

eurac
research

Classification Accuracy Assessment

Alexander Jacob
Eurac Research



Overview

- "A classification is not complete until its accuracy is assessed."
- Classification Error Matrix (confusion matrix, contingency table)
 - Compares, on a category-by-category basis, the relationship between known reference data (ground truth) and the result of the classification.

Overview

- Often, the ground truth data is listed in the columns and the result from the classification in the rows of the matrix, but could be the other way around too. All non-diagonal elements thus represents errors.
 - Error of omission (exclusion, user accuracy)
 - Error of commission (inclusion, producer accuracy)

Accuracy Estimation

Reference Data

Classified Data

Confusion Matrix with Producer's and User's Accuracy in %													
Name	ID	1	2	3	4	5	6	7	8	9	10	Row total	User
High density	1											4272	88,0
Low density	2											2107	60,6
Bare field	3											1060	79,2
Low vegetation	4											1075	76,2
Forest	5											1009	61,8
Golf course	6											1024	79,5
Grass / Pasture	7											1084	53,3
Water	8											3677	96,2
Airport	9											1110	90,9
Major Roads	10											1015	52,2
Colum total		4230	2172	927	1309	718	975	1327	3576	1223	976	17433	
Producer's Acc.		88,9	58,8	90,6	62,6	86,9	83,5	43,6	98,9	82,5	54,3		
Average accuracy	73,78	%											
Overall accuracy	79,09	%											
Kappa Coefficient	0,756	#											

- Confusion matrix
- Overall accuracy
- User / producers accuracy
- Kappa value

Accuracy Estimation

Reference Data

Classified Data

Confusion Matrix with Producer's and User's Accuracy in %													
Name	ID	1	2	3	4	5	6	7	8	9	10	Row total	User
High density	1		222	4	0	0	0	166	0	94	26	4272	88,0
Low density	2	261		0	154	32	13	263	0	0	107	2107	60,6
Bare field	3	87	6		22	0	0	2	7	27	69	1060	79,2
Low vegetation	4	0	177	0		43	4	0	0	0	32	1075	76,2
Forest	5	0	180	0	196		9	0	0	0	0	1009	61,8
Golf course	6	0	59	0	51	0		96	0	0	4	1024	79,5
Grass / Pasture	7	3	157	0	58	0	135		0	56	97	1084	53,3
Water	8	15	4	29	7	0	0	0		0	85	3677	96,2
Airport	9	0	0	0	0	0	0	75	0		26	1110	90,9
Major Roads	10	104	90	54	2	19	0	147	32	37		1015	52,2
Colum total		4230	2172	927	1309	718	975	1327	3576	1223	976		
Producer's Acc.		88,9	58,8	90,6	62,6	86,9	83,5	43,6	98,9	82,5	54,3		
Average accuracy	73,78	%											
Kappa Coefficient	0,756	#											

- Confusion matrix
- Overall accuracy
- User / producers accuracy
- Kappa value

Accuracy Estimation

Reference Data

Classified Data

Confusion Matrix with Producer's and User's Accuracy in %													
Name	ID	1	2	3	4	5	6	7	8	9	10	Row total	User
High density	1	3760	222	4	0	0	0	166	0	94	26	4272	88,0
Low density	2	261	1277	0	154	32	13	263	0	0	107	2107	60,6
Bare field	3	87	6	840	22	0	0	2	7	27	69	1060	79,2
Low vegetation	4	0	177	0	819	43	4	0	0	0	32	1075	76,2
Forest													
Golf course	6	0	59	0	51	0	814	96	0	0	4	1024	79,5
Grass / Pasture	7	3	157	0	58	0	135	578	0	56	97	1084	53,3
Water	8	15	4	29	7	0	0	0	3537	0	85	3677	96,2
Airport	9	0	0	0	0	0	0	75	0	1009	26	1110	90,9
Major Roads	10	104	90	54	2	19	0	147	32	37	530	1015	52,2
Colum total		4230	2172	927	1309	718	975	1327	3576	1223	976	17433	
Producer's Acc.		88,9	58,8	90,6	62,6	86,9	83,5	43,6	98,9	82,5	54,3		
Average accuracy	73,78	%											
Overall accuracy	79,09	%											
Kappa Coefficient	0,756	#											

- Confusion matrix
- Overall accuracy
- User / producers accuracy
- Kappa value

Accuracy Estimation

Reference Data

Classified Data

Confusion Matrix with Producer's and User's Accuracy in %													
Name	ID	1	2	3	4		6	7	8	9	10	Row total	User
High density	1	3760	222	4	0		0	166	0	94	26	4272	88,0
Low density	2	261	1277	0	154		13	263	0	0	107	2107	60,6
Bare field	3	87	6	840	22		0	2	7	27	69	1060	79,2
Low vegetation	4	0	177	0	819		4	0	0	0	32	1075	76,2
Forest	5	0	180	0	196		9	0	0	0	0	1009	61,8
Golf course	6	0	59	0	51		814	96	0	0	4	1024	79,5
Grass / Pasture	7	3	157	0	58		135	578	0	56	97	1084	53,3
Water	8	15	4	29	7		0	0	3537	0	85	3677	96,2
Airport	9	0	0	0	0		0	75	0	1009	26	1110	90,9
Major Roads	10	104	90	54	2		0	147	32	37	530	1015	52,2
Colum total		4230	2172	927	1309		975	1327	3576	1223	976	17433	
Producer's Acc.		88,9	58,8	90,6	62,6		83,5	43,6	98,9	82,5	54,3		
Average accuracy	73,78	%											
Overall accuracy	79,09	%											
Kappa Coefficient	0,756	#											

- Confusion matrix
- Overall accuracy
- User / producers accuracy
- Kappa value

Accuracy Estimation

Reference Data

Classified Data

Confusion Matrix with Producer's and User's Accuracy in %													
Name	ID	1	2	3	4	5	6	7	8	9	10	Row total	User
High density	1		222									4272	88,0
Low density	2	261		0	154	32	13	263	0	0	107	2107	60,6
Bare field	3	87	6		22	0	0	2	7	27	69	1060	79,2
Low vegetation	4	0	177	0		43	4	0	0	0	32	1075	76,2
Forest	5	0	180	0	196		9	0	0	0	0	1009	61,8
Golf course	6	0	59	0	51	0		96	0	0	4	1024	79,5
Grass / Pasture	7	3	157	0	58	0	135		0	56	97	1084	53,3
Water	8	15	4	29	7	0	0	0		0	85	3677	96,2
Airport	9	0	0	0	0	0	0	75	0		26	1110	90,9
Major Roads	10	104	90	54	2	19	0	147	32	37		1015	52,2
Column total		4230	2172	927	1309	718	975	1327	3576	1223	976		
Producer's Acc.		88,9	58,8	90,6	62,6	86,9	83,5	43,6	98,9	82,5	54,3		
Average accuracy	73,78	%											
Overall accuracy	79,09	%											

- Confusion matrix
- Overall accuracy
- User / producers accuracy
- Kappa value

$$K = \frac{\text{Yellow Box} - \text{Green Box}}{\text{Light Green Box} - \text{Dark Green Box}}$$

Confusion matrix as heatmap

