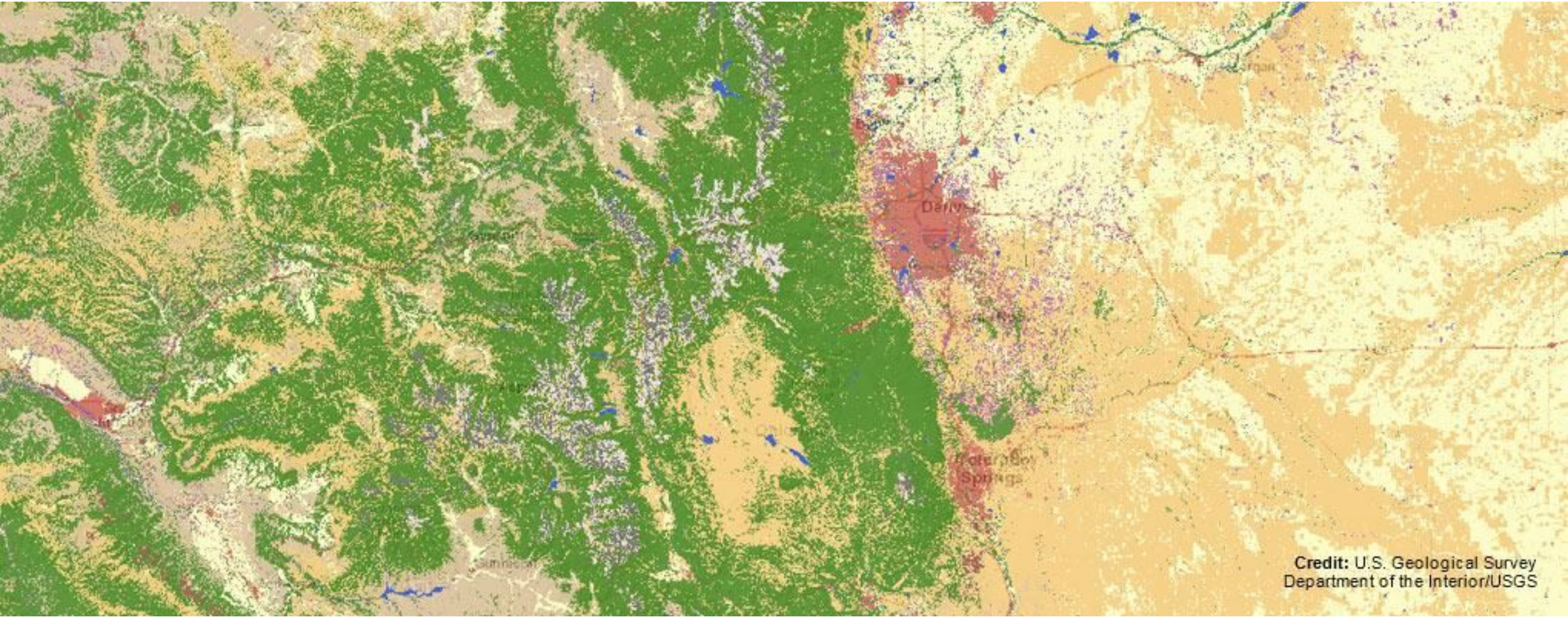
EDS 223: Geospatial Analysis & Remote Sensing
Week 9



Welcome!

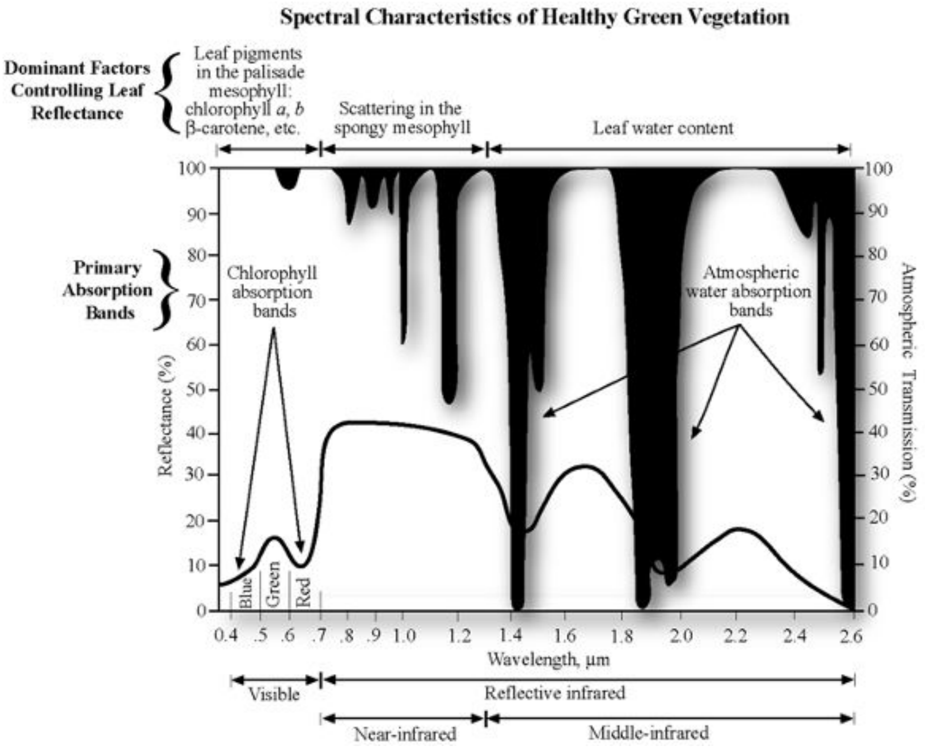
- **Assignments**
  - Assignment 4 due December 9
    - Sorry for the typos! Will distribute revised copy
  - Portfolio due December 15
    - Come to office hours for help/guidance
- **Next week**
  - Active remote sensing
  - Course wrap-up
- **Course-evaluations**

Vegetation indices

$$\textbf{NDVI} = \frac{(\text{NIR - red})}{(\text{NIR + red})}$$

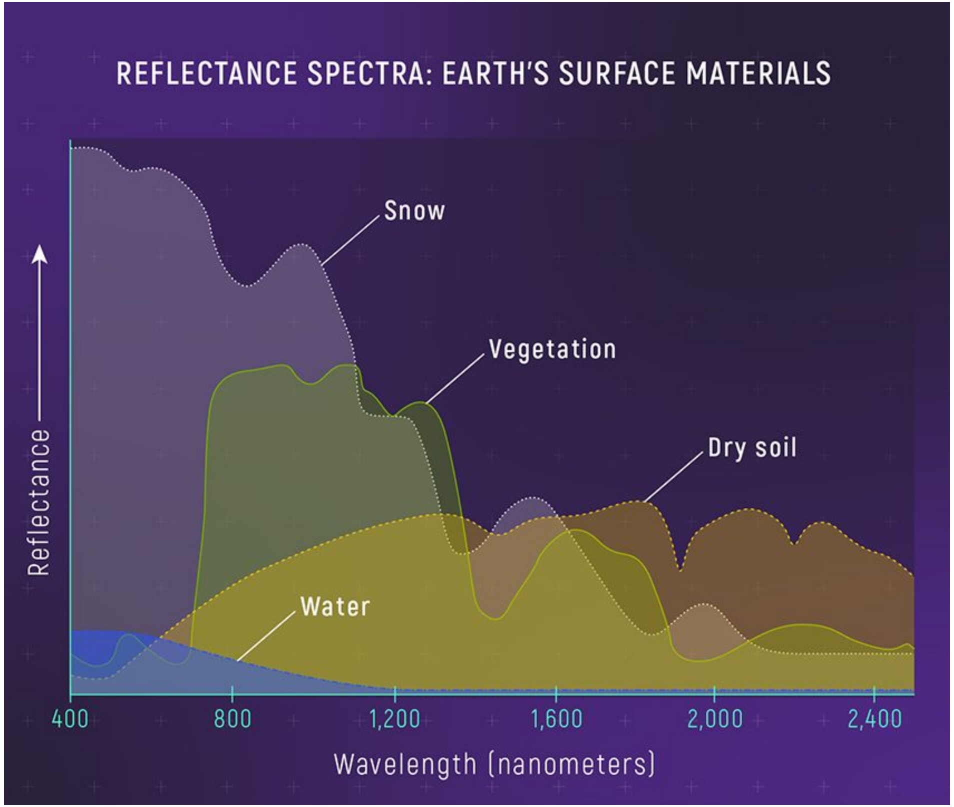
$$\textbf{NMDI} = \frac{(\text{NIR - SWIR})}{(\text{NIR + SWIR})}$$

Factors controlling leaf reflectance



Source: Jensen 2007

Image classification

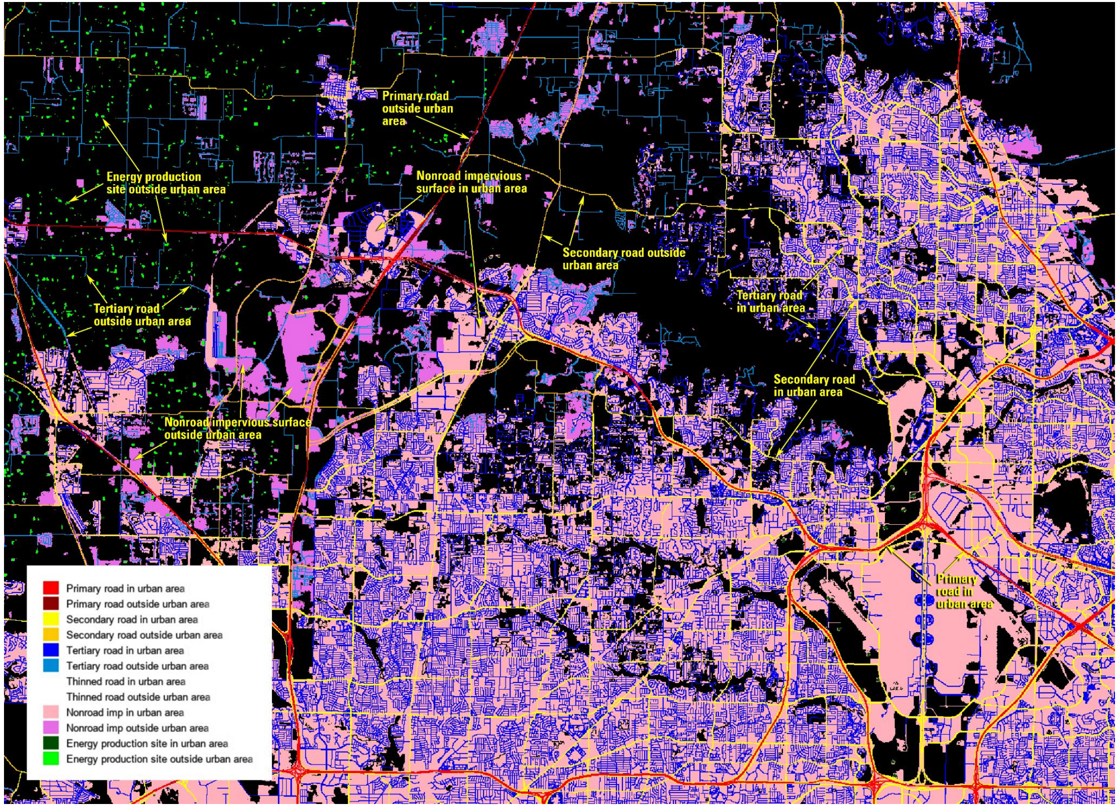


Source: NASA, Leah Hustak

Land cover classification

Source: USGS

Land cover classification



Source: USGS

Image classification

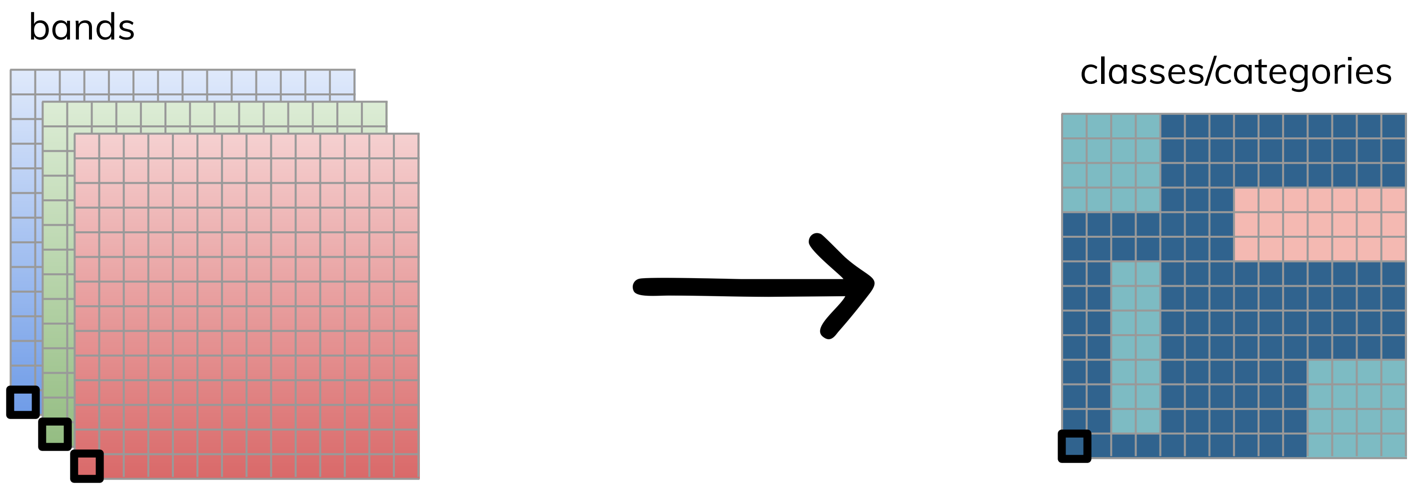
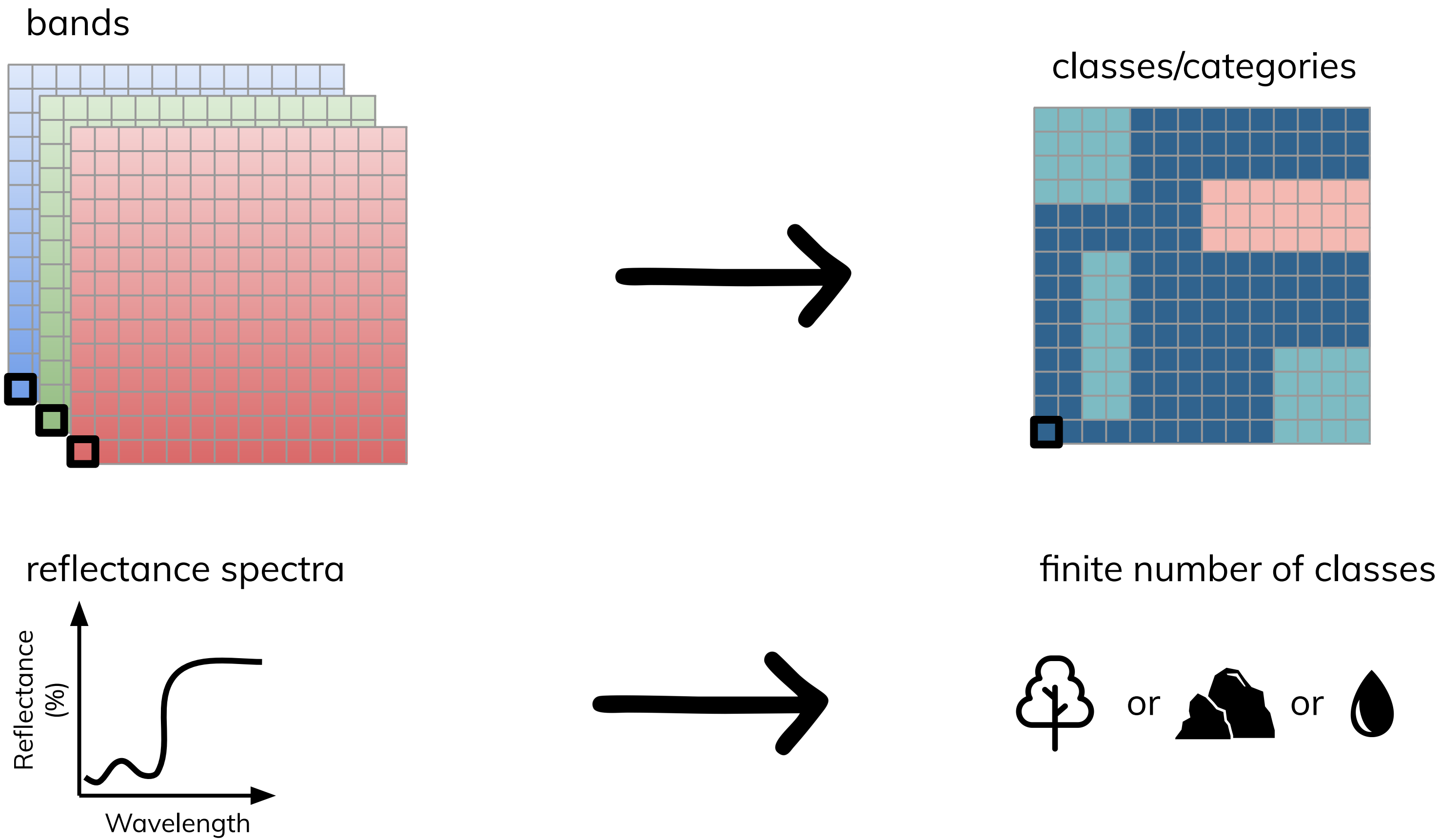
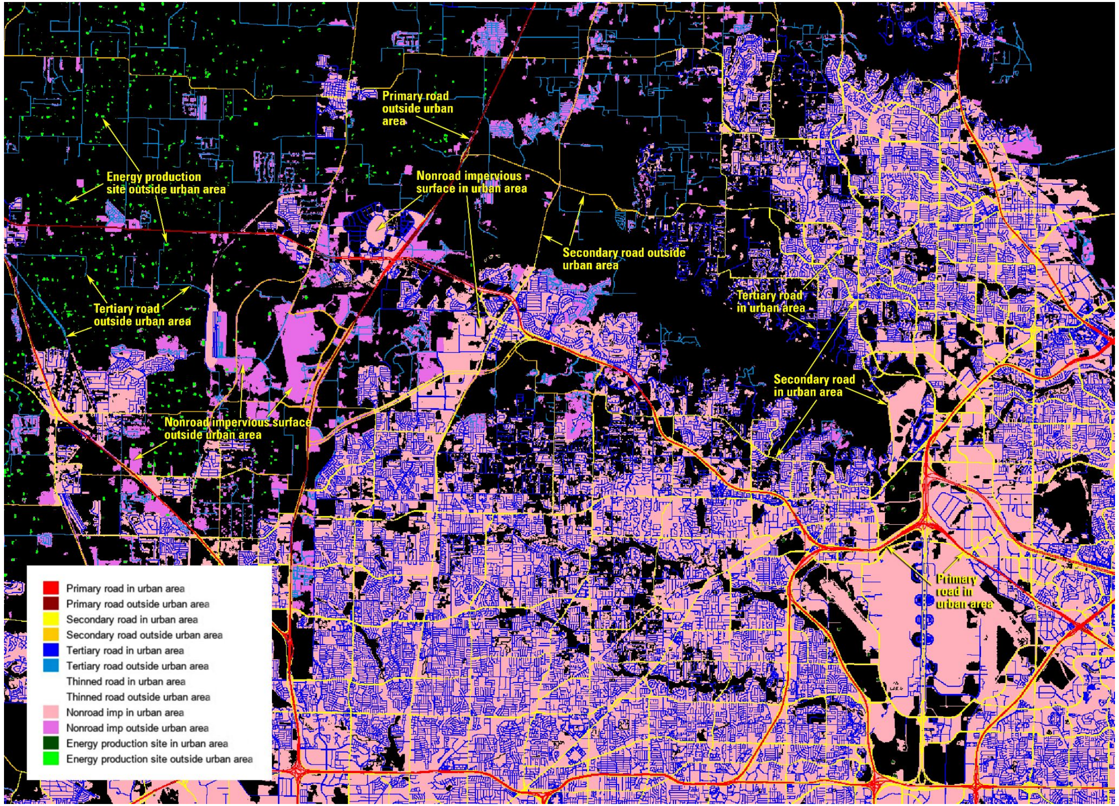


Image classification

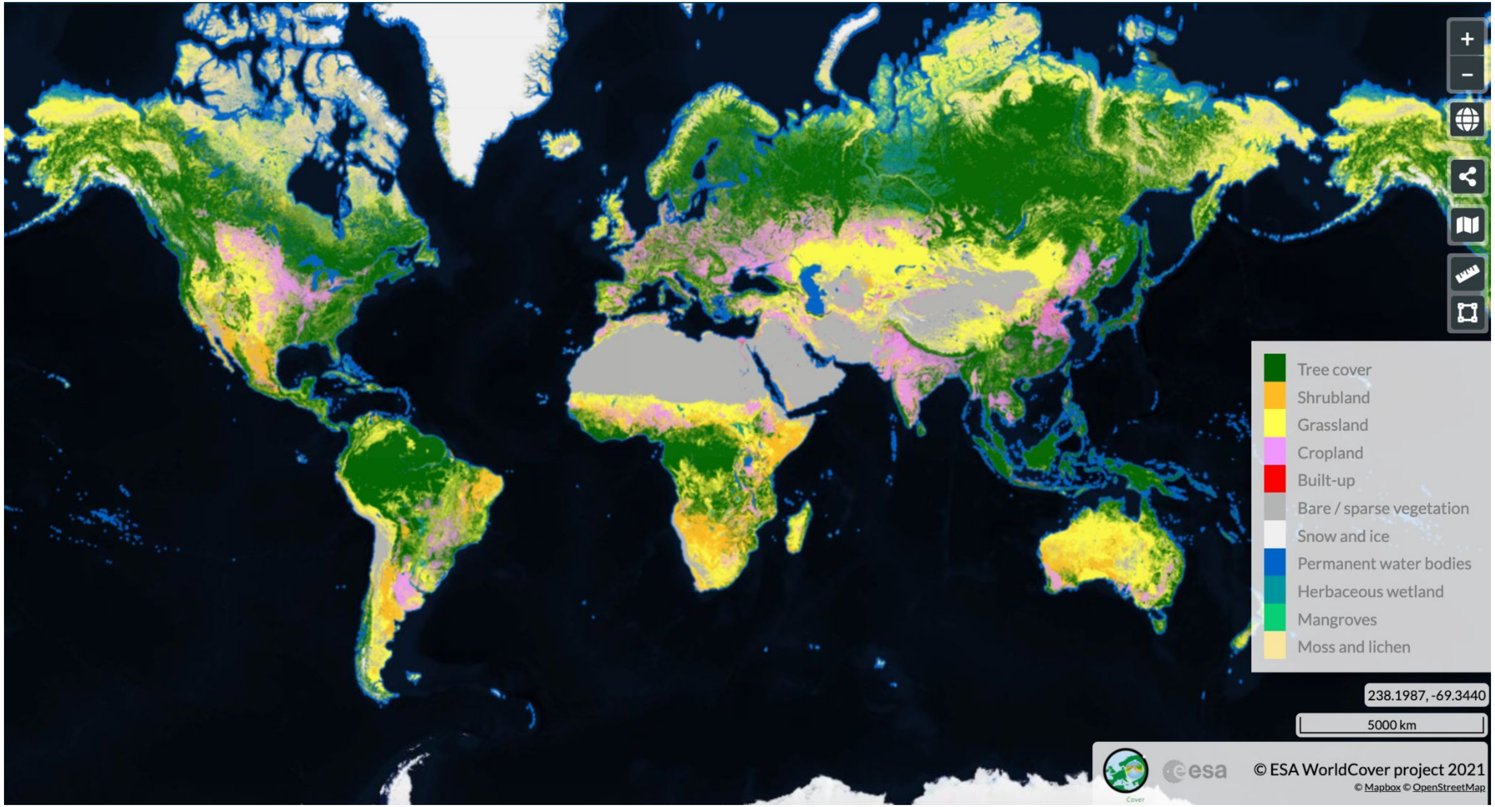


Land cover classification

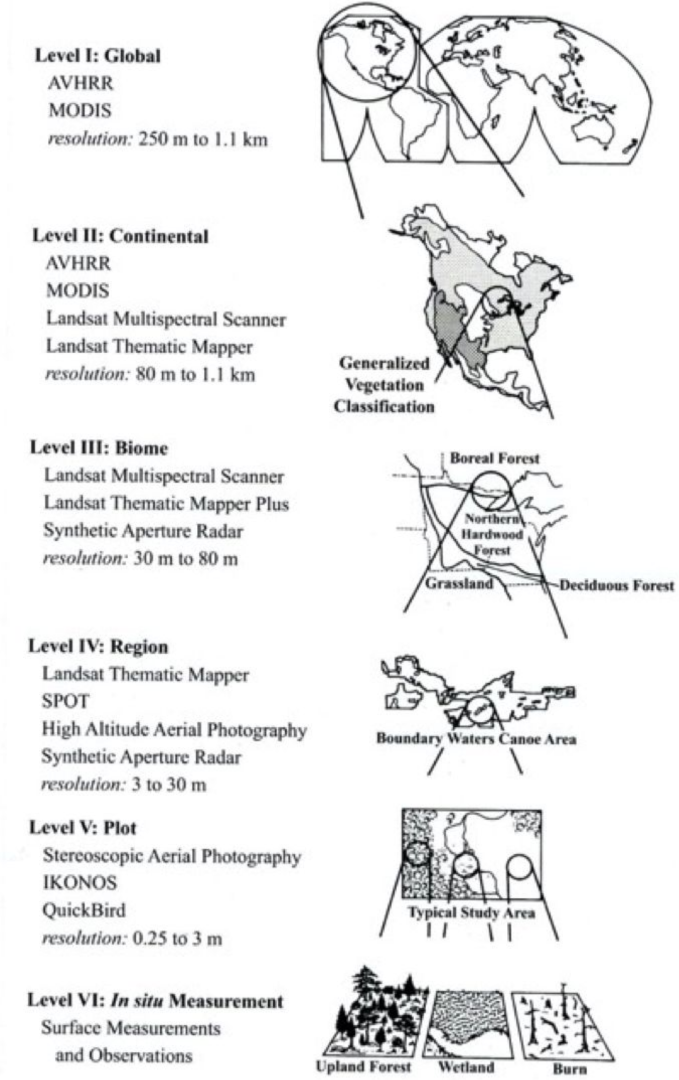


Source: USGS

Land cover classification



Classification scales



Source: Jensen 2007

Land cover vs. land use

Land cover	Land use

Land cover vs. land use

Land cover	Land use
Refers to the type of natural and artificial materials present on a landscape	Refers to the human use of landscapes

Land cover vs. land use

Land cover	Land use
Refers to the type of natural and artificial materials present on a landscape	Refers to the human use of landscapes
E.g. forest, sand, water, cement	E.g. protected area, industrial, residential

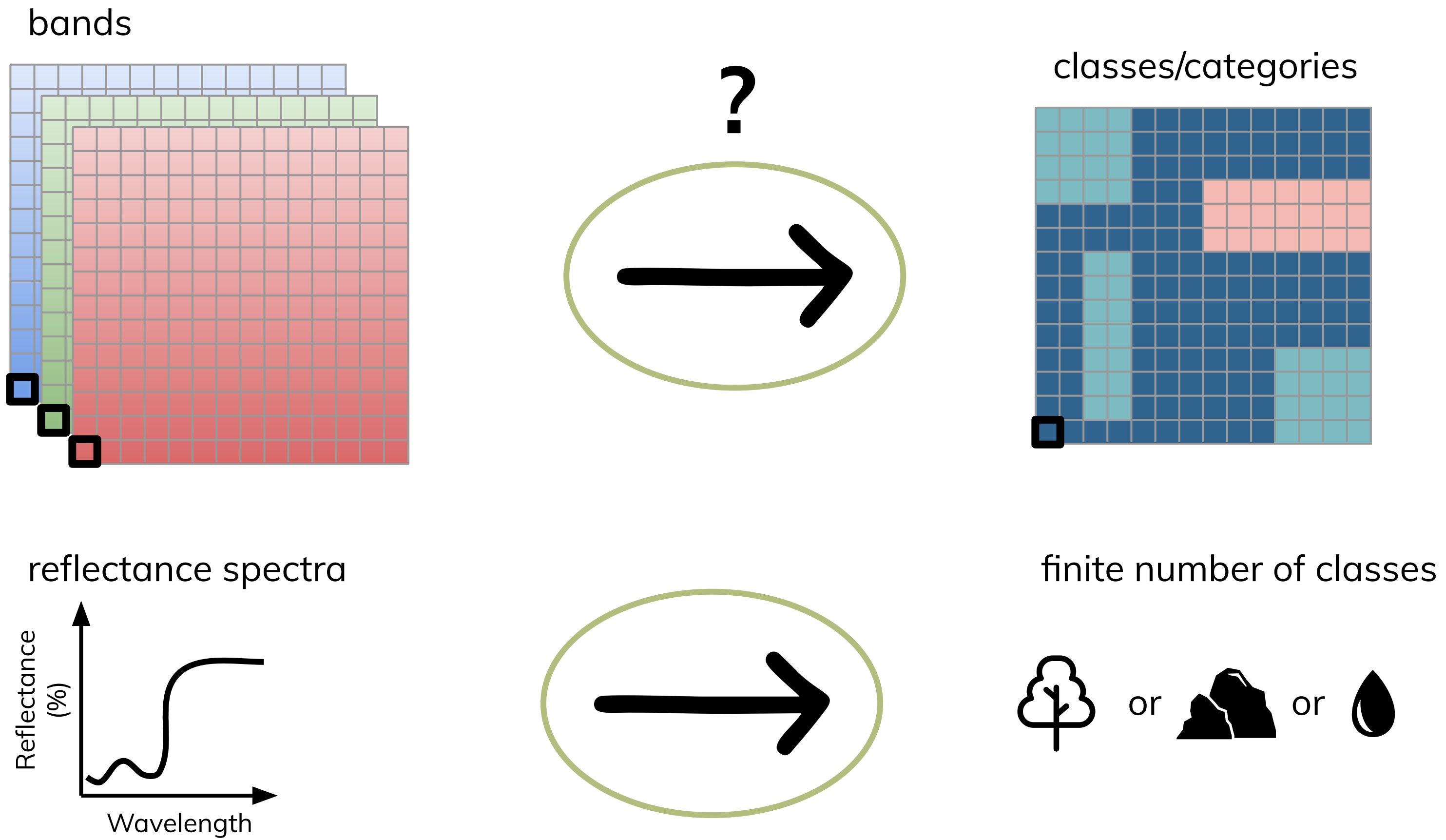
Land cover vs. land use

Land cover	Land use
Refers to the type of natural and artificial materials present on a landscape	Refers to the human use of landscapes
E.g. forest, sand, water, cement	E.g. protected area, industrial, residential
Able to observe	Abstract/intangible, requires deductive reasoning

Land cover vs. land use

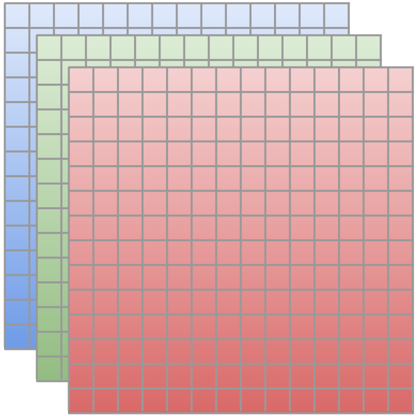
Land cover	Land use
Refers to the type of natural and artificial materials present on a landscape	Refers to the human use of landscapes
E.g. forest, sand, water, cement	E.g. protected area, industrial, residential
Able to observe	Abstract/intangible, requires deductive reasoning

Land cover classification



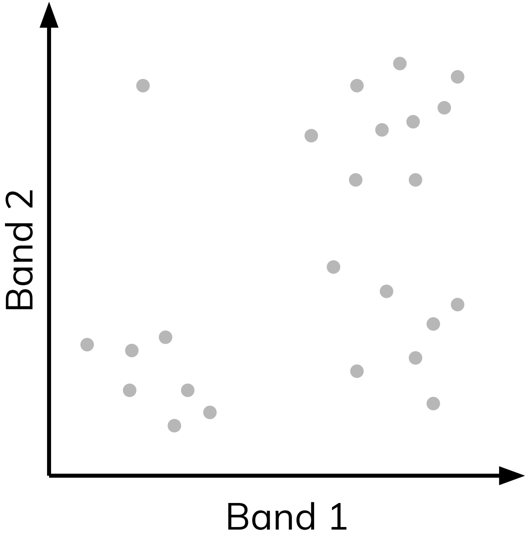
Land cover classification

Geographic space



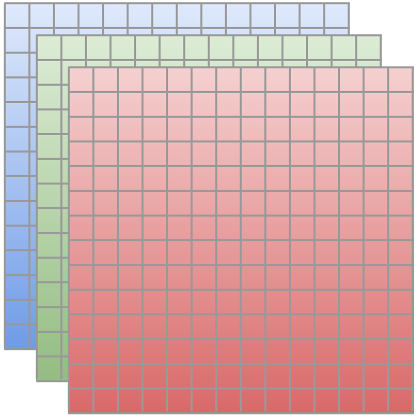
Feature space

Points are pixels



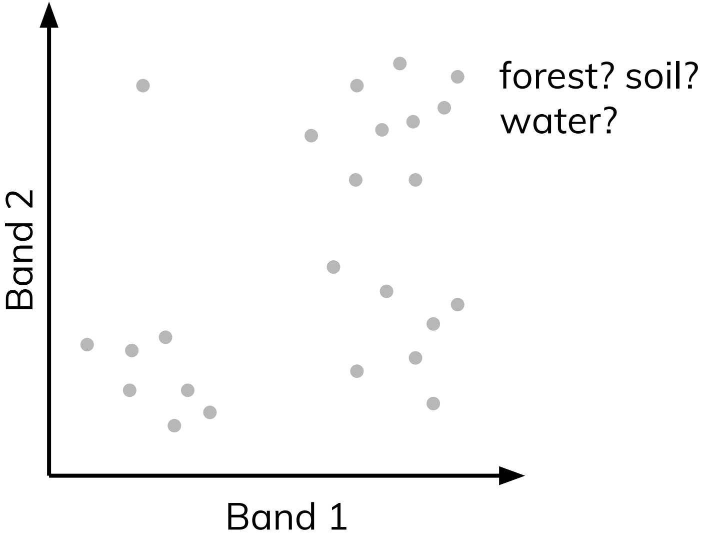
Land cover classification

Geographic space



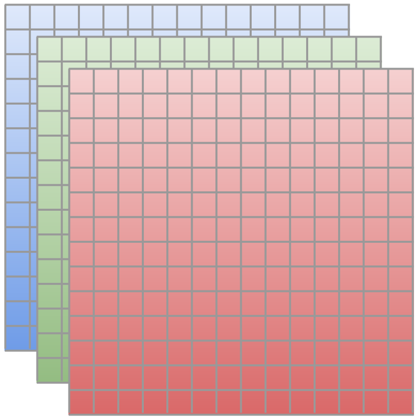
Feature space

Points are pixels



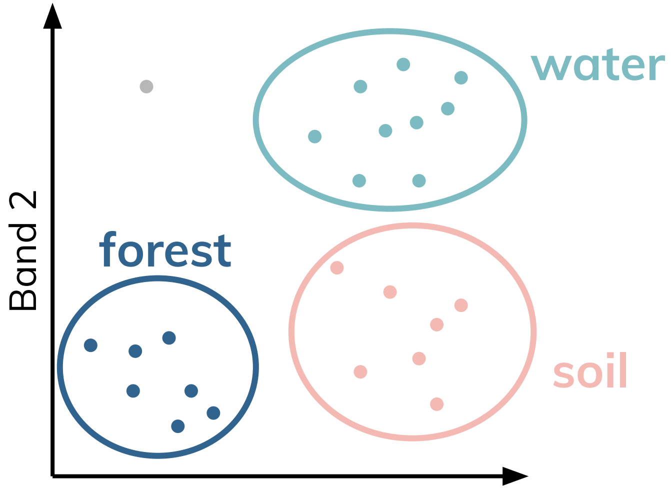
Land cover classification

Geographic space

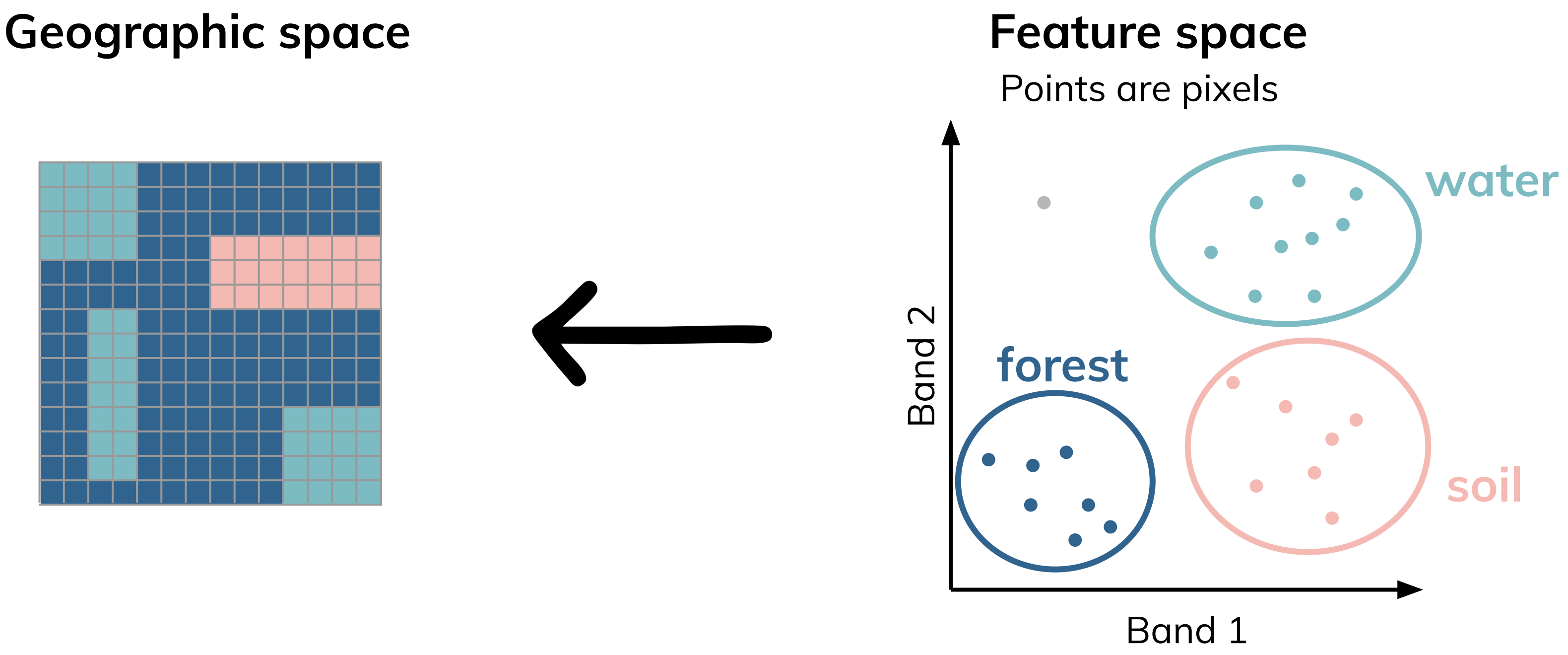


Feature space

Points are pixels

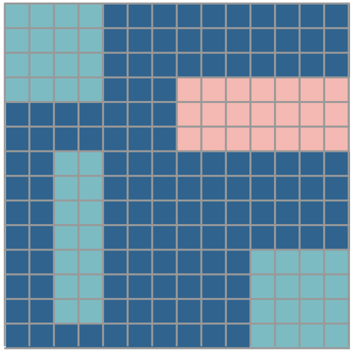


Land cover classification



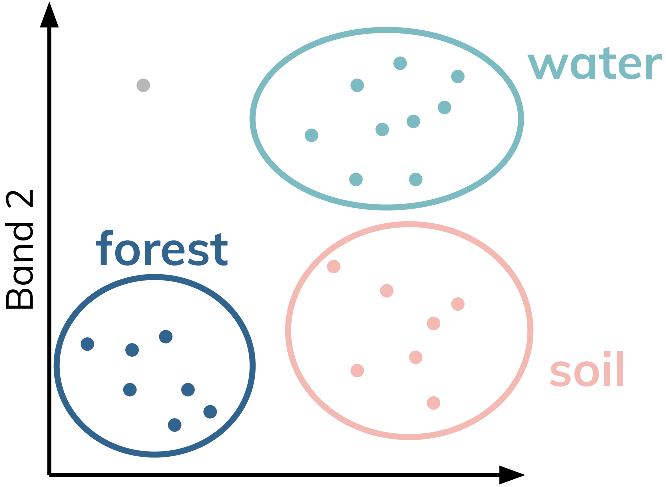
Land cover classification

Geographic space



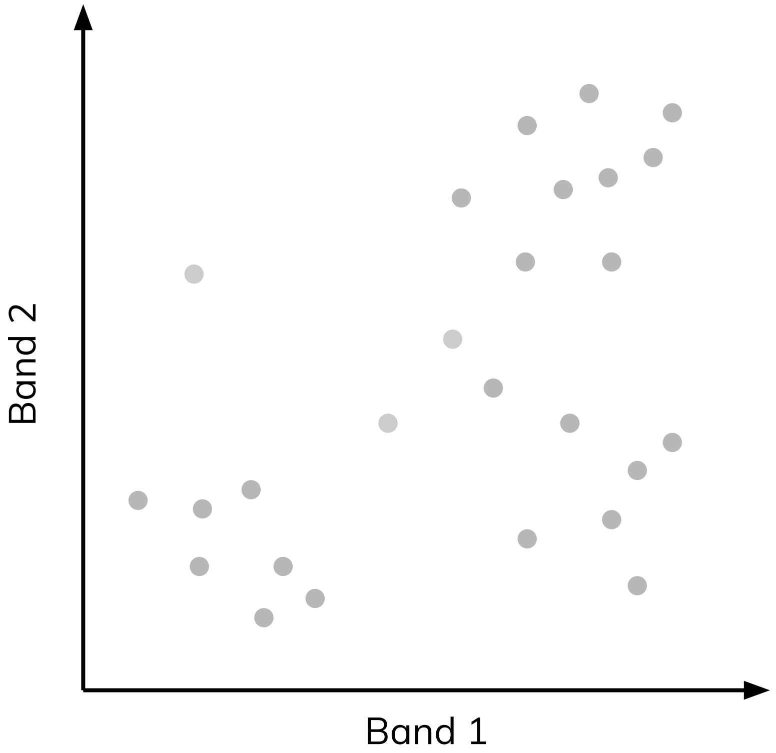
Feature space

Points are pixels



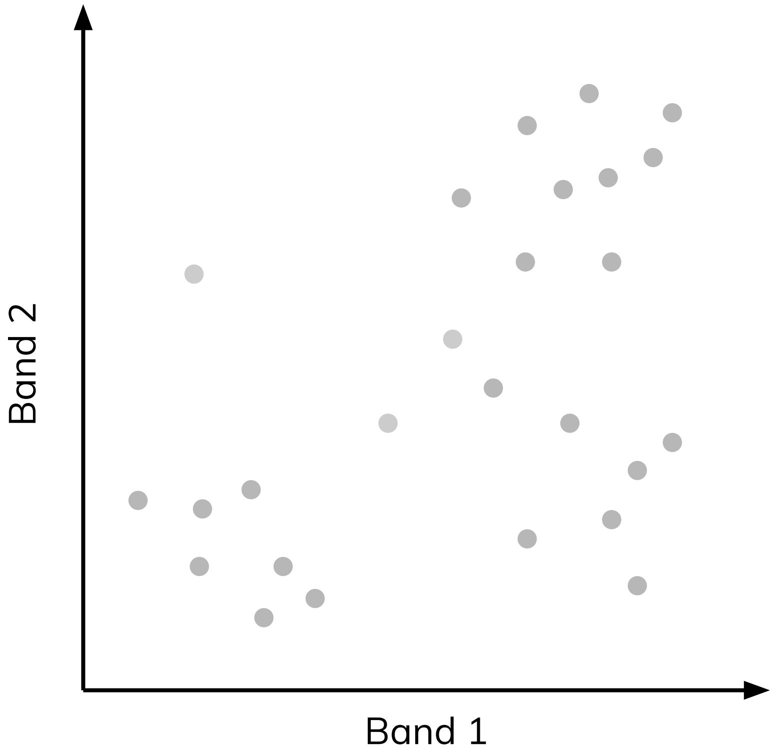
Lots of ways to assign pixels to groups!

How to group pixels into land cover types



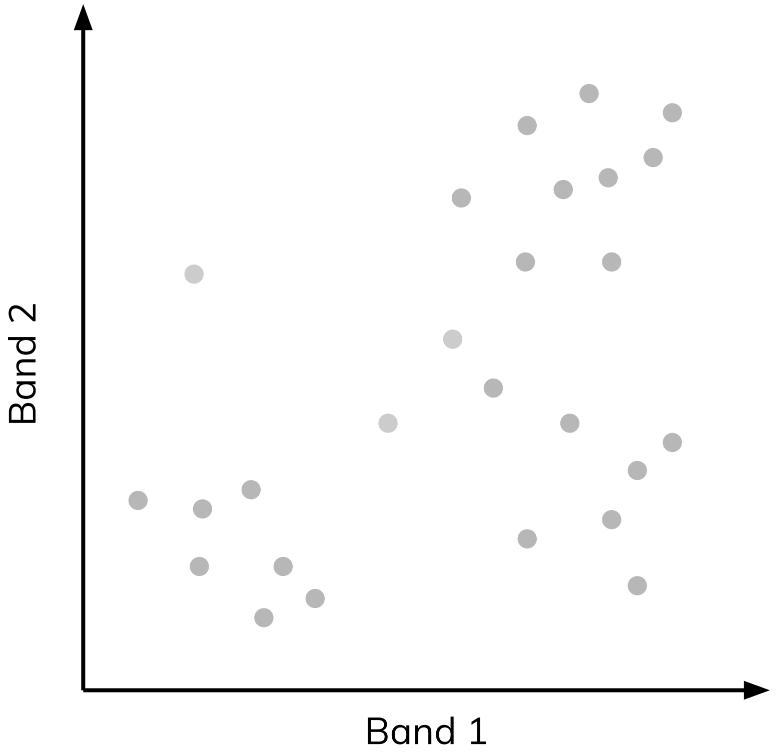
How to group pixels into land cover types

- Pick a number of groups



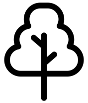
How to group pixels into land cover types

- Pick a number of groups

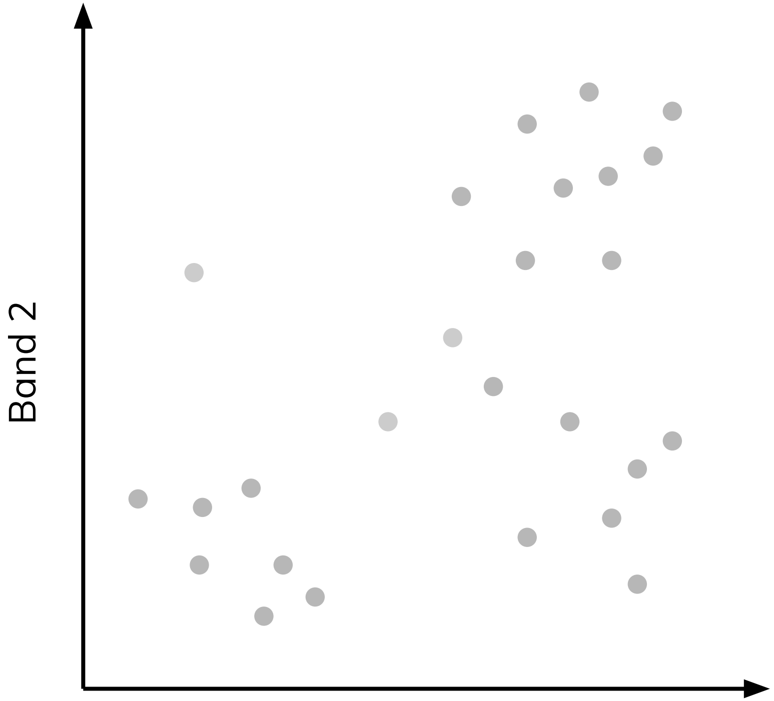


How to group pixels into land cover types

- Pick a number of groups

 or  or 

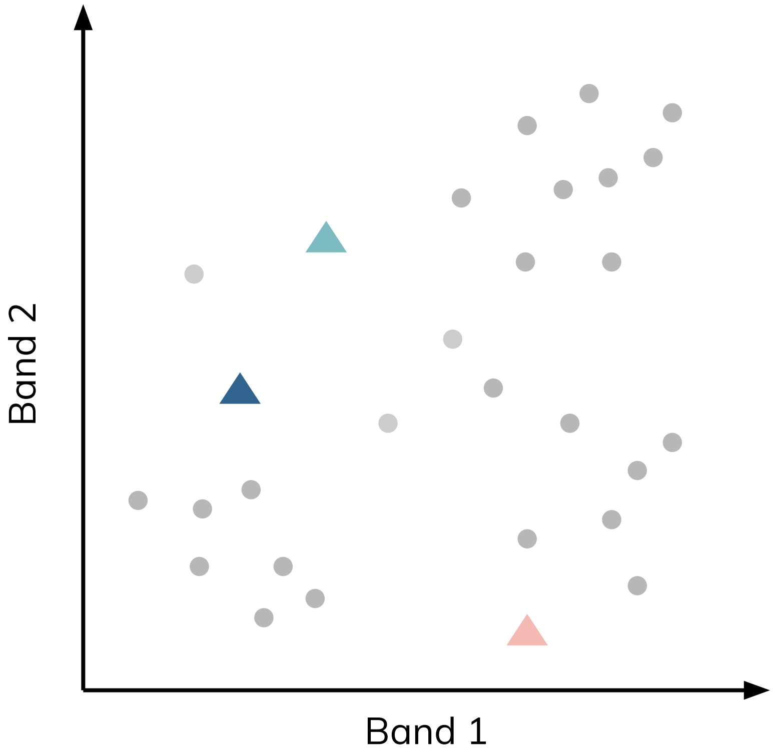
- Make a guess about where those groups are in feature space



How to group pixels into land cover types

- Pick a number of groups

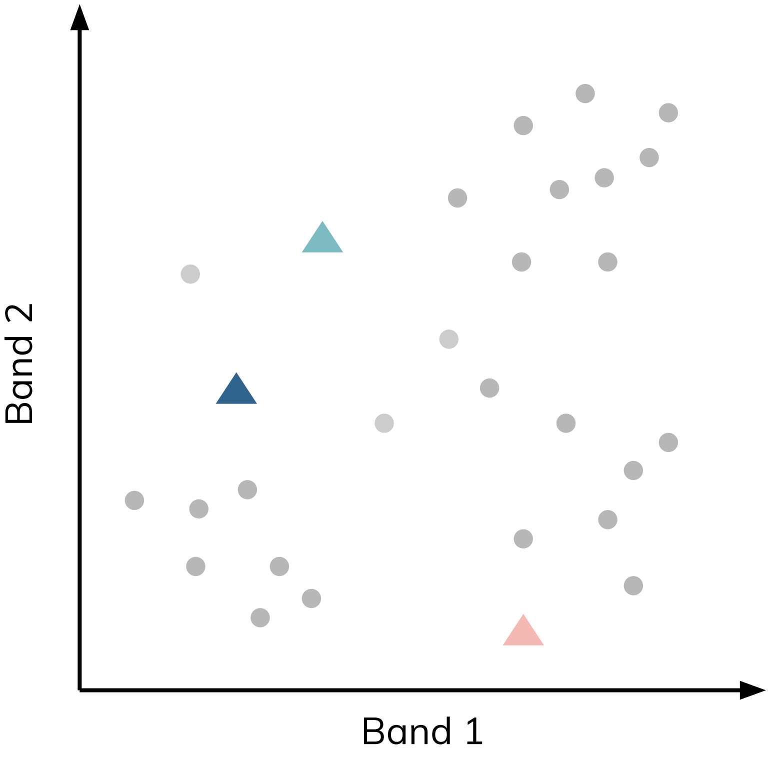
- Make a guess about where those groups are in feature space



How to group pixels into land cover types

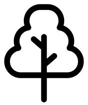
- Pick a number of groups

- Make a guess about where those groups are in feature space
- Assign each point to the closest group

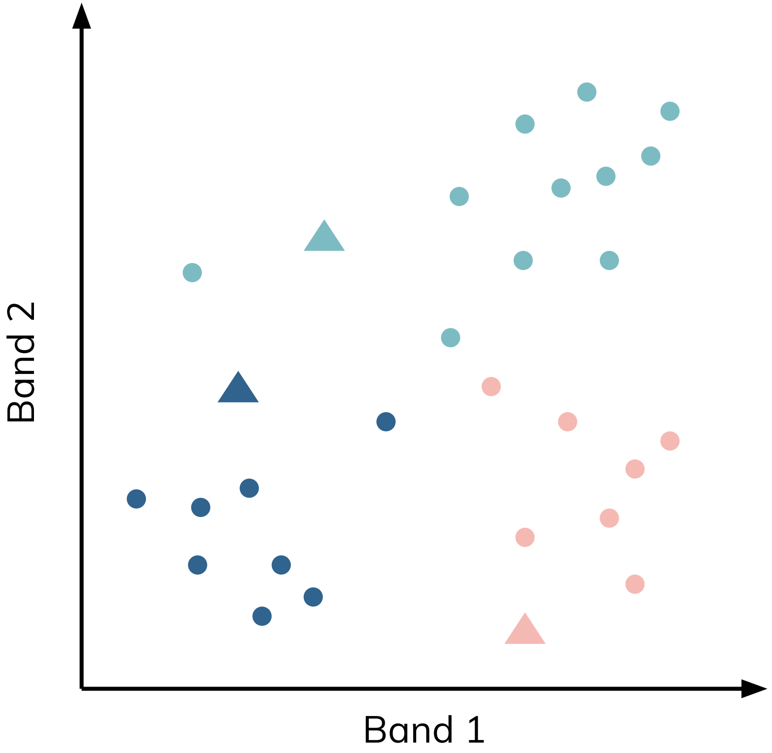


How to group pixels into land cover types

- Pick a number of groups

 or  or 

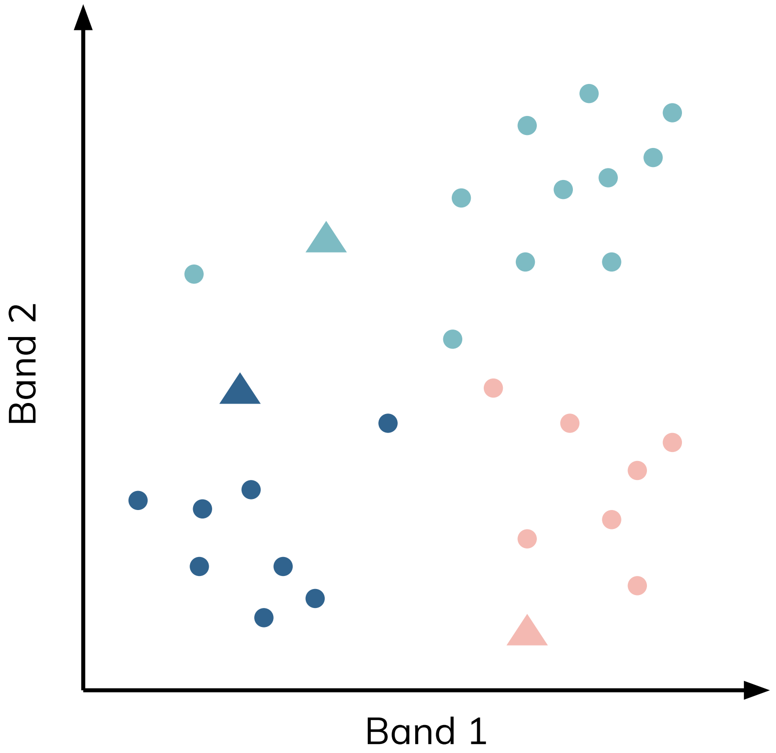
- Make a guess about where those groups are in feature space
- Assign each point to the closest group



How to group pixels into land cover types

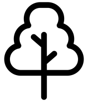
- Pick a number of groups

- Make a guess about where those groups are in feature space
- Assign each point to the closest group
- Move group centers to better represent groups

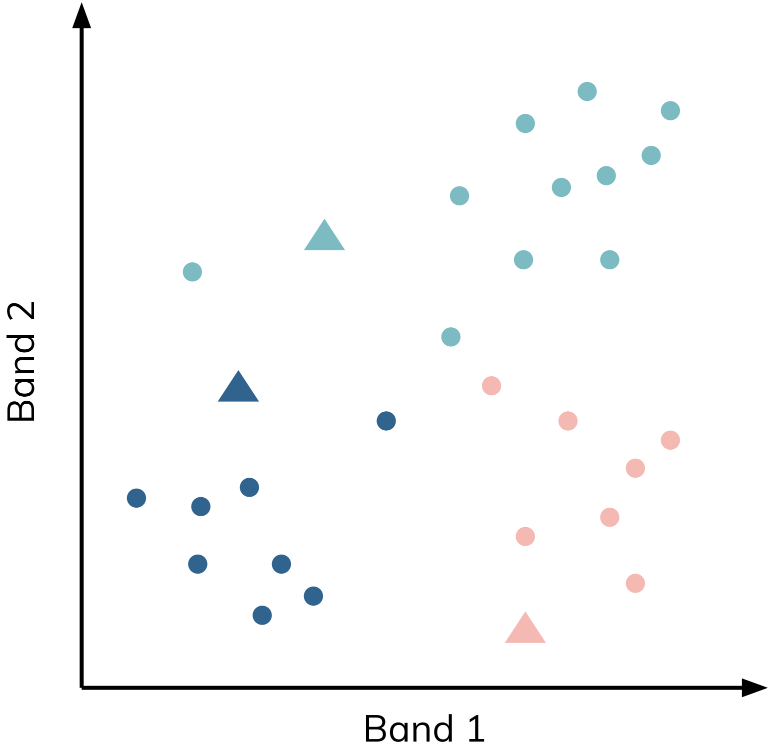


How to group pixels into land cover types

- Pick a number of groups

 or  or 

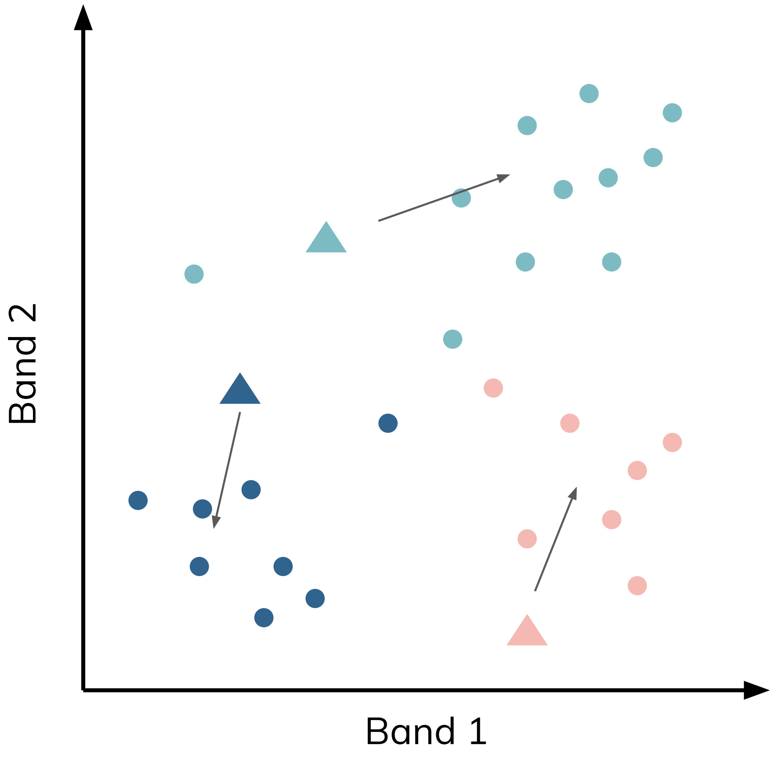
- Make a guess about where those groups are in feature space
- Assign each point to the closest group
- Move group centers to better represent groups
    - Use the mean!



How to group pixels into land cover types

- Pick a number of groups

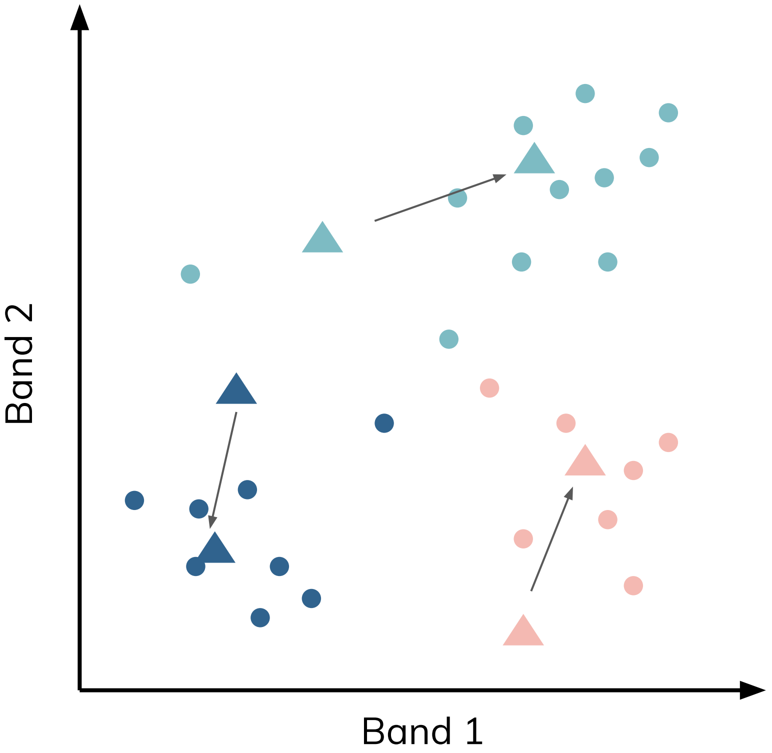
- Make a guess about where those groups are in feature space
- Assign each point to the closest group
- Move group centers to better represent groups
    - Use the mean!



How to group pixels into land cover types

- Pick a number of groups

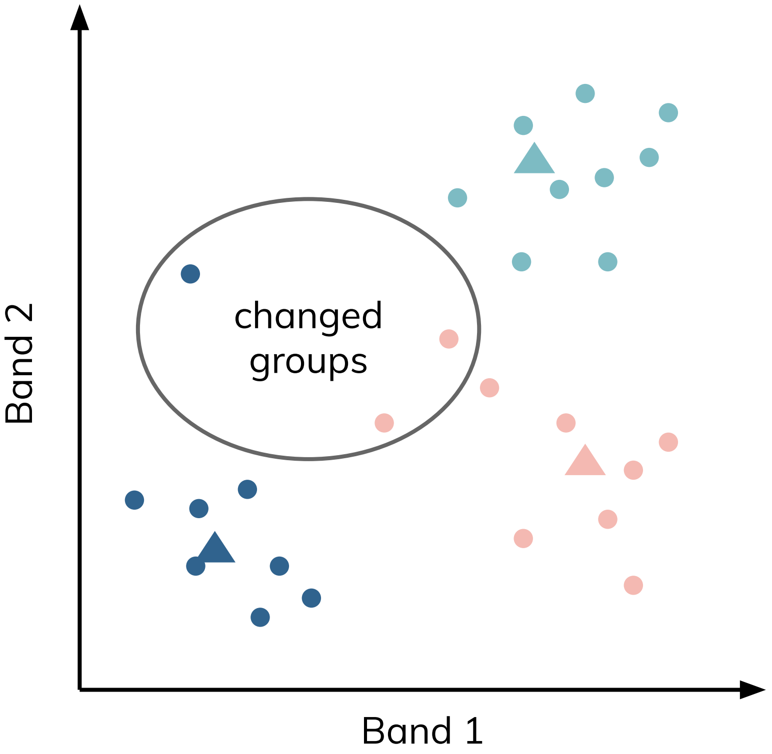
- Make a guess about where those groups are in feature space
- Assign each point to the closest group
- Move group centers to better represent groups
    - Use the mean!



How to group pixels into land cover types

- Pick a number of groups

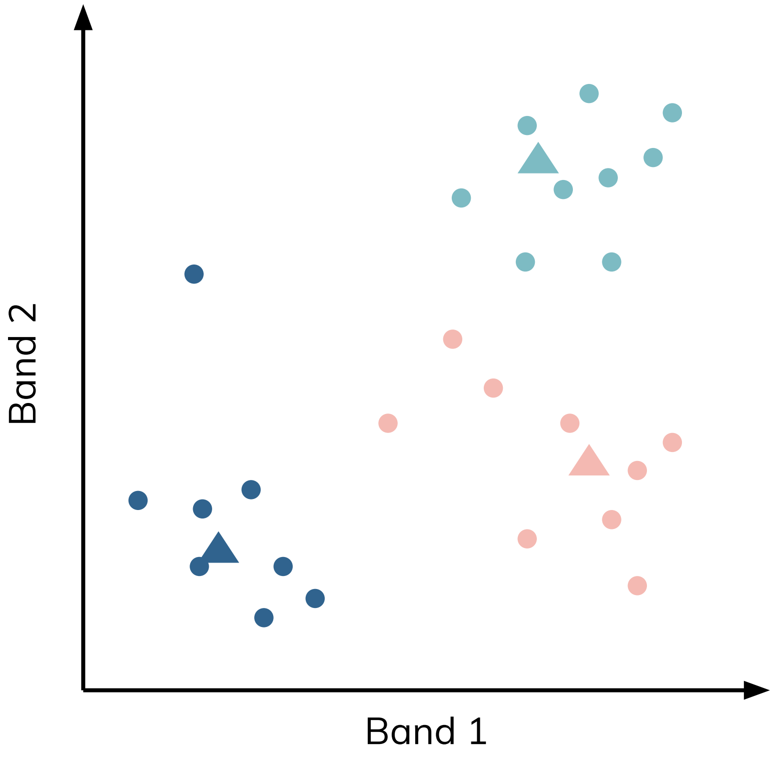
- Make a guess about where those groups are in feature space
- Assign each point to the closest group
- Move group centers to better represent groups
  - Use the mean!
- Update groups



How to group pixels into land cover types

- Pick a number of groups

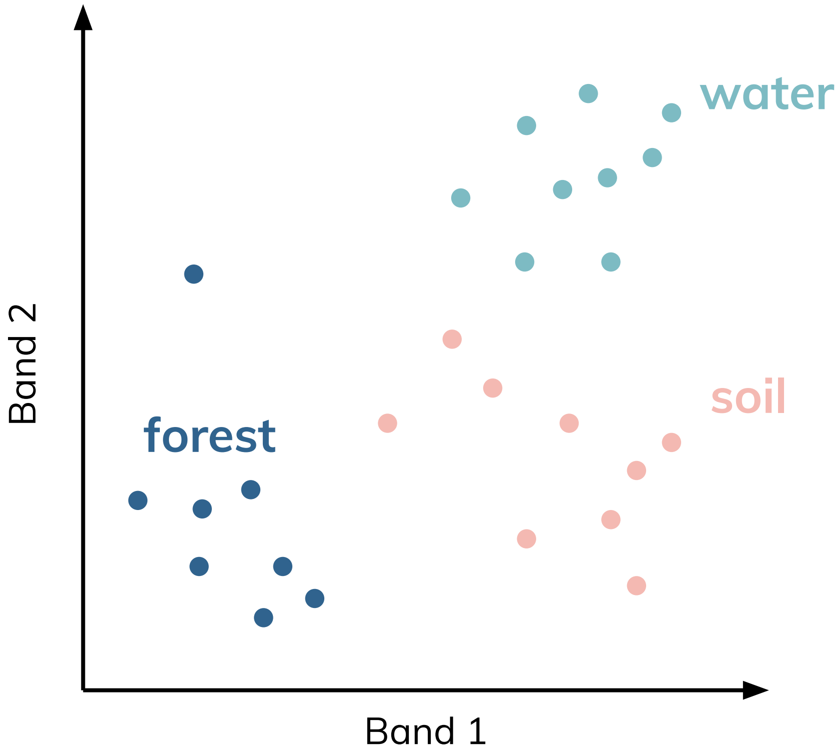
- Make a guess about where those groups are in feature space
- Assign each point to the closest group
- Move group centers to better represent groups
  - Use the mean!
- Update groups
- Keep going until distances are minimized



How to group pixels into land cover types

- Pick a number of groups

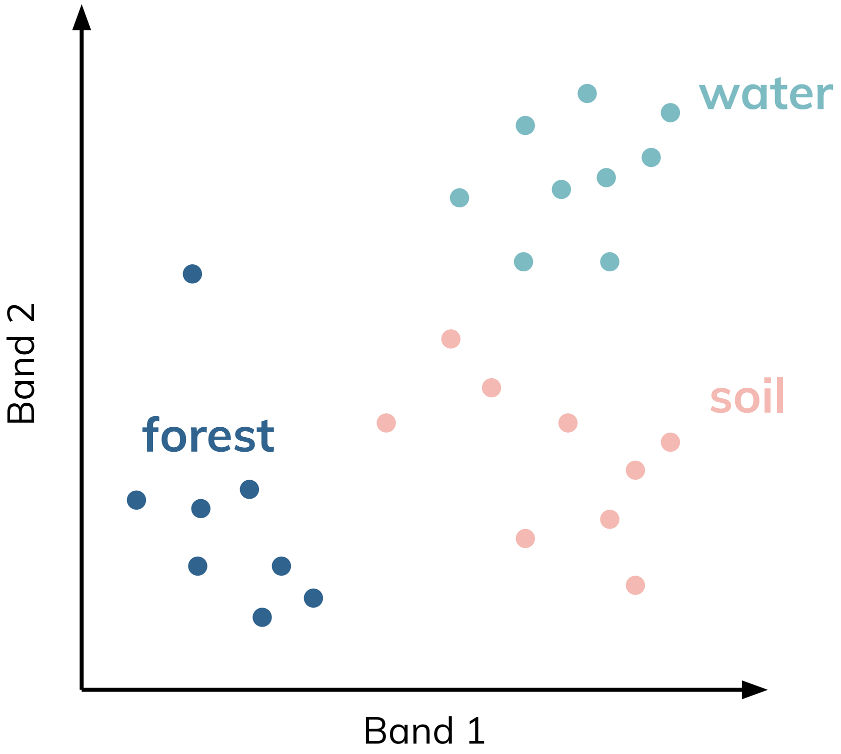
- Make a guess about where those groups are in feature space
- Assign each point to the closest group
- Move group centers to better represent groups
    - Use the mean!
- Update groups
- Keep going until distances are minimized
- Figure out what the groups are



k-means clustering

- Pick a number of groups

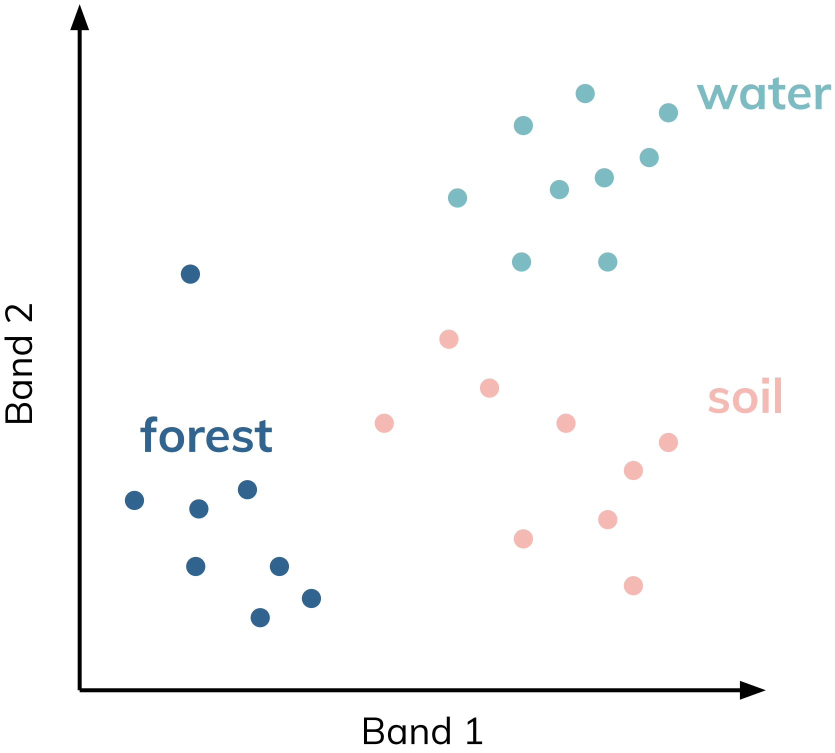
- Make a guess about where those groups are in feature space
- Assign each point to the closest group
- Move group centers to better represent groups
    - Use the mean!
- Update groups
- Keep going until distances are minimized
- Figure out what the groups are



k-means clustering

- Pick a number of groups

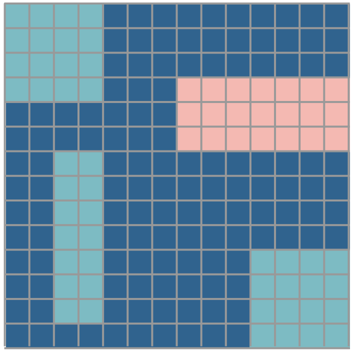
- Make a guess about where those groups are in feature space
- Assign each point to the closest group
- Move group centers to better represent groups
    - Use the mean!
- Update groups
- Keep going until distances are minimized
- Figure out what the groups are

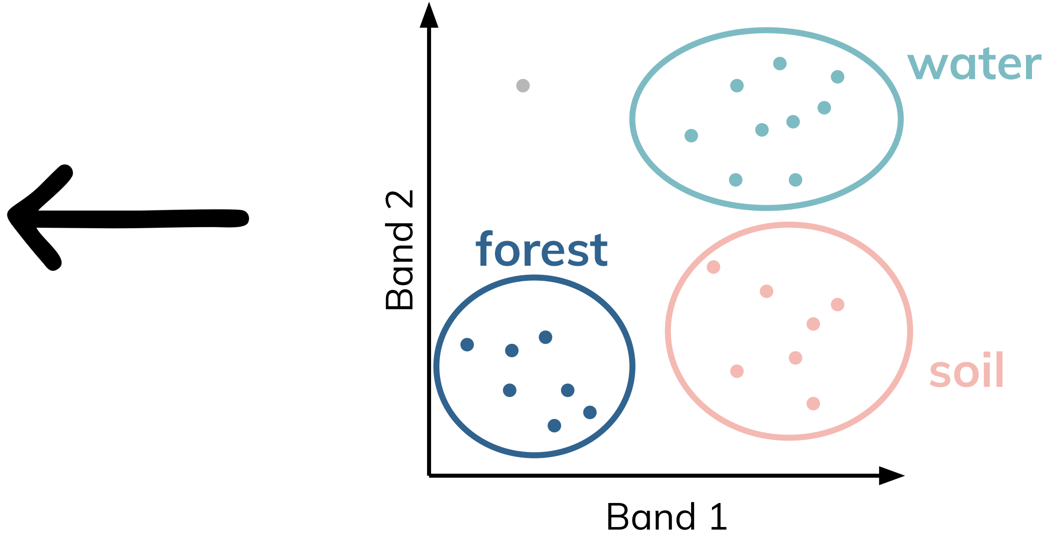


k is the number of groups (or clusters)

clusters are based on the group mean

k-means clustering

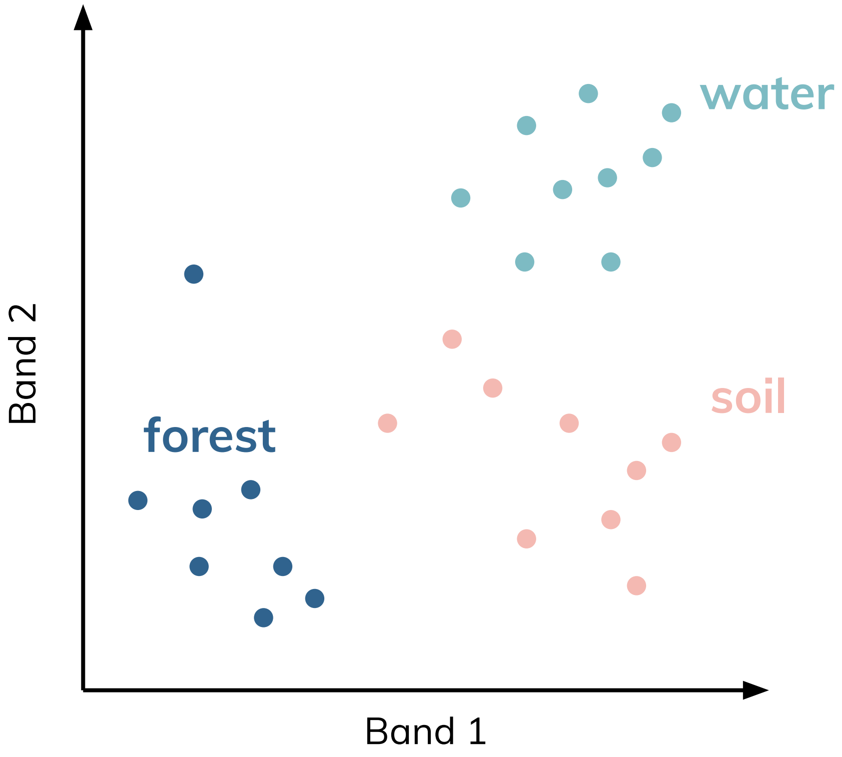




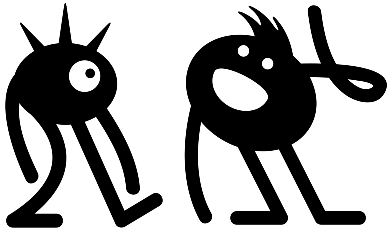
k-means clustering

- Pick a number of groups

- Make a guess about where those groups are in feature space
- Assign each point to the closest group
- Move group centers to better represent groups
    - Use the mean!
- Update groups
- Keep going until distances are minimized
- Figure out what the groups are



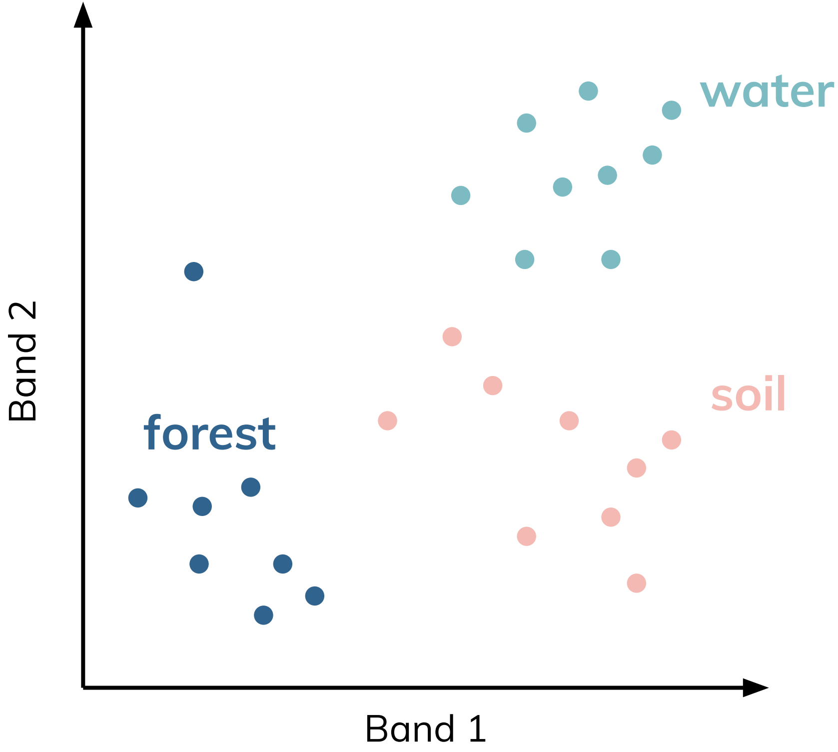
Pros/Cons



k-means clustering

- Pick a number of groups

- Make a guess about where those groups are in feature space
- Assign each point to the closest group
- Move group centers to better represent groups
  - Use the mean!
- Update groups
- Keep going until distances are minimized
- Figure out what the groups are



Pros

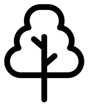
- Only needed remote sensing data
- Explored how similar different areas are

Cons

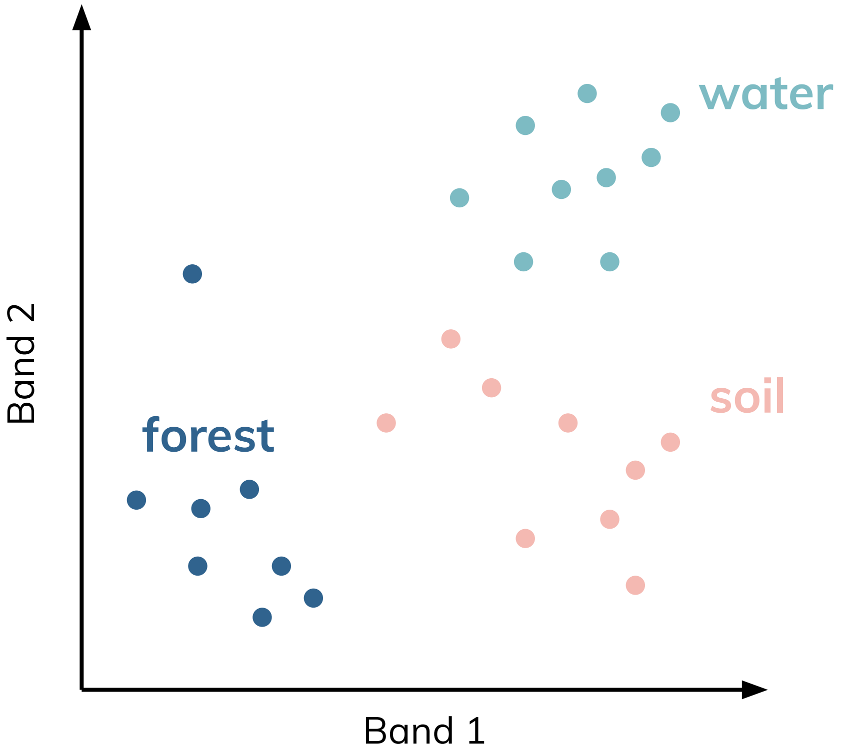
- Clusters might not always work out this well
- Number of groups was arbitrary
- Starting guess on clusters might impact results
- Needed to figure out what the clusters meant

k-means clustering

- Pick a number of groups

 or  or 

- Make a guess about where those groups are in feature space
- Assign each point to the closest group
- Move group centers to better represent groups
    - Use the mean!
- Update groups
- Keep going until distances are minimized
- Figure out what the groups are



Pros

- **Only needed remote sensing data**
- Explored how similar different areas are

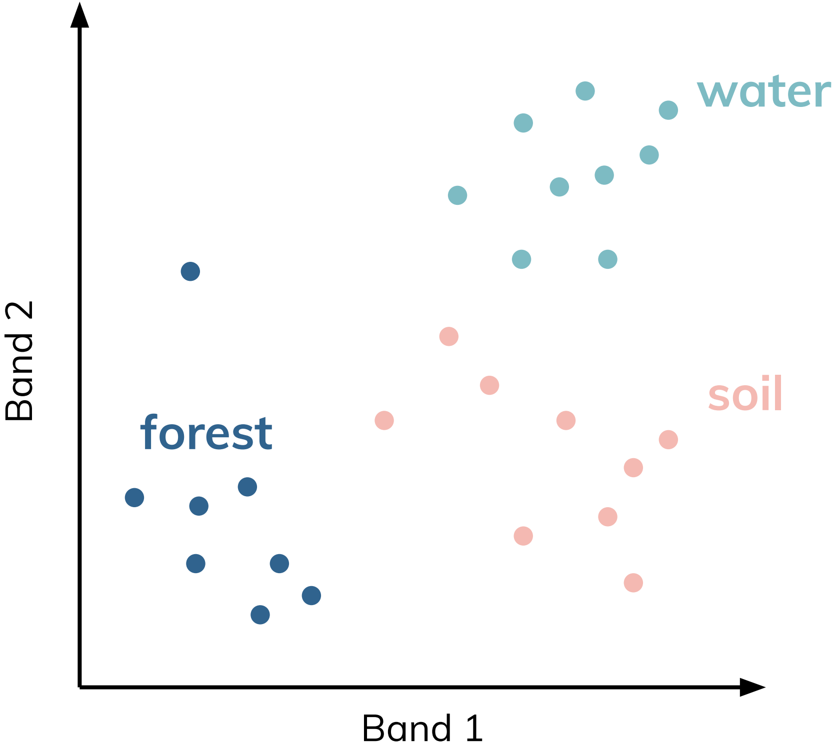
Cons

- Clusters might not always work out this well
- Number of groups was arbitrary
- Starting guess on clusters might impact results
- Needed to figure out what the clusters meant

k-means clustering

- Pick a number of groups

- Make a guess about where those groups are in feature space
- Assign each point to the closest group
- Move group centers to better represent groups
  - Use the mean!
- Update groups
- Keep going until distances are minimized
- Figure out what the groups are



Pros

- **Only needed remote sensing data**
- Explored how similar different areas are

Cons

- Clusters might not always work out this well
- Number of groups was arbitrary
- Starting guess on clusters might impact results
- **Needed to figure out what the clusters meant**

Image classification

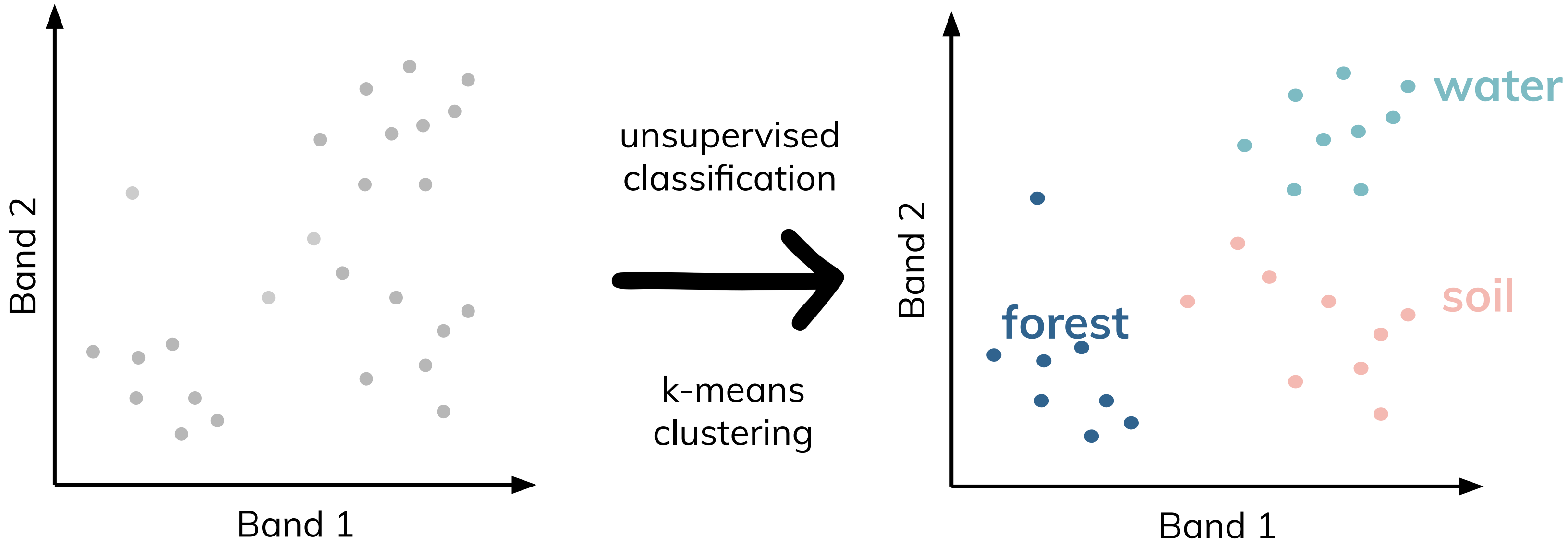


Image classification

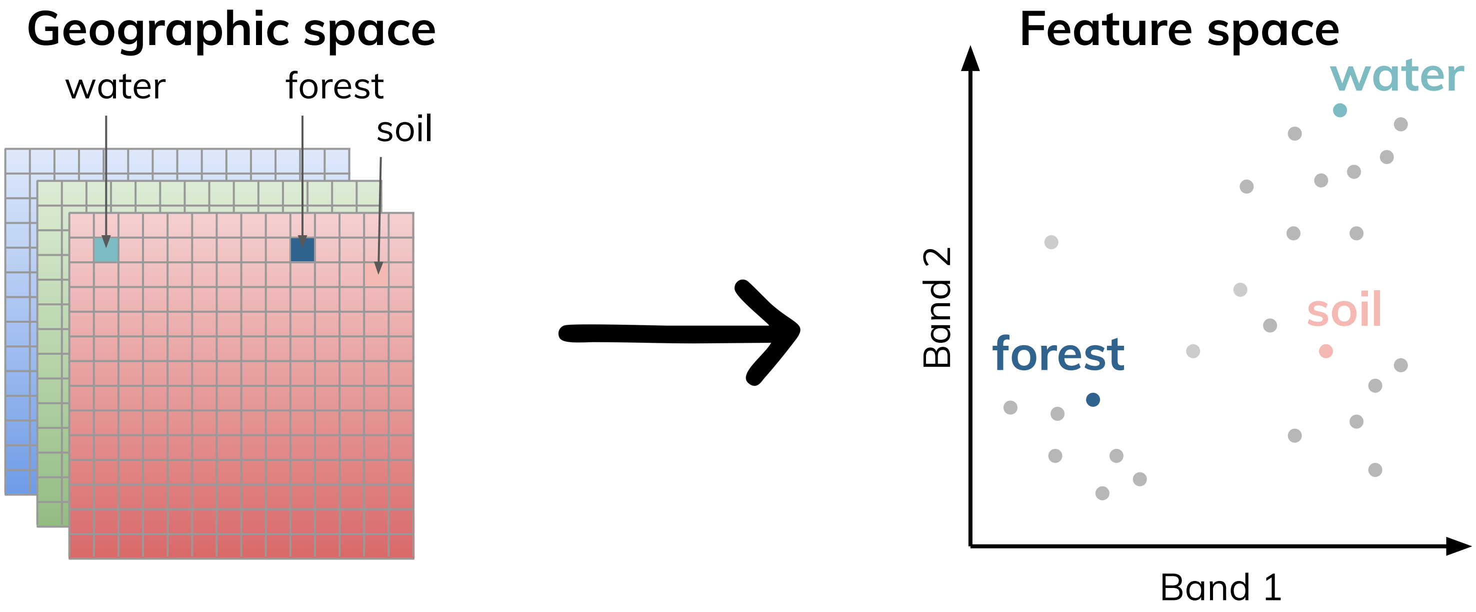
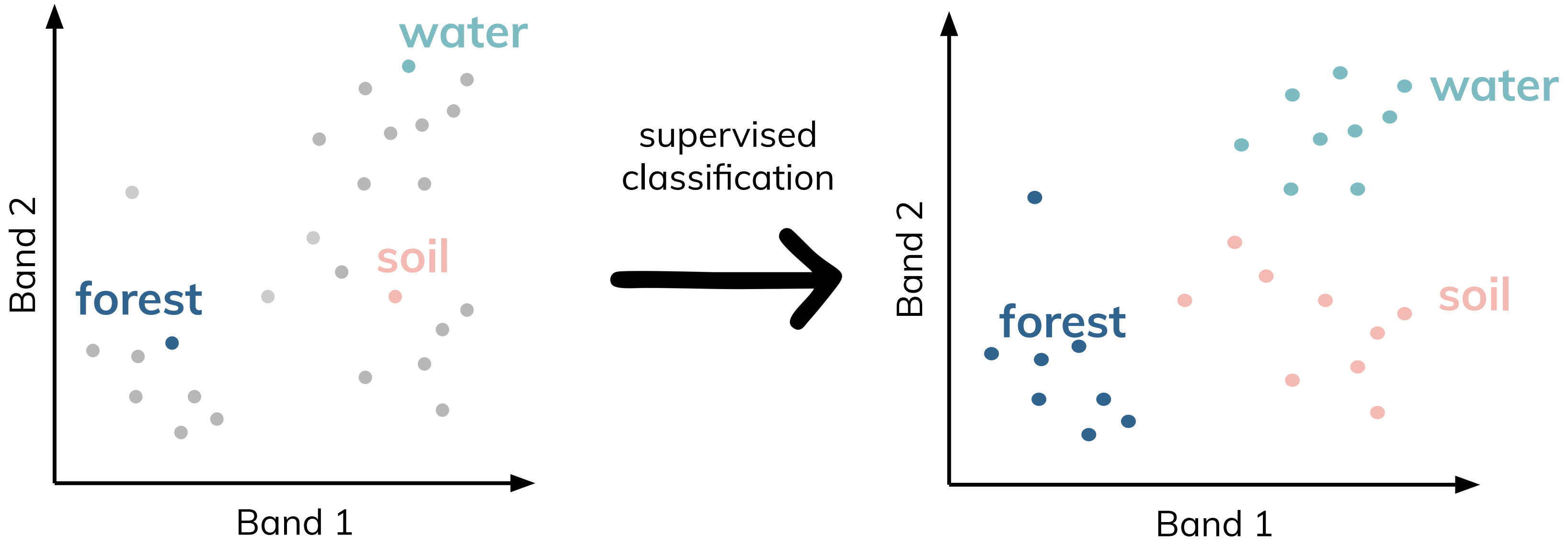
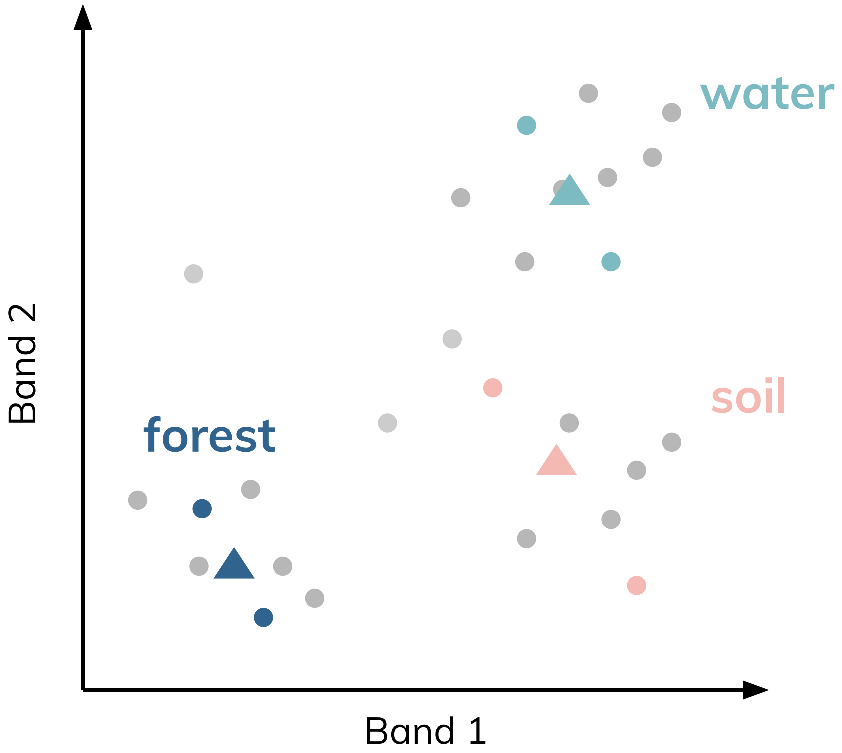


Image classification



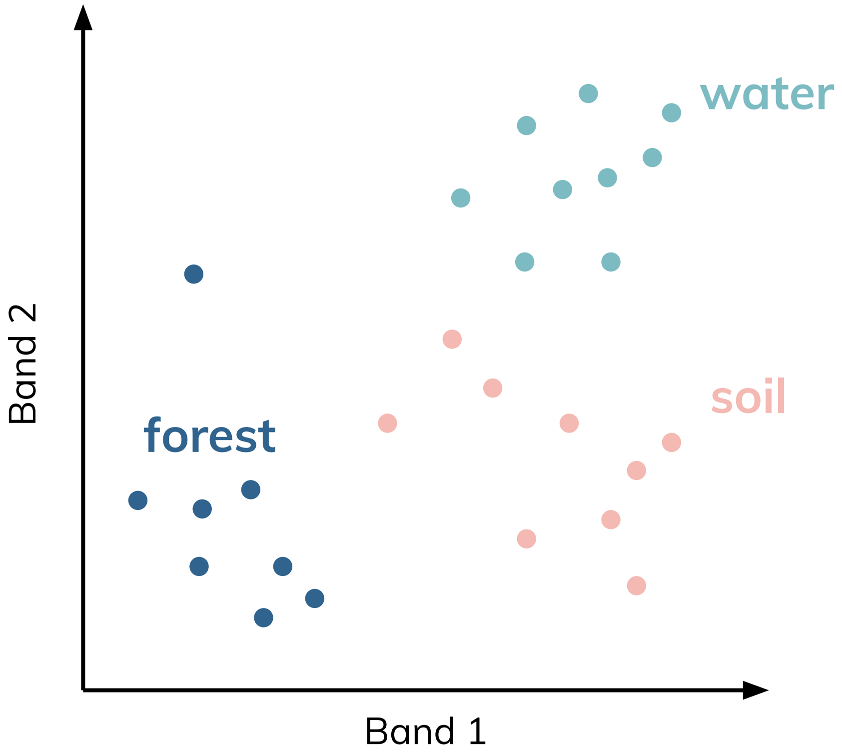
Supervised classification

- Find means for each group based on known points



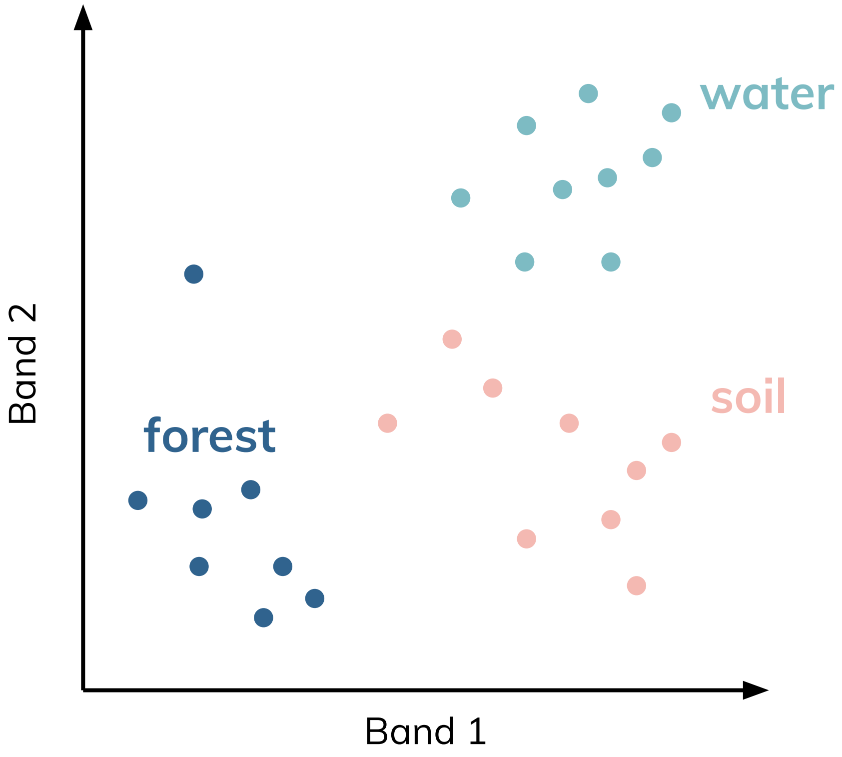
Supervised classification

- Find means for each group based on known points
- Assign each point to the closest group



Minimum distance to mean algorithm

- Find means for each group based on known points
- Assign each point to the closest group

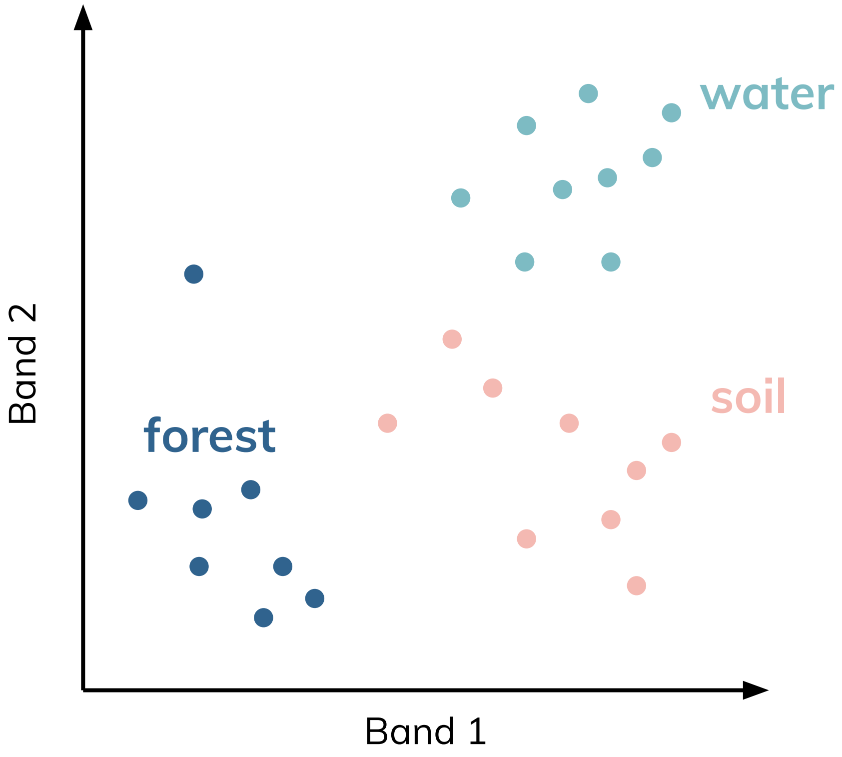


Pros
- fast/easy

Cons
- only uses means, not other statistical differences between classes

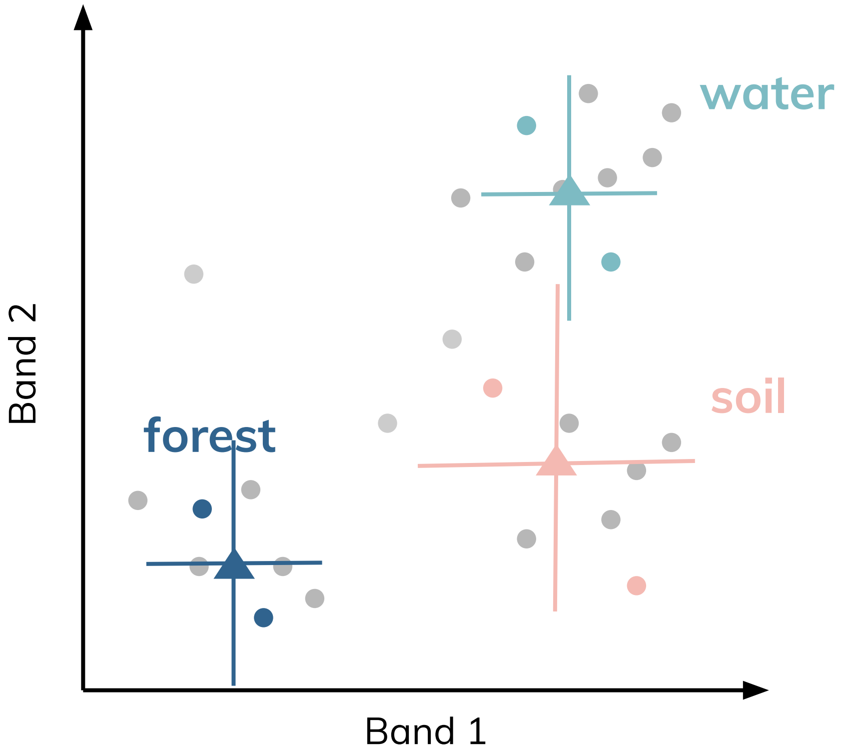
Minimum distance to mean algorithm

- Find means for each group based on known points
- Assign each point to the closest group



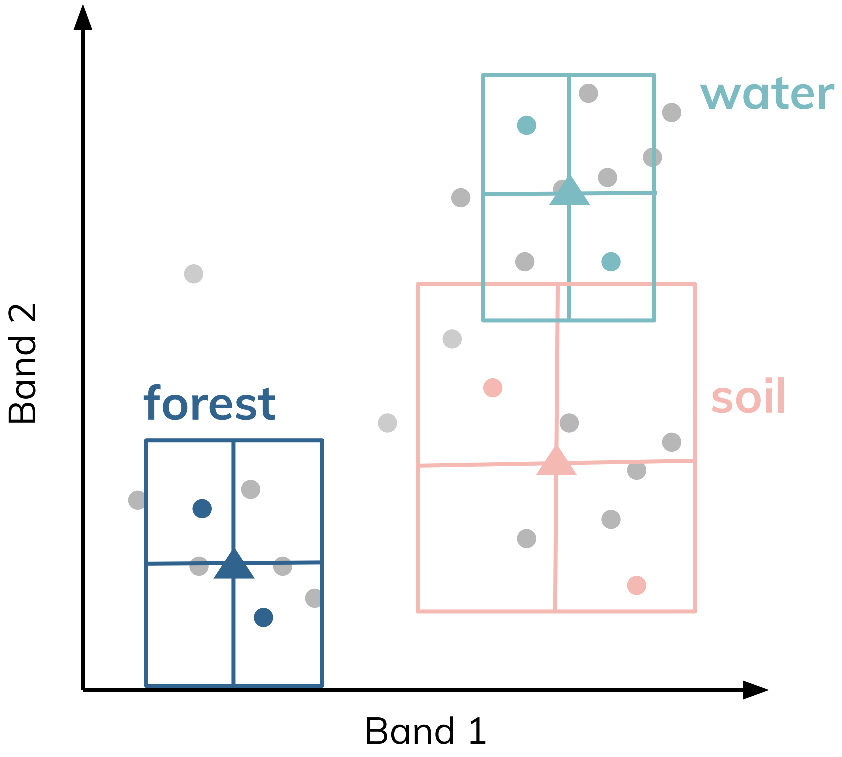
Supervised classification

- Find means and standard deviations for each group based on known points



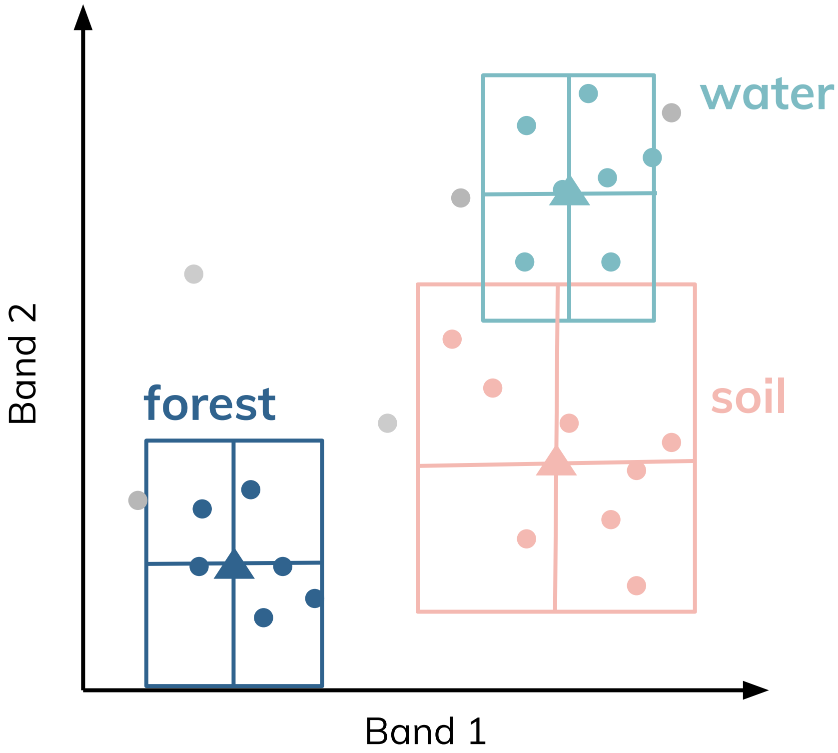
Supervised classification

- Find means and standard deviations for each group based on known points



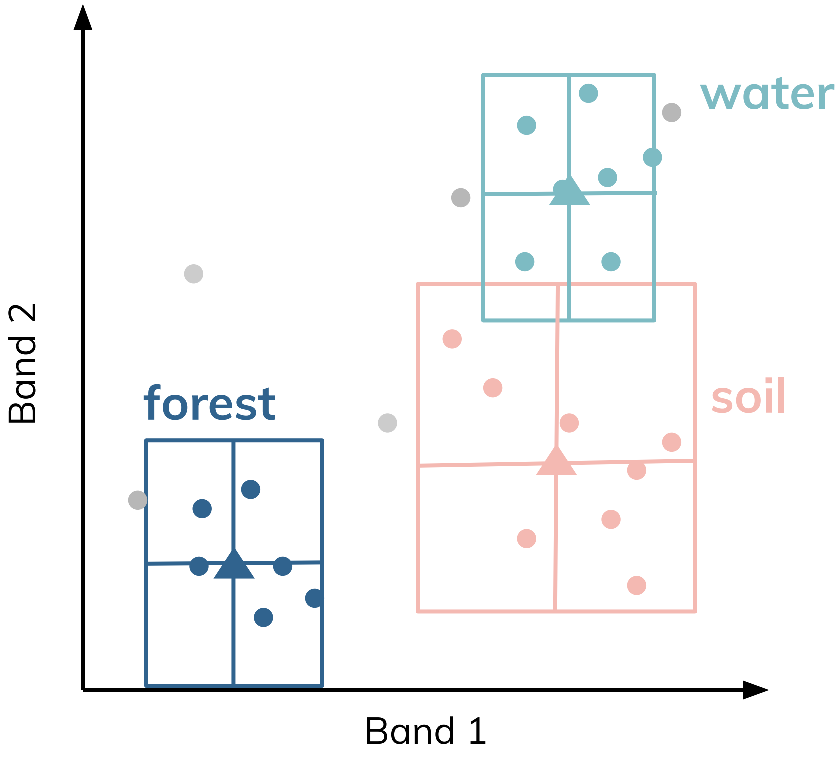
Supervised classification

- Find means and standard deviations for each group based on known points
- Assign points to groups



Parallelipiped

- Find means and standard deviations for each group based on known points
- Assign points to groups



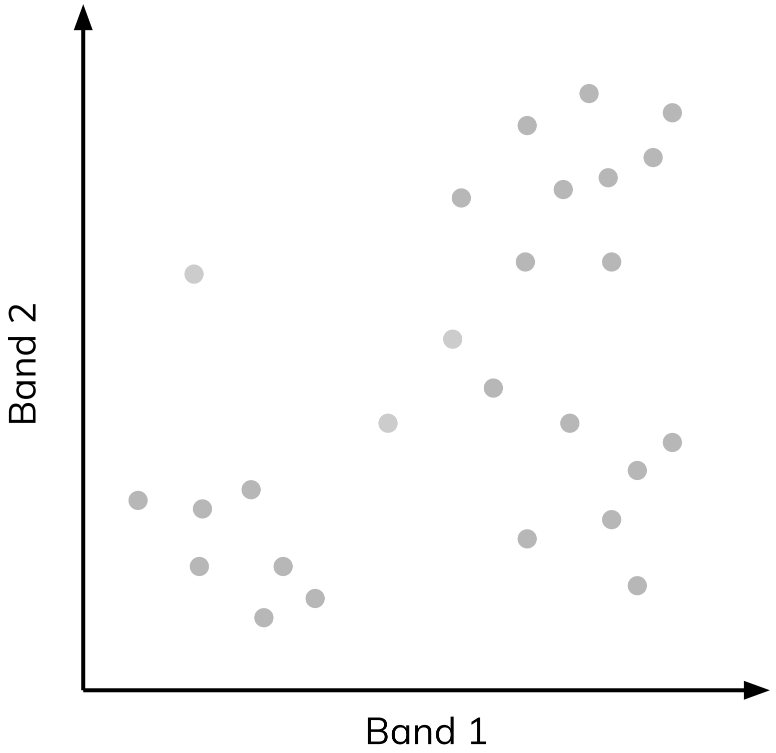
Pros

- fast/easy
- More realistic than just using the mean

Cons

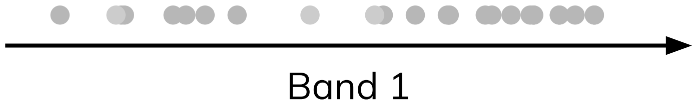
- Unclassified pixels
- Overlapping classes

Supervised classification



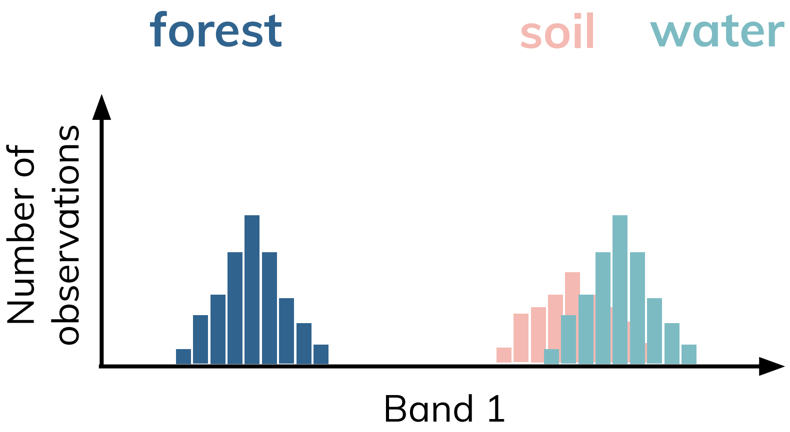
Supervised classification

Unknown pixels:

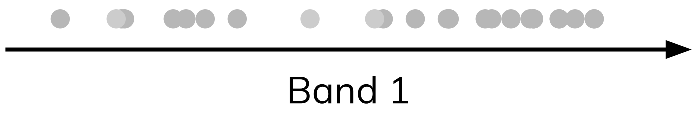


Supervised classification

Known pixels:

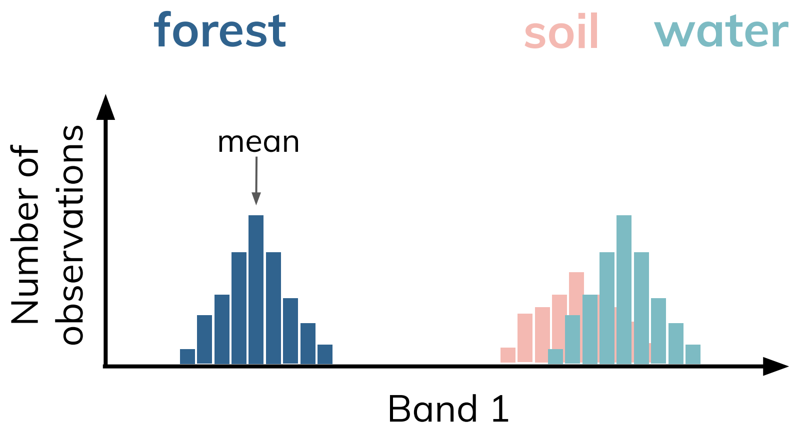


Unknown pixels:

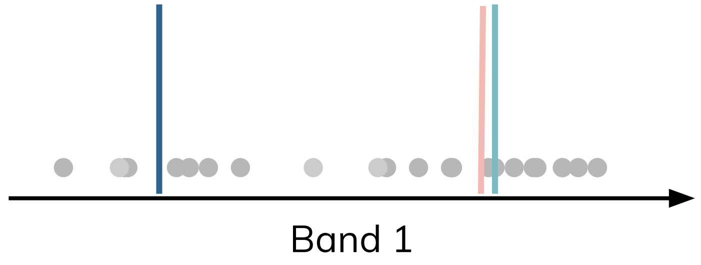


Supervised classification

Known pixels:



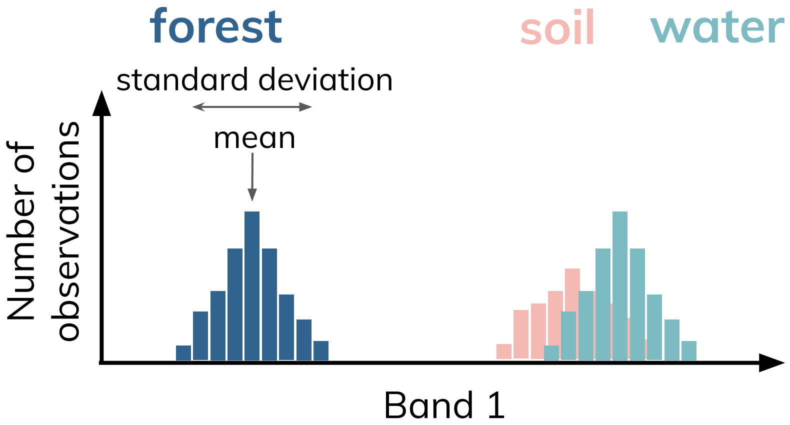
Unknown pixels:



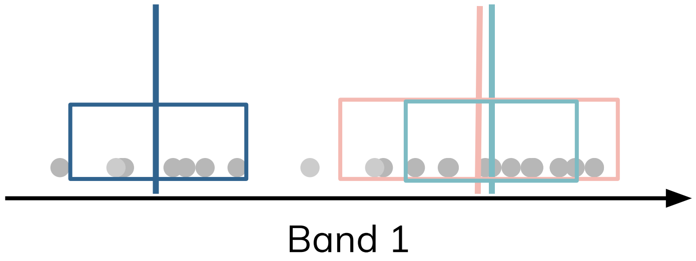
minimum distance to mean

Supervised classification

Known pixels:



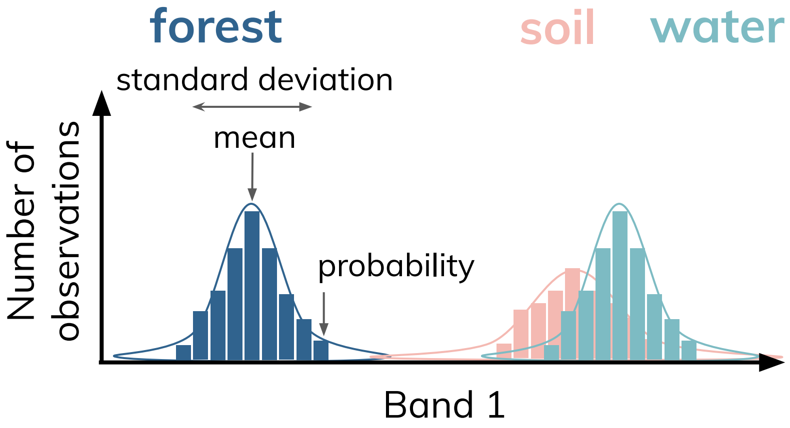
Unknown pixels:



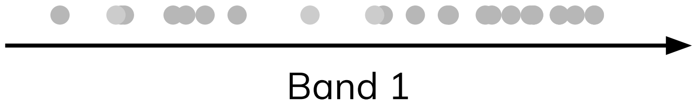
parallelipiped

Supervised classification

Known pixels:

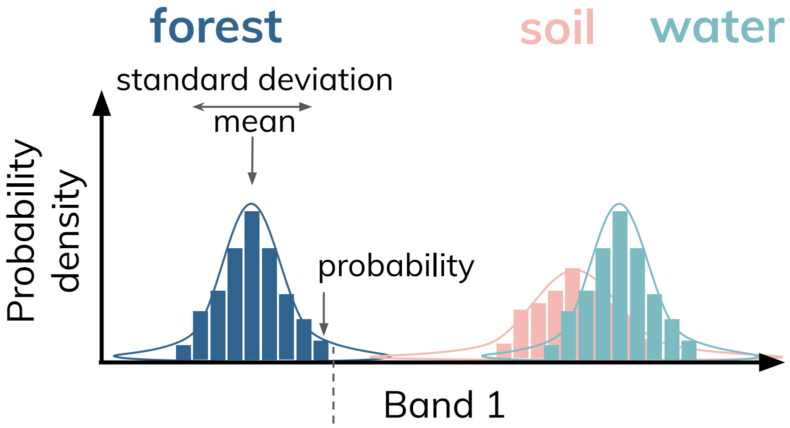


Unknown pixels:

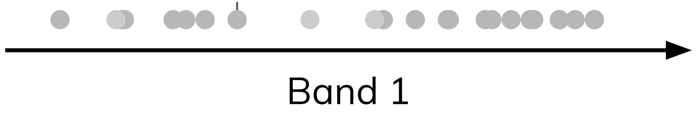


Supervised classification

Known pixels:

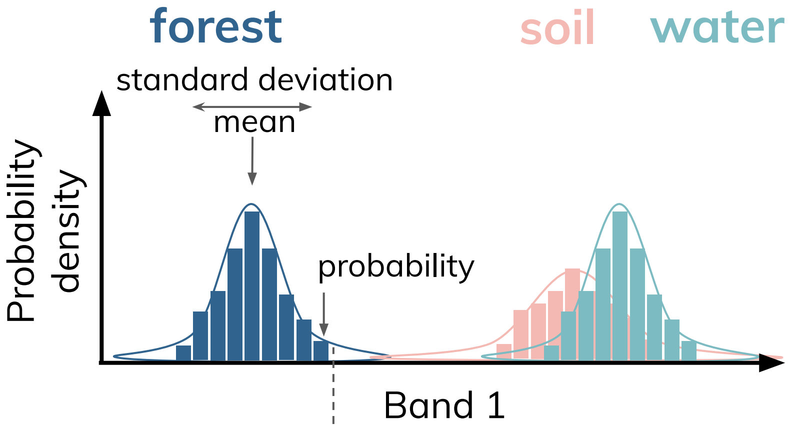


Unknown pixels:

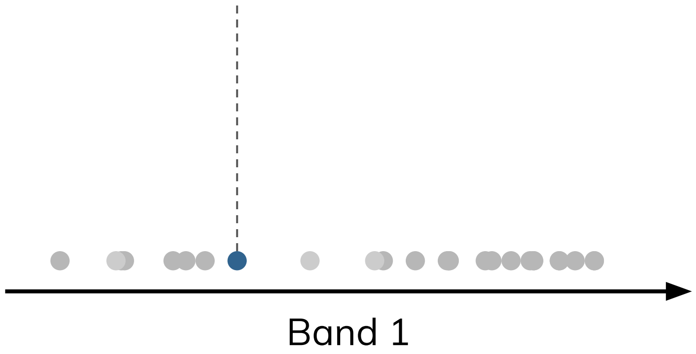


Supervised classification

Known pixels:



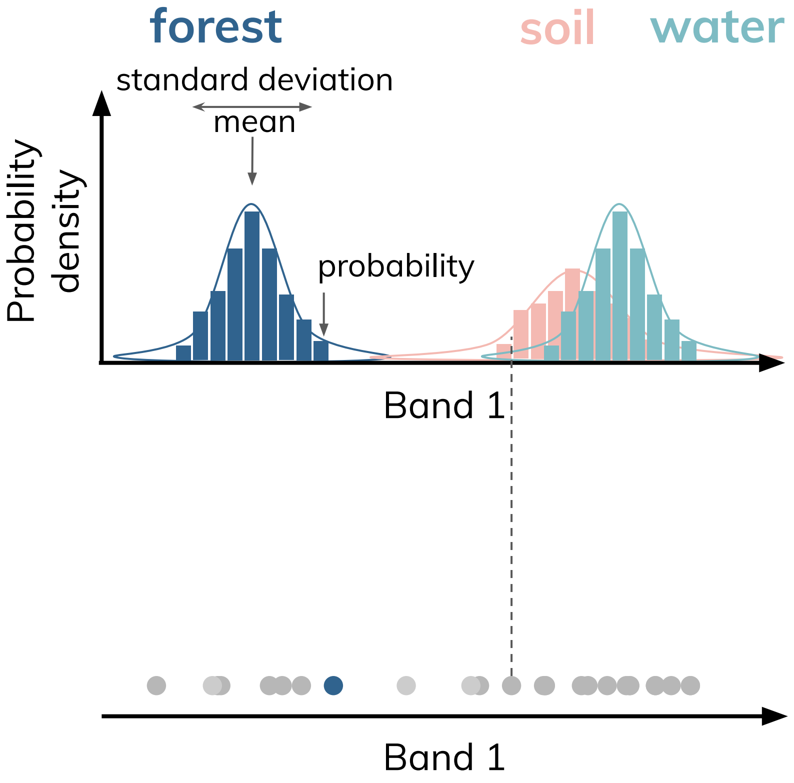
Unknown pixels:



Supervised classification

Known pixels:

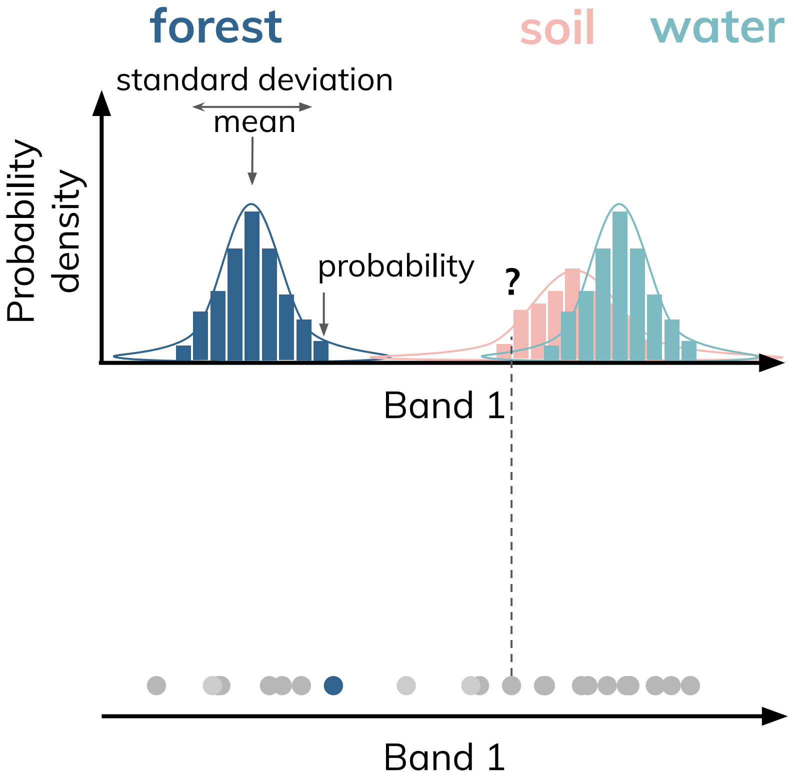
Unknown pixels:



Supervised classification

Known pixels:

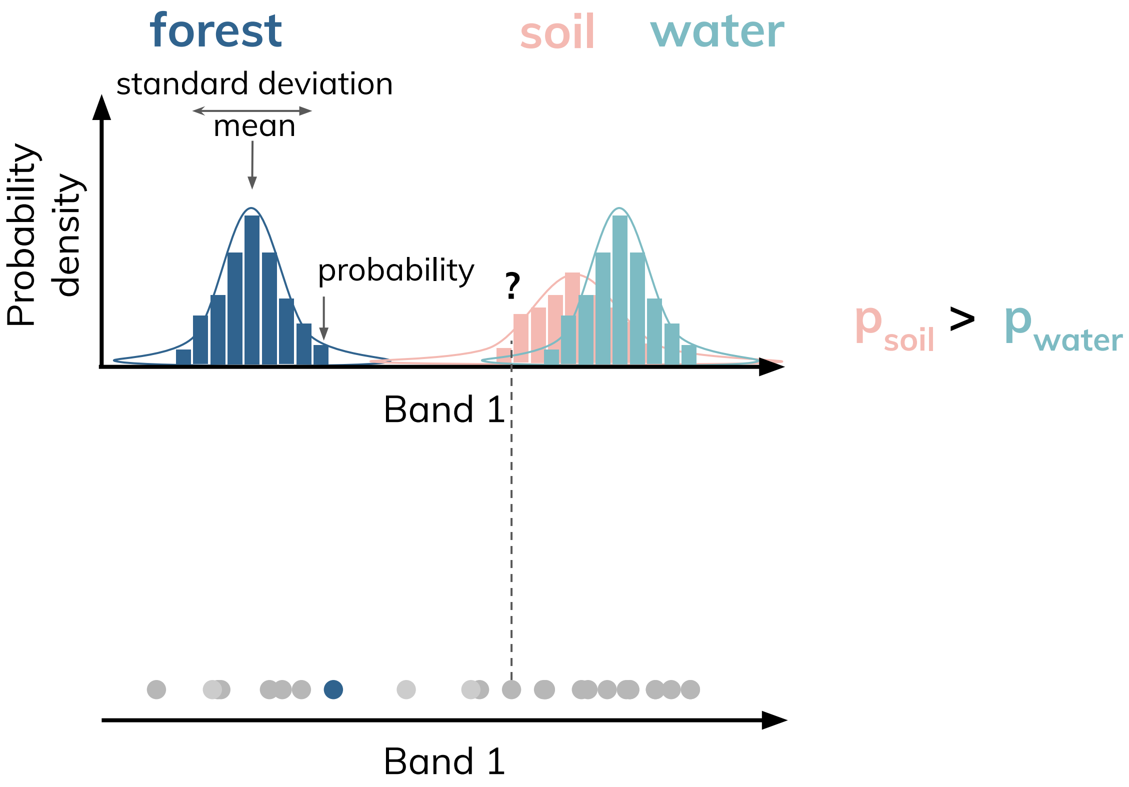
Unknown pixels:



Supervised classification

Known pixels:

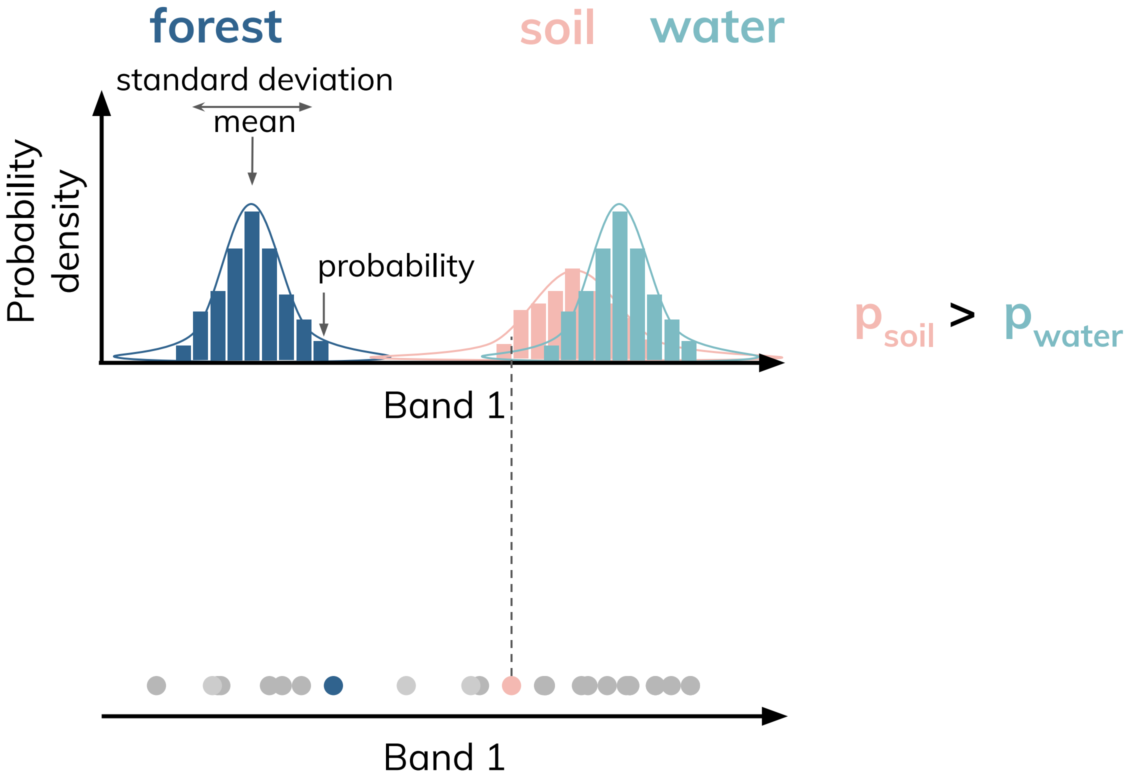
Unknown pixels:



Supervised classification

Known pixels:

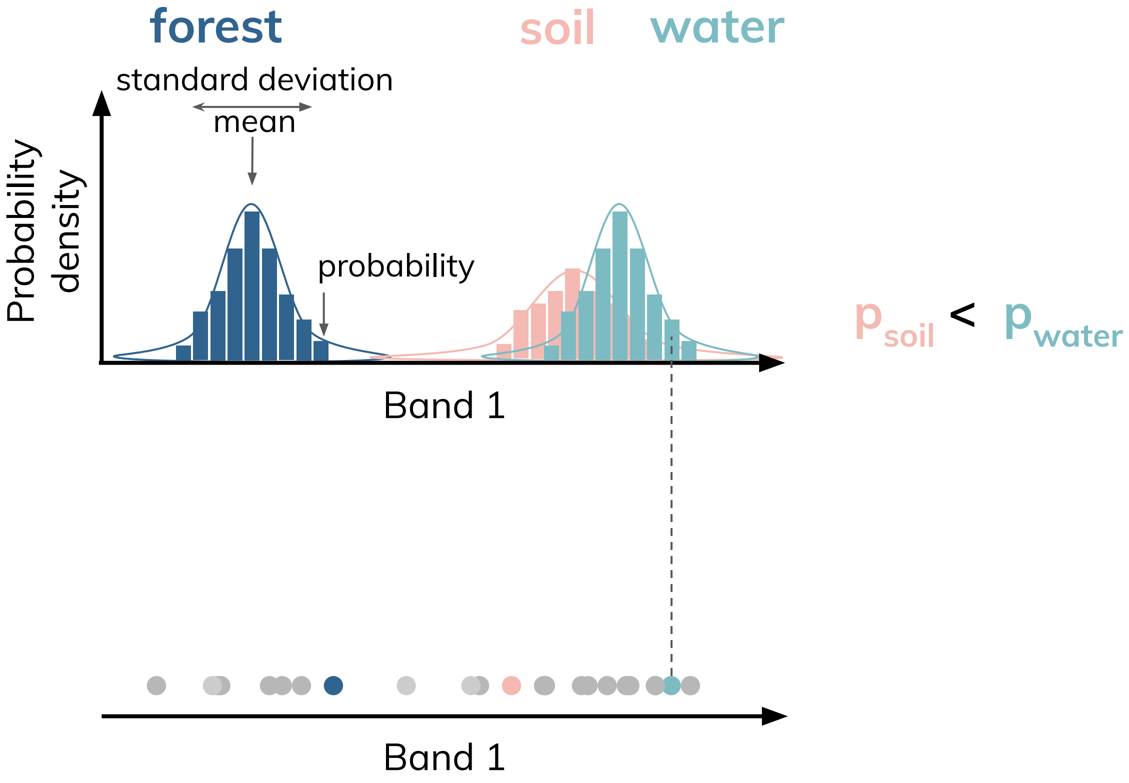
Unknown pixels:



Supervised classification

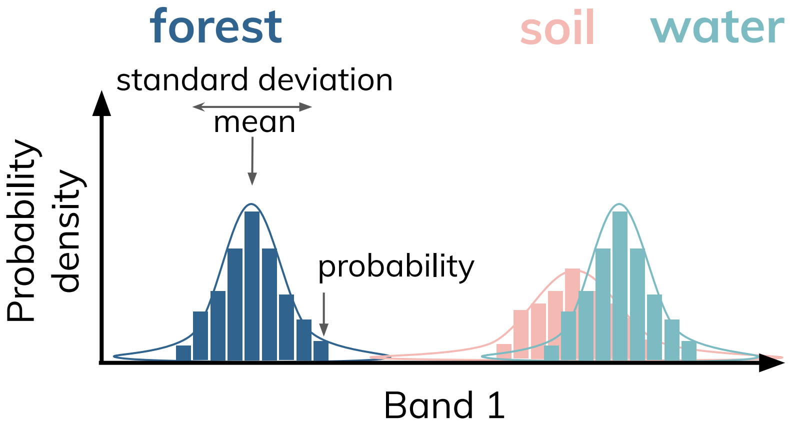
Known pixels:

Unknown pixels:



Maximum likelihood

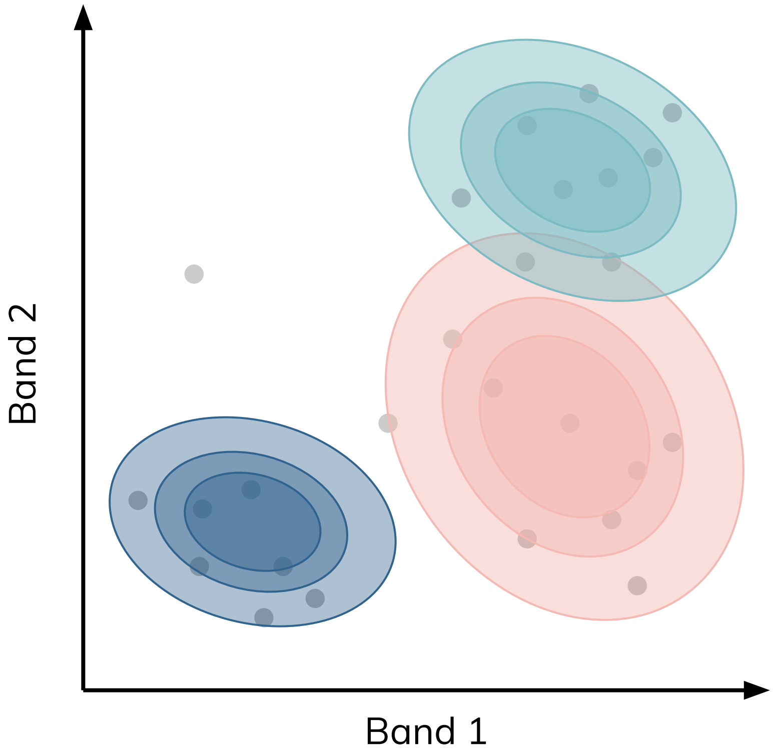
Known pixels:



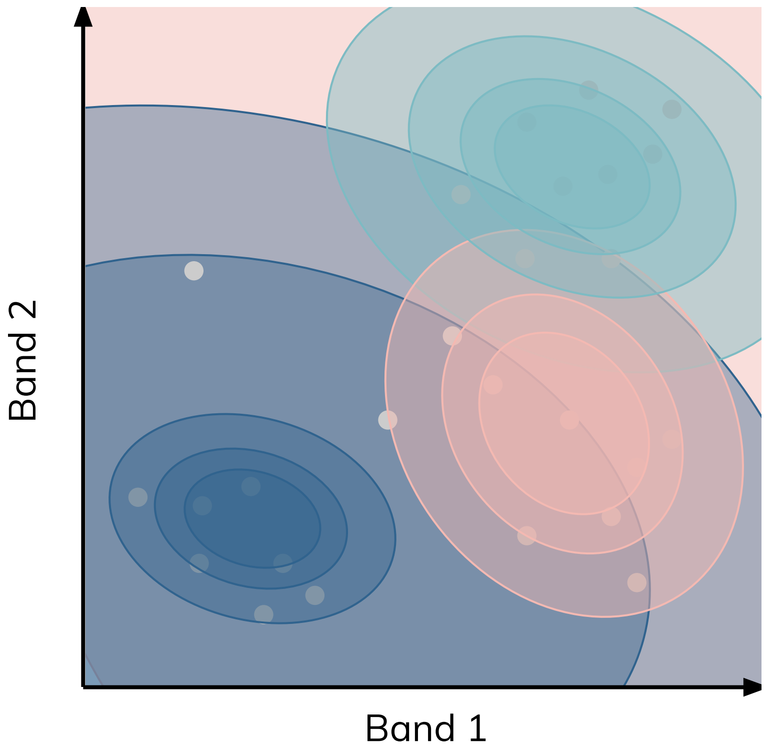
Unknown pixels:

maximum likelihood

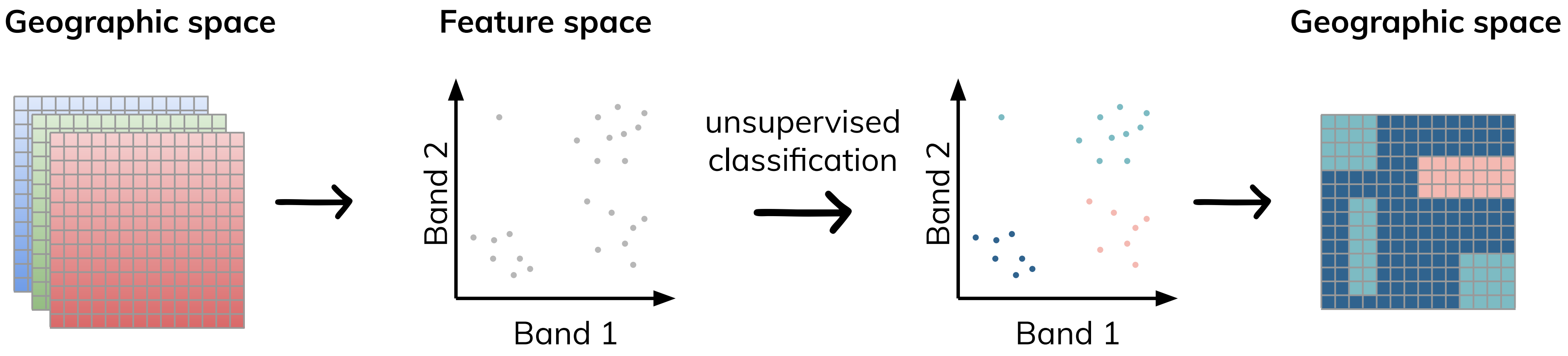
Maximum likelihood



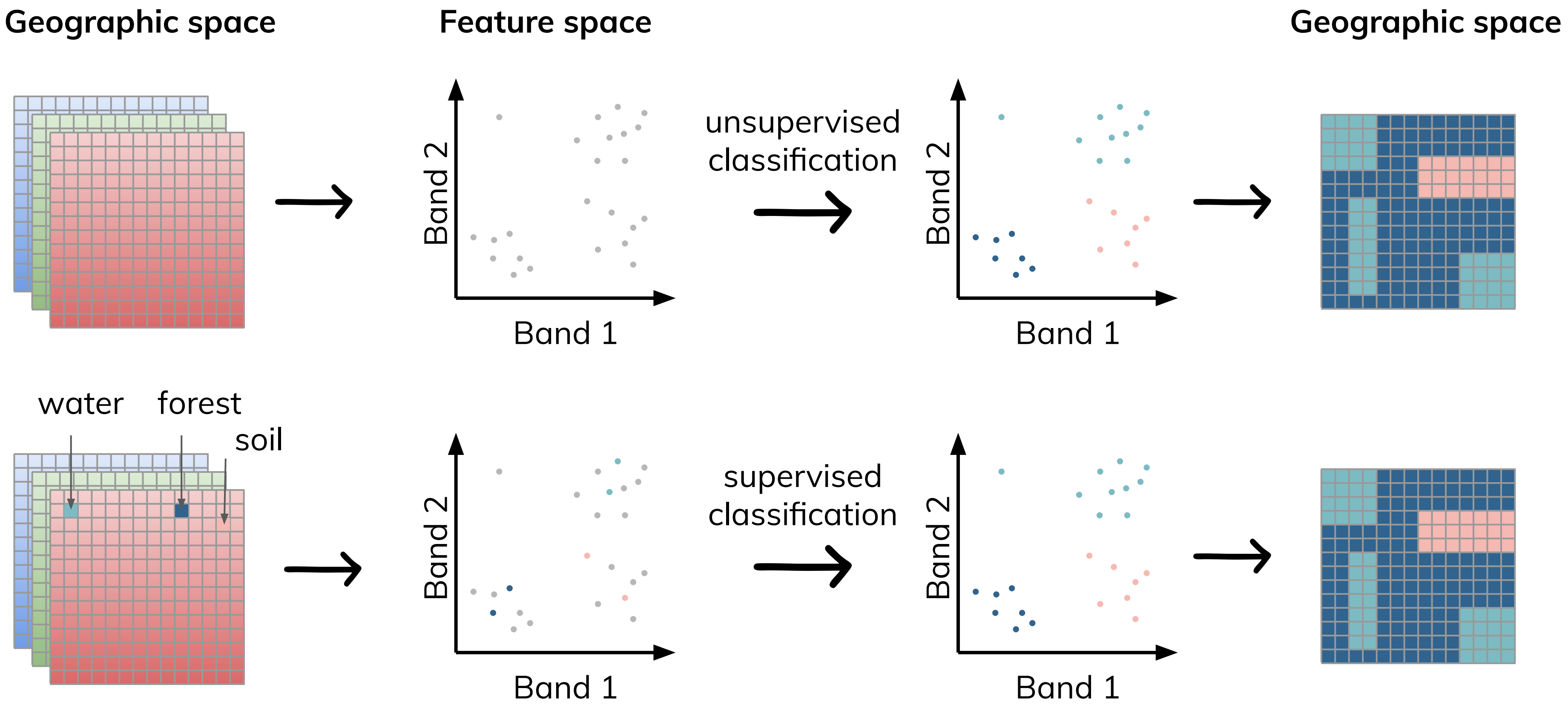
Maximum likelihood



(un)supervised classification



(un)supervised classification



Classification approaches

Unsupervised	Supervised

Classification approaches

Unsupervised	Supervised
• Algorithm identifies groups of pixels with similar spectra	• Algorithm identifies groups of pixels with similar spectra

Classification approaches

Unsupervised	Supervised
• Algorithm identifies groups of pixels with similar spectra • User assigns meaning to resulting classes	• Algorithm identifies groups of pixels with similar spectra • User provides examples for desired groupings

Classification approaches

Unsupervised

- Algorithm identifies groups of pixels with similar spectra
- User assigns meaning to resulting classes
- Bulk of analyst’s work comes after the classification process

Supervised

- Algorithm identifies groups of pixels with similar spectra
- User provides examples for desired groupings
- Bulk of analyst’s work comes before the classification process

Classification approaches

Unsupervised

- Algorithm identifies groups of pixels with similar spectra
- User assigns meaning to resulting classes
- Bulk of analyst's work comes after the classification process
- Pros:
    - No prior knowledge of area required
    - Human error is minimized
    - Relatively fast/easy
    - Unique spectral classes are produced
- Cons:
    - Spectral classes may not represent features on the ground
    - Does not consider spatial relationships
    - Can be time-consuming to interpret
    - Spectral properties may vary over time/images

Supervised

- Algorithm identifies groups of pixels with similar spectra
- User provides examples for desired groupings
- Bulk of analyst's work comes before the classification process

Classification approaches

Unsupervised

- Algorithm identifies groups of pixels with similar spectra
- User assigns meaning to resulting classes
- Bulk of analyst’s work comes after the classification process
- Pros:
    - No prior knowledge of area required
    - Human error is minimized
    - Relatively fast/easy
    - Unique spectral classes are produced
- Cons:
    - Spectral classes may not represent features on the ground
    - Does not consider spatial relationships
    - Can be time-consuming to interpret
    - Spectral properties may vary over time/images

Supervised

- Algorithm identifies groups of pixels with similar spectra
- User provides examples for desired groupings
- Bulk of analyst’s work comes before the classification process
- Pros:
    - Spectral classes represent features on the ground
    - Training areas are reusable
- Cons:
    - Information classes may not match spectral classes
    - Difficulty and cost of selecting training sites

Classification approaches

Unsupervised

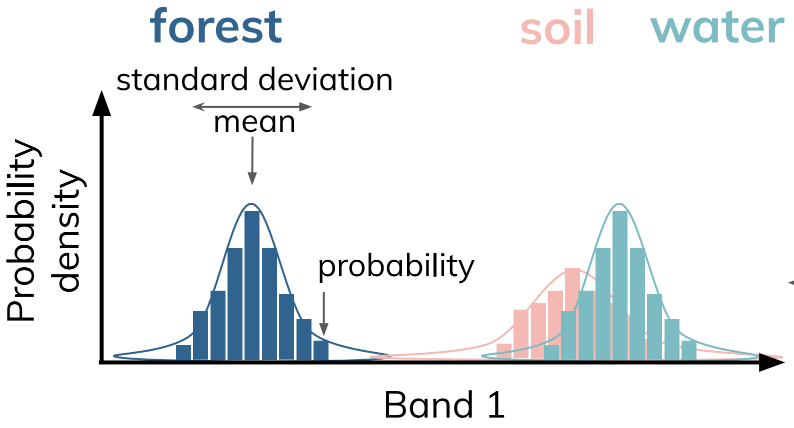
- Algorithm identifies groups of pixels with similar spectra
- User assigns meaning to resulting classes
- Bulk of analyst's work comes after the classification process
- Pros:
    - No prior knowledge of area required
    - Human error is minimized
    - Relatively fast/easy
    - Unique spectral classes are produced
- Cons:
    - Spectral classes may not represent features on the ground
    - Does not consider spatial relationships
    - Can be time-consuming to interpret
    - Spectral properties may vary over time/images

Supervised

- Algorithm identifies groups of pixels with similar spectra
- User provides examples for desired groupings
- **Bulk of analyst's work comes before the classification process**
- Pros:
    - Spectral classes represent features on the ground
    - Training areas are reusable
- Cons:
    - Information classes may not match spectral classes
    - Difficulty and coast of selecting training sites

Maximum likelihood

Known pixels:

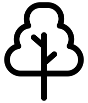


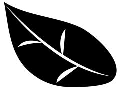
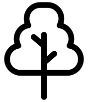
Training data

Unknown pixels:

maximum likelihood

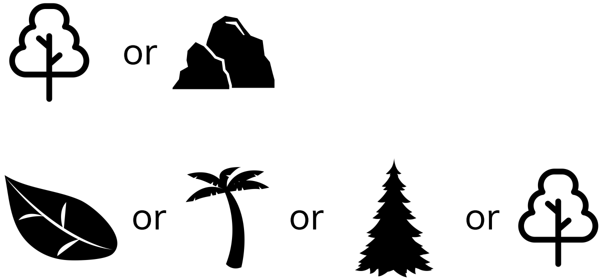
Supervised classification: training data

Classification scheme:  or 

 or  or  or 

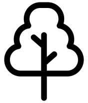
Supervised classification: training data

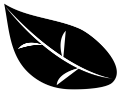
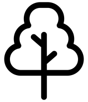
Classification scheme:



Does the resolution match your scheme? (spatial/temporal/spectral/radiometric)

Supervised classification: training data

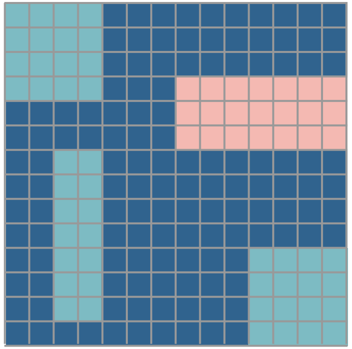
Classification scheme:  or 

 or  or  or 

Does the resolution match your scheme? (spatial/temporal/spectral/radiometric)

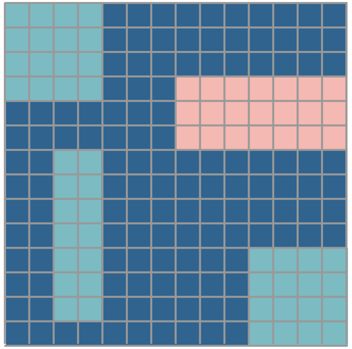
Does your training data capture the heterogeneity of each class?

Testing how we did!



How accurate is this map?

Testing how we did!



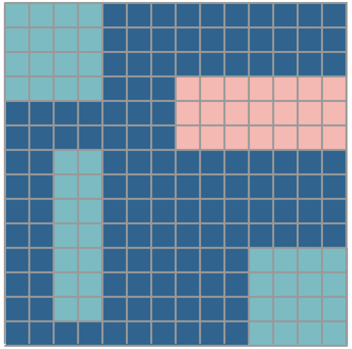
How accurate is this map?

"True answer"

Our guess based on remote sensing data

	forest	soil	water
forest			
soil			
water			

Testing how we did!



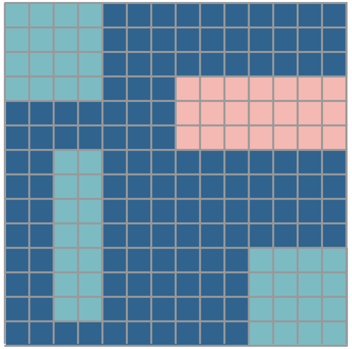
How accurate is this map?

"True answer"

Our guess based on remote sensing data

	forest	soil	water
forest	25	0	0
soil			
water			

Testing how we did!



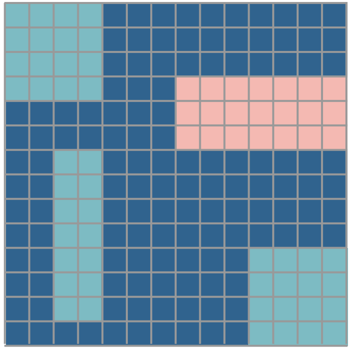
How accurate is this map?

"True answer"

Our guess based on remote sensing data

	forest	soil	water
forest	25	0	0
soil	0	12	0
water	0	0	18

Testing how we did!



How accurate is this map?

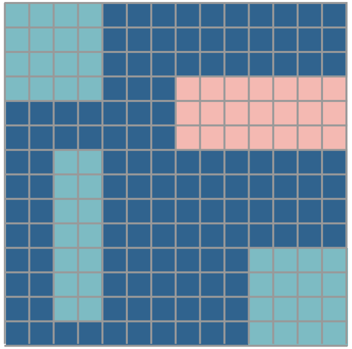
"True answer"

Our guess based on remote sensing data

	forest	soil	water
forest	25	0	0
soil	0	12	0
water	0	0	18

Accuracy = sum of correct matches ÷ total number of cells

Testing how we did!



How accurate is this map?

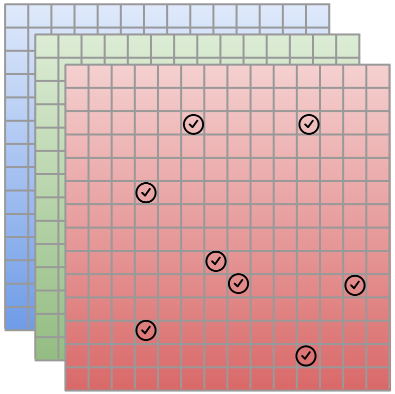
"True answer"

Our guess based on remote sensing data

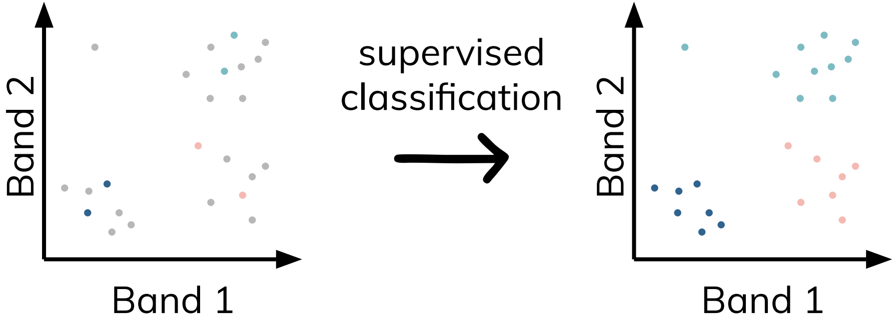
	forest	soil	water
forest	25	0	0
soil	0	12	0
water	0	0	18

Accuracy = sum of correct matches ÷ total number of cells

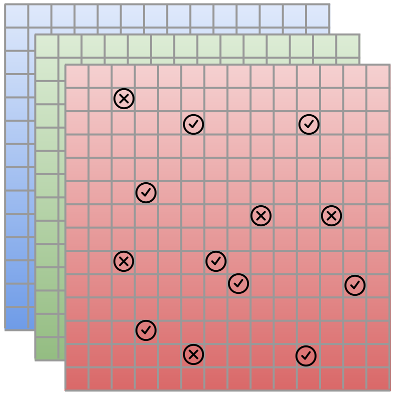
Testing how we did!



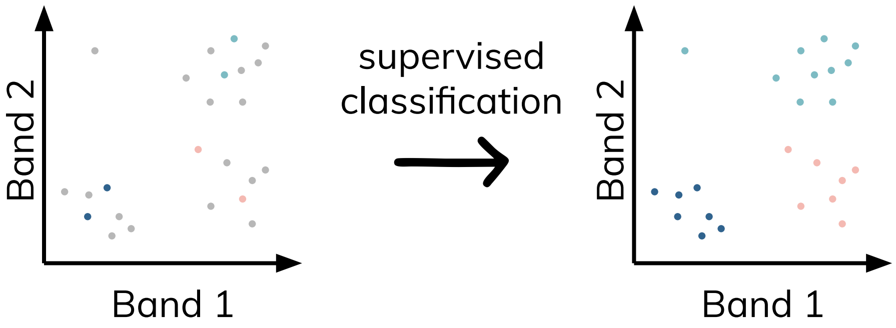
training:



Testing how we did!



training:



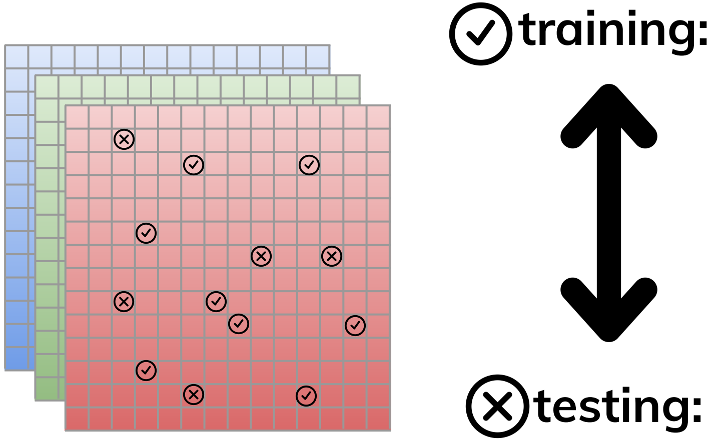
testing:

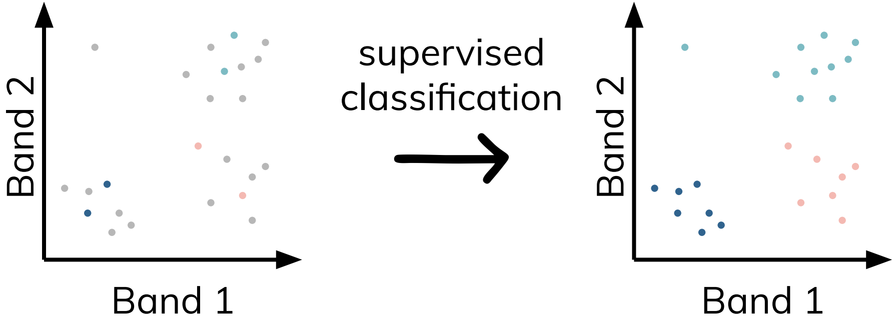
"True answer"

Our guess based on remote sensing data

	forest	soil	water
forest			
soil			
water			

Testing how we did!



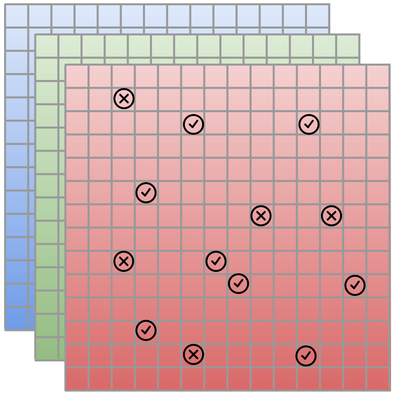


"True answer"

Our guess based on remote sensing data

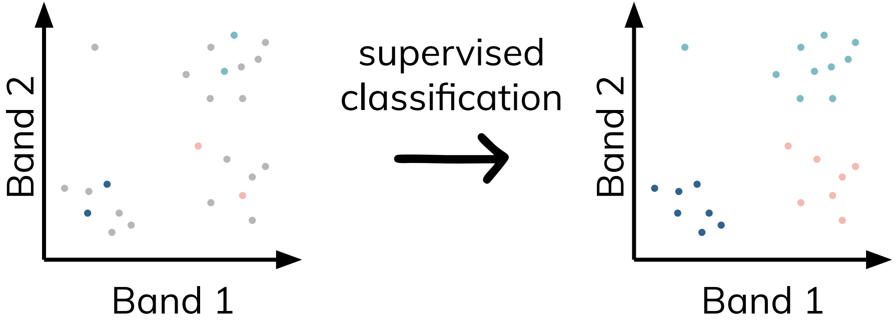
	forest	soil	water
forest			
soil			
water			

Cross-validation



training:

testing:



"True answer"

Our guess based on remote sensing data

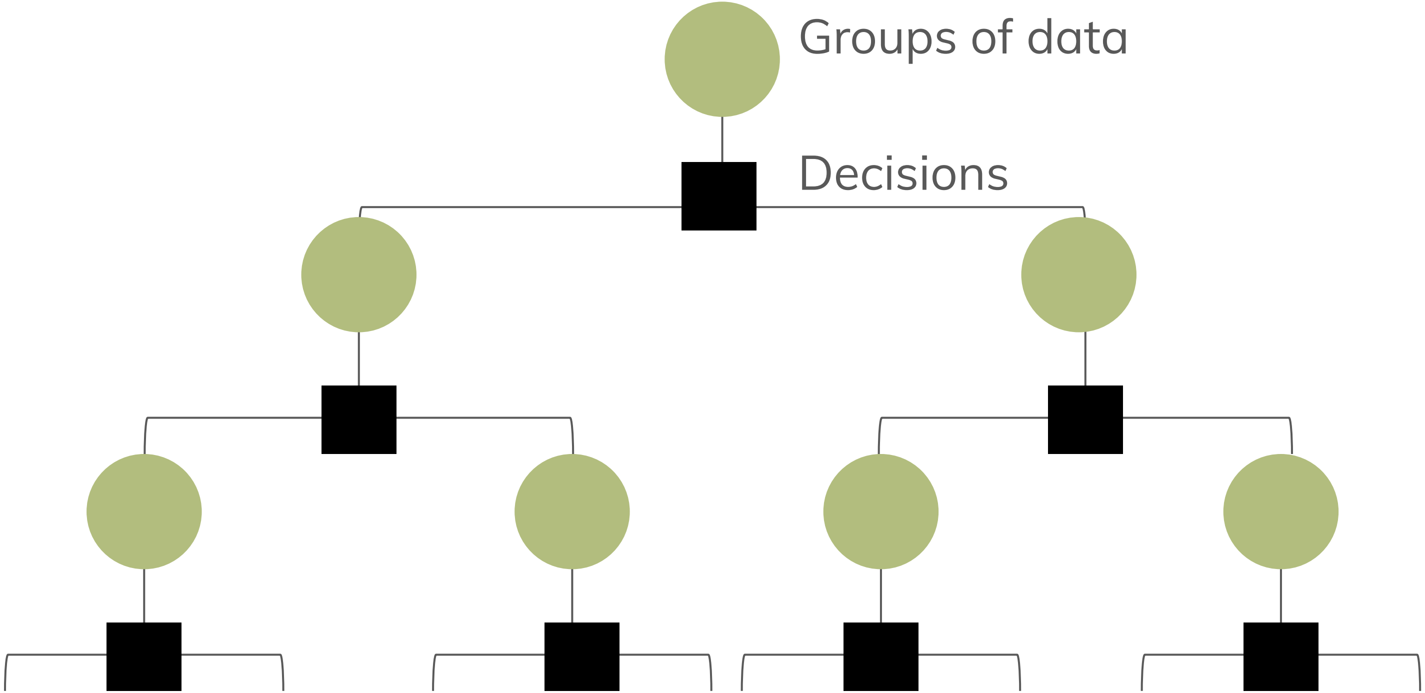
	forest	soil	water
forest			
soil			
water			

Summary

- **To classify objects, we try to separate them based on spectral features**
- **Lots of different ways to do this!**
  - Unsupervised approaches don’t require any information upfront, but is hard to interpret
  - Supervised approaches are easier to interpret, but require information upfront

Today’s classification task

Decision trees



Classification and Regression Trees (CART) algorithm

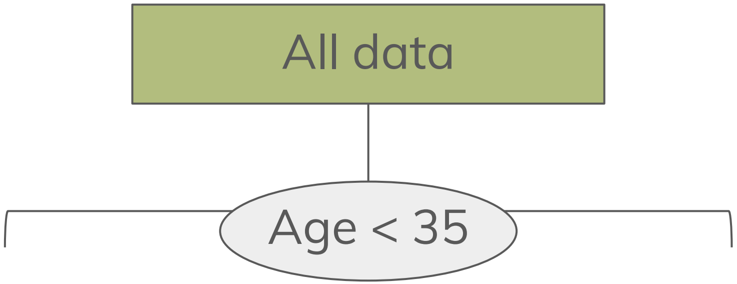
Gini Impurity:

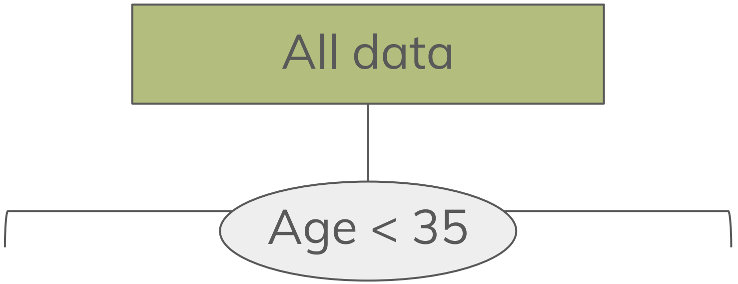
$$G = 1 - \sum_{i=1}^{c} p_i^2$$

c is the number of classes
p_i is the probability of a randomly chosen element in the node being labeled as class i

Classification and Regression Trees (CART) algorithm

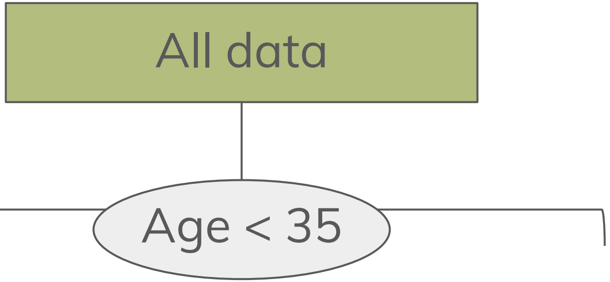
Age	Income	Buy?
30	20,000	Yes
40	50,000	Yes
20	30,000	No
50	60,000	No
60	80,000	Yes


All data
Age < 35



Age	Income	
30	20,000	Yes
20	30,000	No

Age	Income	
40	50,000	Yes
50	60,000	No
60	80,000	Yes



Age	Income	
30	20,000	Yes
20	30,000	No

Age	Income	
40	50,000	Yes
50	60,000	No
60	80,000	Yes

$$G_{left} = 1 - (\tfrac{1}{2})^2 - (\tfrac{1}{2})^2 = 0.5$$

All data

Age < 35

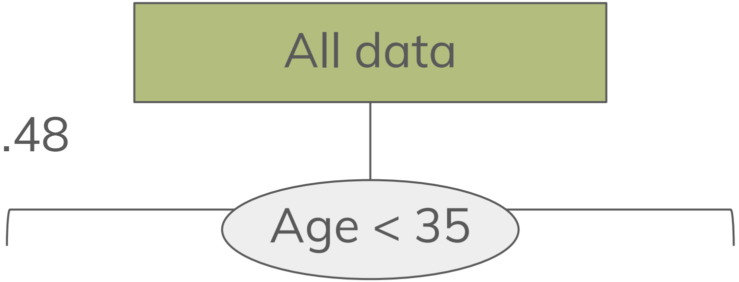
Age	Income	
30	20,000	Yes
20	30,000	No

Age	Income	
40	50,000	Yes
50	60,000	No
60	80,000	Yes

$G_{left} = 1 - (\frac{1}{2})^2 - (\frac{1}{2})^2 = 0.5$

$G_{right} = 1 - (\frac{2}{3})^2 - (\frac{1}{3})^2 \approx 0.444$

$$G_{split} = (2/5)G_{left} + (3/5)G_{right} \approx 0.48$$



Age	Income	
30	20,000	Yes
20	30,000	No

Age	Income	
40	50,000	Yes
50	60,000	No
60	80,000	Yes

$$G_{left} = 1 - (1/2)^2 - (1/2)^2 = 0.5$$

$$G_{right} = 1 - (2/3)^2 - (1/3)^2 \approx 0.444$$

Big ask!

ESCIs due December 8!