Evaluation of supervised classification by class and classification characteristics

Antoine Masse, Danielle Ducrot, Philippe Marthon

CESBIO - Centre d'Etudes Spatiales de la Biosphère,
18, avenue Edouard Belin, bpi 2801, F-31401 Toulouse cedex 9.
Phone: +33 (0)5 61 55 85 31
e-mail: antoine.masse@cesbio.cnes.fr, danielle.ducrot@cesbio.cnes.fr, philippe.marthon@enseeiht.fr

Abstract : As acquisition technology progresses, remote sensing data contains an ever increasing amount of information: optical and radar images, low, high and very high-resolution, hypertemporal hyperspectral images, derived images, and physical or ancillary data (databases, Digital Elevation Model (D.E.M), Geographical Information System (G.I.S.)). Future projects in remote sensing will give high repeatability of acquisition like Venμs (CNES) which may provide data every 2 days with a resolution of 5.3 meters on 12 bands (420nm-900nm) and Sentinel-2 (ESA) 13 bands, 10-60m resolution and 5 days. With such data, supervised classification gives excellent results in term of accuracy indices (like Overall Accuracy, Kappa coefficient). In this paper, we present advantages and disadvantages of existing indices and propose a new index to evaluate supervised classification using all the information available from the confusion matrix. In addition to accuracy, a new feature is introduced in this index: fidelity. For example, a class could have a high accuracy (low omission error) but could be over-represented with other classes (high commission error). The new index reflects accuracy and correct representation of classes (fidelity) using commission and omission errors. Environment applications are in land cover and land use and the goal is to have the best classification for all classes, whether the biggest (corn, trees) or the lightest (rivers, hedges). The tests are performed on Formosat-2 images (every 2 days, 8 meters resolution on 4 bands) in the area of Toulouse (France). Tests used to validate the new index by demonstrating benefits of its use through various thematical studies.

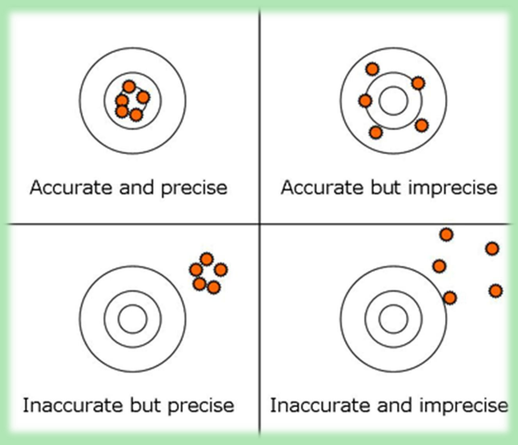
Overview and novelties for supervised classification evaluation

- A class is defined by expert as a thematical object spatially and spectrally by reference data.
  - spectral part is resumed by statistical information like means, covariance and extremum
  - spatial characteristics are size, representation on the map and confusion with others.
- These spatial information are resumed by confusion matrix

Class indices (indices for evaluation of one class)

Class accuracy

- → Faithful measurement or representation of the truth.
- → Correctness.
- → Degree of agreement between a measured value and the accepted value for that measurement.
- → Diagonal value of confusion matrix.



Class representation

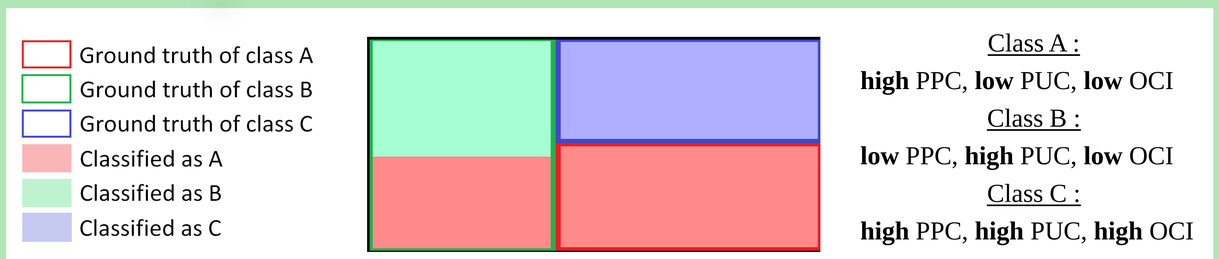
- → Faithful measurement of reproduction.
- → Fidelity.
- → Reproducibility or repeatability.
- → Degree to which repeated measurements show the same results.
- → Based on non-diagonal value of confusion matrix.

PPC index — Producer accuracy index, consider omission error
no consideration of commission error

PUC index — User accuracy index, consider commission error
no consideration of omission error

We introduce a new index to combine both of **omission and commission** errors:

OCI — Omission and commission index, consider both of omission and commission errors



Class A : **high** PPC, **low** PUC, **low** OCI
Class B : **low** PPC, **high** PUC, **low** OCI
Class C : **high** PPC, **high** PUC, **high** OCI

Comments:

We can distinguish 2 types of class characteristics:

- Dispersible classes which have high omission and low commission errors (example: B)
- Absorbing classes which have low omission and high commission errors (example: A)

Classification indices (indices for evaluation of all classes)

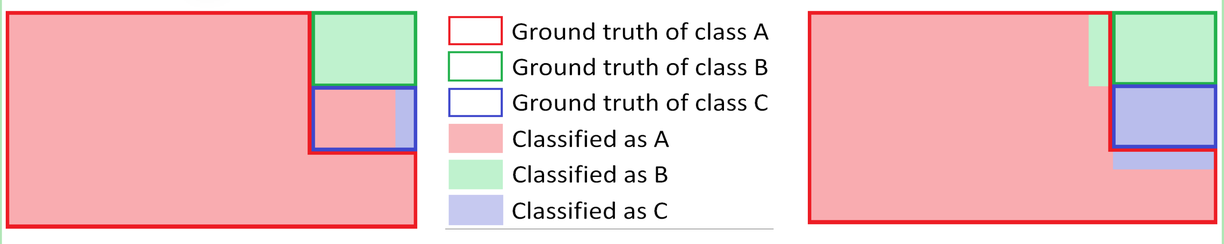
OA index — accuracy index
only considers diagonal terms of the confusion matrix

AA index — accuracy index with normalization of class size
only considers diagonal terms of the confusion matrix

kappa index — accuracy index and uses sums of rows and columns
no explicit calculation of the precision and combination with accuracy

We introduce a new index which is a compromise between **high overall accuracy** and **low commission errors** for each class is index derived from OCI: Overall Omission and Commission Index (OOCI), which is OCI mean.

OOCI — combination of two characteristics: class overrepresentation (a class characteristic) accuracy (a classification characteristic)

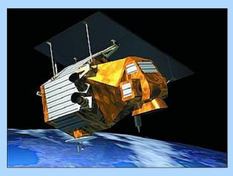


high OA, **low** AA, **high** Kappa, **low** OOCI

high OA, **middle** AA, **high** Kappa, **low** OOCI

ACKNOWLEDGEMENTS: G. Dedieu and O. Hagolle for Formosat 2 correction, J.F. Dejoux and C. Marais-Sicre for ground truths.

APPLICATIONS AND DISCUSSION



Satellite data: Formosat-2 (NSPO, Taiwan) images
high temporal revisit (2 days)
high spatial resolution (8m)
16 dates in **2009** (02/15, 03/17, 03/21, 03/30, 05/03, 06/23, 07/01, 07/12, 07/26, 08/05, 08/14, 08/22, 08/30, 09/06, 09/24 and 10/16).

Ground truth data: Agricultural site located in the South-West of Toulouse, France.

A selection algorithm (presented in [Masse, A. et al, "Tools for multitemporal analysis and classification of multisource satellite imagery", IEEE Analysis of Multi-temporal Remote Sensing Images]) automatically selects the best dataset of images that maximizes each classification quality index

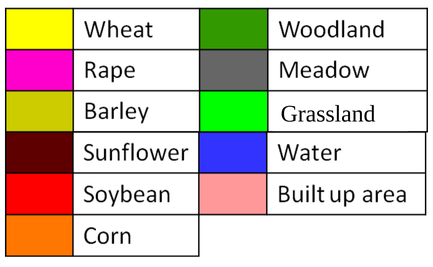
Results

Extracted image from classification of the best dataset for all classes

➢ Over representation of sunflower (brown) to the detriment of built (pink) and trees (green)

Extracted image from classification of the best dataset for "Sunflower" class

➢ Good representation and precision of sunflower (brown)



Sunflower class evaluation	Indices values
PPC	99,1 %
PUC	79,5 %
OCI	**78,8 %**
Classification evaluation	
OA	96,3 %
AA	94,1 %
Kappa	92,3 %
OOCI	**84,1 %**

Sunflower class evaluation	Indices values
PPC	99 %
PUC	95,2 %
OCI	**94,1 %**
Classification evaluation	
OA	95,6 %
AA	94,1 %
Kappa	90,8 %
OOCI	**88,7 %**

CONCLUSION AND PROSPECTIVE

- ➢ Evaluation of supervised classification is very useful for numerical validation
- ➢ Existing indices do not evaluate the same characteristic and cannot evaluate multiple characteristic.
- ➢ We introduce new indices to obtain a more complete classification evaluation and a global point of view on class or classification characteristic.
- ➢ New indices are simple but robust and accurate (kappa based indices introduced by Pontius keep disadvantage of kappa index).
- ➢ Many applications and algorithms could easily integrate these new indices like classification selection or classification fusion.
- ➢ Future works will be spent on unsupervised classification evaluation and interpretation with spectral and temporal information. The goal is to break away from spatial dependence of statistical samples and confusion matrices.

Class index	Recommendation	Classification index	Recommendation
PUC	To detect class deficit.	**OA**	To evaluate accuracy
PPC	To detect class excess.	**AA**	To evaluate accuracy when class sizes are different.
OCI	To detect balanced class without deficit and excess.	**Kappa**	Not recommend but to evaluate agreement by taking into account chance part.
		OOCI	To evaluate classification accuracy and class representation.

The nice laboratory

The nice author