Accuracy Assessment

Goals:

- Assess how well a classification worked
- Understand how to interpret the usefulness of someone else's classification

Accuracy Assessment

- Overview
  - Collect reference data: “ground truth”
    - Determination of class types at specific locations
  - Compare reference to classified map
    - Does class type on classified map = class type determined from reference data?

Accuracy Assessment: Reference Data

- Some possible sources
  - Aerial photo interpretation
  - Ground truth with GPS
  - GIS layers

Accuracy Assessment: Reference Data

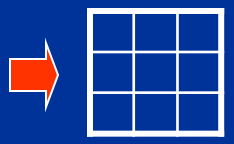
- Issue 1: Choosing reference source
  - Make sure you can actually extract from the reference source the information that you need for the classification scheme
    - I.e. Aerial photos may not be good reference data if your classification scheme distinguishes four species of grass. You may need GPS'd ground data.

Accuracy Assessment: Reference data

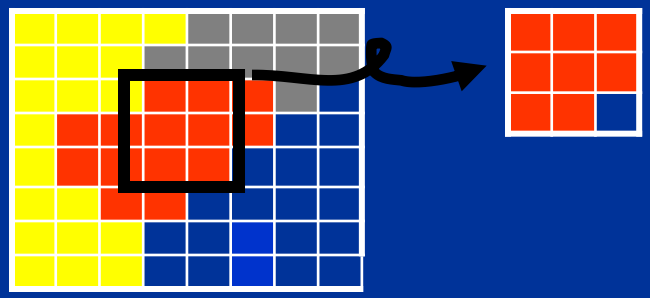
- Issue 2: Determining size of reference plots
  - Match spatial scale of reference plots and remotely-sensed data
    - I.e. GPS'd ground plots 5 meters on a side may not be useful if remotely-sensed cells are 1km on a side. You may need aerial photos or even other satellite images.

Accuracy Assessment: Reference Data

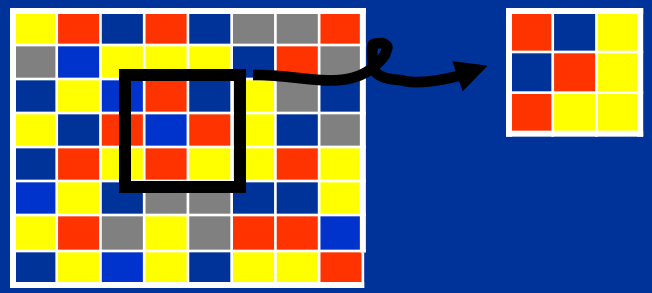
- Issue 2: Determining size of reference plots
  - Take into account spatial frequencies of image
  - E.G. For the two examples below, consider photo reference plots that cover an area 3 pixels on a side



Example 1: Low spatial frequency Homogeneous image



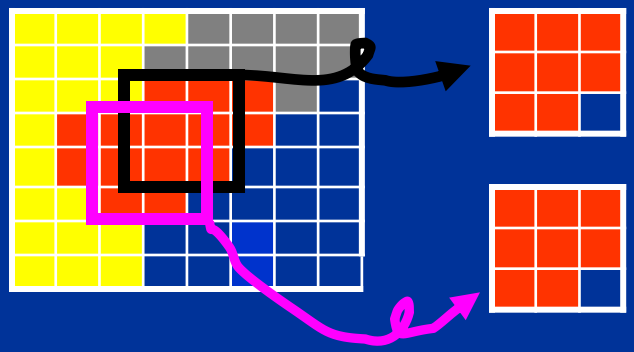
Example 2: High spatial frequency Heterogenous image



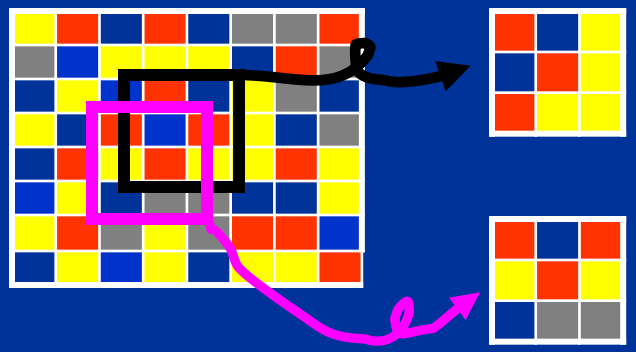
Accuracy Assessment: Reference Data

- Issue 2: Determining size of reference plots
  - HOWEVER, also need to take into account accuracy of position of image and reference data
  - E.G. For the same two examples, consider the situation where accuracy of position of the image is +/- one pixel

Example 1: Low spatial frequency



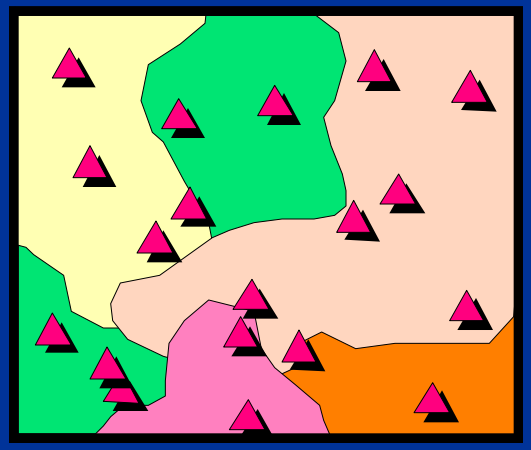
Example 2: High spatial frequency



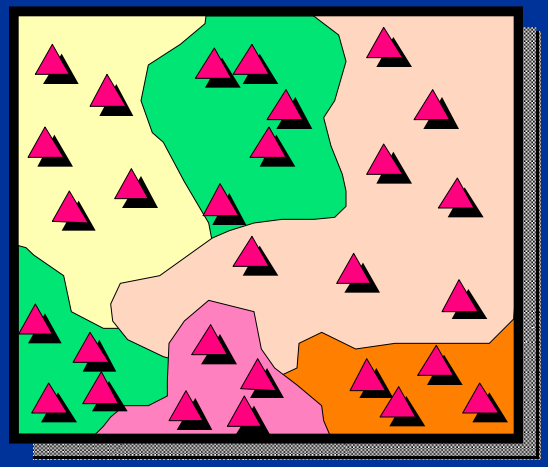
Accuracy Assessment: Reference Data

- Issue 3: Determining position and number of samples
  - Make sure to adequately sample the landscape
  - Variety of sampling schemes
    - Random, stratified random, systematic, etc.
  - The more reference plots, the better
    - You can estimate how many you need statistically
    - In reality, you can never get enough
    - Lillesand and Kiefer: suggest 50 *per class* as rule of thumb

Sampling Methods

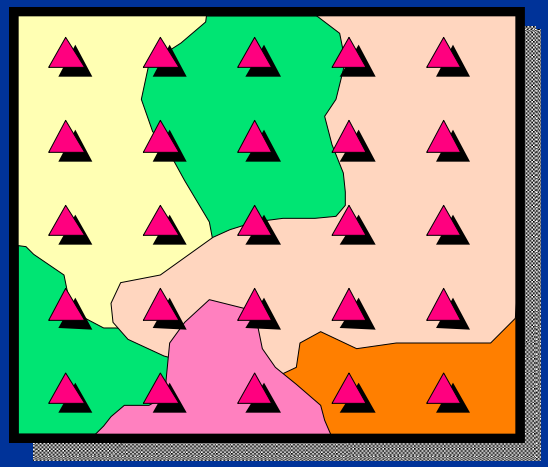


Simple Random Sampling: observations are randomly placed.

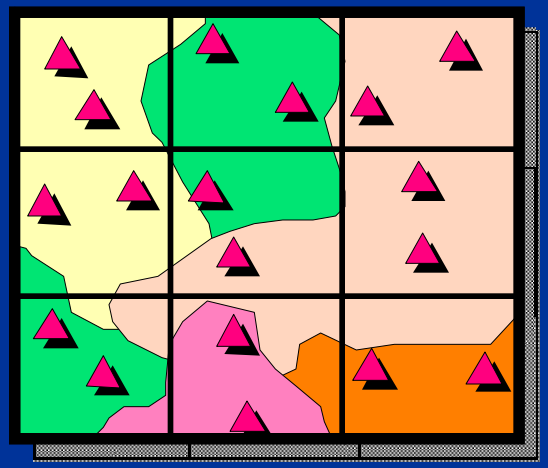


Stratified Random Sampling : a minimum number of observations are randomly placed in each category.

Sampling Methods

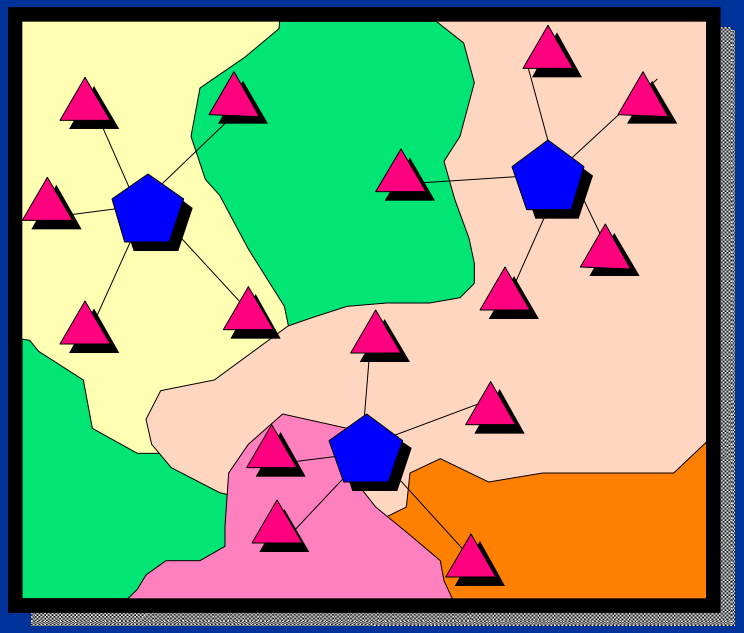


Systematic Sampling: observations are placed at equal intervals according to a strategy.



Systematic Non-Aligned Sampling a grid provides even distribution of randomly placed observations.

Sampling Methods



Cluster Sampling: Randomly placed “centroids” used as a base of several nearby observations. The nearby observations can be randomly selected, systematically selected, etc...

Accuracy Assessment: Reference data

- Having chosen reference source, plot size, and locations:
  - Determine class types from reference source
  - Determine class type claimed by classified map
- Compare them!

Accuracy Assessment: Compare

- Example:

Reference Plot ID Number	Class determined from **reference source**	Class claimed on **classified map**	Agreement?
1	Conifer	Conifer	Yes
2	Hardwood	Conifer	No
3	Water	Water	Yes
4	Hardwood	Hardwood	Yes
5	Grass	Hardwood	No
6	Etc….		

Accuracy Assessment: Compare

How to summarize and quantify?

Accuracy Assessment: Error matrix

- Summarize using an error matrix

	Class types determined from **reference source**				
Class types determined from **classified map**	# Plots	Conifer	Hardwood	Water	**Totals**
	Conifer	50	5	2	57
	Hardwood	14	13	0	27
	Water	3	5	8	16
	Totals	67	23	10	100

Accuracy Assessment: Total Accuracy

- Quantifying accuracy
  - Total Accuracy: Number of correct plots / total number of plots

	Class types determined from **reference source**			
Class types determined from **classified map**	# Plots	Conifer	Hardwood	Water	**Totals**
	Conifer	50	5	2	57
	Hardwood	14	13	0	27
	Water	3	5	8	16
	Totals	67	23	10	100

$$Accuracy_{Total} = \frac{50+13+8}{100} * 100 = 71\%$$

Diagonals represent sites classified correctly according to reference data

Off-diagonals were mis-classified

Accuracy Assessment: Total Accuracy

- Problem with total accuracy:
  - Summary value is an average
    - Does not reveal if error was evenly distributed between classes or if some classes were really bad and some really good
- Therefore, include other forms:
  - User's accuracy
  - Producer's accuracy

User's and producer's accuracy and types of error

- User's accuracy corresponds to error of commission (inclusion):
  - f.ex. 1 shrub and 3 conifer sites included erroneously in grass category
- Producer's accuracy corresponds to error of omission (exclusion):
  - f.ex. 7 conifer and 1 shrub sites omitted from grass category

Accuracy Assessment: User's Accuracy

- From the perspective of the user of the classified map, how accurate is the map?
  - For a given class, how many of the pixels on the map are actually what they say they are?
  - Calculated as:

Number correctly identified in a given map class /

Number claimed to be in that map class

Accuracy Assessment: User's Accuracy

	Class types determined from **reference source**			
Class types determined from **classified map**	# Plots	Conifer	Hardwood	Water	**Totals**
	Conifer	50	5	2	57
	Hardwood	14	13	0	27
	Water	3	5	8	16
	Totals	67	23	10	100

Example: Conifer

$$Accuracy_{User's,Conifer} = \frac{50}{57} * 100 = 88\%$$

Accuracy Assessment: Producer's Accuracy

- From the perspective of the maker of the classified map, how accurate is the map?
  - For a given class in reference plots, how many of the pixels on the map are labeled correctly?
  - Calculated as:

Number correctly identified in ref. plots of a given class /

Number actually in that reference class

Accuracy Assessment: Producer's Accuracy

	Class types determined from **reference source**			
Class types determined from **classified map**	# Plots	Conifer	Hardwood	Water	**Totals**
	Conifer	50	5	2	57
	Hardwood	14	13	0	27
	Water	3	5	8	16
	Totals	67	23	10	100

Example: Conifer

$$Accuracy_{producers,Conifer} = \frac{50}{67} * 100 = 75\%$$

Accuracy Assessment: Summary so far

	Class types determined from **reference source**					
Class types determined from **classified map**	# Plots	Conifer	Hardwood	Water	**Totals**	**User's Accuracy**
	Conifer	50	5	2	57	88%
	Hardwood	14	13	0	27	48%
	Water	3	5	8	16	50%
	Totals	67	23	10	100	
Producer's Accuracy		75%	57%	80%		Total: 71%

Accuracy Assessment: *Kappa*

- *Kappa* statistic
- Estimated as $\hat{K}$
- Reflects the difference between actual agreement and the agreement expected by chance
- Kappa of 0.85 means there is 85% better agreement than by chance alone

$$\hat{K} = \frac{\text{observed accuracy} - \text{chance agreement}}{1 - \text{chance agreement}}$$

Accuracy Assessment: *Kappa*

$$\hat{K} = \frac{\text{observed accuracy} - \text{chance agreement}}{1 - \text{chance agreement}}$$

- Observed accuracy determined by diagonal in error matrix
- Chance agreement incorporates off-diagonal
  - Sum of [Product of row and column totals for each class]
  - See Chapter 7 (p. 574) in Lillesand and Kiefer for computational formula

Accuracy Assessment: *Kappa*

$\hat{K} = 0.46$

	Class types determined from **reference source**				
Class types determined from **classified map**	# Plots	Conifer	Hardwood	Water	**Totals**	**User's Accuracy**
	Conifer	50	5	2	57	88%
	Hardwood	14	13	0	27	48%
	Water	3	5	8	16	50%
	Totals	67	23	10	100	
Producer's Accuracy		75%	57%	80%		Total: 71%

Accuracy Assessment: *Kappa*

- Other uses of *kappa*
  - Compare two error matrices
  - Weight cells in error matrix according to severity of misclassification
  - Provide error bounds on accuracy

Accuracy Assessment: Quantifying

- Each type of accuracy estimate yields different information
- If we only focus on one, we may get an erroneous sense of accuracy

Accuracy Assessment: Quantifying

- Example: Total accuracy was 71%, but User's accuracy for hardwoods was only 48%

	Class types determined from **reference source**					
Class types determined from **classified map**	# Plots	Conifer	Hardwood	Water	**Totals**	**User's Accuracy**
	Conifer	50	5	2	57	88%
	Hardwood	14	13	0	27	48%
	Water	3	5	8	16	50%
	Totals	67	23	10	100	
Producer's Accuracy		75%	57%	80%		Total: 71%

Accuracy Assessment: Quantifying

- What to report?
  - Depends on audience
  - Depends on the objective of your study
  - Most references suggest full reporting of error matrix, user's and producer's accuracies, total accuracy, and *Kappa*

Accuracy Assessment: Interpreting

- Why might accuracy be low?
  - Errors in reference data
  - Errors in classified map

Accuracy Assessment: Interpreting

- Errors in reference data
  - Positional error
    - Better rectification of image may help
  - Interpreter error
  - Reference medium inappropriate for classification

Accuracy Assessment: Interpreting

- Errors in classified map
  - Remotely-sensed data cannot capture classes
    - Classes are land use, not land cover
    - Classes not spectrally separable
    - Atmospheric effects mask subtle differences
    - Spatial scale of remote sensing instrument does not match classification scheme

Accuracy Assessment: Improving Classification

- Ways to deal with these problems:
  - Land use/land cover: incorporate other data
    - Elevation, temperature, ownership, distance from streams, etc.
    - Context
  - Spectral inseparability: add spectral data
    - Hyperspectral
    - Multiple dates
  - Atmospheric effects: Atmospheric correction *may* help
  - Scale: Change grain of spectral data
    - Different sensor
    - Aggregate pixels

Accuracy Assessment: Improving Classification

- Errors in classified map
  - Remotely-sensed data should be able to capture classes, but classification strategy does not draw this out
    - Minority classes swamped by larger trends in variability
      - Use HIERARCHICAL CLASSIFICATION scheme
      - In Maximum Likelihood classification, use Prior Probabilities to weigh minority classes more

Accuracy Assessment: Summary

- Choice of reference data important
  - Consider interaction between sensor and desired classification scheme
- Error matrix is foundation of accuracy assessment
- All forms of accuracy assessment should be reported to user
- Interpreting accuracy in classes can yield ideas for improvement of classification

References

- Lillesand and Kiefer, Chapter 7
- Congalton, R. G. and K. Green. 1999. *Assessing the accuracy of remotely sensed data: Principles and practices*. Lewis Publishers, Boca Raton.
- Congalton, R.G. 1991. A review of assessing the accuracy of classification of remotely sensed data. Remote Sensing of Environment 37:35-46