

Exercise #2

Building a Thematic Error Matrix from a Random Sample

(an exercise to accompany the AmericaView Video Lessons on Assessing the Accuracy of Remotely Sensed Data)

This exercise accompanies the AmericaView Accuracy Assessment Lessons prepared by Dr. Russell G. Congalton, University of New Hampshire and Director of NewHampshireView.

Objective: To randomly sample a map to create a thematic error matrix and see the impact of the sampling scheme.

Instructions: You will select **80 random samples** to compare the map to the reference data and create an error matrix as follows:

1. You have been provided with a map that was generated from remotely sensed data and a reference data set. (The goal here is to see the impact of random sampling on your assessment and so we will not worry about the sample unit size here)
2. Use a random number table or better yet, a computer application, to generate 80 random samples providing an x and a y coordinate. The x coordinate must range between 0 – 24 and the y coordinate between 0 – 37. Any duplicate pairs should be discarded.
3. Plot the locations of the x, y samples on both map and the reference data and then use this information to create a table like this:

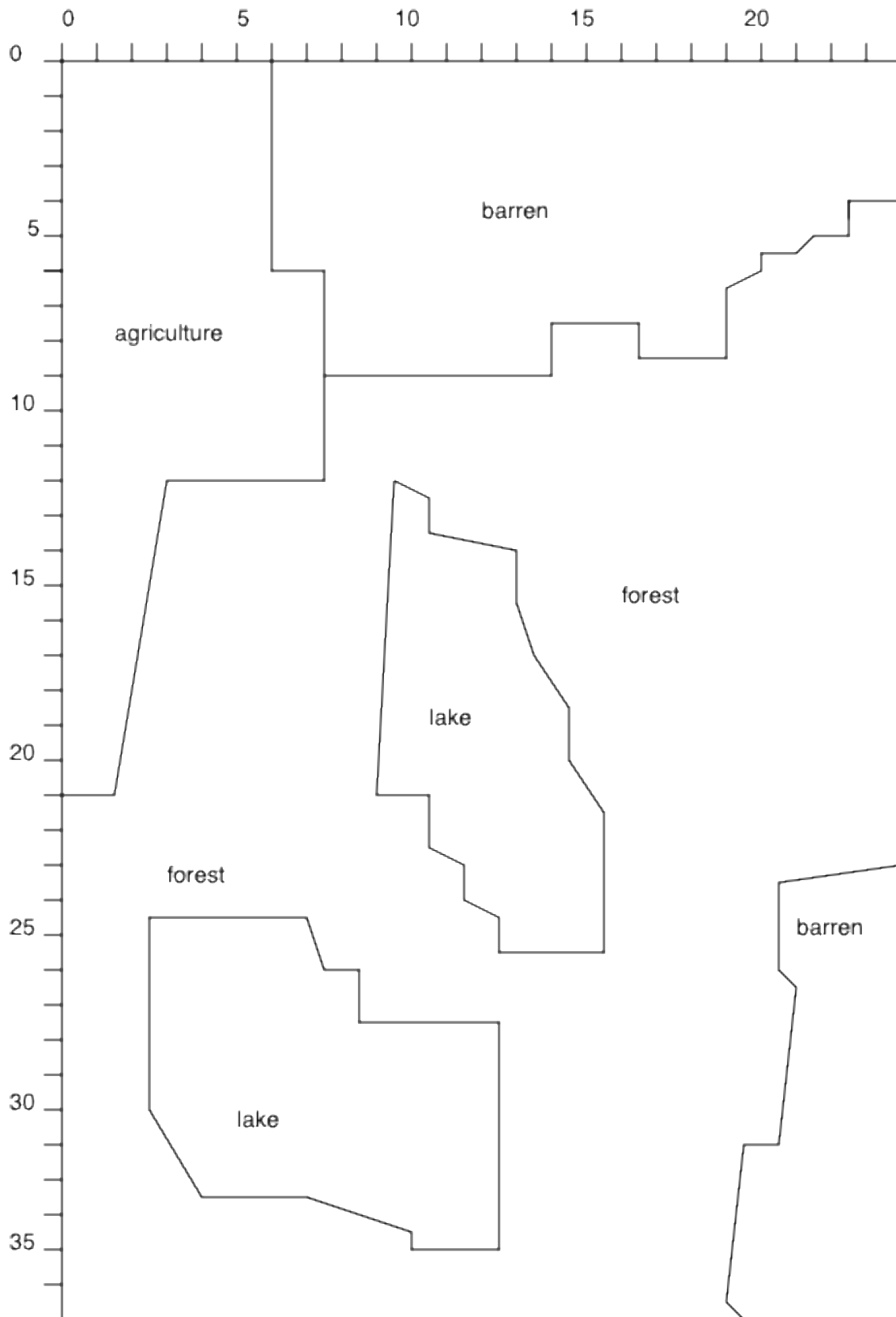
Coordinate (x,y)	Map Label	Reference Label
5, 10	barren	barren
13, 5	forest	lake
30, 5	lake	lake
Etc.	Etc.	Etc.

4. From this table, conduct the thematic accuracy assessment by compiling an error matrix for this 4 land cover class map (the matrix should be 4 x4 (Agriculture, Barren, Forest, and Lake) with the reference data on the x axis and the map on the y axis).
5. Once the error matrix is compiled, compute the overall, producer's and user's accuracies.

Questions to think about:

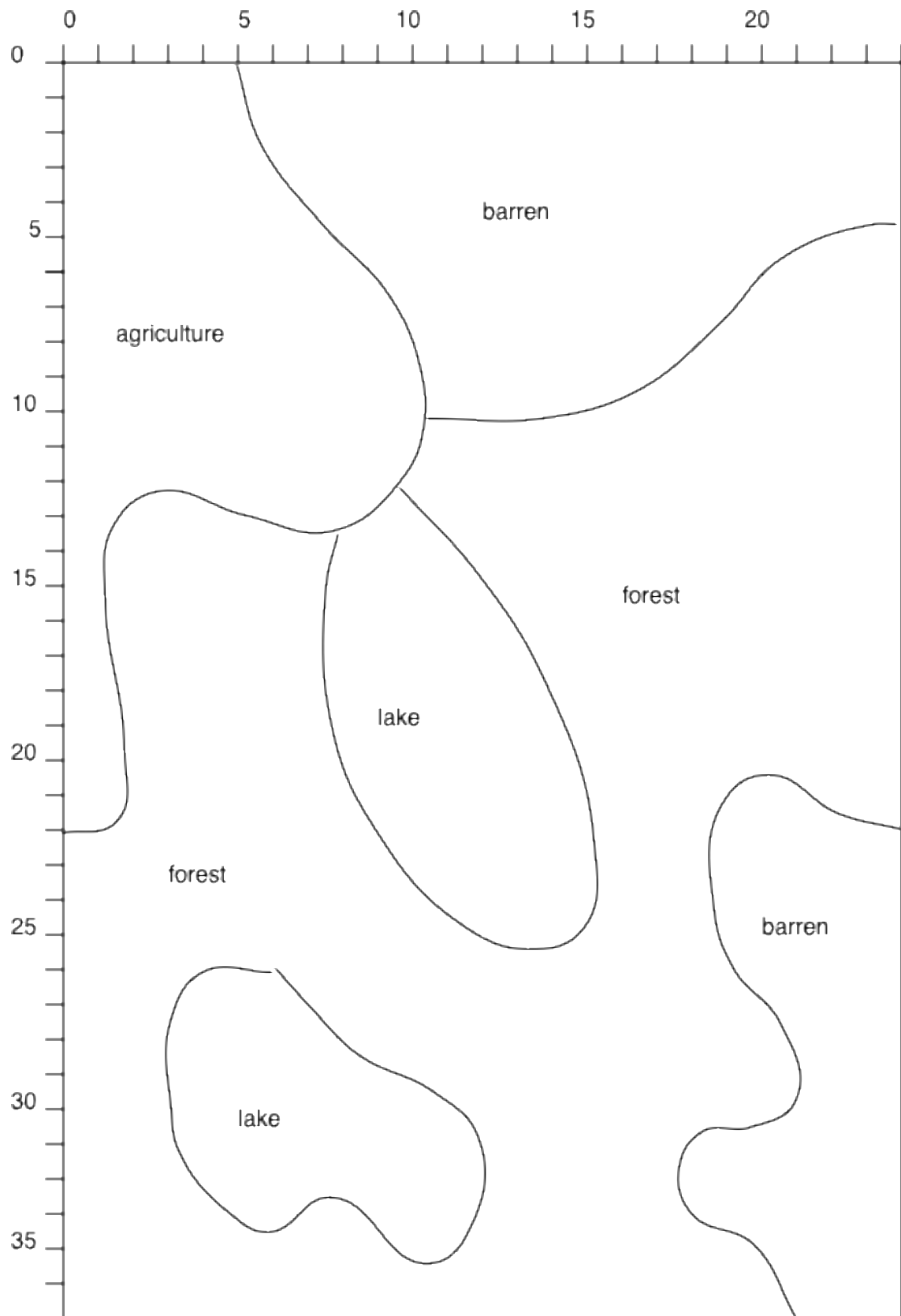
1. What did you do when a sample fell on the boundary between two map or reference classes? What should you do to make sure this does not happen during your actual assessment?
2. Examine the number of samples that you have for each map class. Are there sufficient samples for this to be a valid assessment? Do some map classes have many more samples than the others? Why is this? What could you do to solve this issue?

Exercise #2 – Random Sampling.



CLASSIFIED DATA

Exercise #2 – Random Sampling.



REFERENCE DATA

Exercise #2 – Random Sampling.