Accuracy Assessment

Slides by Umi Sharma

A brief note

- **"A classification is not complete until its accuracy is assessed."**
- Accuracy assessment checks **how well a classification worked**
- Understand how to interpret the *usefulness of other's classification result*.
  - Collect reference data: "ground truth"
  - Compare reference data to classified map
  - Check for **"Does class type on classified map = class type determined from reference data?"**
- It is done by preparation of a classification **error matrix**

  aka **confusion matrix** or a **contingency table**

Slides by Urmi Sharma

Error Matrix

- Error matrices compare, on a *category-by-category basis*, the relationship between **known reference data (ground truth)** and the **corresponding results** of an automated classification.
- This matrix list out the known cover types used for **training/reference data** (*columns*) versus the pixels actually **classified** into each land cover category by the classifier (*rows*).
- It tells about various classification errors.
- *Types* of classification errors:

1. Error of **Omission** (*exclusion*)

Errors of omission **refer to reference sites** that were left out (or omitted) from the correct class in the classified map.

2. Error of **Commission** (*inclusion*)

Errors of omission are in relation to the classified results. These refer sites that are classified as to reference sites that were left out (or omitted) from the correct class in the classified map.

		Reference Data			
		Water	Forest	Urban	Total
Classified Data	Water	21	6	0	**27**
	Forest	5	31	1	**37**
	Urban	7	2	22	**31**
	Total	**33**	**39**	**23**	***95***

Slides by Urmi Sharma

Sampling Considerations

- Sample selected **test pixels** or areas are **geographically representative** of the data set under analysis.
- The cover types present are determined through **ground verification** (or other **reference data**) and compared to the classification data.
- Depending upon the application, the appropriate sample unit might be *individual pixels, clusters of pixels, or polygons*.

Sample size

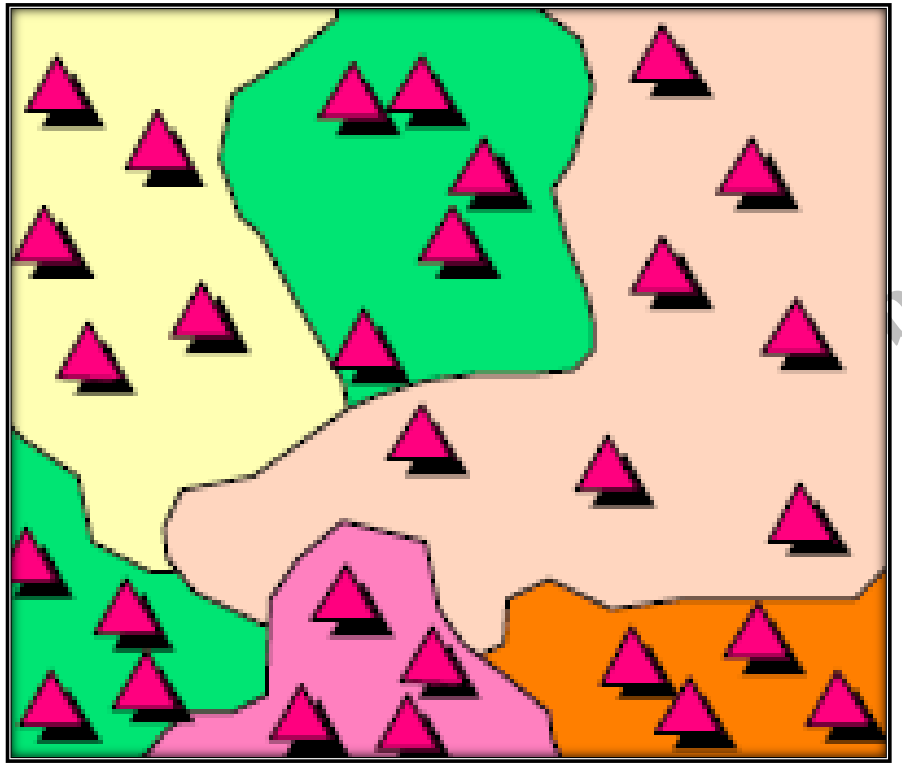
- As a broad guideline, it has been suggested that a **minimum of 50 samples of each land use category** be included in the error matrix. *(50 samples per class as rule of thumb)*
- The number of *samples for each category can be adjusted* based on the relative importance of that category for a particular application.
- Also, sampling might be allocated with respect to the variability within each category (i.e., *more samples taken in more variable categories*).

Slides by Urmi Sharma

Sampling Methods

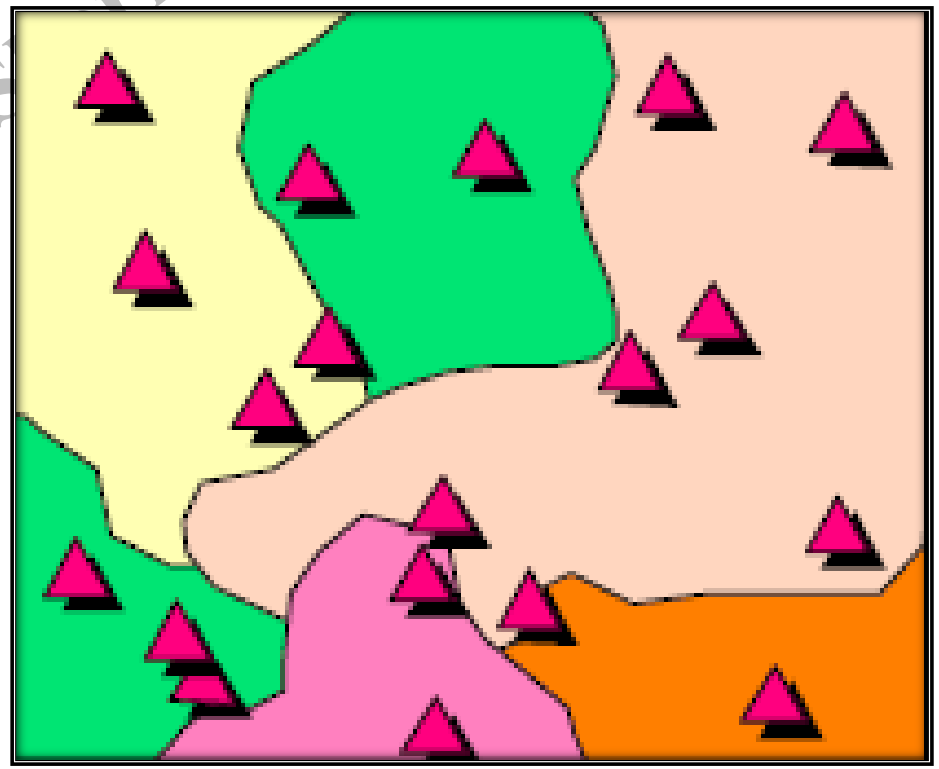
Simple Random Sampling

- observations are randomly placed.



Stratified Random Sampling

- a minimum number of observations are randomly placed in each category.



by Urmi Sharma

Systematic Sampling:

- observations are placed at *equal intervals* according to a strategy.

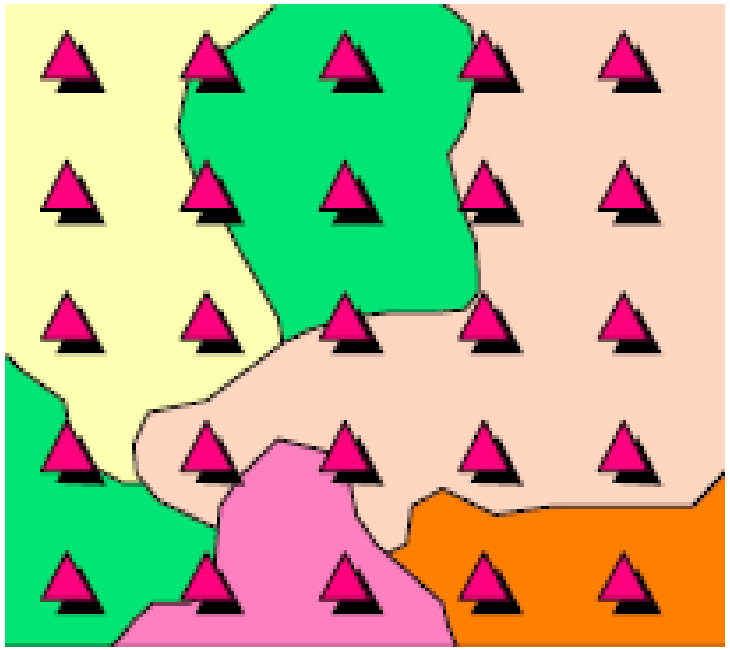
Systematic Non-Aligned Sampling :

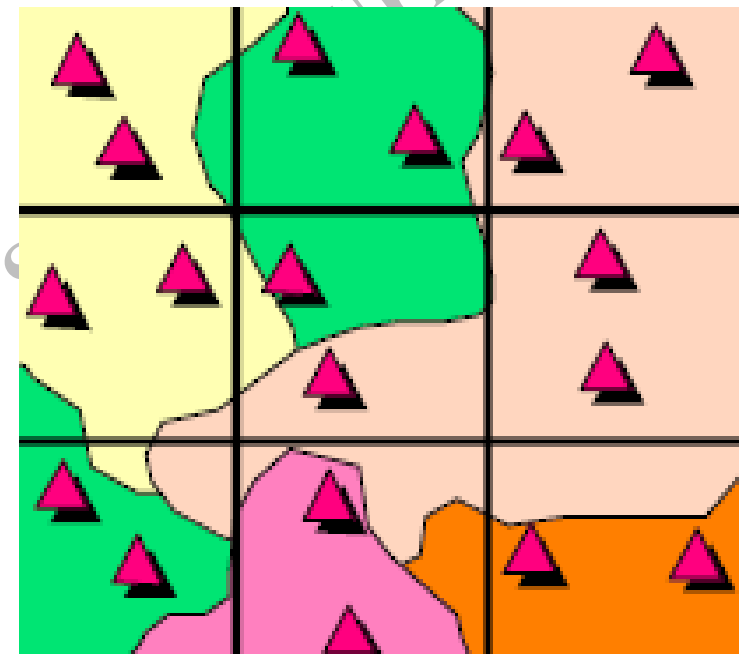
- a grid provides even distribution of randomly placed observations.

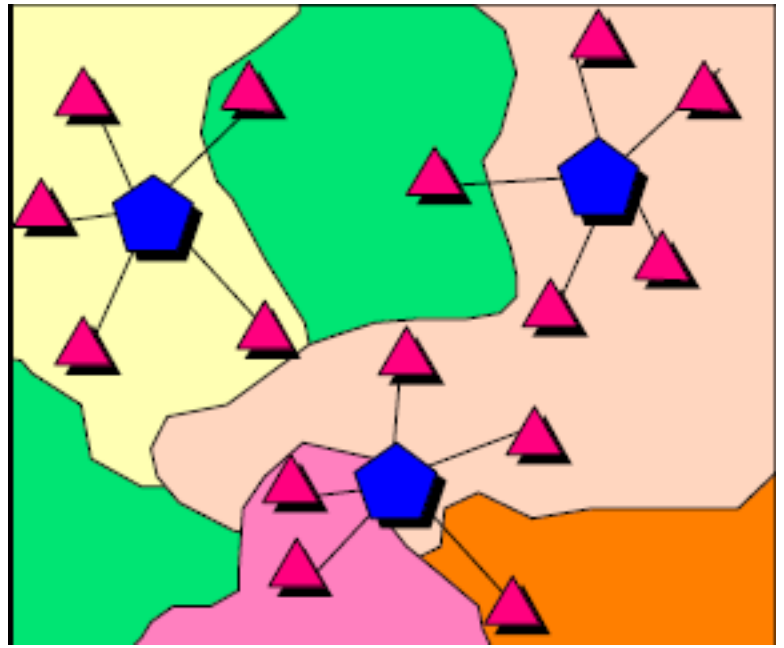
Cluster Sampling

- Randomly placed "centroids" used as a base of several nearby observations.
- The nearby observations can be randomly selected, systematically selected, etc...

Urmi Sharma







Understanding the matrix terminology

- **Diagonal elements**= **True** classified pixels
- All **nondiagonal elements** of the matrix **represent errors of omission or commission**.
- ***Omission errors*** correspond to **nondiagonal column elements**
- ***Commission errors*** are represented by **nondiagonal row elements**
  - ✓Omission: 16 pixel's that should have been classified as "sand" were omitted from that category.
  - ✓**Commission:** 38 "urban" pixels plus 79 "hay" pixels were improperly included in the "corn" category.

	Training Set Data (Known Cover Types)[a]						
Classification data	W	S	F	U	C	H	Row Total
W	480	0	5	0	0	0	485
S	0	52	0	20	0	0	72
F	0	0	313	40	0	0	353
U	0	16	0	126	0	0	142
C	0	0	0	38	342	79	459
H	0	0	38	24	60	359	481
Column total	480	68	356	248	402	438	1992

Producer's Accuracy
W = 480/480 = 100%
S = 052/068 = 76%
F = 313/356 = 88%
U = 126/248 = 51%
C = 342/402 = 85%
H = 359/438 = 82%

User's Accuracy
W = 480/485 = 99%
S = 052/072 = 72%
F = 313/353 = 87%
U = 126/142 = 89%
C = 342/459 = 74%
H = 359/481 = 75%

Overall accuracy = (480 + 52 + 313 + 126 + 342 + 359)/1992 = 84%

[a] W, water; S, sand; F, forest; U, urban; C, corn; H, hay.

Slides by Ur
Scanned with CamScanner

Various accuracies are calculated using error matrix

1. Overall Accuracy

- ✓ by dividing the *total number of correctly classified pixels* (i.e., the sum of the elements along the major diagonal) by the *total number of reference pixels*.
- ✓ the **accuracies of individual categories** is also calculated.
- ✓ by *dividing the number of correctly classified pixels in each category* by *either the total number of pixels in the corresponding row or column*.

Slides by Urmi Sharma

Overall accuracy = (480 + 52 + 313 + 126 + 342 + 359) / 1992 = 84%

Various accuracies …

2. Producer's Accuracies

- ✓ From the **perspective of the producer** 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?**
- ✓ Computed by dividing the *number of correctly classified pixels in each category* (on the major diagonal) by *the number of training set pixels used for that category* (the **column total**).
- ✓ This figure indicates **how well reference sites of the given cover type are classified**.
- ✓ It also corresponds to **error of omission (*Sand: 100-76 = 23%*)**

Slides by Urmi Sharma

Producer's Accuracy

W = 480/480 = 100%

S = 052/068 = 76%

F = 313/356 = 88%

U = 126/248 = 51%

C = 342/402 = 85%

H = 359/438 = 82%

Various accuracies …

3. User's accuracies

- ✓ 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?*
- ✓ computed by *dividing the number of correctly classified pixels in each category* by the *total number of pixels that were classified in that category* (the **row total**).
- ✓ This figure is a **measure of commission error**
- ✓ It indicates the *probability* that a **pixel classified into a given category actually represents** that category on the ground.

User's Accuracy

W = 480/485 = 99%
S = 052/072 = 72%
F = 313/353 = 87%
U = 126/142 = 89%
C = 342/459 = 74%
H = 359/481 = 75%

Slides by Urmi Sharma

Evaluating Classification Error Matrices

- ✓ Overall accuracy = 65%
- ✓ Forest = 84% (producer) (*Map is well classified for forest applications*)
- ✓ But user's accuracy for forest class is only 60%.
- That is, only 60 percent of the areas identified as "forest" within the classification are truly of that category.
- ✓ The only highly reliable associated with this classification from both a producer's and a user's perspective is **"water"**.
  - *Producer = 97%*
  - *User = 94%*

Pixels

	Reference Data[a]						
	W	S	F	U	C	H	Row Total
Classification data							
W	226	0	0	12	0	1	239
S	0	216	0	92	1	0	309
F	3	0	360	228	3	5	599
U	2	108	2	397	8	4	521
C	1	4	48	132	190	78	453
H	1	0	19	84	36	219	359
Column total	233	328	429	945	238	307	2480

Overall accuracy = (226 + 216 + 360 + 397 + 190 + 219) /2480 = 65%

PRODUCER'S ACCURACY

W = 226/233 = 97%
S = 216/328 = 66%
F = 360/429 = 84%
U = 397/945 = 42%
C = 190/238 = 80%
H = 219/307 = 71%

USER'S ACCURACY

W = 226/239 = 94%
S = 216/309 = 70%
F = 360/599 = 60%
U = 397/521 = 76%
C = 190/453 = 42%
H = 219/359 = 61%

Slides by Urm

Kappa statistic

- Reflects the difference between *actual agreement* and the *agreement expected by chance*
- **Kappa of 0.85 means there is 85% of the classified map is true not by chance**

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

Slides by Sharma

- The kappa statistic takes into account **chance agreement**, is defined as:

(observed agreement – expected agreement)/(1 – expected agreement)

- Kappa value ranges from ***0 to 1***
  - ✓ ***A value of 0 indicated that the classification is no better than a random classification.***
  - ✓ ***A value close to 1 indicates that the classification is significantly better than random.***
  - ✓ ***A negative number indicates the classification is significantly worse than random.***

$$\hat{k} = \frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{i+} \cdot x_{+i})}{N^2 - \sum_{i=1}^{r} (x_{i+} \cdot x_{+i})} \tag{7.13}$$

where

r = number of rows in the error matrix
x_{ii} = number of observations in row i and column i (on the major diagonal)
x_{i+} = total of observations in row i (shown as marginal total to right of the matrix)
x_{+i} = total of observations in column i (shown as marginal total at bottom of the matrix)
N = total number of observations included in matrix

$$\sum_{i=1}^{r} x_{ii} = 226 + 216 + 360 + 397 + 190 + 219 = 1608$$

$$\sum_{i=1}^{r} (x_{i+} \cdot x_{+i}) = (239 \cdot 233) + (309 \cdot 328) + (599 \cdot 429)$$

$$+ (521 \cdot 945) + (453 \cdot 238) + (359 \cdot 307) = 1{,}124{,}382$$

$$\hat{K} = \frac{2480(1608) - 1{,}124{,}382}{(2480)^2 - 1{,}124{,}382} = 0.57$$

Reference Data	Classes	Tree Canopy	Impervious Surfaces	Grass & Low-Lying Vegetation	Bare Soils	Open Water	Row Total	Producer's Accuracy	Errors of Omission
	Tree Canopy	494	8	9	0	0	511	96.67%	3.33%
	Impervious	1	71	0	0	0	72	98.61%	1.39%
	Grass/Vegetation	7	1	150	0	0	158	94.94%	5.06%
	Bare Soils	0	0	0	4	0	4	100.00%	0.00%
	Water	0	0	0	0	5	5	100.00%	0.00%
	Column Total	502	80	159	4	5	750		
	User's Accuracy	98.41%	88.75%	94.34%	100.00%	100.00%		Overall Accuracy	96.53%
	Errors of Commission	1.59%	11.25%	5.66%	0.00%	0.00%		Kappa Coefficent	0.9291

Slides b

Confusion Matrix Exercise

Class	Water	Forest	Agriculture	Urban	Bare Soil	Row Total
Water	13	2	0	0	0	
Forest	0	39	11	0	0	
Agriculture	0	17	26	0	0	
Urban	0	1	11	5	0	
Bare Soil	0	0	0	0	0	
Column Total						

Slides by Urmi Sharma

Confusion Matrix Exercise…

Calculate the following:

1. Overall accuracy
2. User's accuracy for each class
3. Producer's accuracy for each class
4. Kappa statistics
5. Compare the user's and producer's accuracy with overall accuracy for each class

Slides by Urmi Sharma