Remote Sensing of Environment

www.elsevier.com/locate/sna

Evaluation of Fuzzy Partitions

Pascal Matsakis, Serge Andréfouët† and Patrick Capolsini†*

The aim of this study is the development of tools dedicated to fuzzy partition evaluation in the field of satellite image classification. While a traditional crisp partition only provides qualitative information, a fuzzy partition represents a large amount of quantitative information. However, such a partition is often evaluated after "defuzzification" (i.e., it is reduced to a crisp partition). The analysis of a traditional confusion matrix, describing the similarities between the computed crisp partition and a control partition, can then be performed. This approach is rather drastic and is far from satisfactory because the quantitative information is lost after the defuzzification. Some methods do not require preliminary defuzzification, but they are not adequate to evaluate nonprobabilistic fuzzy partitions (i.e., fuzzy partitions such that the sum of the membership degrees is not necessarily equal to 1). To solve these issues, we consider the evaluation of any *fuzzy partition μ as the evaluation of a still fuzzy new partition: the plausibilistic closure of μ. This approach comes from the theory of evidence. It allows definition of a set of original tools (plausibility matrices, credibility matrices, and overlap degrees) dedicated to fuzzy partition evaluation. A concrete application illustrates our theoretical work and a tutorial is provided in appendix.*

INTRODUCTION

The evaluation of the accuracy of a partition arises naturally in the field of image classification. It is a mandatory stage, although commonly overlooked (Hammond and Verbyla, 1996). Evaluating a partition generally consists of comparing it with a control partition and measuring the degree of association of two variables that may be binary, qualitative, ordinal, or quantitative. Traditional products of image classification are crisp partitions (i.e., hard, ordinary, classical partitions), thus only the case of qualitative variables has really held interest. The literature on the analysis of confusion matrices is quite abundant. It generally describes methods that summarize the matrices by a single index, hence allowing different classifications to be comparable (Congalton, 1991; Gong and Howarth, 1990; Ma and Redmond, 1995; Zhuang et al., 1995). This abundance of literature is justified by the numerous issues related to accuracy assessment, issues for which clarifications (Stehman, 1997; Stehman and Czaplewski, 1998; Congalton and Green, 1999; Stehman, 2000) are still welcome and necessary after three decades of satellite image classification.

To complicate even more this domain, the last decade has seen the emergence of new approaches and products (i.e., fuzzy partitions) intended to overcome the intrinsic limitations of crisp partitions (e.g., Bezdek and Pal, 1992; Key et al., 1989; Wang, 1990; Simpson and Keller, 1995; Bastin, 1997; Raffy, 1997; Smith et al., 2000). Indeed, traditional crisp partitions may be inappropriate for representing heterogeneous, complex areas, and fuzzy partitions have proven useful to postclassification processing, the detection of high-risk confusion zones, and the correction of flagrant misclassification (Harris, 1985; Maselli et al., 1994; Andréfouët et al., 2000). They also open up new prospects for describing the spatial structure of ecological systems without reducing them to an assemblage of pixels assigned to arbitrary pure classes (Palubinskas et al., 1995; Foody, 1996; Burrough et al., 1997; Andréfouët and Roux, 1998). Such achievements are possible because fuzzy partitions provide ordinal quantitative membership degrees, while crisp partitions only provide qualitative assignments. Despite these characteristics, fuzzy partitions have often been condemned to undergo immediate "defuzzification" before their assessment (Foody and Trodd, 1993). Re-

* Computer Engineering and Computer Science Department, University of Missouri-Columbia, 201 Engineering Building West, Columbia, MO.

† Laboratoire de Géosciences Marines et Télédétection, Université de la Polynésie Française, BP 6570 Faaa-Aéroport, Tahiti, French Polynesia

The present address of S. Andréfouët is Department of Marine Science, Remote Sensing and Biological Oceanography, University of South Florida.

Address correspondence to S. Andréfouët, University of Florida, Dept. of Marine Science, 140 7th Avenue S., St. Petersburg, FL 33701, USA. E-mail: serge@carbon.marine.usf.edu

Received 10 June 1998; revised 13 April 2000.

REMOTE SENS. ENVIRON. 74:516–533 (2000)

655 Avenue of the Americas, New York, NY 10010

0034-4257/00/$–see front matter
PII S0034-4257(00)00143-7

ducing the evaluation of a fuzzy partition to the evaluation of a crisp one is unsatisfactory because it implies the loss of all quantitative information.

The approaches suggested for fuzzy partition evaluations are reviewed by Foody (1996) and summarized hereafter. Indices derived from fuzzy set theory were proposed by Gopal and Woodcock (1994), but only to compare a crisp partition with a fuzzy control partition. Entropy, or relative entropy, was proven useful to fuzzy partition evaluation (Finn, 1993; Maselli et al., 1994; Foody, 1995; Zhang and Kirby, 1997), but the control partition had to be crisp (i.e., the target had to be quite homogeneous). Ultimately, for heterogeneous areas, the evaluation of a fuzzy partition could be based on a comparison with a fuzzy control partition (Gopal and Woodcock, 1994; Deer, 1996). In this case, measures of closeness between fuzzy partitions have been proposed. They generally come from information theory, which provides measures of closeness between two probability distributions (Higashi and Klir, 1983; Foody, 1995; Zhu, 1997; Dubois and Prade, 1999). Indeed, in remote sensing, the fuzzy partitions were usually probability distributions. For each pixel, the sum of the membership degrees equals 1, because membership degrees closely match land cover proportions inside a pixel (Foody and Cox, 1994). In some cases, however, fuzzy partitions are not probabilistic. The sum of membership degrees derived from neural networks may not total 1 (Foody, 1996). Uncertainty measures different than probabilities can be selected to take advantage of other mathematical tools and theories, such as possibility theory (Desachy et al., 1996; Andréfouët et al., 2000; Foody, 2000). Then, probabilistic measures of closeness between fuzzy partitions, such as cross-entropy, are not necessarily adequate, and measures with a broader range of validity must be proposed.

In this paper, we aim to consider the evaluation of any fuzzy partition μ, probabilistic or not. For this, we define a new fuzzy partition, the plausibilistic closure of μ. This new partition can be characterized independently of the model of uncertainty (i.e., probability, possibility, evidence) used for μ. Our approach comes from the theory of evidence (Dempster, 1967; Shafer, 1976). This theory has been used in image classification for the construction of crisp and fuzzy partitions (Lee et al., 1987; Peddle, 1995; Desachy et al., 1996), but not for their evaluation, although "measures of closeness based on information uncertainty may be the most appropriate to use in classification evaluation" (Foody, 1996). First, the required theoretical background and the fundamental idea of plausibilistic closure are introduced. Then, three original tools for the evaluation of plausibilistic closure are proposed: plausibility matrices, credibility matrices, and overlap degrees. Eventually, an example illustrates how partitions of different types are evaluated and characterized. The considered fuzzy partitions come from probabilistic and possibilistic supervised fuzzy classifications of a SPOT HRV XS image of the rim of Tikehau atoll in French Polynesia. Mathematical proofs are provided in the appendix, and a tutorial guides the reader through the computations required by the analysis.

BASIC NOTIONS AND PRELIMINARY DEFINITIONS

Within this paper, N and C denote two integers such that $N \geqslant C \geqslant 2$, and E denotes a set of N elements numbered from 1 to N: $E=\{e_j\}_{j\in 1\ldots N}$. Typically, E is a satellite image and each e_j is a pixel. To help the reader with the theoretical developments below, a tutorial example in Appendix A illustrates every definition and computation scheme.

It is also useful to redefine some mathematical symbols. For instance, the expression "$\forall i \in 0 \ldots N$, $\exists j \in N \ldots 2N \backslash N=(i+j)/2$" means that for any element ($\forall$) i of $0 \ldots N$ (i.e., for any integer i greater than or equal to 0 and less than or equal to N), there exists at least one element ($\exists$) j of $N \ldots 2N$ such that ($\backslash$) N equals half of $i+j$. Finally, for any set A, $|A|$ denotes the number of elements of A.

Crisp Partitions

Definition 1 (Fig. 1a and Appendix A1)
A crisp C-partition of E, or crisp partition into C classes of E is a C-tuple $(E_i)_{i\in 1\ldots C}$ of subsets of E such that [see Eqs. (1), (2), and (3)]:

$$\cup_{i\in 1\ldots C} E_i = E \tag{1}$$

$$\forall i \in 1 \ldots C,\ E_i \neq \varnothing \tag{2}$$

$$\forall i \in 1 \ldots C,\ \forall j \in 1 \ldots C,\ i \neq j \Rightarrow E_i \cap E_j = \varnothing \tag{3}$$

where $\varnothing$ is the empty set. E_i is the class i.

This definition coincides with the standard mathematical definition of a partition. Equation (2) states that each class contains at least one element of E. Likewise, Eqs. (1) and (3) state that each element of E belongs (entirely) to a class and does not belong (at all) to the others.

Fuzzy Partitions

A pixel may cover more than one discrete land cover class. Moreover, the land cover may be continuous. A crisp partition cannot represent these facts. A fuzzy partition is a solution.

Definition 2 (Fig. 1c and Appendix A1)
A fuzzy C-partition of E, or fuzzy partition into C classes of E, is a C-tuple $(\mu_i)_{i\in 1\ldots C}$ of functions from E into [0,1] such that [see Eqs. (4) and (5)]:

$$\forall j \in 1 \ldots N,\ \exists i \in 1 \ldots C \backslash \mu_{ij} > 0 \tag{4}$$

$$\forall i \in 1 \ldots C,\ \exists j \in 1 \ldots N \backslash \mu_{ij} > 0 \tag{5}$$

where μ_{ij} denotes the image of element e_j of E by the

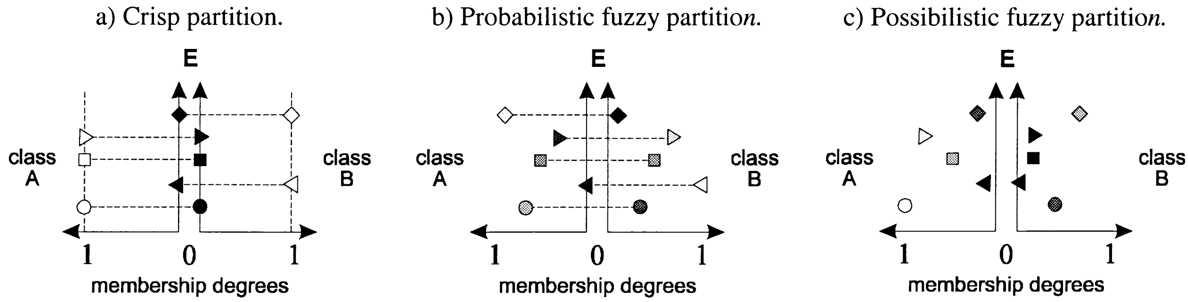

Figure 1. Representation of crisp, fuzzy probabilistic and fuzzy possibilistic two-partitions of a set E containing five elements $\{\Diamond, \triangleright, \Box, \triangleleft, \bigcirc\}$. The left part of each diagram gives the membership degrees in class A, and the right part the membership degrees in class B. The lighter an element is, the higher its membership degree into the considered class. Black stands for 0, white stands for 1. Vertical discontinuous lines symbolize the constraint: $\max_i \mu_{ij}=1$. Horizontal discontinuous lines symbolize the constraint: $\Sigma_i\ \mu_{ij}=1$. A crisp partition is constrained in both directions. A possibilistic partition doesn't have any constraint.

function μ_i. The fuzzy subset of E defined by μ_i (Zadeh, 1965) is the class i.

μ_i is the membership of E attached to class i, and μ_{ij} is the membership degree of e_j in class i. If all the functions μ_i take their values in {0,1}, they define crisp subsets E_i of E (i.e., nonfuzzy, hard, ordinary, classical subsets of E), and Eqs. (4) and (5) respectively may be rewritten as Eqs. (1) and (2). Hence, a crisp partition is a particular fuzzy partition. Krishnapuram and Keller (1993) interpret μ_{ij} as the possibility (Zadeh, 1978) that element e_j of E belongs to class i (Krishnapuram and Keller, 1993; Barni et al., 1996; Krishnapuram and Keller, 1996). According to this interpretation, a fuzzy partition is actually a possibilistic fuzzy partition. On adding the condition "$\forall j \in 1 \ldots N,\ \Sigma_{i \in 1 \ldots C}\ \mu_{ij}=1$" to Eqs. (4) and (5), we obtain a probabilistic fuzzy partition (Definition 3), like that of Ruspini (1969) and Bezdek (1981). Hence, a probabilistic fuzzy partition is a particular possibilistic fuzzy partition.

Definition 3 (Fig. 1b and Appendix A1)
The fuzzy C-partition $\mu=(\mu_i)_{i\in 1 \ldots C}$ of E is known as probabilistic if and only if [see Eq. (6)]:

$$\forall j \in 1 \ldots N,\ \Sigma_{i\in 1 \ldots C}\mu_{ij}=1 \qquad (6)$$

For any element j of $1 \ldots N$, the function from $1 \ldots C$ into [0,1] that associates μ_{ij} with any i is then a distribution of probability. μ_{ij} may be interpreted as the probability that element e_j of E belongs to class i.

"Defuzzification" of a Fuzzy Partition

It is usual practice to associate a crisp partition to any fuzzy partition.

Definition 4 (Appendix A1)
Let $\mu=(\mu_i)_{i\in 1 \ldots C}$ be a fuzzy C-partition of E. We say that a crisp C-partition $\nu=(\nu_i)_{i\in 1 \ldots C}$ of E is the result of a "defuzzification" of μ if and only if [see Eq. (7)]:

$$\forall i \in 1 \ldots C,\ \forall j \in 1 \ldots N,\ \nu_{ij}=1 \Rightarrow \mu_{ij}=\max_{k\in 1 \ldots C}\mu_{kj} \qquad (7)$$

Element e_j of E is assigned to the class to which it belongs "the most" (i.e., the class that is the most probable, or the most possible).

Plausibility and Credibility Measures

This paper proposes new tools for the evaluation of fuzzy partitions. Their development is based on the theory of evidence (Dempster, 1967; Shafer, 1976). Let Ω be the set of choices possible to take a decision and let $P(\Omega)$ be the set of all possible subsets of Ω. An event is an element of $P(\Omega)$ (i.e., a subset of Ω). In particular, Ω is the certain event and the empty set $\varnothing$ is the impossible event. A basic probability function is a function m from $P(\Omega)$ into [0,1] such that (see Eq. (8)]:

$$\Sigma_{A\in P(\Omega)}m(A)=1 \quad \text{and} \quad m(\varnothing)=0 \qquad (8)$$

In the theory of evidence, a mass of belief $m(A)$ is assigned to any set A. The total mass of belief to be distributed amounts to 1. $m(A)$ quantifies "the belief that one commits *exactly* to A, *not* the *total* belief that one commits to A" (Shafer, 1976). The uncertainty of the occurrence of event A is measured by means of two functions: a plausibility measure Pl (also called upper probability function) and a credibility measure Cr (also called belief function, or lower probability function). They are functions from $P(\Omega)$ into [0,1] such that [see Eqs. (9) and (10):

$$\forall A \in P(\Omega),\ Pl(A)=\Sigma_{B\in P(\Omega)\backslash B\cap A\neq\varnothing}m(B) \qquad (9)$$

$$\forall A \in P(\Omega),\ Cr(A)=\Sigma_{B\in P(\Omega)\backslash B\subset A}m(B) \qquad (10)$$

$Pl(A)$ assesses the degree to which the available information does not contradict A, and $Cr(A)$ assesses the degree to which this information supports A. In the field of image classification, the decision to make is "to which class should this pixel be assigned?" Ω is therefore the set of classes $1 \ldots C$, and each of the uncertainties we are interested in concerns the attribution of the considered pixel to a given class. So, only the values $(Cr(\{i\})$ and

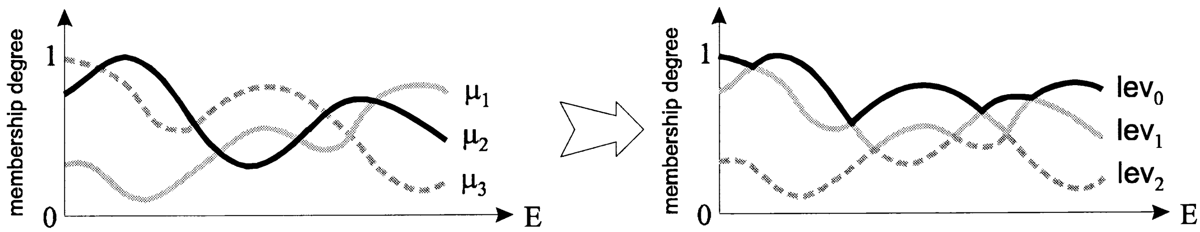

Figure 2. Levels of fuzzy three-partition μ. lev is a new fuzzy partition.

$Pl(\{i\})$, where i belongs to $1 \ldots C$, are relevant, and not the entire set of combinations of classes. A simple numerical example is provided in Appendix A2.

PLAUSIBILISTIC CLOSURE AND RELATED TOOLS

Within this section, $\mu=(\mu_i)_{i\in 1 \ldots C}$ denotes a fuzzy C-partition of E.

Plausibilistic Closure

Partitions can be crisp or fuzzy, possibilistic or probabilistic, and of various origins. Moreover, a fuzzy partition represents a huge amount of quantitative information. Only a fraction of this information may be relevant. The plausibilistic closure of μ is a partition that stems from the search of μ's intrinsic qualities, without any external references. It breaks free from the nature of the data and focuses on the most significant information. The key notion is the notion of levels, defined later.

Definition 5 (Fig. 2 and Appendix A3)
Let us consider an element j of $1 \ldots N$. There is at least one permutation (arrangement) τ of $1 \ldots C$ such that: $\forall k \in 1 \ldots C-1$, $\mu_{\tau(k)j} \geqslant \mu_{\tau(k+1)j}$. Let us assume then, for any element of k of $0 \ldots C-1$ [see Eq. (11)]:

$$\text{lev}_{kj}=\mu_{\tau(k+1)j} \tag{11}$$

These values lev_{kj} are independent of the choice of τ (if several permutations answer the expressed criterion). The lev_k functions map E into [0,1], thus defining the levels of partition μ. lev_k is level k.

The formal definition 5 implies that for any pixel e_j, it is possible to sort the C values $\mu_{i=1 \ldots C,j}$ per decreasing order and that the result is a new fuzzy partition, the partition lev. Note that the closer the level 1 is to level 0 (in the sense of Hamming distance for instance, see Klir and Yuan, 1995), the fuzzier the partition, the greater the conflicts, and the more uncertain the classification. Conversely, the farther away level 1 is from level 0, the crisper the partition, the wider the consensus, and the more certain the classification. Helpful measures of fuzziness are covered extensively in Dubois and Prade (1999). The self-analysis is carried out by referring to level 1.

Let's define μ^*_{ij} such that μ^*_{ij} is the proportion of elements of E (i.e., percentage of pixels in the image E) that have a membership degree in level 1 lower than μ_{ij} [see Eq. (12)]:

$$\mu^*_{ij}=|\{\ell \in 1 \ldots N \backslash \text{lev}_{1\ell}<\mu_{ij}\}|/N \tag{12}$$

It will be assumed, even if this point actually needs some clarifications (Matsakis, 1998), that the μ^*_{ij} values define a fuzzy C-partition of E. We now propose consideration of the evaluation of μ as that of this new partition μ^*. Three main points justify this proposition.

1. In the field of satellite image classification, evaluating a partition generally consists of comparing it to a control partition. So, let μ and ν be two fuzzy C-partitions of E, each being of any origin. μ may result from a probabilistic fuzzy segmentation or from a probabilistic fuzzy classification, or from the drawing of control zones by an analyst. The same applies for ν. Therefore, two values like μ_{ij} and ν_{ij}, for given integers i and j, are not necessarily of the same nature, and they should not be compared directly. On the other hand, μ^*_{ij} and ν^*_{ij} are both ratios. Considering the comparison of μ and ν as that of μ^* and ν^* avoids dealing with partitions of different nature.
2. Typically, any element e_j of E is associated to a vector of a given Euclidean space. In the case of a SPOT-HRV X image, for example, the space is generally the (XS1, XS3, XS3) radiometric space. In the same way, any class i is associated to a prototype (it can be a vector of the previous space). A distance d_{ij} then permits measuring the degree to which the vector associated to e_j agrees with the prototype of class i. The membership degree μ_{ij} is deduced from this measure d_{ij} with $\mu=\varphi(d)$, where φ denotes a decreasing function. It can be shown that μ^* is independent of the choice of φ (Matsakis, 1998). Therefore, considering the evaluation of μ as that of μ^* leads to evaluation of μ according to its intrinsic qualities only.
3. Let e_j be a pixel of E. If one must assign e_j to one of the C classes, he would naturally choose to assign it to the class i such that: $\text{lev}_{0j}=\mu_{ij}$. It is expected that level 0 "gathers" most of the pixels;

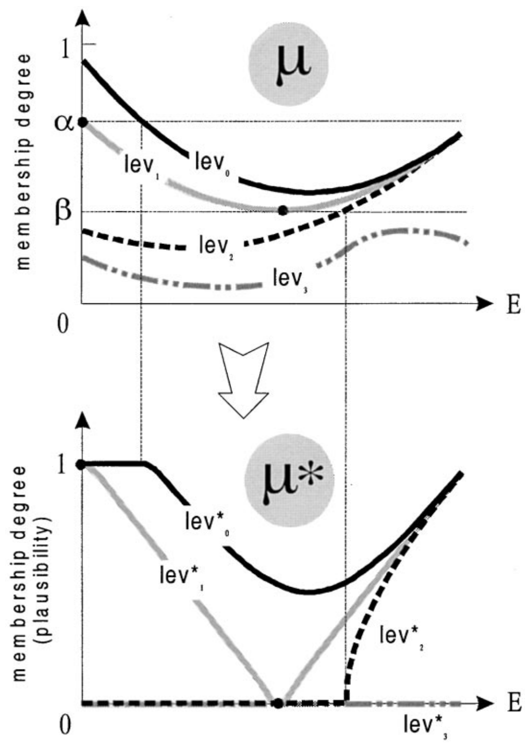

Figure 3. Plausibilistic closure μ^* of a fuzzy partition μ (representation by levels). α denotes the maximum membership degree in level 1 ($\alpha=\max_j \text{lev}_{1j}$), and β denotes the minimum membership degree in level 1. Any membership degree μ_{ij} greater than α is "set" to 1 (the proportion of elements of E that have, in level 1, a membership degree lower than μ_{ij}). Likewise, Any membership degree lower than or equal to β is "set" to 0. The other membership degrees are "stretched" between 0 and 1.

otherwise, the classification algorithm should be questioned. This is what justifies the "defuzzification" process. If ever class i was not the correct one, e_j would then be assigned to class k such that $\text{lev}_{1j}=\mu_{kj}$, and so on. Of course, level 1 is expected to gather most of the "stray" pixels. Even though consultation of lev_2 or lev_3 cannot be absolutely rejected, there is no reasonable need to consult all the lower levels. Changing μ into μ^* expresses this point of view, which will be corroborated further in the Application section. A huge amount of information lies hidden within partition μ: it is quite usual to deal with 10, 15, or even 20 classes, and as many levels. The new partition μ^* focuses on the most significant information (Fig. 3).

The formal definition of the plausibilistic closure μ^* of a partition μ follows.

Definition 6 (Appendix A3)
Let μ^* be the fuzzy C-partition of E defined by (Eq. (13)]:

$$\forall i\in 1\ldots C,\ \forall j\in 1\ldots N,\ \mu^*_{ij}=|\{\ell\in 1\ldots N \backslash \text{lev}_{1\ell}<\mu_{ij}\}|/N \tag{13}$$

where μ^* is the plausibilistic closure of μ.

Then, we define two measures, plausibility and credibility, associated with the plausibilistic closure.

Definition 7 (Appendix A4)
Let j be an element of $1\ldots N$. There is at least one permutation τ of $1\ldots C$ such that [see Eq. (14)]:

$$\forall k\in 1\ldots C-1,\ \mu^*_{\tau(k)j}\geqslant\mu^*_{\tau(k+1)j} \tag{14}$$

Integer j is associated to two functions pl_j and cr_j from $1\ldots C$ into [0,1] such as in Eqs. (15) and (16):

$$\forall i\in 1\ldots C,\ pl_j(i)=\mu^*_{ij} \tag{15}$$

$$crj[\tau(1)]=\mu^*_{\tau(1)j}-\mu^*_{\tau(2)j}\quad\text{and}\quad\forall k\in 2\ldots C,\ cr_j[\tau(k)]=0 \tag{16}$$

where $pl_j(i)$ is the plausibility, according to μ, that e_j is a member of class i, and $cr_j(i)$ is the credibility, according to μ, that e_j is a member of class i. It is independent of the choice of τ.

Propositions 1 and 2 below (proofs given in the Appendices B and C) justify the terminology adopted in Definitions 6 and 7. If the evaluation of μ can be reduced to that of μ^*, can the evaluation of μ^* be reduced in turn to that of a third partition, $(\mu^*)^*$, and so on? Proposition 1 answers this question and motivates the use of the term closure (Definition 6). Proposition 2 justifies the reference to the theory of evidence (Definition 7). Let us add that if μ is crisp, then $\mu^*=\mu$ (i.e., a crisp partition is the plausibilistic closure of itself) (Matsakis, 1998).

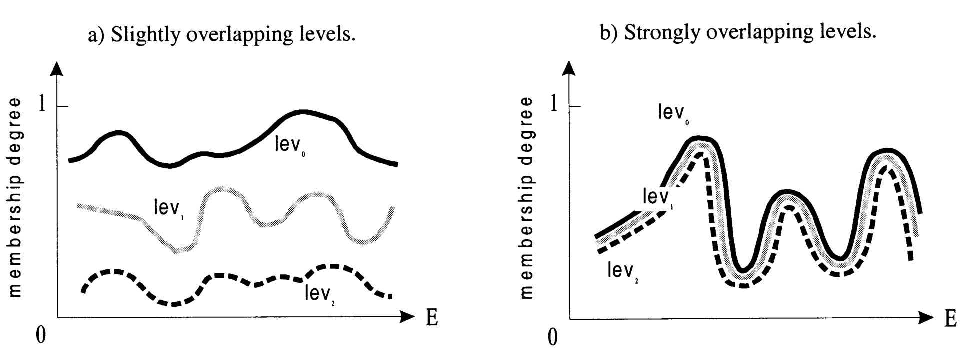

Figure 4. Two extreme cases of fuzzy partitions, with poor and high degree of overlapping between the different levels.

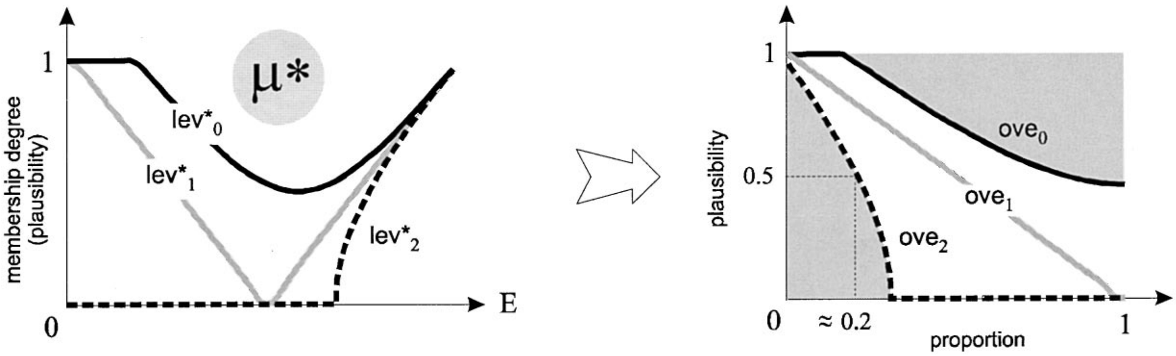

Figure 5. Overlapping construction is done using the plausibilistic closure of the considered partition. Each level of μ^* is "reorganized": its points are arranged in order of decreasing plausibilities. The graph on the right shows for instance that, in level 2, about 20% of the elements of E receive a plausibility greater than 1/2. The overlap degree of level 0 is quite high: in the upper triangular part, the grey zone delimited by ove_0 is large. The overlap degree of level 2 is lower, but not negligible.

Proposition 1 (Appendix A3)

$$(\mu^*)^* = \mu^* \quad (17)$$

Proposition 2 (Appendix A4)
Let j be an element of $1 \ldots N$. There is a basic probability function m_j, defined on the set of the subsets of $1 \ldots C+1$, such that (see [Eq. (18)]:

$$\forall i \in 1 \ldots C,\ Pl_j(\{i\}) = pl_j(i) \quad \text{and} \quad Cr_j(\{i\}) = cr_j(i) \quad (18)$$

where Pl_j and Cr_j are respectively the plausibility and credibility measures associated with m_j.

Next, we show through examples how the notion of plausibilistic closure allow the development of tools dedicated to the evaluation of fuzzy partitions.

Overlapping

Figure 4 illustrates the levels of two fuzzy partitions. The opinion expressed by the partition on left side of the figure is incomparably less ambiguous than that expressed by the right partition. These examples show the structural differences that we intend to interpret using the tools presented hereafter. These tools quantify how the levels of μ overlap with the reference level lev_1.

Definition 8 (Appendix A6)
Let k be an element of $0 \ldots C-1$ and let lev_k^* be level k of partition μ^*. There is at least one permutation τ of $1 \ldots N$ such that $\forall \ell \in 1 \ldots N-1$, $\text{lev}^*_{k\tau(\ell)} \geq \text{lev}^*_{k\tau(\ell+1)}$. This permutation corresponds to an arrangement of decreasing plausibilities of the points of level k. Now, let us consider the functions ove_k from [0,1] into [0,1] such that [see Eqs. (19) and (20)]:

$$\forall u \in [0,1] \quad u \leq 1/N \Rightarrow \text{ove}_k(u) = \text{lev}^*_{k\tau(1)} \quad (19)$$

$$\forall u \in [0,1], \quad \forall \ell \in 2 \ldots N, \quad (\ell-1)/N < u \leq \ell/N \Rightarrow \text{ove}_k(u) = \text{lev}^*_{k\tau(\ell)} \quad (20)$$

ove_k is independent of the choice of permutation τ: it is the overlapping of level k of μ.

Table 1. Mean ± Standard Deviation (in Digital Counts) of the Training Zones in Each Class for the Three Wavebands XS1, XS2, and XS3 of a SPOT-HRV Multispectral Image

Class	*N*	*XS3*	*XS2*	*XS1*
1. Deep water	503	7.30±2.02	33.08±6.88	95.24±15.45
2. Hoa	299	8.04±1.63	45.85±8.11	89.53±9.05
3. Kopara pond	44	19.70±6.41	33.95±4.31	49.27±3.69
4. Reef flat	593	12.10±5.04	59.67±10.14	92.40±10.73
5. Vegetation	417	86.42±9.41	24.50±3.68	43.23±3.53
6. Conglomerate	122	87.80±4.37	80.66±5.99	96.00±6.58
7. Soil	148	62.70±6.21	46.92±9.33	63.57±8.45
8. Coral rubble	109	117.29±8.32	131.00±11.27	150.71±10.15
9. Residual hoa	50	76.34±8.58	72.62±7.35	88.82±6.94
10. Intertidal conglomerate	332	24.21±14.18	36.82±6.63	52.49±6.74
11. Laguna	178	9.69±1.49	78.88±5.94	127.09±5.08
12. Intertidal reef flat	279	71.80±12.46	68.94±10.66	79.84±9.03
13. Intertidal coral rubble	126	41.22±34.21	103.78±28.09	133.75±14.91

N is the number of training pixels. "Kopara" is a vernacular Polynesian name for microbial mats. "Hoa" is the name for transversal spillways.

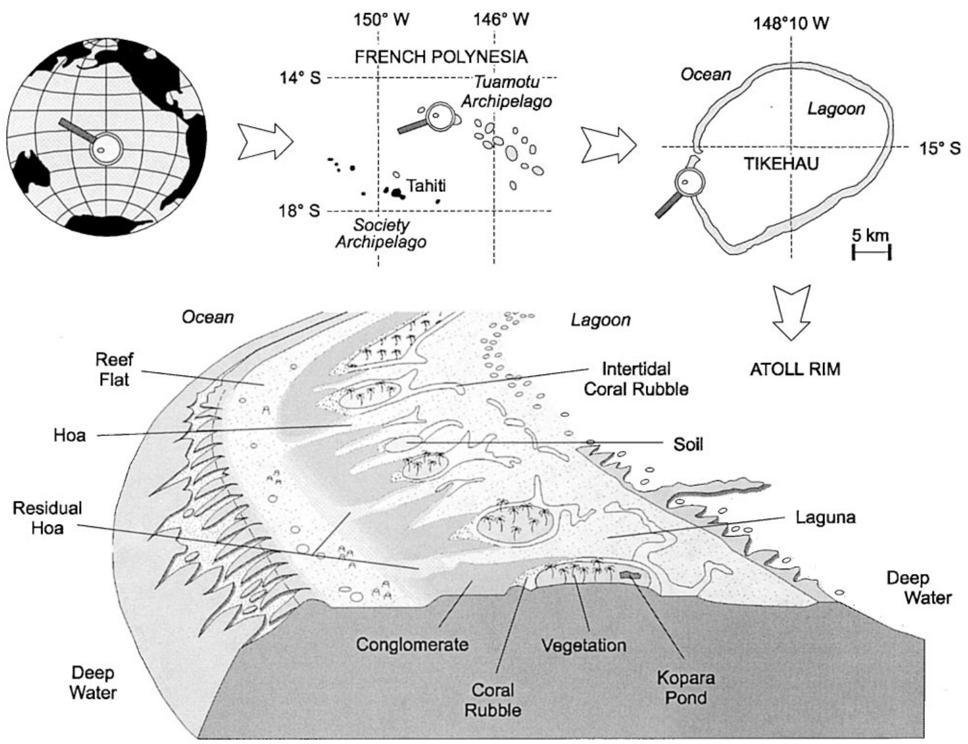

Figure 6. Location of Tikehau atoll. Location of the various classes inside a typical portion of atoll rim (bloc diagram from Battistini et al., 1975).

Overlapping verifies: $\forall k \in 0 \ldots C-2$, $\text{ove}_k \geq \text{ove}_{k+1}$. ove_k are decreasing functions normalized by N, the number of individuals (or pixels) (Fig. 5). This set of functions authorizes overlap degrees to be defined (see Definition 9).

Definition 9 (Appendix A6)
The overlap degree of level 0 is [see Eq. (21)]:

$$\int_0^1 \frac{\text{ove}_0(u)}{1-\text{ove}_1(u)} du \tag{21}$$

Now, let k be an element of $1 \ldots C-1$. The overlap degree of level k is (see Eq. (22)]:

$$\int_0^1 \frac{\text{ove}_k(u)}{\text{ove}_1(u)} du \tag{22}$$

considering that if $\text{ove}_1(u)$ is null then the ratio $\text{ove}_k(u)/\text{ove}_1(u)$ is null too.

Overlap degrees are values ranging between 0 and 1 (Definition 9). They are all equal to 0 if μ is a crisp partition and equal to 1 if μ is a fuzzy partition like the one reported in Fig. 4b. Using Proposition 1, it is easy to show that the overlapping of level k of μ is also the overlapping of level k of μ^*, and that the overlap degrees associated with μ are equal to the overlap degrees associated with μ^*.

Table 2. Confusion Matrix Obtained with PGK Classifier

	Confusion Matrix PGK Reference													
Class	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*Total*
1	**242**	42	0	0	0	0	0	0	0	0	0	0	0	284
2	21	**97**	0	1	0	0	0	0	0	1	0	0	0	120
3	0	0	**8**	0	0	0	0	0	0	49	0	0	0	57
4	0	19	0	**112**	0	0	0	0	0	6	0	0	0	137
5	0	0	0	0	**182**	0	0	0	0	0	0	0	0	182
6	0	0	0	0	0	**43**	0	0	3	0	0	0	0	46
7	0	0	0	0	7	0	**47**	0	7	0	0	0	0	61
8	0	0	0	0	0	2	0	**43**	0	0	0	0	0	45
9	0	0	0	0	0	6	0	0	**15**	0	0	2	0	23
10	0	0	15	0	0	0	0	0	0	**248**	0	0	0	263
11	0	0	0	0	0	0	0	0	0	0	**4**	0	0	4
12	0	0	0	0	0	0	2	0	11	3	0	**87**	0	103
13	7	20	0	64	0	16	0	1	0	0	47	4	**16**	175
N	270	178	23	177	189	67	49	44	36	307	51	93	16	**1500**

N is the number of control pixels. Kappa: 73.16%; Tau: 74.29%; PA: 76.27%; NPA: 84.11%.

Table 3. Confusion Matrix Obtained with FGG Classifier

Class	*Confusion Matrix FGG Reference* *1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*Total*
1	**241**	33	0	0	0	0	0	0	0	0	0	0	0	274
2	29	**131**	0	3	0	0	0	0	0	1	0	0	0	164
3	0	0	**13**	0	0	0	0	0	0	155	0	0	0	168
4	0	14	0	**152**	0	0	0	0	0	6	1	0	0	173
5	0	0	0	0	**185**	0	0	0	0	0	0	0	0	185
6	0	0	0	0	0	**58**	0	0	6	0	0	2	0	66
7	0	0	0	0	4	0	**49**	0	9	0	0	2	0	64
8	0	0	0	0	0	7	0	**44**	0	0	0	0	1	52
9	0	0	0	0	0	2	0	0	**14**	0	0	3	0	19
10	0	0	10	0	0	0	0	0	2	**142**	0	0	0	154
11	0	0	0	8	0	0	0	0	0	0	**46**	0	0	54
12	0	0	0	0	0	0	0	0	5	3	0	**86**	0	94
13	0	0	0	14	0	0	0	0	0	0	4	0	**15**	33
N	270	178	23	177	189	67	49	44	36	307	51	93	16	**1500**

N is the number of control pixels. Kappa: 75.79%; Tau: 76.60%; PA: 78.40%; NPA: 83.50%.

Plausibility and Credibility Matrices

Overlapping aims to provide a self-characterization of a fuzzy partition. Hereafter, we propose tools to evaluate fuzzy partitions still based on the plausibilistic closure, but according to an external reference. This external reference could be a crisp or a fuzzy control partition, obtained, for instance, by field survey, image interpretation, and softcopy digitizing. Eventually, it is preferable to handle both crisp and fuzzy control partitions. However, here, we only develop the case of a crisp control partition. Indeed, the theoretical development for the case of fuzzy reference partitions goes beyond the scope of the present paper, which aims to introduce the main ideas that led to the development of new tools. The understanding of these ideas does not require the analysis of all the possible cases. The general case of the evaluation of fuzzy partition according to another fuzzy control partition will be described elsewhere (Matsakis and Andréfouët, in preparation).

Let $\nu=(F_i)_{i\in 1...C}$ be a crisp C-partition of a subset F of E. Let us assume that ν is a control partition. ν thus corresponds to the "reference sample data" of traditional accuracy assessment. The study of ν allows to state if the proposition "the pixel e_j of F belongs to class i" is TRUE or FALSE, whereas the study of μ allows assigning an uncertainty quantified by the couple $[cr_j(i), pl_j(i)]$. It is expected that μ will support the picture given by ν as often as possible and with the firmest possible belief, hence the tools proposed in this section.

Definition 10 (Appendix A5)

The plausibility matrix associated with couple (μ,ν) is the matrix of size $C\times C$ defined by Eq. (23):

$$\forall k\in 1\ldots C,\ \forall i\in 1\ldots C,\ m_{ki}=\textstyle\sum_{j\in 1\ldots N \backslash e_j\in F_i} pl_j(k) \qquad (23)$$

In the same way, the credibility matrix associated with couple (μ,ν) is the matrix of size $C\times C$ defined by Eq. (24):

$$\forall k\in 1\ldots C,\ \forall i\in 1\ldots C,\ m_{ki}=\textstyle\sum_{j\in 1\ldots N \backslash e_j\in F_i} cr_j(k) \qquad (24)$$

The plausibility and credibility matrices are not the result of counting the co-occurrence of two qualitative variables (i.e., contingency tables). Therefore, they are not confusion matrices. But, if partition μ is crisp, both plausibility and credibility matrices coincide with the confusion matrix associated with couple (μ,ν) (Matsakis, 1998). Finally, it is easy to show (using Proposition 1) that the plausibility and credibility matrices associated with (μ,ν) are also the plausibility and credibility matrices associated with (μ^*,ν^*).

Table 4. Ordinal Information (in percentage) Associated with Partition PGK

Class	*Ordinal Information PGK* *1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*All*
lev_0	89.63	54.49	34.78	63.28	96.30	64.18	95.92	97.73	41.67	80.78	7.84	93.55	100.00	76.27
lev_1	9.26	45.51	56.52	33.33	3.70	17.91	4.08	2.27	27.78	16.61	84.31	4.30		20.53
lev_2	1.11		8.70	3.39		17.91			19.44	1.63	7.84	2.15		2.73
lev_3									11.11	0.00				0.27
lev_4–lev_{12}										0.98				0.20
N	270	178	23	177	189	67	49	44	36	307	51	93	16	1500

The test pixels of class 1 "join" level 0 in 89.63% of the cases. *N* is the number of control pixels.

Table 5. Ordinal Information (in percentage) Associated with Partition FGG

Class	*Ordinal Information FGG*													
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*All*
lev_0	89.26	73.60	56.52	85.88	97.88	86.57	100.00	100.00	38.89	46.25	90.20	92.47	93.75	78.40
lev_1	10.74	26.40	43.48	11.30	2.12	10.45			30.56	50.81	1.96	4.30	6.25	19.33
lev_2				2.82		1.50			16.67	1.63	3.92	3.23		1.47
lev_3						0.00			8.33	0.00	0.00			0.20
lev_4–lev_{12}						1.50			5.56	1.30	3.92			0.20
N	270	178	23	177	189	67	49	44	36	307	51	93	16	1500

The test pixels of class 1 "join" level 0 in 89.26% of the cases. *N* is the number of control pixels.

Note that the plausibility and credibility matrices are sample size-dependent, as are the confusion matrices. Of course, each entry can be divided by the number of control pixels in each class to get probability that can be readily compared form one matrix to another, as is typically done for confusion matrices.

APPLICATION

Data

Two partitions were considered to illustrate our developments. They came from supervised fuzzy classification algorithms. One algorithm derived from the unsupervised "PCM-Possibilistic C-Means" algorithm proposed by Krishnapuram and Keller (1993) and based on work by Gustafson and Kessel (1979); the other is derived from the unsupervised "FCM-Fuzzy C-Means" (Bezdek, 1981) algorithm proposed by Gath and Geva (1989). Henceforth, we used the acronyms PGK ("Possible Gustafon and Kessel") and FGG ("Fuzzy Gath and Geva") to differentiate our supervised algorithms from their unsupervised ancestors (PCM and FCM). Both algorithms used the Mahalanobis distance and required that class be previously characterized by a prototype made of a mean vector and a covariance matrix.

The fuzzy partitions provided a classification into $C=13$ classes (Table 1) to describe the morphology of a typical atoll rim of the Pacific Ocean (Andréfouët et al., in press). The 13 classes described the deep water areas (class 1); the shallow water areas (classes 2, 4, 11, and 12); the vegetation, soils, and ponds (classes 3, 5, 7, and 9); the different carbonate objects (classes 6, 8, 10, and 13) (Fig. 6). Some classes were considered only because of their geomorphologic relevance and to obey standards in the description of atoll structures (Battistini et al., 1975). We used a set E containing $N=160{,}345$ pixels coming from a SPOT-HRV XS multispectral image of Tikehau atoll (Intes et al., 1995), in French Polynesia. The control partition used for the computation of the various matrices (confusion, plausibility, and credibility matrices) included 1,500 pixels of a subset F of E. In the field, training zones were profiles of several hundred meters long and 60 meters wide that crossed both homogeneous and transition zones. Profiles were randomly positioned along the atoll rim. Mean and standard deviation of the prototype vectors appear in Table 1. Some classes have similar statistical parameters (e.g., "kopara pond" and "intertidal conglomerate"), hence the system was not *a priori* optimized for spectral discrimination of the classes (Chuvieco and Congalton, 1988). The control partition de-

Table 6. Plausibility Matrix Associated with Partition PGK

Class	*Plausibility Matrix PGK*												
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
1	**250.95**	115.15	4.83	11.91	0	0	0	0	0	54.24	1.37	0	0
2	157.69	**167.19**	0.94	46.48	0	0	0	0	0	26.44	4.59	0	0
3	0	0	**16.85**	0	0	0	1.12	0	0.58	204.13	0	1.72	0
4	14.87	48.39	4.14	**158.68**	0	0	0	0	0	68.70	14.53	0	0.06
5	0	0	0.14	0	**141.85**	0	4.66	0	0.37	4.19	0	0.03	0
6	0	0	0	0	0	**48.33**	2.86	0.62	10.34	0.22	0	14.09	0.01
7	0.09	0.06	0.47	0.73	20.73	2.12	**48.34**	0	18.20	10.09	0	44.37	0.44
8	0	0	0	0	0	16.35	0.75	**42.55**	1.36	0	0	1.27	0.66
9	0.97	0.17	0.83	0.60	0.02	41.79	23.18	0.51	**28.99**	17.78	0	62.07	0.43
10	0.99	1.70	20.72	9.46	0.71	0.94	8.31	0	5.97	**291.14**	0.06	23.02	0.33
11	3.47	16.00	0	18.55	0	0	0	0	0	0.03	**37.74**	0	0
12	1.60	0.98	3.33	6.05	9.31	13.40	39.85	0.51	27.39	72.64	0.07	**87.74**	1.52
13	127.78	134.32	1.07	135.62	4.48	34.28	5.01	23.56	7.98	19.50	50.02	18.45	**13.61**
N	270	178	23	177	189	67	49	44	36	307	51	93	16

The sum of the plausibilities that pixels of class 1 belong to class 2 was estimated at 157.69 (first column, second row). *N* is the number of control pixels.

Table 7. Plausibility Matrix Associated with Partition FGG

	Plausibility Matrix FGG												
Class	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
1	**264.66**	85.99	1.00	0	0	0	0	0	0	11.31	0	0	0
2	132.97	**169.62**	0	26.06	0	0	0	0	0	15.04	0.22	0	0
3	0	0	**19.26**	0	0	0	0	0	0	202.23	0	0	0
4	3.36	32.83	0	**170.29**	0	0	0	0	0	31.53	5.12	0	0
5	0	0	0	0	**188.02**	0	1.15	0	0	0	0	0	0
6	0	0	0	0	0	**63.88**	0	0	9.62	0	0	11.79	0
7	0	0	0	0	14.31	0	**49.00**	0	14.53	0	0	33.18	0
8	0	0	0	0	0	14.24	0	**44.00**	0	0	0	0	1.00
9	0	0	0	0	0	32.28	13.50	0	**25.24**	1.39	0	52.30	0
10	0	0	19.41	0	0	0	0	2.62	**264.63**	0	8.86	0	
11	0	3.32	0	14.32	0	0	0	0	0	0	**47.53**	0	0
12	0	0	0.28	0	0	0	23.25	0	19.93	14.18	0	**91.09**	1.50
13	11.51	22.71	0	56.26	0	7.33	0	0.80	0	0	28.22	4.42	**15.47**
N	270	178	23	177	189	67	49	44	36	307	51	93	16

The sum of the plausibilities that pixels of class 1 belong to class 2 was estimated at 132.97 (first column, second row). *N* is the number of control pixels.

rived from our training profiles could be both crisp or fuzzy, depending on the inclusion of the transition zones in the data set. As a beginning, for illustration of the use of the new evaluation tools, we used only a crisp control partition and focused only on the homogeneous zones of the atoll rim. PGK is possibilistic (Definition 2), and FGG is probabilistic (Definition 3). Therefore, partitions created by FGG and PGK were of different nature, but were evaluated regardless of their nature through their plausibilistic closure. Intentionally using two different partitions, we intended to demonstrate that the plausibilistic closure respects their intrinsic properties and that overlapping, plausibility, and credibility matrices are adequate to express their particularities.

Confusion Matrices

Table 2 reports the confusion matrix associated with the PGK partition. This matrix resulted from the comparison of two crisp partitions; the control partition and the partition obtained by defuzzification of PGK. Four coefficients were computed: the PA ("Percentage Agreement"), NPA ("Normalized Percentage Agreement" as defined by Zhuang et al., 1995), Kappa and Tau (Story and Congalton, 1986; Congalton, 1991; Ma and Redmond, 1995) coefficients. The confusion matrix and the coefficients associated with partition FGG are shown in Table 3. The pairwise differences observed between the eight coefficients were not significant (Ma and Redmond, 1994). At the class level, PGK was most successful for some classes (class 10), while FGG performed better for others (classes 2, 4, and 11). Most of the time, the performances were similar (classes 1, 3, 5, 7, 8, 9, 12, and 13). Overall, it is difficult to distinguish the best partition among the two crisp partitions achieved by defuzzification.

Ordinal Information

The ordinal information corroborated a point of view expressed previously. Let us consider the first column of

Table 8. Normalized Plausibility Matrix Associated with Partition FGG

	Normalized Plausibility Matrix FGG												
Class	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
1	**98.02**	48.31	4.35	0	0	0	0	0	0	3.68	0	0	0
2	49.25	**95.29**	0	14.72	0	0	0	0	0	4.90	0.43	0	0
3	0	0	**83.74**	0	0	0	0	0	0	65.87	0	0	0
4	1.24	18.44	0	**96.21**	0	0	0	0	0	10.27	10.04	0	0
5	0	0	0	0	**99.48**	0	2.35	0	0	0	0	0	0
6	0	0	0	0	0	**95.34**	0	0	26.72	0	0	12.68	0
7	0	0	0	0	7.57	0	**100.00**	0	40.36	0	0	35.68	0
8	0	0	0	0	0	21.25	0	**100.00**	0	0	0	0	6.25
9	0	0	0	0	0	48.18	27.55	0	**70.11**	0.45	0	56.24	0
10	0	0	84.39	0	0	0	0	0	7.28	**86.20**	0	9.53	0
11	0	1.87	0	8.09	0	0	0	0	0	0	**93.20**	0	0
12	0	0	1.22	0	0	0	47.45	0	55.36	4.62	0	**97.95**	9.38
13	4.26	12.76	0	31.79	0	10.94	0	1.820	0	0	55.33	4.75	**96.69**

Values are expressed in percentages. On average, a text pixel of class 1 (first column) was assigned the plausibility 0.49 (49.25%) of belonging to class 2 (second row).

Table 9. Credibility Matrix Associated with Partition PGK

	Credibility Matrix PGK												
Class	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
1	**80.00**	3.27	0	0	0	0	0	0	0	0	0	0	0
2	1.45	**9.14**	0	0.01	0	0	0	0	0	0	0	0	0
3	0	0	**0.07**	0	0	0	0	0	0	0.25	0	0	0
4	0	1.37	0	**28.92**	0	0	0	0	0	0.81	0	0	0
5	0	0	0	**121.68**	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	**4.35**	0	0	0.19	0	0	0	0
7	0	0	0	0	0.58	0	**8.50**	0	0.74	0	0	0	0
8	0	0	0	0	0	0.28	0	**19.18**	0	0	0	0	0
9	0	0	0	0	0	0.41	0	0	**3.73**	0	0	0.01	0
10	0	0	3.95	0	0	0	0	0	0	**55.98**	0	0	0
11	0	0	0	0	0	0	0	0	0	0	**0.01**	0	0
12	0	0	0	0	0	0	0.01	0	0.89	0.99	0	**19.39**	0
13	0.51	1.96	0	11.89	0	1.88	0	0.19	0	0	10.79	0.28	**11.47**
N	270	178	23	177	189	67	49	44	36	307	51	93	16

The sum of the credibilities that pixels of class 1 belong to class 2 was estimated at 1.45 (first column, second row). *N* is the number of control pixels.

Table 4 (partition PGK). In 89.63% of the cases, the control pixels of class 1 "joined" level 0, and defuzzification of PGK would lead to the correct classification of these pixels. In the other cases, defuzzification of PGK would lead to a wrong classification. But, as expected, level 1 itself gathered most of the "stray" pixels, and, even if some still went astray to level 2 (1.11% of the cases), none got lost in the lower levels. As a general rule, if a control pixel did not join level 0, then it went astray to level 1. For partition FGG (Table 5), this rule was transgressed only in 2.27% of the cases (1.47+0.20+0.60). For PGK, it was in 3.20% of the cases (2.73+0.27+0.20).

Plausibility and Credibility Matrices

The diagonal gathered the greatest values in any matrices. Generally, the partition PGK agreed, like FGG, with the control partition. However, PGK and FGG displayed different features:

- The out-of-diagonal values of the plausibility matrix associated with partition PGK (Table 6) were generally greater than the corresponding values in the plausibility matrix associated with partition FGG (Table 7). Table 8 illustrates a normalized plausibility matrix, where each entry was divided by the number of control pixels to get probability. Here, Tables 6 and 7 can be compared readily because the same set of control pixels was used.
- The diagonal values of the credibility matrix associated with partition PGK (Table 9) were much lower than the corresponding values in the credibility matrix associated with partition FGG (Table 10). Here, Tables 9 and 10 can be compared readily because the same set of control pixels was used.

As a general rule, PGK considered that several classes were plausible for a given pixel. Consequently, even the

Table 10. Credibility Matrix Associated with Partition FGG

	Credibility Matrix FGG												
Class	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
1	**133.38**	6.08	0	0	0	0	0	0	0	0	0	0	0
2	5.34	**55.51**	0	0.41	0	0	0	0	0	0.10	0	0	0
3	0	0	**3.59**	0	0	0	0	0	0	35.97	0	0	0
4	0	2.30	0	**95.81**	0	0	0	0	0	1.04	0.07	0	0
5	0	0	0	0	**174.69**	0	0	0	0	0	0	0	0
6	0	0	0	0	0	**25.12**	0	0	1.58	0	0	0.56	0
7	0	0	0	0	0.98	0	**25.47**	0	3.26	0	0	0.13	0
8	0	0	0	0	0	2.83	0	**43.20**	0	0	0	0	0.53
9	0	0	0	0	0	0.28	0	0	**5.46**	0	0	0.30	0
10	0	0	3.74	0	0	0	0	0	0.07	**66.34**	0	0	0
11	0	0	0	3.31	0	0	0	0	0	0	**22.70**	0	0
12	0	0	0	0	0	0	0	0	1.28	1.28	0	**29.34**	0
13	0	0	0	2.71	0	0	0	0	0	0	1.08	0	**13.50**
N	270	178	23	177	189	67	49	44	36	307	51	93	16

The sum of the credibilities that pixels of class 1 belong to class 2 was estimated at 5.34 (first column, second row). *N* is the number of control pixels.

Table 11. Overlap Degrees (in percentage) Associated with Partitions PGK and FGG

	Level												
	0	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
PGK overlap degree	26.61	98.63	50.74	19.80	8.57	3.20	1.53	0.60	0.28	0.08	0	0	0
FGG overlap degree	0.05	78.52	13.86	2.45	0.23	0	0	0	0	0	0	0	0

most plausible class to which the pixel could be assigned did not receive very high credibility. Inspect, for instance, the case of classes 3 and 10, which are spectrally very close (Table 1). Plausibilities were similar for both algorithms, but credibilities were very different and were higher for FGG. It may be asked whether partition PGK is not excessively cautious and partition FGG excessively self-confident. In the case of classes 3 and 10, there is no reason that justifies the high credibilities achieved by FGG. Indeed, only contextual knowledge could tell if a pixel belonged or not to the class 3 (kopara pond) or class 10 (intertidal conglomerate) (Andréfouët and Roux, 1998).

Overlapping

Overlap is a self-evaluation of the entire fuzzy partition, without any external reference. Overlapping was unrelated to the 1,500 control pixels and was instead computed from the entire set of 160,345 pixels. Table 11 and Fig. 7b show that the levels of the FGG fuzzy partition overlapped very slightly. Level 0, for instance, was almost completely separated from level 1: its overlap degree was only 0.05%. The overlap degree for level 2 was much higher, 13.86%. It could be estimated, observing Fig. 7b, that only 10% of the pixels in level 2 received a plausibility greater than 0.50 (only 30% receive a nonzero plausibility). A two-dimensional representation of partition FGG would be close to Fig. 4a. For PGK, it would be halfway between Figs. 4a and 4b. The overlapping figures confirmed visually the constraints of the probabilistic partition (Definition 3), which generated low degree of overlap, as well as the lack of constraints (Definition 2) of a possible partition that could generate higher degrees of overlap.

Remarks

The low degree of overlap and high level of credibility stated by FGG, even in case of spectral confusion, is explained by the probabilistic nature of the FGG algorithm. Such patterns did not clearly reveal the ambiguities between classes. Conversely, the overlap figure obtained for PGK, poor credibility, and high plausibility suggested that many classes were spectrally similar. This is an accurate conclusion, consistent with the presence of classes of similar nature (shallow water) separated only to obey geomorphological considerations (e.g., the class 2 “hoa,” transversal to the rim, and the class 4, “reef flat,” longitudinal). With regard to the characteristics of some classes (spectrally close, but differentiated by the analyst for functional or structural reasons), PGK appeared better than FGG since it did not arbitrarily separate classes of similar nature.

Actually, when comparing partitions to decide which one is better for a given application, it is necessary to keep in mind the primary objectives that led to their creations. These objectives may be divergent and lead to antagonistic criteria. If the partition has to describe as precisely as possible the true nature of the ground, its complexity and heterogeneity, even if that entails mixing the classes (and thus bears testimony to the artificial character of these classes), then a good partition has high degrees of overlap, high values within the plausibility matrix, diagonal and out-of-diagonal as well. Such partitions allows the location of high-risk confusion zones and their corrections, for instance, in the light of contextual knowledge (Andréfouët et al., 2000). If the partition has to represent the ground while preserving a predefined class

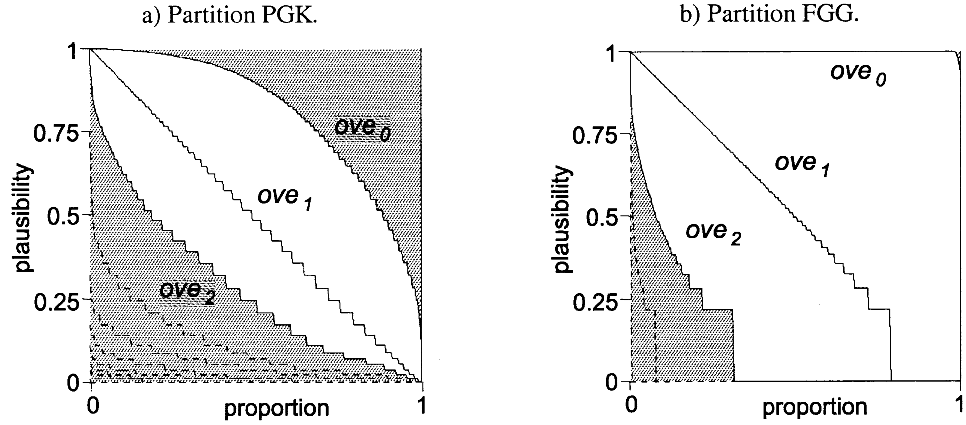

Figure 7. Overlapping graphs for partitions PGK and FGG.

Table 12.

μ	1	2	3	4	5	6	7	8	9	10
μ_1	0.4	0.5	0.4	1	0.3	0	0.2	0.9	0.5	0.7
μ_2	0.8	0.4	0.4	0.2	0.6	0.9	0.3	0.6	0.7	0.4
μ_3	0.1	0.7	0.4	0.8	0.3	0.5	0.5	0.1	0.7	0.3

Table 13.

ν	1	2	3	4	5	6	7	8	9	10
ν_1	0	0	1	1	0	0	0	1	0	1
ν_2	1	0	0	0	1	1	0	0	0	0
ν_3	0	1	0	0	0	0	1	0	1	0

system, even if that entails reducing the vision on the ground reality, then a good partition has low degrees of overlap, high diagonal values within the credibility matrix, and low out-of-diagonal values. A similar criterion is implicitly adopted when conventional cartography is done. Whatever the criterion selected for the comparison of several partitions, we show that it can be assessed through overlapping, plausibility, and credibility matrices.

CONCLUSION

Fuzzy partition evaluation is an intricate problem that still requires many developments. In this paper, we have presented an approach employing elements of artificial intelligence and information theory. We particularly asked ourselves which material could serve as a basis for the evaluation of a fuzzy partition. Three main preoccupations guided our research: (1) judging a partition by its intrinsic qualities only; (2) breaking free from the nature of the data, probabilistic or possibilistic; and (3) focusing on the most significant information located in the highest levels of the partition. The notion of plausibilistic closure of a fuzzy partition was introduced. Our theoretical developments led to the proposition of three new tools to evaluate a fuzzy partition. First, overlapping provides information on the entire structure of the partition without any external reference. Then, plausibility and credibility matrices, whose construction requires the use of a control partition, allow a class-by-class analysis according to measures derived from the theory of evidence. Because of these tools, major differences between partitions of various natures can be detected to select which partition is preferable for a given application. Overlapping, plausibility, and credibility matrices that we introduced here represent a novel step toward the development of methods dedicated to the evaluation of fuzzy partition.

The numerous constructive comments of the reviewers were of great value in improving the initial manuscript. We are also grateful to GEOIMAGE company (Sophia-Antipolis, France) and Lionel Laurore, who provided a license of their image processing software. We also thank IFREMER (Dr. Yann Morel) and Territoire de Polynésie Française for providing the SPOT image used in this study. "Programme National sur les Récifs Coralliens" and "Institut de Recherche pour le Developpment" provided financial support to S. A. for fieldwork in Tikehau.

APPENDIX A (TUTORIAL)

In this tutorial example, E is a set $\{e_1,e_2,e_3,e_4,e_5,e_6,e_7,e_8,e_9,e_{10}\}$ of $N=10$ elements.

Crisp and Fuzyy Partitions

Each value of the Table 12 belongs to the interval [0,1], and each column and each row contain at least one non-null element. Table 12 therefore defines a (possibilistic) fuzzy partition. Each value corresponds to a membership degree. The notation μ_{ij} suits the definitions, but $\mu_{i,j}$ is preferable here. The membership degree of e_{10} in class 2 is thus $\mu_{2,10}=0.4$. Table 13 defines another fuzzy partition into $C=3$ classes of E.

Contrary to μ, this second partition ν is a probabilistic fuzzy partition (adding together the three elements of any column gives 1), and the membership degrees $\nu_{i,j}$ are all equal to 0 or 1. ν is therefore a crisp three-partition of E. It could have been defined by the three-tuple (E_1,E_2,E_3) as well, with $E_1=\{e_3,e_4,e_8,e_{10}\}$, $E_2=\{e_1,e_5,e_6\}$, and $E_3=\{e_2,e_7,e_9\}$. As a matter of fact, ν is the result of a "defuzzification" of μ: a membership degree $\nu_{i,j}$ equal to 1 indicates that the related membership degree $\mu_{i,j}$ is the maximum value of the column concerned. For instance, $\nu_{2,6}=1$ and $\max_{k\in 1\ldots 3}\mu_{k,6}=\max\{0,0.9,0.5\}=0.9=\mu_{2,6}$. The partition ν expresses that element e_6 essentially belongs to class 2 according to μ.

Plausibility and Credibility Measures

Let Ω be the set 1 . . . 4, and let m be the function from $\mathcal{P}(\Omega)$ into [0,1] defined by $m(\{2\})=0.7$, $m(\{1,2\})=0.2$, $m(\{4\})=0.1$, and $m(A)=0$ for any other subset A of Ω. We have [see Eq. A1)]:

Table 14.

lev	1	2	3	4	5	6	7	8	9	10
lev_0	0.8	0.7	0.4	1	0.6	0.9	0.5	0.9	0.7	0.7
lev_1	0.4	0.5	0.4	0.8	0.3	0.5	0.3	0.6	0.7	0.4
lev_2	0.1	0.4	0.4	0.2	0.3	0	0.2	0.1	0.5	0.3

Table 15.

μ^*	1	2	3	4	5	6	7	8	9	10
μ_1^*	0.2	0.5	0.2	1	0	0	0	1	0.5	0.8
μ_2^*	0.9	0.2	0.2	0	0.7	1	0	0.7	0.8	0.2
μ_3^*	0	0.8	0.2	0.9	0	0.5	0.5	0	0.8	0

Table 16.

*lev**	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
lev_0^*	0.9	0.8	0.2	1	0.7	1	0.5	1	0.8	0.8
lev_1^*	0.2	0.5	0.2	0.9	0	0.5	0	0.7	0.8	0.2
lev_2^*	0	0.2	0.2	0	0	0	0	0	0.5	0

Table 18.

cr	cr_1	cr_2	cr_3	cr_4	cr_5	cr_6	cr_7	cr_8	cr_9	cr_{10}
1	0	0	0	0.1	0	0	0	0.3	0	0.6
2	0.7	0	0	0	0.7	0.5	0	0	0	0
3	0	0.3	0	0	0	0	0.5	0	0	0

$$\sum_{A\in\mathcal{P}(\Omega)} m(A)=m(\{2\})+m(\{1,2\})+m(\{4\})=1 \quad \text{and} \quad m(\emptyset)=0 \tag{A1}$$

m is therefore a basic probability function.

Let Pl and Cr be the plausibility and credibility measures associated with m. We have [see Eqs. (A2), (A3), (A4), (A5), (A6), and (A7)]:

$$Pl[\{1\})=\sum_{B\in\mathcal{P}(\Omega)\backslash B\cap\{1\}\neq\emptyset} m(B)=m(\{1,2\})=0.2 \tag{A2}$$

$$Pl[\{2\})=\sum_{B\in\mathcal{P}(\Omega)\backslash B\cap\{2\}\neq\emptyset} m(B)=m(\{2\})+m(\{1,2\})=0.9 \tag{A3}$$

$$Pl[\{3\})=\sum_{B\in\mathcal{P}(\Omega)\backslash B\cap\{3\}\neq\emptyset} m(B)=0 \tag{A4}$$

$$Cr[\{1\})=\sum_{B\in\mathcal{P}(\Omega)\backslash B\subset\{1\}} m(B)=0 \tag{A5}$$

$$Cr[\{2\})=\sum_{B\in\mathcal{P}(\Omega)\backslash B\subset\{2\}} m(B)=m(\{2\})=0.7 \tag{A6}$$

$$Cr[\{3\})=\sum_{B\in\mathcal{P}(\Omega)\backslash B\subset\{3\}} m(B)=0 \tag{A7}$$

Levels and Plausibilistic Closure

The levels of partition μ are shown in Table 14. For each column of the μ table (see above), the membership degrees are arranged in decreasing order (from top to bottom). The plausibilistic closure of μ is shown in Table 15. The 10 lev_1 values are $\text{lev}_{1,1}=0.4$, $\text{lev}_{1,2}=0.5$, $\text{lev}_{1,3}=0.4$, $\text{lev}_{1,4}=0.8$, etc. Two of these values, $\text{lev}_{1,5}=0.3$ and $\text{lev}_{1,7}=0.3$, are lower than $\mu_{1,1}=0.4$; therefore, $\mu^*_{1,1}=2/10=0.2$. All these values except $\text{lev}_{1,4}=0.8$ are lower than $\mu_{2,1}=0.8$; therefore, $\mu^*_{2,1}=9/10=0.9$. None of these values are lower than $\mu_{3,5}=0.3$; therefore, $\mu^*_{3,5}=0/10=0$, etc. Table 16 shows the levels of partition μ^*. We let the reader compute the plausibilistic closure of μ^* and check that $(\mu^*)^*=\mu^*$ (see Proposition 1).

Plausibilities and Credibilities According to μ

The plausibility (Table 17) according to μ that e_1 is a member of class 1 is $pl_1(1)=\mu^*_{1,1}=0.2$. The plausibility according to μ that e_1 is a member of class 2 is $pl_1(2)=\mu^*_{2,1}=0.9$. The plausibility according to μ that e_5 is a member of class 3 is $pl_5(3)=\mu^*_{3,5}=0$, etc.

Now, let us consider for instance the first column of the μ^* table. The maximum value is $\mu^*_{2,1}=0.9$ (according to μ^*, and μ, the element e_1 essentially belongs to class 2). The value $\mu^*_{1,1}=0.2$ comes after. Therefore, the credibility, according to μ, that e_1 is a member of class 2 is $cr_1(2)=\mu^*_{2,1}-\mu^*_{1,1}=0.9-0.2=0.7$ (Table 18). The other credibilities cr_1 and $cr_1(3)$ are null.

Table 17.

pl	pl_1	pl_2	pl_3	pl_4	pl_5	pl_6	pl_7	pl_8	pl_9	pl_{10}
1	0.2	0.5	0.2	1	0	0	0	1	0.5	0.8
2	0.9	0.2	0.2	0	0.7	1	0	0.7	0.8	0.2
3	0	0.8	0.2	0.9	0	0.5	0.5	0	0.8	0

The basic probability function m defined as above is related to the membership degrees $\mu_{i,1}$ concerning element e_1 (see Proposition 2). The first column of the pl and cr tables can be retrieved from the plausibility and credibility measures Pl and Cr associated with m [see Eqs. (A8) and (A9):

$$pl_1(1)=Pl(\{1\})=0.2,\ pl_1(2)=Pl(\{2\})=0.9,\ pl_1(3)=Pl(\{3\})=0 \tag{A8}$$

$$cr_1(1)=Cr(\{1\})=0,\ cr_1(2)=Cr(\{2\})=0.7,\ cr_1(3)=Cr(\{3\})=0 \tag{A9}$$

Plausibility and Credibility Matrices

The crisp partition ν presented earlier is the result of a "defuzzification" of μ. Let us also consider ν a test partition. In that particular case, the subset F is E itself, F_1 is $\{e_3,e_4,e_8,e_{10}\}$, F_2 is $\{e_1,e_5,e_6\}$, and F_3 is $\{e_2,e_7,e_9\}$. The plausibility and credibility matrices associated with couple (μ,ν) are defined by the Tables 19 and 20.

For instance, the value 1.1 that appears in the plausibility matrix, 1st row, 2nd column, is the sum of the values located in the second row of the pl table and that are related to the elements e_3,e_4,e_8, and e_{10} of class 1. We have: 1.1=0.2+0+0.7+0.2.

Overlapping

Let us consider the following points: (0.1,1.0), (0.2,1.0), (0.3,1.0), (0.4,0.9), (0.5,0.8), (0.6,0.8), (0.7,0.8), (0.8,0.7), (0.9,0.5), and 1.0,0.2). The x-coordinates of these points are evenly distributed between 0 (excluded) and 1 (included). The y-coordinates are the 10 lev_0^* values arranged in decreasing order. The points define a scaled function ove_0, the overlapping of level 0 of μ (Fig. 8).

The overlap degree of level 0 is about 28% [see Eq. (A10)]:

Table 19.

Plaus.	*1*	*2*	*3*
1	3.0	1.1	1.1
2	0.2	2.6	0.5
3	1.0	1.0	2.1
N	4	3	3

Table 20.

Cred.	*1*	*2*	*3*
1	1.0	0	0
2	0	1.9	0
3	0	0	0.8
N	4	3	3

$$\int_0^1 \frac{1-\mathrm{ove}_0(u)}{1-\mathrm{ove}_1(u)}du = \frac{1-1.0}{1-0.9}0.1+\frac{1-1.0}{1-0.8}0.1+\frac{1-1.0}{1-0.7}0.1 + \frac{1-0.9}{1-0.5}0.1+\frac{1-0.8}{1-0.5}0.1+\frac{1-0.8}{1-0.2}0.1 + \frac{1-0.8}{1-0.2}0.1+\frac{1-0.7}{1-0.2}0.1+\frac{1-0.5}{1-0.0}0.1 + \frac{1-0.2}{1-0.0}0.1 = 0.2775 \quad \text{(A10)}$$

The overlap degree of level 1 is 80% and the overlap degree of level 2 is about 11% [see Eq. (A11)]:

$$\int_0^1 \frac{\mathrm{ove}_1(u)}{\mathrm{ove}_1(u)}du = 1\times 0.8 + 0\times 0.2 = 0.80 \quad \text{and}$$

$$\int_0^1 \frac{\mathrm{ove}_2(u)}{\mathrm{ove}_1(u)}du \approx 0.11 \quad \text{(A11)}$$

Remember that if $\mathrm{ove}_1(u)$ is null then the ratio $\mathrm{ove}_k(u)/\mathrm{ove}_1(u)$ is considered null too.

APPENDIX B (PROOF OF PROPOSITION 1)

For any k of $0 \ldots C-1$, let lev_k be the level k of μ and lev_k^* be the level k of μ^*. Let also f and f^* be the mappings from [0,1] into [0,1] defined by Eqs. (A12) and (A13):

$$\forall t \in [0,1],\ f(t) = |\{\ell \in 1 \ldots N \backslash \mathrm{lev}_{1\ell} < t\}|/N \quad \text{(A12)}$$

$$\forall t \in [0,1],\ f^*(t) = |\{\ell \in 1 \ldots N \backslash \mathrm{lev}_{1\ell}^* < t\}|/N \quad \text{(A13)}$$

According to Definition 6 of the plausibilistic closure of a partition [see Eqs. (A14) and (A15)]:

$$\forall i \in 1 \ldots C,\ \forall j \in 1 \ldots N,\ \mu_{ij}^* = f(\mu_{ij}) \quad \text{(A14)}$$

$$\forall i \in 1 \ldots C,\ \forall j \in 1 \ldots N,\ (\mu^*)_{ij}^* = f^*(\mu_{ij}^*) \quad \text{(A15)}$$

For any j of $1 \ldots N$ let τ_j be a permutation of $1 \ldots C$ such that $\forall k \in 1 \ldots C-1$, $\mu_{\tau_j(k)j} \geq \mu_{\tau_j(k+1)j}$. We have [see Eqs. (A16) and (A17)]:

$$\forall j \in 1 \ldots N,\ \forall k \in 1 \ldots C-1,\ \mu_{\tau_j(k)j} \geq \mu_{\tau_j(k+1)j} \quad \text{(A16)}$$

$$\forall j \in 1 \ldots N,\ \forall k \in 0 \ldots C-1,\ \mathrm{lev}_{kj} = \mu_{\tau_j(k+1)j} \quad \text{(A17)}$$

where Eq. (A17) results from Definition 5 concerning levels.

Generally, f is not a strictly increasing function. But we have [see Eq. (A18)]:

$$\forall i \in 1 \ldots C,\ \forall j \in 1 \ldots N,\ \forall k \in 1 \ldots N,\ \mathrm{lev}_{1k} < \mu_{ij} \Leftrightarrow f(\mathrm{lev}_{1k}) < f(\mu_{ij}) \quad \text{(A18)}$$

We will prove this point now.

So, let i be an element of $1 \ldots C$, and j and k two elements of $1 \ldots N$. Alternately applying Eq. (A12) and Eq. (A17) we have [see Eqs. (A19) and (A20)]:

$$f(\mathrm{lev}_{1k}) = f[\mu_{\tau_k(2)k}] = |[\ell \in 1 \ldots N \backslash \mathrm{lev}_{1\ell} < \mu_{\tau_k(2)k}]|/N = |\{\ell \in 1 \ldots N \backslash \mu_{\tau_\ell(2)\ell} < \mu_{\tau_k(2)k}\}|/N \quad \text{(A19)}$$

$$f(\mu_{ij}) = |\{\ell \in 1 \ldots N \backslash \mathrm{lev}_{1\ell} < \mu_{ij}\}|/N = |\ell \in 1 \ldots N \backslash \mu_{\tau_\ell(2)\ell} < \mu_{ij}\}|/N \quad \text{(A20)}$$

Let us prove first that f is an increasing function. Let t_1 and t_2 be two elements of [0,1] such that $t_1 < t_2$. This assumption and Eq. (A12) give [see Eqs. (A21) and (A22)]:

$$f(t_2) = |\{\ell \in 1 \ldots N \backslash \mathrm{lev}_{1\ell} < t_2\}|/N = (|\{\ell \in 1 \ldots N \backslash \mathrm{lev}_{1\ell} < t_1\}| + |\{\ell \in 1 \ldots N \backslash t_1 \leq \mathrm{lev}_{1\ell} < t_2\}|)/N \quad \text{(A21)}$$

$$f(t_2) = f(t_1) + |\ell \in 1 \ldots N \backslash t_1 \leq \mathrm{lev}_{1\ell} < t_2\}|/N \quad \text{(A22)}$$

Consequently, $f(t_1) \leq f(t_2)$; f is then an increasing function. So $\mathrm{lev}_{1k} \geq \mu_{ij} \Rightarrow f(\mathrm{lev}_{1k}) \geq f(\mu_{ij})$.

Let us assume that $\mathrm{lev}_{1k} < \mu_{ij}$, that is, according to Eq. (A17): $\mu_{\tau_k(2)k} < \mu_{ij}$. This assumption and also Eqs. (A19) and (A20) give [see Eqs. (A23) and (A24)]:

$$f(\mu_{ij}) = (|\{\ell \in 1 \ldots N \backslash \mu_{\tau_\ell(2)\ell} < \mu_{\tau_k(2)k}\}|$$

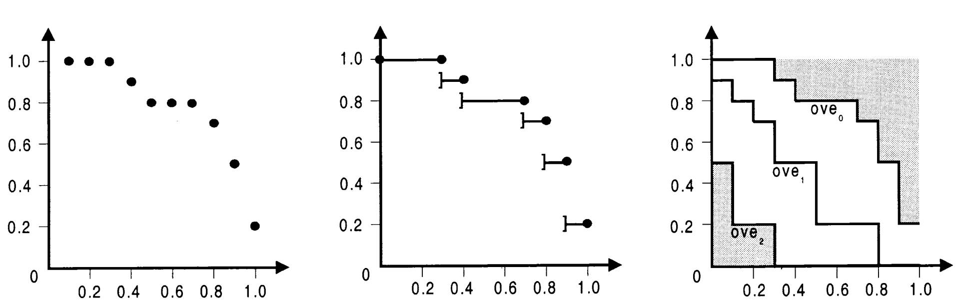

Figure 8.

$$+|\{\ell\in 1\ldots N\backslash\mu_{\tau_k(2)k}\leq\mu_{\tau_\ell(2)\ell}<\mu_{ij}\}|)/N \quad (A23)$$

$$f(\mu_{ij})=f(\text{lev}_{1k})+|\{\ell\in 1\ldots N\backslash\mu_{\tau_k(2)k}\leq\mu_{\tau_\ell(2)\ell}<\mu_{ij}\}|/N \quad (A24)$$

Now, $|\{\ell\in 1\ldots N\backslash\mu_{t_k(2)k}\leq\mu_{\tau_\ell(2)\ell}<\mu_{ij}\}|>0$. Indeed, $k\in\{\ell\in 1\ldots N\backslash\mu_{\tau_k(2)k}\leq\mu_{\tau_\ell(2)\ell}<\mu_{ij}\}$. Consequently, $f(\mu_{ij})>f(\text{lev}_{1k})$.

We proved that $\text{lev}_{1k}<\mu_{ij}\Rightarrow f(\text{lev}_{1k})<f(\mu_{ij})$. On the other hand, we also proved that $\text{lev}_{1k}\geq\mu_{ij}\Rightarrow f(\text{lev}_{1k}\geq f(\mu_{ij})$. In other words, $f(\text{lev}_{1k})<f(\mu_{ij})\Rightarrow\text{lev}_{1k}<\mu_{ij}$. In conclusion, $\text{lev}_{1k}<\mu_{ij}\Leftrightarrow f(\text{lev}_{1k})<f(\mu_{ij})$.

Now, the proof can be achieved. Let j be an element of $1\ldots N$. Eq. (A16) and the monotony of f give $\forall k\in 1\ldots C-1$, $f(\mu_{\tau_j(k)j}\geq f(\mu_{\tau_j(k+1)j})$. In other words, according to Eq. (A14), $\forall k\in 1\ldots C-1$, $\mu^*_{\tau_j(k)j}\geq\mu^*_{\tau_j(k+1)j}$. Definition 5 concerning levels thus gives $\forall k\in 0\ldots C-1$, $\text{lev}^*_{kj}=\mu^*_{\tau_j(k+1)j}$. Using Eqs. (A14) and (A17), we finally obtain the following result [see Eq. (A25)]:

$$\forall j\in 1\ldots N,\ \forall k\in 0\ldots C-1,\ \text{lev}^*_{kj}=\mu^*_{\tau_j(k+1)j}=f[\mu_{\tau_j(k+1)j}]$$
$$=f(\text{lev}_{kj}) \quad (A25)$$

Let i be an element of $1\ldots C$ and j an element of $1\ldots N$. Successively applying Eqs. (A15), (A13), (A25), (A18), (A12), and (A14), we have [see Eqs. (A26) and (A27)]:

$$(\mu^*)^*_{ij}=f^*(\mu^*_{ij})=|\{\ell\in 1\ldots N\backslash\text{lev}^*_{1\ell}<\mu^*_{ij}\}|/N$$
$$=|\{\ell\in 1\ldots N\backslash f(\text{lev}_{1\ell})<f(\mu_{ij})\}|/N \quad (A26)$$

$$(\mu^*)^*_{ij}=|\{\ell\in 1\ldots N\backslash\text{lev}_{1\ell}<\mu_{ij}\}|/N=f(\mu_{ij})=\mu^*_{ij} \quad (A27)$$

We thus have $(\mu^*)^*=\mu^*$.

APPENDIX C (PROOF OF PROPOSITION 2)

Let $(p_i)_{i\in 1\ldots C}$ be a decreasing sequence of elements of [0,1] and let $(c_i)_{i\in 1\ldots C}$ be the sequence of elements of [0,1] defined by $c_1=p_1-p_2$ and $\forall i\in 2\ldots C$, $c_i=0$.

Proving Proposition 2 implies proving that there is a basic probability function m, defined on the set of subsets of $1\ldots C+1$, such that «$\forall i\in 1\ldots C$, $Pl(\{i\})=p_i$ and $Cr(\{i\})=c_i$», where Pl and Cr are, respectively, the plausibility and credibility measures associated with m. Let us assume that [see Eq. (A28)]:

$$\forall i\in 1\ldots C-1,\ m(1\ldots i)=p_i-p_{i+1} \quad \text{and}$$
$$m(1\ldots C)=p_C \quad \text{and} \quad m(\{C+1\})=1-p_1 \quad (A28)$$

For any other subset A of $1\ldots C+1$, let us state $m(A)=0$. We thus define a function m from the set of the subsets of $1\ldots C+1$ into [0,1]. We have [see Eqs. (A29), (A30), (A31), and (A32)]:

$$m(\varnothing)=0 \quad (A29)$$

$$\textstyle\sum_{A\subset 1\ldots C+1}m(A)=m(\{C+1\})+\sum_{i\in 1\ldots C}m1\ldots i)=(1-p_1)$$
$$+[\textstyle\sum_{i\in 1\ldots C-1}(p_i-p_{i+1})]+p_C=1 \quad (A30)$$

$$\forall i\in 1\ldots C,\ \textstyle\sum_{A\subset 1\ldots C+1\backslash A\cap(i)\neq\varnothing}m(A)=\sum_{j\in i\ldots C}m(1\ldots j)$$
$$=[\textstyle\sum_{j\in i\ldots C-1}(p_j-p_{j+1})]+p_C$$
$$=p_i \quad (A31)$$

$$\textstyle\sum_{A\subset 1\ldots C+1\backslash A\subset\{1\}}m(A)=m(\{1\})=p_1-p_2=c_1 \quad \text{and}$$
$$\forall i\in 2\ldots C,\ \textstyle\sum_{A\subset 1\ldots C+1\backslash A\subset\{i\}}m(A)=0=c_i \quad (A32)$$

Thus m perfectly answers the problem posed.

Note:

Let us assume that the elements of E actually distribute themselves in a "crisp" way among the C classes considered. In practice, only a noisy set of data is likely to be available. So, even in that case, the existence of elements reasonably not assignable to any of the C classes is to be expected. One mean of answering this problem consists in introducing a $(C+1)^{th}$ class, in charge of collecting the noise (Dave, 1992). This is what is done here: m is defined on the set of subsets of $1\ldots C+1$, not of $1\ldots C$. Resorting to a "noise" class is justified by the not very constraining character of the fuzzy partition definition (Definition 2). For example, we may easily have $\sum_{i\in 1\ldots C}p_i<1$. In that case, no basic probability function defined on the set of subsets of $1\ldots C$ exists that $\forall i\in 1\ldots C$, $Pl(\{i\})=p_i$.

Only one class is credible (or at least, can be): the most plausible one. It is possible to reinforce the credibility of other classes? Let us redefine for example $(c_i)_{i\in 1\ldots C}$ in the following way [see Eq. (A33)]:

$$c_1=p_1-p_2 \quad \text{and}$$
$$\forall i\in 2\ldots C-1,\ c_i=\min\{1-p_1,c_{i-1},p_i-p_{i+1}\}$$
$$\text{and} \quad c_C=\min\{1-p_1,c_{C-1},p_C\} \quad (A33)$$

This new sequence seems to correspond to a "minimum" reinforcement of credibilities. But if it is adopted, no basic probability function m, defined on the set of subsets of $1\ldots C+1$, is able to vouch for $\forall i\in 1\ldots C$, $Pl(\{i\})=p_i$ and $Cr(\{i\})=c_i$.

REFERENCES

Andréfouët, S., Claereboudt, M., Matsakis, P., Pagès, J., and Dufour, P. (in press), Typology of atolls rims in Tuamotu archipelago (French Polynesia) at landscape scale using SPOT-HRV images. *Int. J. Remote Sensing.*

Andréfouët, S., and Roux, L. (1998), Characterization of ecotones using membership degrees computed with a fuzzy classifier. *Int. J Remote Sens.* 19:3205–3211.

Andréfouët, S., Roux, L., Chancerelle, Y., and Bonneville, A. (2000), A fuzzy possibilistic scheme of study for objects with indeterminate boundaries: Application to French Polynesian reefscapes. *IEEE Trans. Geosci. Remote Sens.* 38:257–270.

Barni, M., Cappellini, V., and Mecocci, A. (1996), Comments on "A possibilistic approach to clustering." *IEEE Trans. Fuzzy, Systems* 4:393–396.

Bastin, L. (1997), Comparison of fuzzy C-means classification, linear mixture modelling and MLC probabilities as tools for unmixing course pixels. *Int. J. Remote Sens.* 18:3629–3648.

Battistini, R., et al. (1975), Eléments de terminologie récifale indopacifique. *Thétys* 7:1–111.

Bezdek, J. C. (1981), *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York.

Bezdek, J. C., and Pal, S. K. (1992), *Fuzzy Models for Pattern Recognition*. IEEE Press, New York.

Burrough, P. A., Gaans, P. F. M., and Hootsmans, R. (1997), Continuous classification in soil survey: Spatial correlation, confusion and boundaries. *Geoderma* 77:115–135.

Chuvieco, E., and Congalton, R. G. (1988), Using clustering analysis to improve the selection of training statistics in classifying remotely sensed data. *Photog. Eng. Remote Sens.* 54:1275–1281.

Congalton, R. G. (1991), A review of assessing the accuracy of classification of remotely sensed data. *Remote Sens. Environ.* 37:35–46.

Congalton, R., and Green, K. (1999), *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices.* Lewis Publishers, New York.

Dave, R. N. (1992), Characterization and detection of noise in clustering. *Pattern Recognition Letters* 12:657–664.

Deer, P. J. (1996), The application of fuzzy sets theory to image processing and remote sensing. In *Proceedings of 8th Australasian Remote Sensing Conference*, Canberra, Australia, pp. 152–168.

Dempster, A. P. (1967), Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics* 38:325–339.

Desachy, J., Roux, L., and Zahzah, E. (1996), Numeric and symbolic data fusion: A soft computing approach to remote sensing images analysis. *Pattern Recognition Letters* 17: 1361–1378.

Dubois, D., and Prade, H. (1999), *Fundamentals of Fuzzy Sets*. Kluwer Academic Publishers.

Finn, J. T. (1993), Use of the average mutual information index in evaluating classification error and consistency. *Int. J. Geogr. Inf. Syst.* 7:349–366.

Foody, J. T. (1995), Cross-entropy for the evaluation of the accuracy of a fuzzy land cover classification with fuzzy ground data. *ISPRS J. Photogramm. Remote Sens.* 50:2–12.

Foody, G. M. (1996), Approaches for the production and evaluation of fuzzy land cover classifications from remotely sensed data. *Int. J Remote Sens.* 17:1317–1340.

Foody, G. M. (2000), Estimation of sub-pixel land cover composition in the presence of untrained classes. *Computer and Geosciences* 26:469–478.

Foody, G. M., and Cox, D. P (1994), Sub-pixel land-cover composition estimation using a linear mixture model and fuzzy membership functions. *Int. J. Remote Sens.* 15:619–631.

Foody, G. M., and Trodd, N. M. (1993), Non-classifactory analysis and representation of heathland vegetation from remotely sensed imagery. *GeoJournal* 29:343–350.

Gath, I., and Geva, A. B. (1989), Unsupervised optimal fuzzy clustering. *IEEE Trans. PAMI* 11:773–781.

Gong, P., and Howarth, P. J. (1990), An assessment of some factors influencing multispectral land-cover classification. *Photogramm. Eng. Remote Sens.* 56:597–603.

Gopal, S., and Woodcock, C. (1994), Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogramm. Eng. Remote Sens.* 60:181–188.

Gutafson, E. E., and Kessel, W. C. (1979), Fuzzy clustering with a fuzzy covariance matrix. In *Proceedings of IEEEC CDC*, San Diego, CA, USA, pp. 761–766.

Hammond, T. O., and Verbyla, D. L. (1996), Optimistic bias in classification accuracy assessment. *Int. J. Remote Sens.* 17:1261–1266.

Harris, R. (1985), Contextual classification post-processing of Landsat data using a probabilistic relaxation model. *Int. J. Remote Sens.* 6:847–866.

Higashi, M., and Klir, G. J. (1983), On the notion of distance representing information closeness: Possibility and probability distributions. *Int. J. General Systems* 9:103–115.

Intes, A., Caillart, B., Charpy-Roubaud, C.-J., Charpy, L., Harmelin-Vivien, M., Galzin, R., and Morize, E. (1995), Tikehau: An atoll of the Tuamotu Archipelago. *Atoll Research Bulletin* 415.

Key, J. R., Maslanik, A., and Barry, R. G. (1989), Cloud classification from satellite data using a fuzzy sets algorithm: A polar example. *Int. J. Remote Sens.* 10:1823–1842.

Klir, G. J., and Yuan, B. (1995), *Fuzzy Sets and Fuzzy Logic, Theory and Applications.* Prentice Hall, Upper Saddle River, NJ.

Krishnapuram, R., and Keller, J. M. (1993), A possibilistic approach to clustering. *IEEE Trans. Fuzzy Systems* 1:98–110.

Krishnapuram, R., and Keller, J. M. (1996), The possibilistic C-Means algorithm: Insights and recommendations. *IEEE Trans. Fuzzy Systems* 4:385–393.

Lee, T., Richards, J. A., and Swain, P. H. (1987), Probabilistic and evidential approaches for multi-source data analysis. *IEEE Trans. Geosci. Remote Sens.* 25:283–292.

Ma, Z., and Redmond, R. L. (1995), Tau coefficients for accuracy assessment of classification of remote sensing data. *Photogramm. Eng. Remote Sens.* 61:435–439.

Maselli, R., Conese, C., and Petkov, L. (1994), Use of entropy for the estimation and graphical representation of the accuracy of maximum likelihood classifications. *ISPRS J. Photogramm. Rem. Sens.* 49:13–20.

Matsakis, P. (1998), *Relations Spatiales Structurelles et Interpretation D'Images*. PhD dissertation, Université de Toulouse III, France.

Palubinskas, G., Lucas, R., Foody, G. M., and Curran, P. J. (1995), An evaluation of fuzzy texture-based classification approaches for mapping regenerating tropical forest classes from Landsat TM. *Int. J. Remote Sens.* 16:747–759.

Peddle, D. R. (1995), Knowledge formulation for supervised evidential classification. *Photogramm. Engin. Remote Sens.* 61:409–417.

Raffy, M. (1997), Quantification of surfaces and spatial resolution: The necessity of fuzzy sets concepts. *Remote Sensing Reviews* 15:257–264.

Ruspini, E. H. (1969), A new approach to clustering. *Information and Control* 15:22–23.

Shafer, G. (1976), *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ.

Simpson, J. J., and Keller, R. H. (1995), An improved fuzzy logic segmentation of sea ice, clouds, and ocean in remotely sensed arctic imagery. *Remote Sens. Envir.* 54:290–312.

Smith, G. R., Woodward, J. C., Heywood, D. I., and Gibbard, P. L. (2000), Interpreting Pleistocene glacial features from SPOT HRV data using fuzzy techniques. *Computers and Geosciences* 26:479–490.

Stehman, S. V. (1997), Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ.* 62:77–89.

Stehman, S. V. (2000), Practical implications of design-based sampling inference for thematic map accuracy assessment. *Remote Sens. Environ.* 72:35–45.

Stehman, S. V., and Czaplewski, R. L. (1998), Design and analysis for thematic map accuracy assessment: Fundamental principles. *Remote Sens. Environ.* 64:331–344.

Story, M., and Congalton, R. G. (1986), Accuracy assessment: A user's perspective. *Photogramm. Eng. Remote Sens.* 52:397–399.

Wang, F. (1990), Improving remote sensing image analysis through fuzzy information representation. *Photogramm Eng. Remote Sens.* 56:1163–1169.

Zadeh, L. A. (1965), Fuzzy sets. *Information and Control* 8:338–353.

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

Zhang, J., and Kirby, R. P. (1997), An evaluation of fuzzy approaches to mapping land cover from aerial photographs. *ISPRS J. Photogramm. Remote Sens.* 52:193–201.

Zhu, A. X. (1997), Measuring uncertainty in class assignment for natural resource maps under fuzzy logic. *Photogramm. Eng. Remote Sens.* 63:1195–1202.

Zhuang, X., Engel, B. A., Xiong, X., and Johannsen, C. J. (1995). Analysis of classification results of remotely sensed data and evaluation of classification algorithms. *Photogramm Eng. Remote Sens.* 61:427–433.