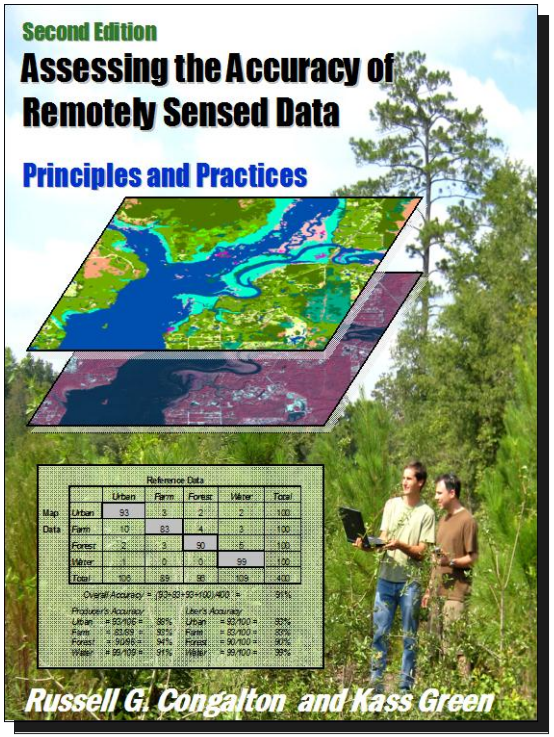
Assessing the Accuracy of Remotely Sensed Data: Doing It Right!

Russ Congalton

University of New Hampshire

Director, New Hampshire View

(russ.congalton@unh.edu)



Presented at AmericaView Meeting
Sioux Falls, SD October 5, 2009

America View Partners Need to Get this Right!

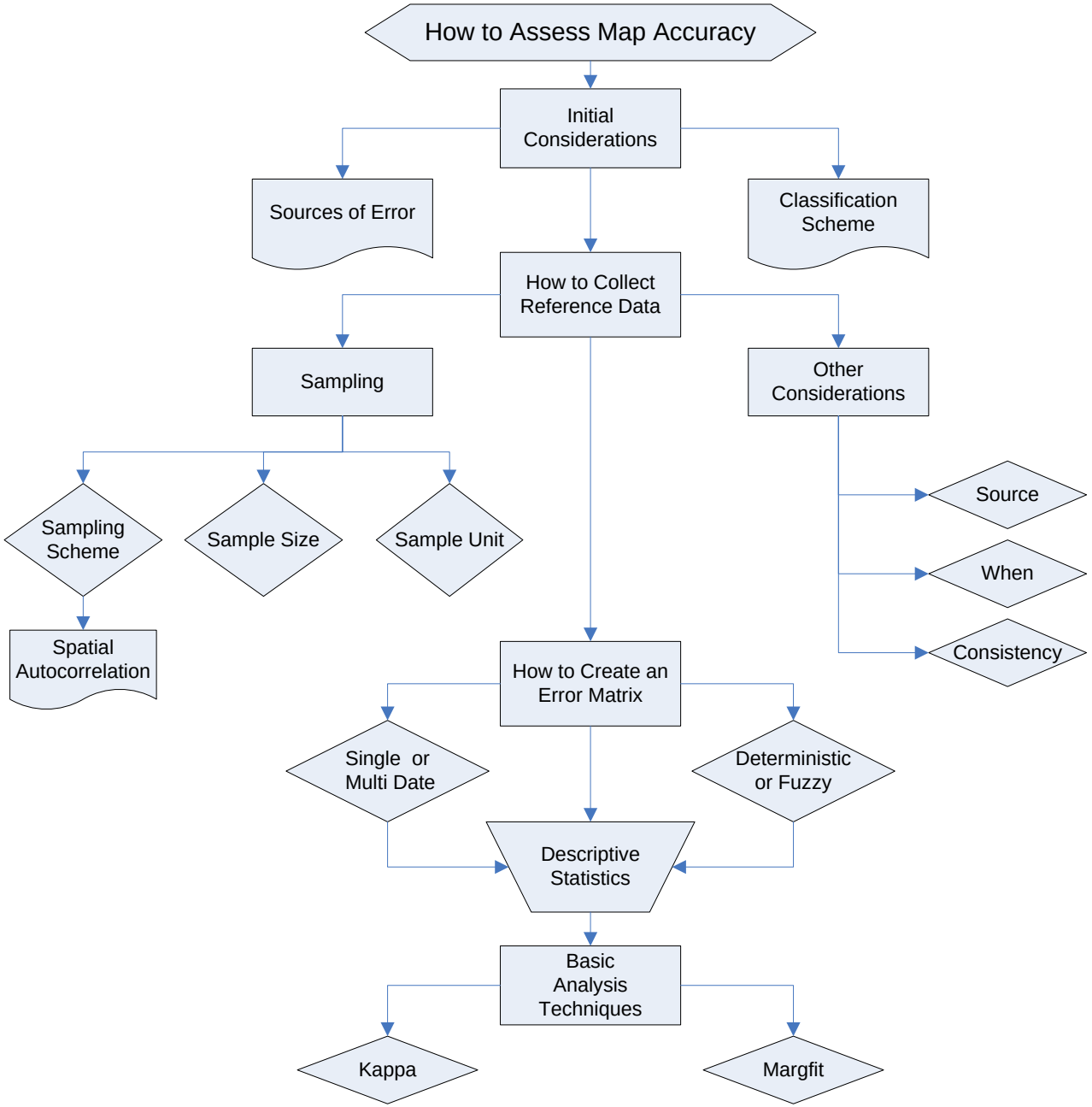
- In my role as editor of PE&RS, I have noticed that most of the papers I get these days do have some type of accuracy assessment.
- This is very good!
- However, most of the assessment are flawed to seriously flawed.
- This is very bad!

Goal of this Presentation

- I do not have time to go over everything.
- I will highlight some of the major issues and give some suggestions
- I am willing, at any time, to work with you or your partners to help them get this right.

No Single Strategy

- Unfortunately, thematic map accuracy assessment does not follow a simple recipe.
- There are many issues and considerations to think about.
- The following flow chart outlines these issues.


How to Assess Map Accuracy
Initial Considerations
Sources of Error
Classification Scheme
How to Collect Reference Data
Sampling
Other Considerations
Sampling Scheme
Sample Size
Sample Unit
Source
When
Consistency
Spatial Autocorrelation
How to Create an Error Matrix
Single or Multi Date
Deterministic or Fuzzy
Descriptive Statistics
Basic Analysis Techniques
Kappa
Margfit

Goal of the Assessment

- Balance statistical validity with practical application.
  - If it is not going to be valid, why do it?
  - If can not afford to do is right, why do it?
  - Must document your process!

6 Important Topics

- Classification Scheme
- Show the Error Matrix
- Sample Unit
- Sample Size
- Sampling Scheme
- Spatial Autocorrelation

1. Classification Scheme

- Key to any mapping project.
  - Must be done at beginning of project.
- Requirements of the Classification Scheme:
  - Meets the user's needs
  - Consists of both labels and rules (definitions) that are
    - Mutually exclusive
    - Totally exhaustive
    - Hierarchical
  - Includes a minimum mapping unit.

2. Show the Error Matrix

Map Data	Reference Data F	W	U	row total
F	28	14	15	57
W	1	15	5	21
U	1	1	20	22
column total	30	30	40	100

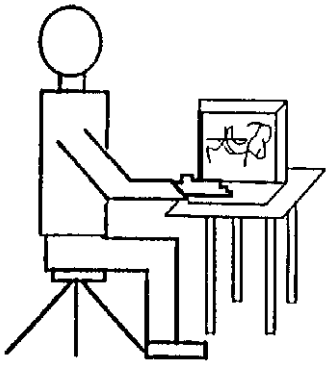
Land Cover Categories

F = Forest

W = Water

U = Urban

OVERALL ACCURACY
= 63/100 = 63%

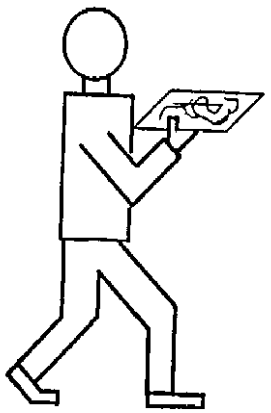


PRODUCER'S ACCURACY

F = 28/30 = 93%
W = 15/30 = 50%
U = 20/40 = 50%

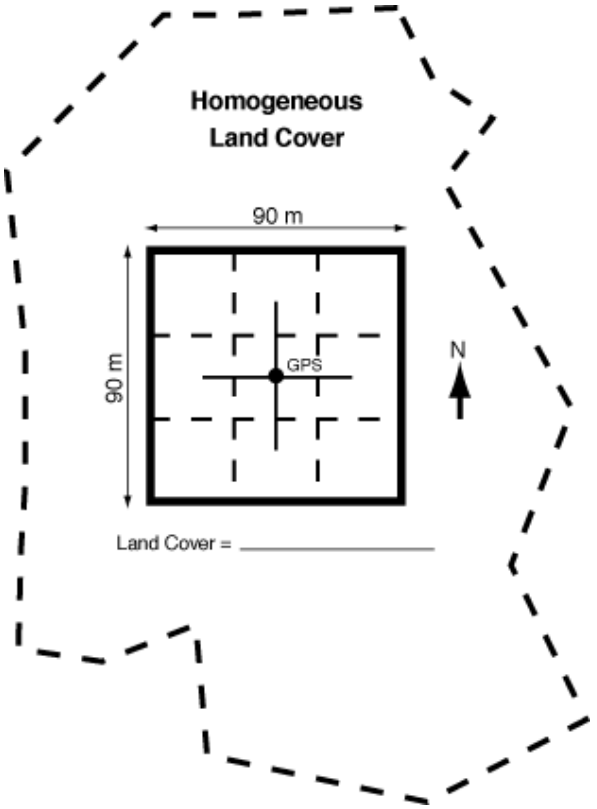
USER'S ACCURACY

F = 28/57 = 49%
W = 15/21 = 71%
U = 20/22 = 91%



3. Sample Unit

- Must consider positional accuracy and mmu
- Historically, 3 choices:
  - Single pixel
  - Cluster of pixels
  - Polygon
- Really only 2 choices:
  - Cluster of pixels
  - Polygon



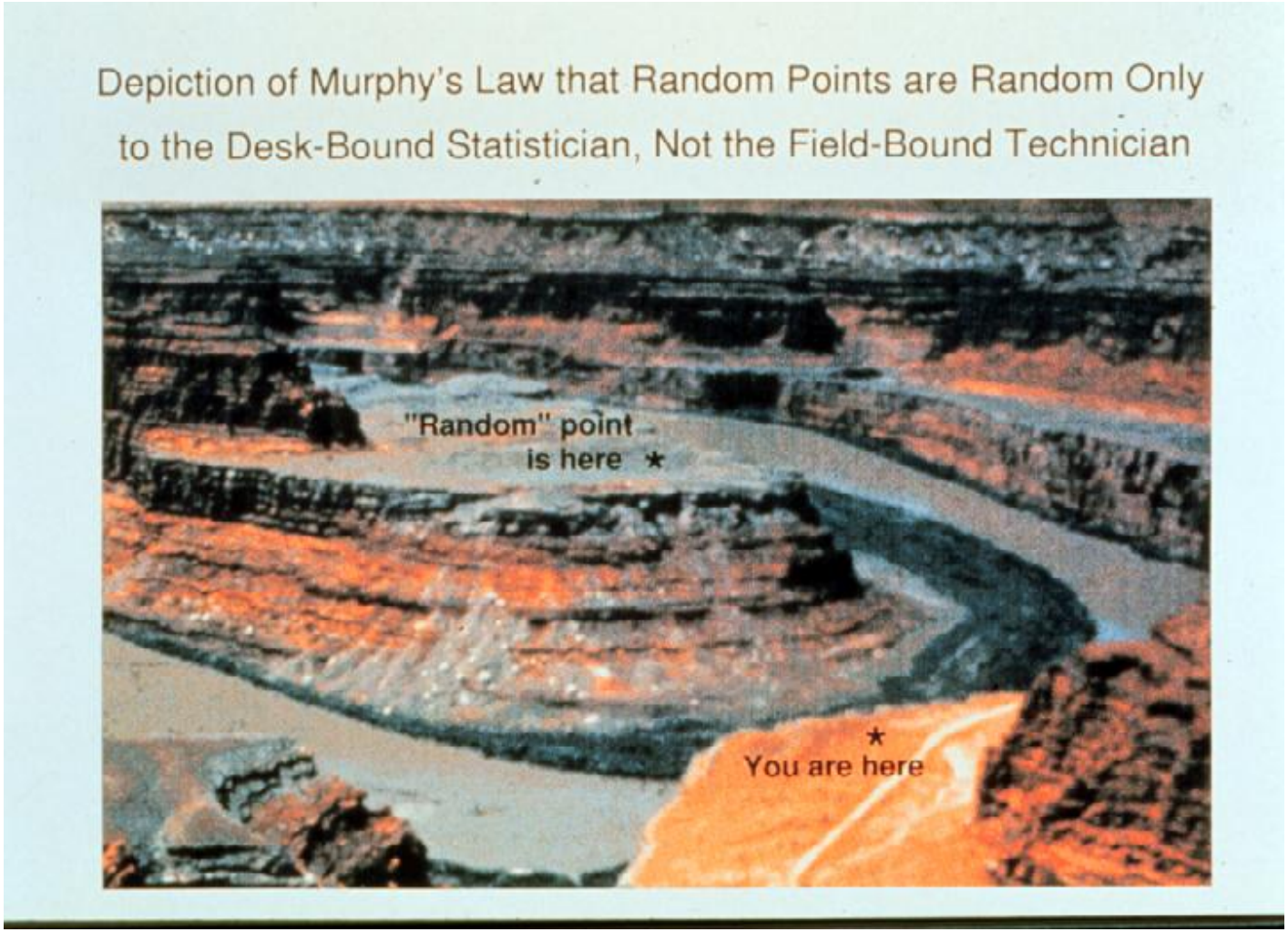
4. Sample Size

- Rule of thumb: 50 sample units per map class

OR

- Equations to compute sample size
  - Binomial: simply right and wrong
  - multinomial **
- Need enough samples to insure good distribution across the map (avoid spatial autocorrelation)
- Samples independent of training data
- If assess map of the state, the results are for the entire state. If need county estimate, need to do another assessment.

5. Sampling Scheme



6. Spatial Autocorrelation

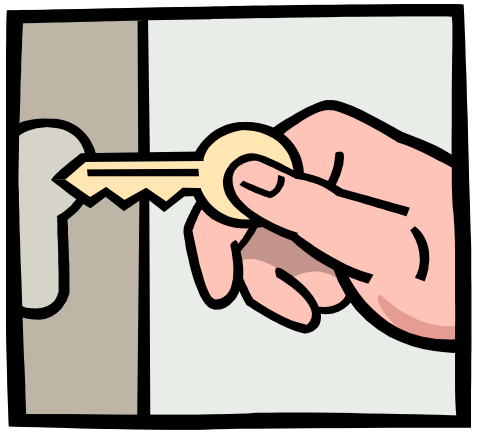
- Spatial autocorrelation occurs when the presence, absence, or degree of a certain characteristic affects the presence, absence, or degree of the same characteristic in neighboring units (Cliff and Ord 1973)
- Samples must be adequately spaced apart or they will be spatially autocorrelated.

A Final Favor

Please use the term – “Reference Data” or “Ground data”

Summary

- If you are going to spend the effort to conduct an AA, please do it right.
- Think through the process before you begin.
- Balance statistical validity with practical application.
- Document your process and show your results.




Thanks