Accuracy assessment

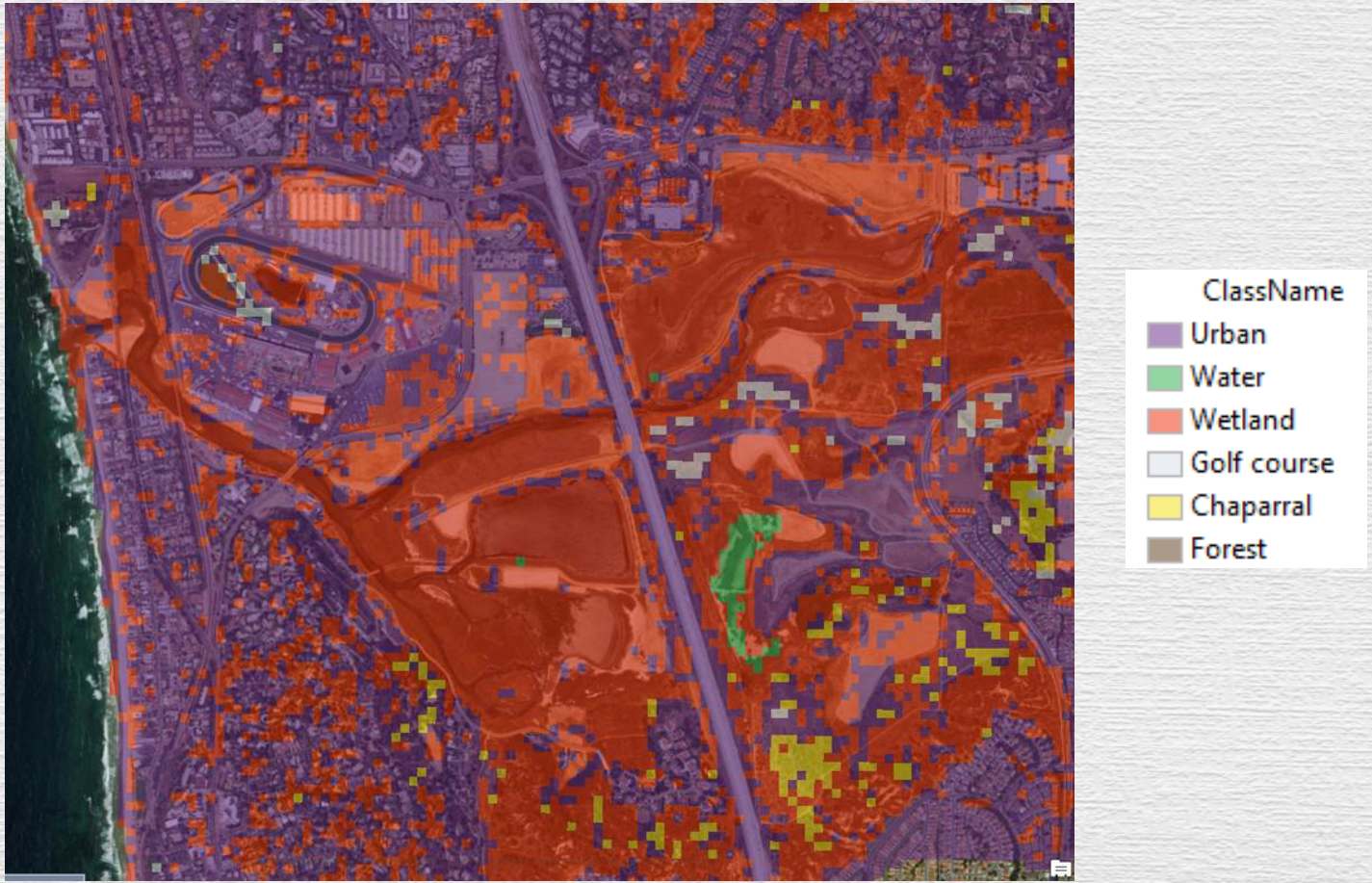
Cover type maps are never completely accurate

- Human error
  - Polygons that don’t correspond well with patches on the ground
  - Assigning cover types incorrectly
- Statistical classification error
  - Pixels with spectral values that are closer to the wrong cover type than to the correct one
- We can ground truth, and fix mistakes that we detect
- But, ground truthing every pixel in the watershed is impractical

Do the lines correspond to changes in land use on the ground?

Are the categories right?

Supervised classification – accurate?



Measuring accuracy

- If you can’t fix every error, you can at least measure the error rate
  - Measure of the reliability of the map
  - Interpretable as the probability that a randomly selected pixel or feature will be correctly classified
- We do this by comparing the actual cover type at a sample of locations to the mapped cover type
- A cross-tabulation of observed cover types and mapped (or **predicted**) cover types is constructed = the **confusion matrix**

Random sample of points

Record actual cover type at each location

Overlay points with cover type map to get predicted cover type

Data table

	A	B	C	D	E	F	G
1	FID	Shape *	FID_rand_s	CID	FID_landus	LU_current	GroundTruth
2	0	Point	15	0	1590	Undeveloped	Undeveloped
3	1	Point	50	0	1590	Undeveloped	Undeveloped
4	2	Point	3	0	1590	Undeveloped	Residential
5	3	Point	4	0	1590	Undeveloped	Undeveloped
6	4	Point	82	0	1590	Undeveloped	Undeveloped
7	5	Point	88	0	109	Beach	Developed
8	6	Point	56	0	2935	Residential	Residential
9	7	Point	73	0	549	Transportation	Transportation
10	8	Point	59	0	1590	Undeveloped	Undeveloped
11	9	Point	37	0	1590	Undeveloped	Undeveloped
12	10	Point	92	0	2952	Residential	Transportation
13	11	Point	86	0	2670	Transportation	Transportation
14	12	Point	25	0	2952	Residential	Residential
15	13	Point	79	0	1590	Undeveloped	Undeveloped
16	14	Point	76	0	549	Transportation	Transportation
17	15	Point	84	0	2670	Transportation	Transportation
18	16	Point	65	0	2972	Residential	Residential
19	17	Point	53	0	549	Transportation	Transportation
20	18	Point	43	0	1590	Undeveloped	Undeveloped

Confusion matrix

Count of LU_current	Column Labels								
Row Labels	Bay or Lagoon	Beach	Commercial	Developed	Recreational	Residential	Transportation	Undeveloped	Grand Total
Bay or Lagoon	9								9
Beach		1							1
Commercial			1						1
Developed		1		9				1	11
Recreational					1				1
Residential						19		2	21
Transportation						1	12		13
Undeveloped								43	43
Grand Total	**9**	**2**	**1**	**9**	**1**	**20**	**12**	**46**	**100**

The actual (ground truthed) cover type is identified in the rows
Mapped (predicted) cover type are in the columns

Correctly classified points are the main diagonal, where observed = predicted (blue box)

Numbers that are not in the main diagonal are mismatches (red circles)

Measures of accuracy of the map

- Overall accuracy = an overall measure of correct classification
  - Sum of the correctly classified points divided by the total
- Producer's accuracy = a measure of how often the actual class is the same as the predicted class
- User's accuracy = a measure of how often the predicted class is correct

Overall accuracy

Count of LU_current	Column Labels								
Row Labels	**Bay or Lagoon**	**Beach**	**Commercial**	**Developed**	**Recreational**	**Residential**	**Transportation**	**Undeveloped**	**Grand Total**
Bay or Lagoon	9								9
Beach		1							1
Commercial			1						1
Developed		1		9				1	11
Recreational					1				1
Residential						19		2	21
Transportation						1	12		13
Undeveloped								43	43
Grand Total	**9**	**2**	**1**	**9**	**1**	**20**	**12**	**46**	**100**

Correctly classified points = 9 + 1 + 1 + 9 + 1 + 19 + 12 + 43 = 95

Overall accuracy = 95/100 = 0.95, or 95%

Producer's accuracy

Count of LU_current	Column Labels								
Row Labels	Bay or Lagoon	Beach	Commercial	Developed	Recreational	Residential	Transportation	Undeveloped	Grand Total
Bay or Lagoon	9								9
Beach		1							1
Commercial			1						1
Developed		1		9				1	11
Recreational					1				1
Residential						19		2	21
Transportation						1	12		13
Undeveloped								43	43
Grand Total	**9**	**2**	**1**	**9**	**1**	**20**	**12**	**46**	**100**

9/9 = 1
1/1 = 1
1/1 = 1
9/11 = 0.82
1/1 = 1
19/21 = 0.90
12/13 = 0.92
43/43 = 1

Assessed for the actual cover types (each row in the table)

Answers the question, what is the chance an actual cover type was mapped correctly?

User's accuracy

Count of LU_current	Column Labels								
Row Labels	Bay or Lagoon	Beach	Commercial	Developed	Recreational	Residential	Transportation	Undeveloped	Grand Total
Bay or Lagoon	9								9
Beach		1							1
Commercial			1						1
Developed		1		9				1	11
Recreational					1				1
Residential						19		2	21
Transportation						1	12		13
Undeveloped								43	43
Grand Total	**9**	**2**	**1**	**9**	**1**	**20**	**12**	**46**	**100**

9/9 = 1	1/2 = 0.5	1/1 = 1	9/9 = 1	1/1 = 1	19/20 = 0.95	12/12 = 1	43/46 = 0.93

Assessed for each column in the table

Answers the question, what is the chance that a cover type on the map is accurate?

Error rates

- The converse of accuracy
- Overall error rate = number of incorrect points divided by the total
- Omission error rate = how often a cover type was predicted to be something else
- Commission error rate = how often a predicted cover type was incorrect

Overall error rate

Count of LU_current	Column Labels								
Row Labels	**Bay or Lagoon**	**Beach**	**Commercial**	**Developed**	**Recreational**	**Residential**	**Transportation**	**Undeveloped**	**Grand Total**
Bay or Lagoon	9								9
Beach		1							1
Commercial			1						1
Developed		1		9				1	11
Recreational					1				1
Residential						19		2	21
Transportation						1	12		13
Undeveloped								43	43
Grand Total	**9**	**2**	**1**	**9**	**1**	**20**	**12**	**46**	**100**

Incorrectly classified points = 1 + 1 + 1 + 2 = 5

Overall error rate = 5/100 = 0.05, or 5%

Also, 1 – overall accuracy = 1 – 0.95 = 0.05, or 5%

Omission error

Count of LU_current	Column Labels									
Row Labels	Bay or Lagoon	Beach	Commercial	Developed	Recreational	Residential	Transportation	Undeveloped	Grand Total	
Bay or Lagoon	9								9	0/9 = 0
Beach		1							1	0/1 = 0
Commercial			1						1	0/1 = 0
Developed		1		9				1	11	2/11 = 0.18
Recreational					1				1	0/1 = 0
Residential						19		2	21	2/21 = 0.10
Transportation						1	12		13	1/13 = 0.08
Undeveloped								43	43	0/43 = 0
Grand Total	**9**	**2**	**1**	**9**	**1**	**20**	**12**	**46**	**100**	

Number of points incorrectly mapped divided by total number of points checked in each cover type – by rows

Measure of how often each actual cover type is mapped incorrectly

Commission error

Count of LU_current	Column Labels								
Row Labels	Bay or Lagoon	Beach	Commercial	Developed	Recreational	Residential	Transportation	Undeveloped	Grand Total
Bay or Lagoon	9								9
Beach		1							1
Commercial			1						1
Developed		1		9				1	11
Recreational					1				1
Residential						19		2	21
Transportation						1	12		13
Undeveloped								43	43
Grand Total	**9**	**2**	**1**	**9**	**1**	**20**	**12**	**46**	**100**

0/9 = 0 1/2 = 0.5 0/1 = 0 0/9 = 0 0/1 = 0 1/20 = 0.05 0/12 = 0 3/46 = 0.07

Number of mis-mapped cover type errors divided by the total number of points checked with that mapped cover type. Calculated by column.

Measure of how often polygons on the map are in error for each cover type

Cover type maps become less accurate over time

- Changes happen on the ground
  - Development
  - Fire
  - Succession
- These changes accumulate over time
- Cover type maps become less accurate indicators of current conditions as time passes
- These inaccuracies may not be errors – may have been accurate at the time of mapping
- But, if you wish to use a map to represent current conditions, then these changes contribute to inaccuracy