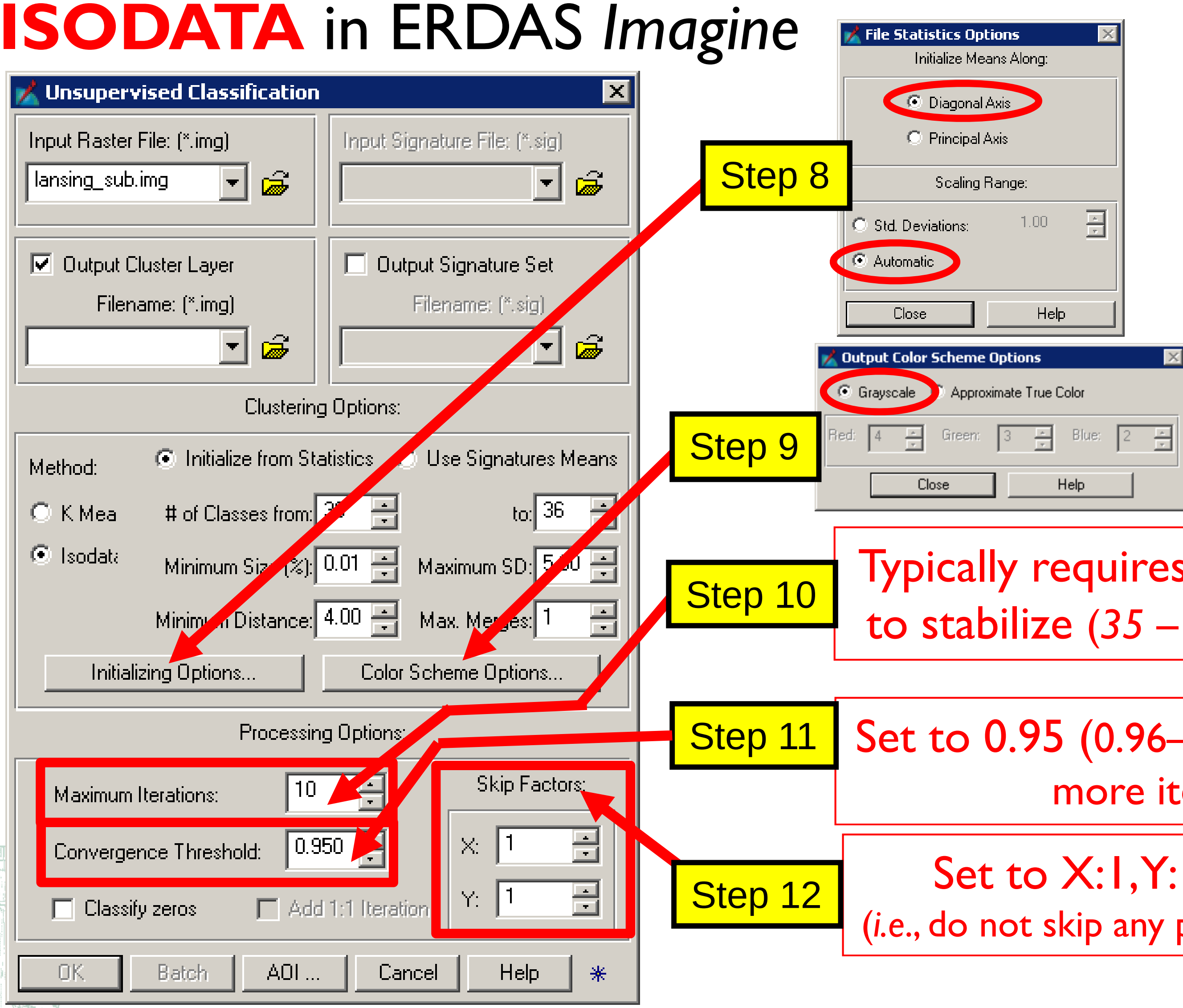
As a general rule, this should be at least 3x the number of legend classes you desire + 3 – 5 more

Minimum Size = 1 – 5%
Maximum SD = 3.0 – 4.0

Minimum Distance = 4 – 5
(for 8-bit images)
Maximum Merges = 10 - 15

Unsupervised Classification

- **ISODATA** in ERDAS *Imagine*



Typically requires 15 – 20 iterations to stabilize (*35 – 40 for 16-bit images*)

Set to 0.95 (0.96– 0.98 will take many more iterations)

Set to X:1,Y:1 (*i.e.*, do not skip any pixels)

Unsupervised Classification

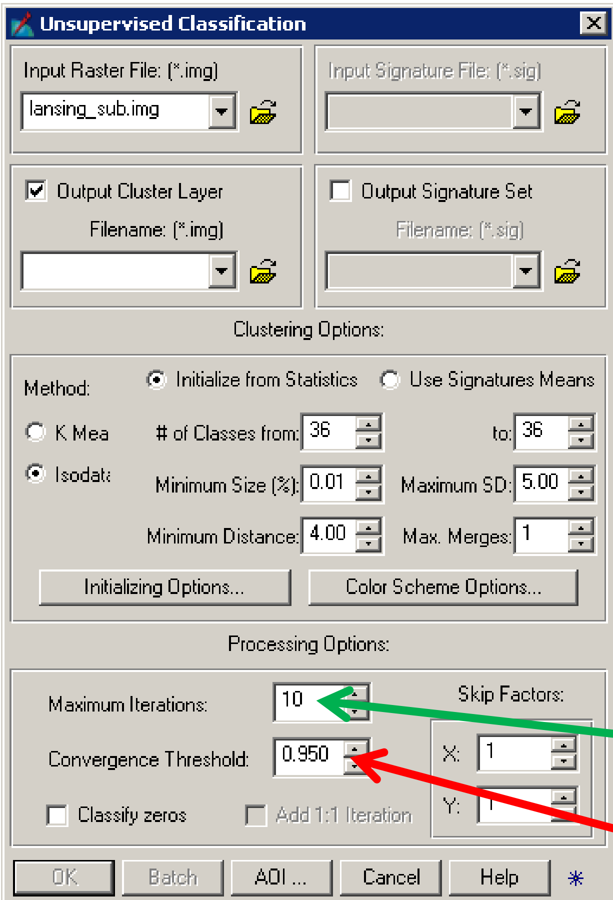
- After each iteration, the normalized percentage of pixels whose assignments are unchanged since the last iteration is determined.

- When this number reaches ***T*** (the convergence threshold), the program terminates.

- As a **QA check**, ensure that the program terminated because it reached the convergence threshold and not because it reached the maximum number of iterations.

Unsupervised Classification

- File → Session → View Session Log



Performing iteration: 1 Convergence: 0.000

Performing iteration: 2 Convergence: 0.260

Performing iteration: 3 Convergence: 0.525

Performing iteration: 4 Convergence: 0.862

Performing iteration: 5 Convergence: 0.920

Performing iteration: 6 Convergence: 0.938

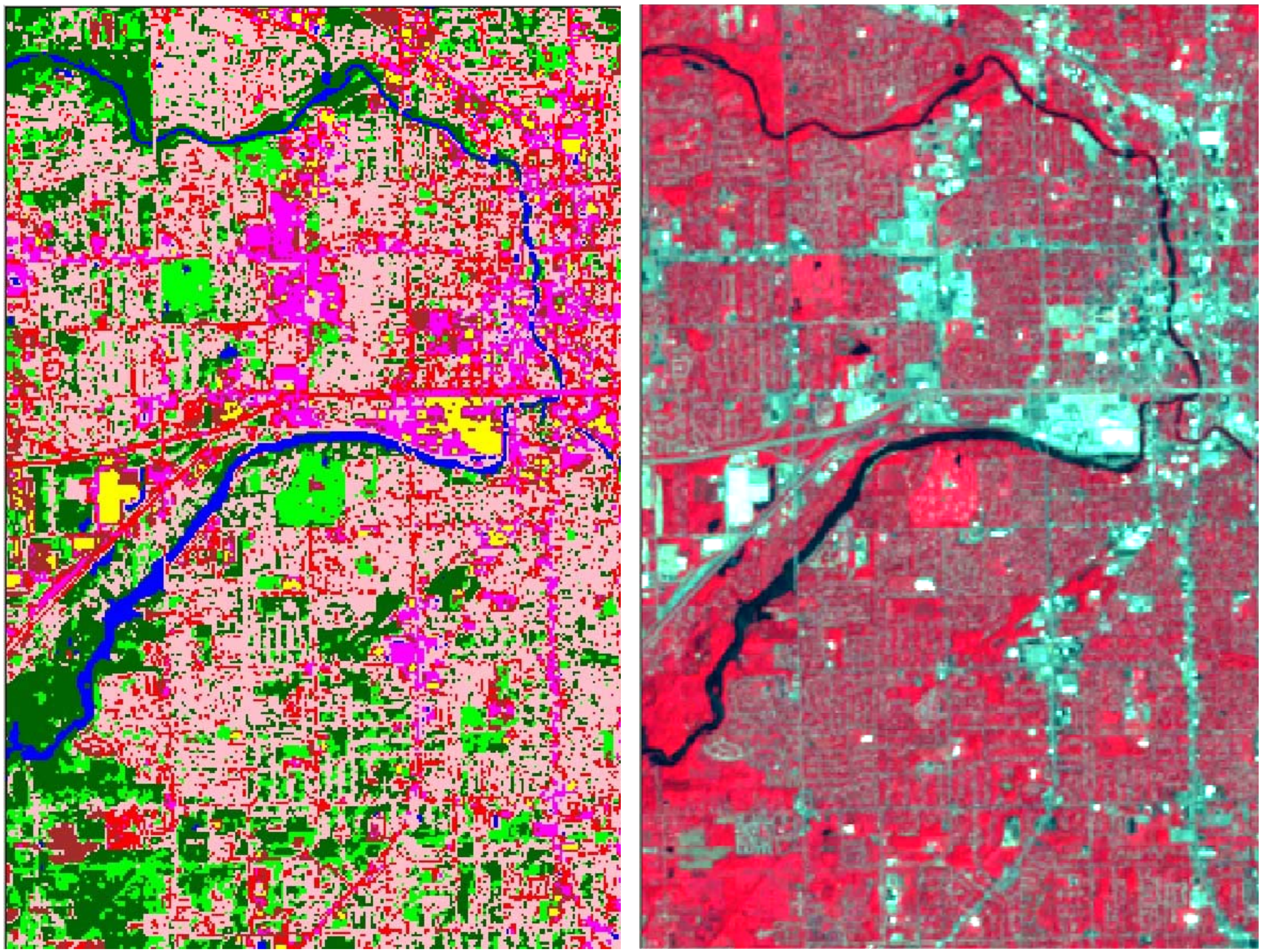
Performing iteration: 7 Convergence: 0.949

Performing iteration: 8 Convergence: $\boxed{0.957}$

Unsupervised Classification

20 Clusters

6 Iterations



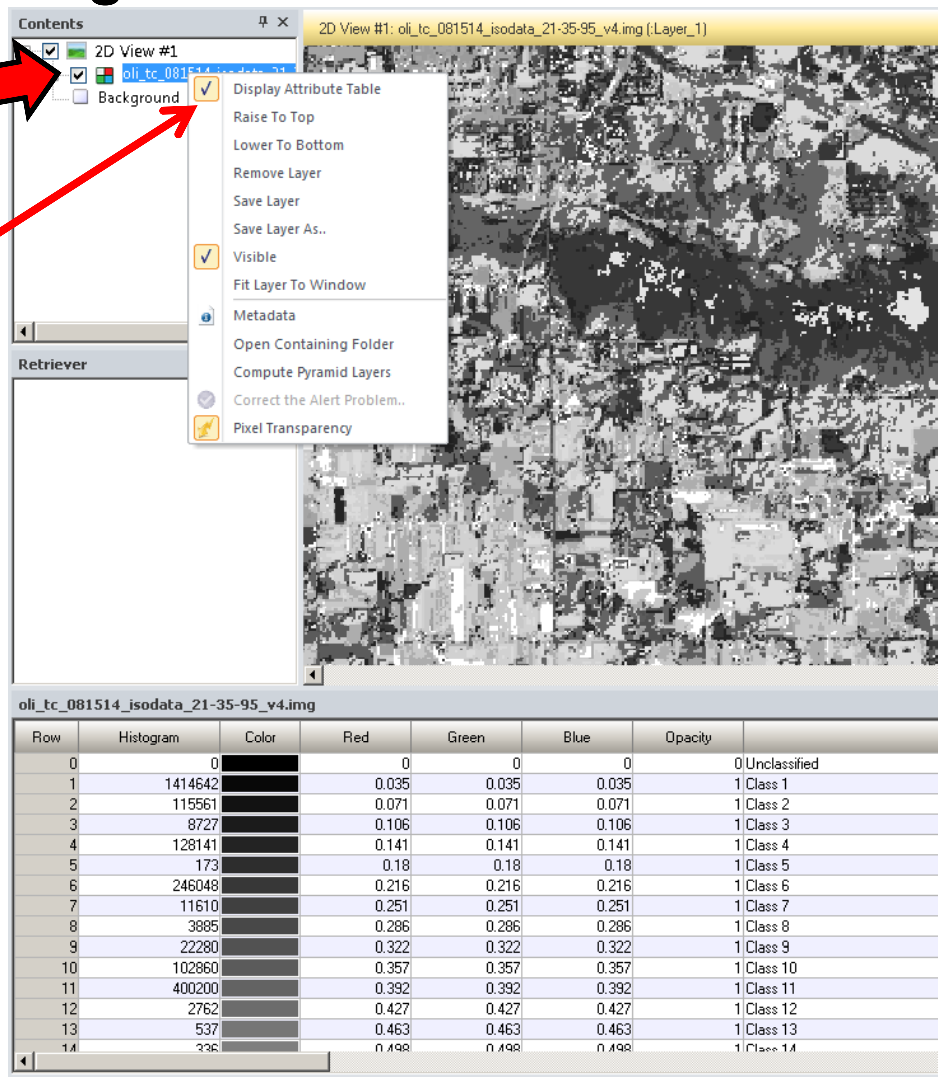
Unsupervised Classification

❑ Cluster Labeling → Attribute Table

Right-click for pop-up menu.

Select "Display Attribute Table"

Attribute Table



Row	Histogram	Color	Red	Green	Blue	Opacity	
0	0		0	0	0	0	Unclassified
1	1414642		0.035	0.035	0.035	1	Class 1
2	115561		0.071	0.071	0.071	1	Class 2
3	8727		0.106	0.106	0.106	1	Class 3
4	128141		0.141	0.141	0.141	1	Class 4
5	173		0.18	0.18	0.18	1	Class 5
6	246048		0.216	0.216	0.216	1	Class 6
7	11610		0.251	0.251	0.251	1	Class 7
8	3885		0.286	0.286	0.286	1	Class 8
9	22280		0.322	0.322	0.322	1	Class 9
10	102860		0.357	0.357	0.357	1	Class 10
11	400200		0.392	0.392	0.392	1	Class 11
12	2762		0.427	0.427	0.427	1	Class 12
13	537		0.463	0.463	0.463	1	Class 13
14	336		0.498	0.498	0.498	1	Class 14

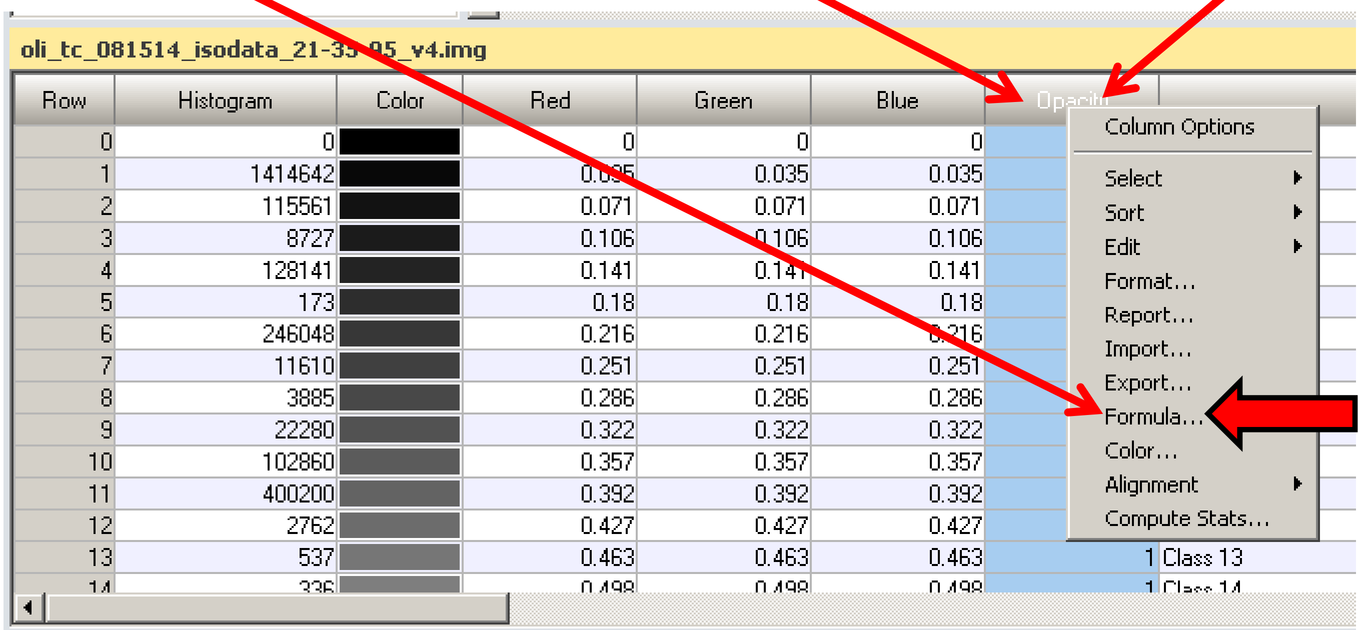
Unsupervised Classification

❑ Cluster Labeling → Attribute Table

1. Left-click on Opacity column heading to select all

2. Right-click on Opacity column heading to open “Column Options” pop-up menu

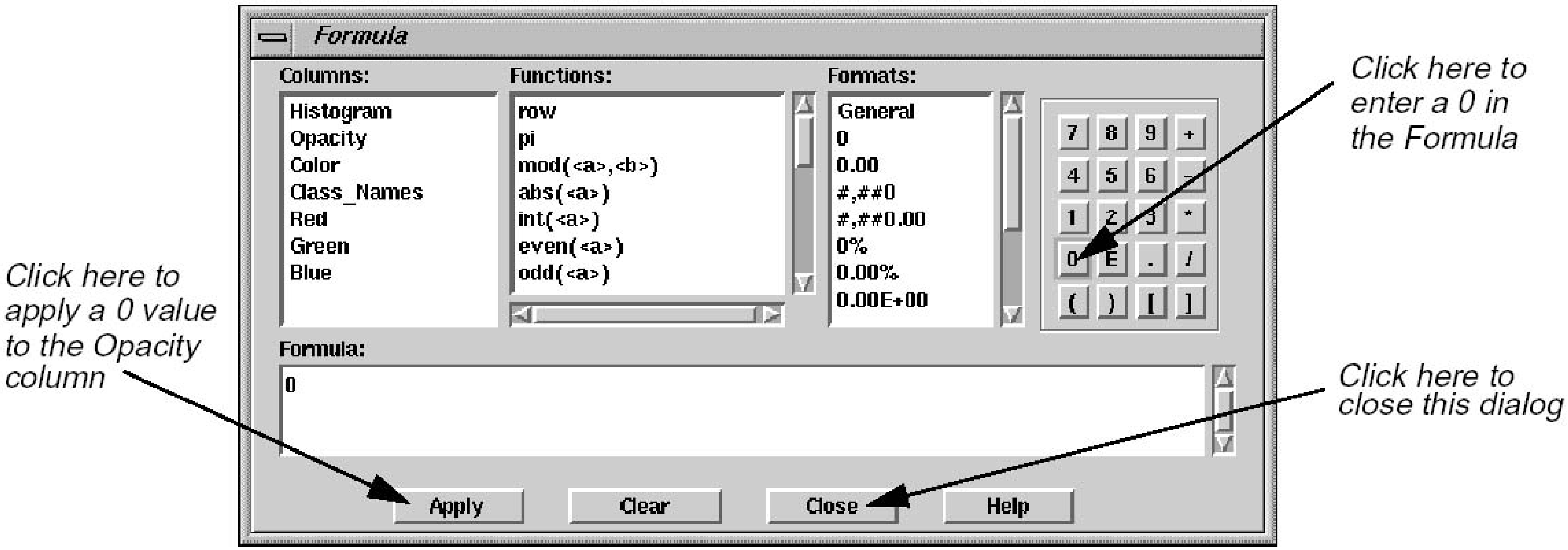
3. Select “Formula” from menu



Row	Histogram	Color	Red	Green	Blue	Opacity	
0	0		0	0	0		
1	1414642		0.035	0.035	0.035		
2	115561		0.071	0.071	0.071		
3	8727		0.106	0.106	0.106		
4	128141		0.141	0.141	0.141		
5	173		0.18	0.18	0.18		
6	246048		0.216	0.216	0.216		
7	11610		0.251	0.251	0.251		
8	3885		0.286	0.286	0.286		
9	22280		0.322	0.322	0.322		
10	102860		0.357	0.357	0.357		
11	400200		0.392	0.392	0.392		
12	2762		0.427	0.427	0.427		
13	537		0.463	0.463	0.463	1	Class 13
14	336		0.498	0.498	0.498	1	Class 14

Unsupervised Classification

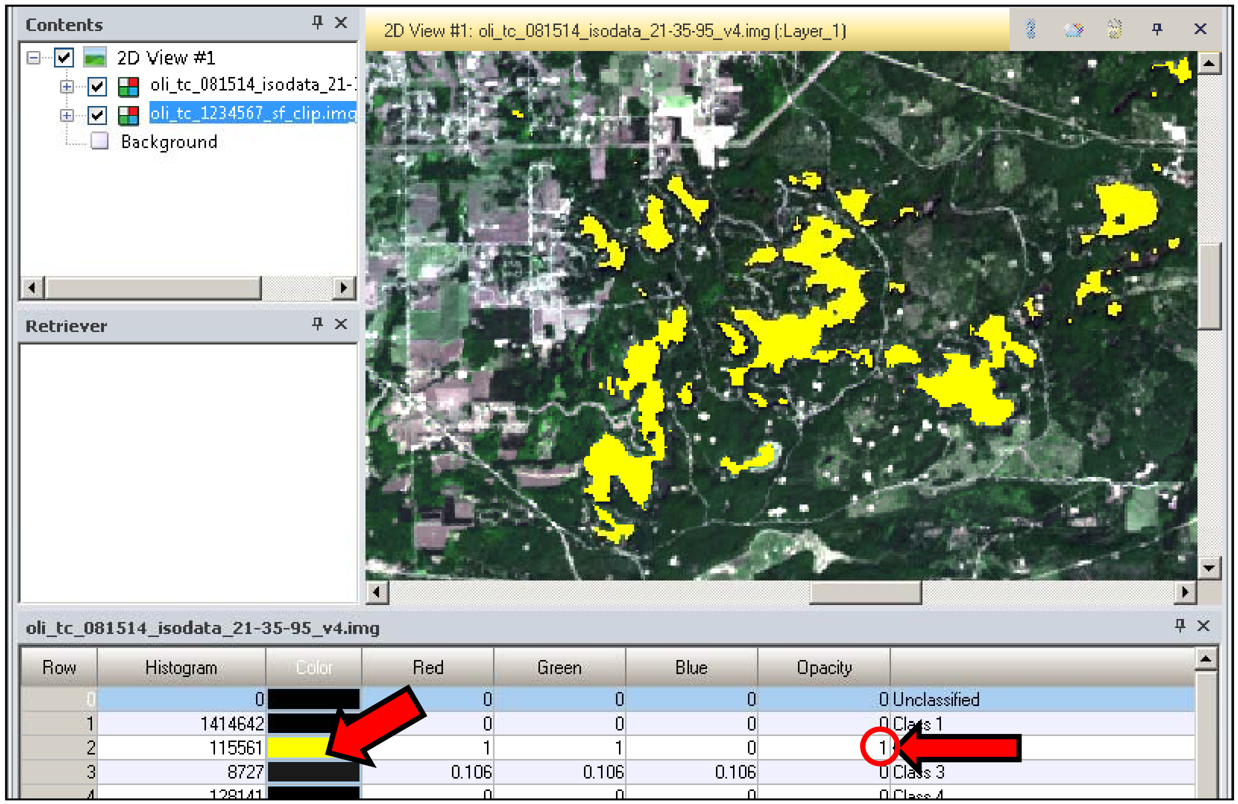
❑ Cluster Labeling → Attribute Table

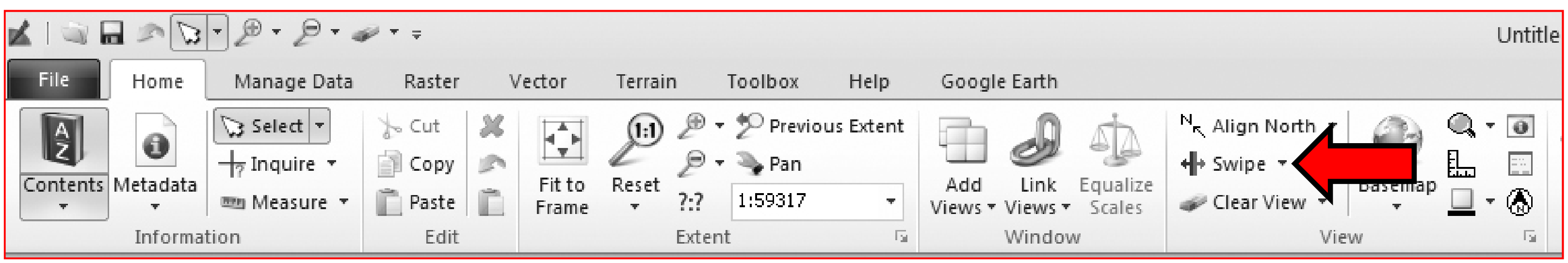


Unsupervised Classification

❑ Cluster Labeling → Attribute Table

Inspect the included LULC classes in each cluster using the SWIPE tool
(HOME tab)





Unsupervised Classification

❑ Cluster Busting

➢ It is common when performing unsupervised classification that the algorithm will generate ***n*** clusters, but you have no confidence in labeling ***q*** of them to an appropriate information class (e.g., 15 out of 30 clusters contain multiple LULC categories).

- ***Step 1*** **-** all the pixels associated with the 15 clusters that are mixed (e.g., clusters 3, 5, 6, 9, 11, 13, 15, 17, 18, 20, 22, 23, 25, 28 and 30) are all ***RECODED*** to a value of 1 and a **binary mask file** is created.

- ***Step 2*** **-** A mask program is run multiplying the binary mask file times the original imagery file. The output of the mask program is a new multiband image file consisting of **only the pixels that could not be adequately labeled during the initial unsupervised classification**.

Unsupervised Classification

❑ Cluster Busting

- ***Step 3*** **-** Perform a new unsupervised classification on the output file from Step 2, perhaps requesting the original number of clusters.

- ***Step 4*** – Repeat Steps 1 and 2, keeping all of new clusters that are dominated by a single LCLU (e.g., 7). There are likely still some mixed-class clusters, but the proportion of these is definitely smaller.

- ***Step 5*** **-** Repeat Step 3 performing a third iterative unsupervised classification. Perhaps an additional 3 good clusters are extracted.

Unsupervised Classification

❑ Cluster Busting

➢ In this hypothetical example, the final cluster map would be composed of :

- 15 good clusters from the initial classification,
- 7 good clusters from the first cluster-busting pass (recoded as values 31 to 38), and
- 3 good clusters from the second cluster-busting pass (recoded as values 39 to 41).

➢ The final cluster map file is re-assembled using a simple GIS maximum-dominate overlay function. The final cluster map is then recoded to create the final classification map.