Multisensor and multitemporal image fusion methods to improve remote sensing image classification

D. Ducrot, A. Masse, C. Marais-Sicre, J. F. Dejoux, F. Baup
Centre d'Etudes Spatiales de la Biosphère
18 av E.Belin - bpi 2801 - 31401 TOULOUSE cedex 9 France
danielle.ducrot@cesbio.cnes.fr, antoine.masse@cesbio.cnes.fr, claire.marais-Sicre@cesbio.cnes.fr

ABSTRACT: *As acquisition technology progresses, remote sensing data contains an ever increasing amount of information. To correctly understand the fundamental advantages of data fusion, let us list the different kinds of remote sensing data: optical and radar images, low, high and very high-resolution, multitemporal hyperspectral images, derived images, and physical or ancillary data (databases, D.E.M, G.I.S.). Data fusion is the joint use of heterogeneous images for decision-making. It is essential to group the available information from various sensor images, temporal data, expert information, etc….*

In this paper, fusion is used for improving classification at two levels: pixel by pixel or after independent "sub classification". This fusion is particularly interesting in the case of imperfect data to obtain more reliable information. Thus, it takes advantage of the best of each data type and overcomes individual limitations of each type. For example, optical and radar sensors are not sensitive to the same kind of information. We used the complementarities in these data to extract more complete information and to make a better distinction between various classes. We present some applications such as fusion between radar and optical images to detect agricultural practices on bare soil (TerraSar-X and Formosat2) which improves overall accuracy for supervised classification by 12%. Radar images lend more precision to bare soil classes and optical imagery more accuracy to active vegetation classes. We also show temporal complementarities with different dates from the same source (Formosat2).

1 FUSION DEFINITION

Fusion is the combination of heterogeneous images to support decision-making [Bloch et al, 2002, 2005]. From a mathematical point of view, fusion is applied when it is not possible to find a common metric for all the images, whatever their source. So fusion is used to blend information available from different sources: sensors, spectral or temporal data, but also external data, such as river or road maps, altitude, slope orientation (GIS), contour and texture. So we can exploit the complementarity of these data to extract more information, and provide a clearer distinction between different classes. We can decide to assign a pixel to the most reliable and accurate class possible.

In this paper, fusion is used in the sense of ***classification*** combination, which requires the right classification method as a first step. The applications will use multisensory and multitemporal classification from remote sensing data in a context of land use.

The various sources generally provide imperfect information i.e. uncertain, imprecise, incomplete, contradictory or inaccurate [Smets, 1999]. Fusion will use benefits of each data type to overcome the individual limitations of each one. For example, optical and radar sensors do not have the same sensitivity with the same information. Contents of parcels are more visible, sharper, and more homogenous with the optical images than in radar imaging, which is very heterogeneous due to speckle. However, the insensitivity to weather conditions of acquisition radar remains a major advantage. So we can exploit the complementarities of these data to extract more information to better distinguish classes.

2 DIFFERENT APPROACHES

In practice, fusion is not a simple problem; it requires the use of specific methods. The main approaches to data fusion are Bayesian probabilistic methods (the oldest), fuzzy sets and possibility theory introduced by Zadeh [Zadeh 78, Dubois and Prades 93] and belief (or evidence) theory of Dempster-Schafer (first Dempster (1960) and renormalized by Shafer (1976)), widely used in image processing.

Methods implemented

A great number of tests showed that the combination of a large amount of temporal images does not necessarily give the best results [Ducrot *et al*, 1998]. We know that too much information can impair meaning and spoil the results. With specific temporal image combinations, classifications provide superior results for some classes and with other combinations, the best results for other classes. To obtain the best classification possible, the most effective solution is to merge classifications performed with various image combinations (temporal and/or different sources) chosen for their discrimination.

The fusion method is introduced in the classification and developed according to supervised or

unsupervised use. Bayesian methods are applied with different laws appropriate to each image batch and decisions made during each pixel assignment, or fusion of several classifications made independently according to their contribution to discrimination.

3 FUSION AND SUPERVISED CLASSIFICATION

For supervised classification, the fusion method which gives the best results is obtained with the use of confusion matrices for each classified image [Chust *et al*, 2004]. These matrices, based on the proportion of Pixels Correctly Classified (PCC), can objectively quantify classification quality with mathematical and statistical criteria. Two main coefficients are calculated from these matrices: MPCC (Mean of PCC) called also Overall Accuracy and Kappa (Congalton (91)); they determine the global quality of the classification and the quality for each class.

$$MPCC = \frac{1}{n}\sum_{i=1}^{n} PCC(i) = \frac{1}{n}\sum_{i=1}^{n}\frac{m_{ii}}{m_i}$$

with n = confusion matrix size, m_{ij} value at the i^{th} row and i^{th} column, mi. sum of row i.

The confusion matrix is computed by crossing the ground truths (lines) and classes found by the classification method (columns).

For each class, this allows the following measurements to be extracted: (1) deficit error (commission: percentage of classified sample pixels in other classes) that involves a loss (100 - PCC) for this class and (2) excess error (omission: percentage of sample pixels that absorbs other classes, 100 – number of incorrectly classified pixels divided by the sum of the column) that indicates an "over-represented" class.

3.1 Automatic and systematic measurements of image combinations

The issue to solve is the choice of the best combinations to obtain the best classification results. It is impossible to analyze each matrix manually for a lengthy time series. These combinations are selected by expert knowledge, with great uncertainty. For example, for seven dates, we get $\sum_{i=1,7} C^i_7 = 127$ possibilities. To find out the best dates for classification, it is necessary to measure the discriminatory power of these dates by calculating MPCC.

We developed an algorithm which provides, automatically and exhaustively, all confusion matrices of all possible date combinations (or spectral combinations). Thus, the choice is objective. The algorithm gives, in ascending order, the MPCC (with check and\or learning samples) of all the combinations, and confusion matrix diagonal elements in a summary table. We can also extract the best combinations by class, which will guide us for the date combination choice to better discriminate classes.

3.2 Automatic fusion

An automatic fusion algorithm was developed. In the classification process, several classifications (called "sub classifications") are made on ***n*** image batches; every batch contains ***m*** spectral and/or temporal images chosen according to the distinction of certain classes. An image batch is taken as a reference (the first "sub classification"), generally giving the best MPCC and kappa. Certain classes are better discriminated with other combinations, which show the relevance of merging these classified images. This principle can be applied within the classification or later in post-processing.

For two "sub classifications" with combinations of different dates, two classes can be in conflict, caused by their attribution to the same pixel. The ambiguity is removed by comparison of confusion matrices of both images which reveal the conflicts between classes.

SPOT (1)	Corn	Sunfl
corn	50.23	30.5
Sunflower	5.6	71.02

ERS (2)	Corn	Sunfl
corn	90.5	10.1
Sunfl	2.8	50

ERS (3)	Corn	Sunfl
corn	77.53	0
Sunflower	42	25

ERS (4)	Corn	Sunfl
corn	60.83	32.6
Sunfl	1.2	50.6

Table 1: confusion matrices extract (two sources)

A class can have a better score in one classification, but include other classes; it is an "over-represented" class. Thus, this class seems correctly classified (its PCC is better), but the classification loses the incorporated classes. For example, in table 1 (classification with SPOT and ERS (radar)), in the 1st classification (1) Corn is confused with Sunflower; the confusion is removed thanks to the fusion with the 2nd (2) without ambiguity. If a pixel is classified as Corn (1st classification) and Sunflower (2nd classification), the fusion with the described rules will affect it to Sunflower. Case (3), Corn seems well classified (90.5%), but it is over-represented and often replaces Sunflower (42.1%°), fusion between 1st and 3rd classifications would be incorrect.

The algorithm avoids this type of confusion.

The 4th seems worst. Therefore, a fusion with this classification decreases slightly the over representation of Corn to the detriment of Sunflower.

The M_{Sk} and M_{Sl} confusion matrices corresponding to classification of two sources S_k (reference source with the best global results (kappa and MPCC)) and S_l (different combination). Sources can be different sensors, dates (multi -spectral, -temporal, -source combinations).

S_k	i	...	j
i	$m_k$ii	...	$m_k$ij
...	...	...	...
j	$m_k$ji		$m_k$jj

S_l	i	…	j
i	$m_l$ii	...	$m_l$ij
...	...	...	...
j	$m_l$ji	...	$m_l$jj

Table 2: S_k and S_l confusion matrices

The diagonal elements $m_k$ii of M_{Sk} and $m_l$ii of M_{Sl}, represent the pixel percentage correctly assigned to class i (degree of confidence for assignment to this class) (table 2). The non-diagonal elements $m_k$ij and $m_l$ij are the percentage of pixels in class ***i*** assigned to class ***j***. Example of certain rules:

If $i \cap j = \varnothing \Leftrightarrow m_k\, i\, j = 0$ and $m_k\, j\, i = 0$, necessary and sufficient condition, that means: no confusion between classes ***i*** and ***j***, the confidence degree is maximum on ***i*** for the source S_k

Else $x \in i \cup j$, pixel $x \in S_k$

If $m_k\, i\, j \neq 0$ **then** there is confusion between ***i*** and ***j***, some pixels of class ***i*** are assigned to class ***j***.

If $m_k\, j\, i \neq 0$ **then** there is confusion between ***i*** and ***j***, but the pixels of class ***j*** are assigned to class ***i***.

If there is confusion between ***i*** and ***j*** in the source S_k ($i \cap j \neq \varnothing$ with a high score for $m_k$ij), the source S_l can raise this ambiguity if $i \cap j = \varnothing$ in S_l or if the score is low for $m_l$ij.

Let a pixel x, classed ***i*** with S_k and ***j*** with S_l. If the class ***i*** is confused with class ***j*** and has a poor score in M_{SK}, ***j*** having a better score (low or no confusion) in M_{SL}, the fusion facilitates class ***j***. x is then assigned to ***j***. A test is performed to minimize a possible loss in another class. To increase the overall accuracy, we introduce the notion of gain to allow the algorithm to decrease the (j,i) confusion even if the (i,j) confusion slightly increases.

3.3 Supervised fusion

Supervised fusion improves classification results by combining another classified image, best judged by the expert, after examining this classification and its confusion matrix (PCCs). Criteria are imposed by the user. Thus, the fusion concerns only certain classes chosen according to the user's knowledge

4 PROBABILISTIC MODELS: FUSION BY BAYESIAN THEORY

The widely spread probabilistic methods characterize uncertainty and, with difficulty, inaccuracy. The Bayesian method makes it possible to perform data combination in an associative way: conjunction, disjunction and averages. Fusion is carried out thanks to the Bayes rule, which requires an estimation of probability. The various terms are computed by training (supervised or random).

$$P(x \in C_i | I_1, \ldots, I_l) = \frac{P(I_1, \ldots, I_l | x \in C_i) P(x \in C_i)}{P(I_1, \ldots, I_l)}$$

To perform data combination according to the theory of Bayes, the probability of the observation model is calculated by supposing data sources to be independent. The different sources infer different and more relevant laws according to the type of sensors. We chose the Gaussian law for optical images and the Gauss-Wishart law [Fukunaga, 1972] for radar images. It is a contextual law taking into account the pixel neighborhood with segmentation (object classification) [Ducrot et al 1998]. These laws being different, the results are no longer comparable, what involves fusion rules.

Method implemented

The process implemented consists of local decision-making first being applied to every image batch separately. We then merge the local decisions in a global decision. This type of fusion is adapted to unsupervised classifications, but requires an adjustment for correspondence to the different unsupervised sub-classification classes

A weight giving more or less credibility to each batch can be associated to these batches.

Given the inaccuracy of these data, why do we keep a single class? The difference between ***n*** first better measurements is unimportant. The class showing the best score is generally very close to the next classes So, there is no valid reason to choose the 1st rather than the 2nd and so on. Several classes are thus acceptable for a pixel if their measurement shows a sufficient degree of confidence.

For every image batch, the ***n*** better classes are chosen, their associated measurement is kept, and the global decision is carried out for these ***n*** classes. For a given pixel, the measurement of the ***n*** classes is ordered in descending order, we would then have, for instance, [0.3865, 0.3864, 0.3862, 0.3861, 0.3858, 0.29,]. In this case, the first five classes of this image batch have very close measurements, so we can select them.

The intersection of acceptable class sets for every image batch is performed to determine the most representative class. The number of classes selected is sufficient: in our applications, an average of 6 classes are kept per image batch and per pixel.

The algorithm is more accurate and faster because it does not use all the classes for the global decision.

5 – APPLICATIONS

5-1 Fusion on a temporal series: 2008 Formosat-2 images (8m*8m) from CESBIO (Center of Spatial Study of the Biosphere) - Toulouse Southwest project

This application concerns an agricultural zone of the Toulouse Southwest (France). Nine Formosat-2 images (Taiwanese satellite, resolution 8m*8m with 4 spectral ranges: Blue (0.45 – 0.52 µm), Green (0.52 – 0.60 µm, Red (0.63 – 0.69 µm), Near Infrared (0.76 – 0.90 µm) are used.

The February image contains some clouds; the classes under clouds are mainly classified as diffuse or industrial Built or Mineral surface classes. This image was used because no spring data was available which provides optimum discrimination of the Rapeseed class.

After examining the best dates for class discrimination, "sub-classifications" with different image batches were successively completed and merged. The confusion matrix method was used.

1. The first image batch with 9 dates: 2008/02/11, 2008/06/19, 2008/10/07, 2008/31/07, 2008/21/08, 2008/25/09, 2008/06/10, 2008/10/10 and 2008/26/10.

2. Second image batch (without February): this sub classification is merged with the previous one. When pixels are classified as Built or Mineral Surface in the first sub classification, the class of the second one replaces them (figure 1). We were able to eliminate clouds blending with Mineral surfaces. This fusion provides better discrimination of the Rapeseed class as well as Built classes because, in February, the presence of many bare soil crops causes confusions between these classes and some crop classes.

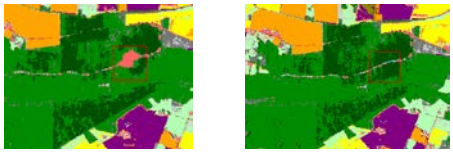

Figure 1: right, 9 date classification (grey frame: cloud corresponds to Built), left this classification merged without February classification → cloud is removed

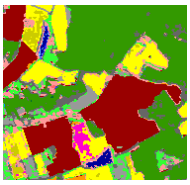

Figure 2: right, previous fusion (figure 1), left,

fusion with NDVI classification; small lakes appear (dark blue), previously classified as mineral surfaces

3. The third image batch is formed with only PIR channels or NDVIs of every date. This combination improves Water, Lake or Gravel pit classes confused with roads, as well as Wheat class which present confusions with mineral class (figure 2, table 3).

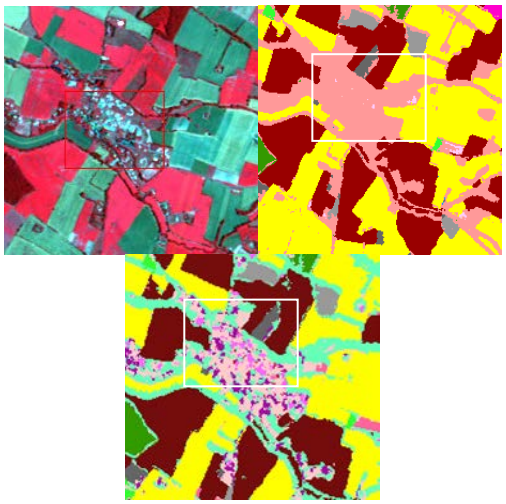

Figure 3: (a) Colored composition 08/07/10 Formosat (b) final classification (c) fusion with unsupervised classification to distinguish hedges and different stage of Built

% corrected classes	fusion final	9 dates	8 dates without february	NDVI 9dates
decideous trees	97.88	96.9	96.4	94.7
coniferous	97.05	96.82	96.15	96.79
eucalyptus.	74.53	73.84	70.41	59
wheat	97.48	96.28	92.44	71.05
rapeseed	99.33	99.11	98.05	86.63
barley	99.06	97.41	95.87	76.09
corn	99.6	99.13	98.75	90.86
sunflower	99.12	96.91	93.08	72.61
sorghum	100	98.17	97.48	96.68
soybeans	97.36	97.42	97.09	94.26
set-aside land	97.75	93.65	93.11	74.49
wild land	90.27	91.64	91.19	71.96
meadow	95.16	92.61	88.73	76.64
river	99.73	97.41	97.28	91.43
lake	99.95	99.68	99.68	94.1
built densely	96.06	96.06	94.98	93.19
built industrial	98.2	98.2	97.48	88.49
gravel pit	99.91	99.83	99.78	93.19
poplars.	98.97	97.87	98.27	96.84
diffuse built	95.47	97.46	95.29	87.5
bicultural	93.32	95.47	92.15	76.78
Sunflower particular.	100	96.91	96.74	95.37
mineral surface	99.82	99.82	99.27	93.96
Sunflower late	100	100	99.95	98.81
MPCC	96.92	96.19	94.98	86.31
Kappa	97.23	96.13	94.31	82.69

Table 3: successive improvements (from right to left) according to the steps described previously

4. A 4th step gives better accuracy for Built/mineral surface classes and hedges with unsupervised classification fusion (figure 3). Improvements are visible through the observation of the image, but they are difficult to quantify because we have no samples for their measurement.

These classifications remove a lot of confusion about certain crops (Sunflower, Corn …). For very detailed

classification, the results for land cover are highly acceptable (table 3).

5.2 - Optical / radar fusion (Spot/Radarsat) Grand Morin basin in Seine and Marne (France)

As part of a research program for the promotion of the Radarsat data (ADRO: Development and Research Opportunity application), the objective was to define an environment indicator to assess the vulnerability of the natural environment to pollution caused by human activity through land use mapping. This was established by several classifications from the following data (1) multitemporal radar data (4 unfiltered and filtered amplitude Radarsat images: 1997 June 28th, 1997 July 22nd, 1997 September 8th, 1997 November 19th); (2) optical data: SPOT image multispectral (XS), 1997 August 13th.

The area investigated is located in the Grand Morin basin in the Seine and Marne department (France) which supplies drinking water to part of the Paris region. This site is mainly fragmented agricultural terrain, with wood and forests.

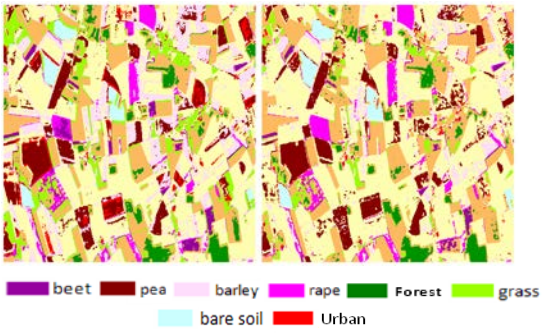

Figure 4: left, SPOT classification, right, SPOT + Radarsat fusion image

Methods and data	Gauss Spot	GW Radarsat non filtered	GW Spot + Radarsat unfiltered	gauss Spot + Pearson Radarsat unfiltered
forest	98.6	94.87	98.05	98.06
corn	85.46	80.52	99.18	97.54
wheat 1	49.55	59.89	88.87	87.79
wheat 2	86.21	62.3	88.89	90.03
barley	59.62	45.56	59.91	60.35
rape	46.02	40.15	46.02	46.02
pea 1	54.65	60.21	55.43	55.81
pea 2	79.78	70.57	80.66	82.97
beet	97.92	86.32	99.14	94.1
herb	97.09	86.25	97.43	97.44
urban	97.5	93.85	100	91
bare soil	99.56	87.49	100	96.84
mean	**79.33**	**72.33**	**84.46**	**83.43**

Table 4: percentage of corrected classified for the Spot & Radarsat classification, fusion Radarsat-Spot with two methods (GW = Gauss-Wishart is better)

The probability fusion process with the Radarsat unfiltered images, classified with the contextual Gauss Wishart law (GW), improves the results of the SPOT image (Gauss law) for the majority of the classes: approximately 5% (table 5). The display of this classification (figure 4) shows the results improved for Forest, Corn, Beet and Pea classes. Blending between the parcels of these classes is removed. The fusion eliminates pixels from Urban-mineral class (generally roads) spread across the image. Contour pixel confusion between fields and grass is reduced, thereby improving the structure of fields. This improvement is important, but does not count towards the percentage of correctly classified which is computed from samples located in the centre of parcels.

5.3 - Optical / radar fusion for the detection of agricultural practices on bare soil

The goal is to estimate the possibility of agricultural practice detection on bare soil, to take into account the soil work in modeling carbon inventories in the southwest region of Toulouse. This study is conducted with radar and optical satellite data and ground truths focusing on soil humidity and roughness. The type of the agricultural soil working plays a major role in carbon inventory on the landscape scale. Indeed, soils represent CO_2 sinks and sources. It has been shown that a ploughed ground stores 3 to 5 tons of CO_2 by hectare per year, less than slightly worked ground or direct sowing. The detection and the monitoring of the various farming practices are used to take into account their CO_2 emission and storage in complex climatic models.

The classes to be considered are as follows: ***(1) Inter-crop***: begins after harvest; no recent or visible soil tillage; crop residues may be visible on soil surface; some greenery may grow, like weeds; ***(2) Stubble disking***: superficial soil tillage to mix crop residues and soil (5 to 15 cm deep) and to destroy weeds; soil surface is irregular with some small clods and slight roughness, biomass may remain on the surface; ***(3) Deep plowing*** (20 to 45 cm deep): 95% bare soil, deeply worked with pronounced roughness (plowing), visible clods; ***(4) Harrowing***: 95%, bare soil; secondary or superficial tillage, medium-sized clods, improper for seeding ***(5) Sowing preparation***: 95%, bare soil, regular surface, soil ready to be sowed, fine work, small clods (***6) Emergence***: germination; plants are visible, cotyledons or first leaves development stages (5 cm height) [Inglada, 2010]. We shall add some classes corresponding to fields in active vegetation in this period: Corn, Rapeseed and Sunflower.

Fusion was performed from **08/10/10** FORMOSAT-2 optical image and **08/10/09** TerraSAR-X radar image**,** chosen because of a 1-day time gap.

TerraSAR-X (DLR aerospace centre and EADS Astrium) has a band X (9.65 GHz) SAR; the data is

generated by a multilook process (3x3m² pixel size and a spatial resolution of 6.5*6.5 m²).

FORMOSAT-2, Taiwanese high-resolution/revisit daily (8m) satellite, can acquire any image within its coverage area every day, Blue 0.45-0.52µm, Green 0.52-0.60µm, Red 0.63-0.69µm, Near Infra Red 0,76-0.90µm, coverage 24 km X 24 km.

A mask is applied to the TerraSAR-X images to eliminate urban and forest zones, lakes, meadows and fallows. Classes are thus attributed only to agricultural zones and will differentiate bare soil types more easily

For one date, the results achieved are: with Terrassar MPCC = 41.92 %, with Formosat-2 MPCC = 76.61 %. After the fusion (using confusion matrices) we obtain MPCC = 83.16 %. The RADAR contribution is: Stubble disking 67.43%, deep plowing 91.72% Harrowing 53.52%; the optical contribution is: Sowing preparation 81.28% Emergence 93.69% Inter-crop 76.35%, Corn 92.54% Rape seed 93.3% Sunflower 96.69%

The contribution of RADAR on the optical supervised classification improves bare soil distribution classes of 12 %. The advantage of fusion in this case is to provide more precision for bare soil classes by the intervention of the RADAR and to improve active vegetation classes thanks to optical data.

4 CONCLUSION

In complex contexts involving different sensors or long temporal series, the objective was to obtain superior classification scores. The most effective solution was to perform classification fusions with diverse date combinations (image batches), selected according to discrimination of classes or sensors.

We presented two main methods: one well suited to supervised cases because it needed confusion matrices. A more general method could be adapted to unsupervised classifications, but, after each image batches classification ("sub-classification"), it involves finding correspondence between classes of these sub unsupervised classifications, which produces different labels for a same class. This type of fusion enables the law to be applied to the different image batches to be selected.

Three different applications were presented: optical temporal series, optical/radar fusion for land cover, and detection of different agricultural bare soil types. In these three cases we must obtain very precise and detailed classification. Using the fusion process developed, we achieved very satisfactory results. We used complementarities of optical and radar data to extract richer information and to obtain an enhanced distinction between the various classes.

Future objectives are to improve the unsupervised method with a Markovian model combined with the probabilistic mode to better integrate temporal changes and the problem of correspondence between classes of different sub-classifications.

REFERENCES

Bloch I. & Maître H., 2002, Fusion d'informations en traitement d'image: spécificités modélisation et combinaison par des méthodes numériques, *Techniques de l'Ingénieur* 2002 ,TE 5 230, 1-26.

Bloch I., 2005, Fuzzy Spatial Relationships for Image Processing and Interpretation: A Review, *Image and Vision Computing 2005* vol. 23, 2, 89-110.

Chust G., Ducrot D. & Pretus J. Ll, 2004, Land cover discrimination potential of radar multitemporal series and optical multispectral images in a Mediterranean cultural landscape, *International Journal of Remote Sensing*, 5, 17, 3513-3528.

Congalton R.G., 1991, A review of assessing the accuracy of classification of remotely sensed data, *Remote Sensing of Environment* 37, 35-46.

Dubois D. & Prades H., 1993, Combination of information in the framework of possibility theory data fusion in robotics and machine intelligence, Academic Press.

Ducrot D., Sery F., Sassier H., Goze S., Planes J.G., 1998, Classification and fusion of optical and radar satellite data for land use extraction, *Fusion of Earth Data98*, Sophia Antipolis, France, 15, 1,5-20.

Fukunaga K., 1972, Introduction to statistical pattern recognition, *Electrical Sciences Series*, Academic press, London (UK).

Inglada J., Beguet B., DejouxJ.-F, Marais Sicre C., Ducrot, D., Huc M. , Hagolle O. , Baup F. , Dedieu G. , 2010, Use of dense time series of high resolution images for change detection and land use classification. *In Recent Advances in Quantitative Remote Sensing*

Shafer G., 1976, A mathematical theory of evidence, *Princeton University Press*. Princeton, NJ.

Smets P., 1999, Imperfect Information: Imprecision, Uncertainity, *report IRIDIA July 27*

Zadeh L.A., 1978, Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems* 1, 3-28