Accuracy measures for fuzzy classification in remote sensing

Daniel Gómez
Faculty of Statistics
Complutense University
Madrid, Spain
dagomez@estad.ucm.es

Javier Montero
Faculty of Mathematics
Complutense University
Madrid, Spain
monty@mat.ucm.es

Greg Biging
College of Natural Resources
University of California
Berkeley, California, U.S.A.
biging@nature.berkeley.edu

Abstract

Over the last decades, many fuzzy classification algorithms have been proposed for image classification, and in particular to classify those images obtained by remote sensing. But relatively little effort has been done to evaluate goodness or effectiveness of such algorithms. Such a problem is most of the times solved by means of a subjective evaluation, meanwhile in the crisp case quality evaluation can be based upon an error matrix, in which the reference data set (the expert classification) and crisp classifiers data set are been compared using specific accuracy measures. In this paper, some of these measures are translated into the fuzzy case, so that more general accuracy measures between fuzzy classifiers and the reference data set can be considered.

Keywords: Accuracy assessment, remote sensing, fuzzy classification.

1 Introduction

The evaluation of effectiveness of classification algorithms is always a must, being itself a part of the classification process: the expert classification can not be avoided meanwhile behavior of a classifier has not been properly tested. In the crisp framework, evaluation of a classifier can be made by comparing pixel by pixel if the classes proposed by the classifier and the expert are the same or not. A good classifier will show a high percentage of coincidences with the expert classification. Remote sensing literature distinguishes three different steps within this accuracy assessment (Stehman et al. [19]): sampling design, measurement design (usually requiring an expert) and data analysis. But standard classical approaches can not be directly translated into very complex images, like the ones considered in remote sensing, where regions may show no particular pattern and large transition mixture zones. In this paper, we will generalize some particular crisp models so that accuracy assessment can be applied in case reference data set is fuzzy.

In the crisp framework, an error matrix is been built from the data reference set (classification given by the expert for some pixels) and the classification given by the classifier. The diagonal of this matrix represents the set of pixels in which the classifier and the data reference set coincide. However, as pointed out by Congalton and Green [4], standard accuracy assessment based on the error matrix can not be directly applied when dealing with fuzzy classifiers or when the expert opinion is given in fuzzy terms.

In order to be able to assess the accuracy of a fuzzy classification, some approaches consider measures based on a linguistic scale of accuracy (see, e.g., [7] and [12] for two different approaches). Alternative approaches also for fuzzy classification are based on a fuzzy error matrix that generalizes the classical crisp error matrix. In particular, Binaghi [3] considers (for each pixel) the degree to which a pixel has been classified in class i by the classifier and in class j by the expert (by means of the *min* operator). A fuzzy error matrix is then obtained from this information. After that the overall, producer and user accuracy for each pixel is obtained normalized by the sum of membership degrees of the expert opinion (see [3]).

With the same aim but with a different point of view, in this work we build a new disagreement measure that takes into account decision maker preferences. Such a measure generalizes the classical measure for crisp experts and crisp classifiers in case all errors are considered equally important by the decision maker, allowing fuzzy experts and fuzzy classifiers, without even imposing a Ruspini's partition as the basic classification system (see [2]).

The paper is organized as follows: first of all (section 2) we review main classical accuracy measures (overall, producer and user measures). In section 3 a new family of disagreement measures for fuzzy classification is proposed, taking into account decision maker preferences, so those classical accuracy measures can be translated into the fuzzy case (section 4). In section 5 we present two possible approaches for the construction of the matrix of weights, followed by some final remarks (section 6).

2 The crisp case

Let us consider a fixed remote sensing image, divided into a set of pixels P, with $T \subset P$ the family of pixels to be tested. Let $A_1, \ldots A_k$ be the set of crisp classes under consideration. The error matrix N is then defined as a frequency matrix, where each element n_{ij} represents the number of pixels that the expert classified a pixel in A_i but the classifier did in A_j.

Definition 2.1 *Given the error matrix $N = (n_{ij})$, the producer accuracy for the class A_i is defined as*

$$\frac{n_{ii}}{\sum_{j=1}^{k} n_{ji}}$$

Definition 2.2 *Given the error matrix $N = (n_{ij})$, the user accuracy for the class A_i can be defined as*

$$\frac{n_{ii}}{\sum_{j=1}^{k} n_{ij}}$$

The user accuracy of a class A_i estimates the probability of right assessment when the user (the classifier) proposes class A_i, while the producer accuracy of a class A_j estimates the probability that a pixel of class A_j is identified by the classifier. The overall accuracy will estimate the probability p_o of right classification (same opinion for both expert and classifier), and can be therefore used to get a global comparison between different classifiers.

Definition 2.3 *Given the error matrix $N = (n_{ij})$, the overall accuracy is defined as*

$$\hat{p}_o = \frac{\sum_{i=1}^{k} n_{ii}}{|T|}$$

being $|T|$ the number of pixels we are testing.

Proposition 2.1 *Assuming independence of right classification between different pixels, the observed overall accuracy $\hat{p}_o$ has a Normal limit distribution, so for a large enough sample its distribution can be approached by a Normal distribution with expected value the probability p_o of right classification and variance*

$$\frac{p_o(1 - p_o)}{|T|}$$

Proof: Straightforward from the Central Limit theorem.

In this way we can easily obtain confidence intervals for right classification.

Example 2.1 *Let us consider the example given in [4],where three crisp classes A_1, A_2*

and A_3 were considered: Forest, Wetland and the Urban Areas, respectively. The producer's accuracy and the user's accuracy for each class can be therefore obtained as the ratio between the diagonal element (number of pixels well classified) and the total in that column or row. If the error matrix is

$$M = \begin{pmatrix} 23 & 9 & 6 \\ 3 & 18 & 5 \\ 4 & 3 & 29 \end{pmatrix}$$

we obtain the accuracy measures given in table 2.1. It can be checked that overall accuracy is $70/100 = 70\%$, *in such a way that* $[0.61, 0.79]$ *is an interval confidence for* p_o *at* 0.05 *significance level.*

Class	Producer	User
Forest	$\frac{23}{30} = 77\%$	$\frac{23}{38} = 61\%$
Wetland	$\frac{18}{30} = 60\%$	$\frac{18}{26} = 69\%$
Urban	$\frac{29}{40} = 73\%$	$\frac{29}{36} = 81\%$

3 The mathematical model

Traditional remote sensing accuracy assessment assumes crisp classes, in such a way that agreement between the classifier (C) and the expert (E) is modelled according to a two-valued model: perfect agreement (0) or total disagreement (1). This restriction implies a strong oversimplification of reality, since the continuum of variation in many landscapes will be difficult to be properly represented. In order to address this issue, we will propose a continuous error measure that summarizes the differences between a crisp reference data set (most expert are still crisp) and a fuzzy classifier.

From a mathematical point of view, a pixel being classified by the expert (E) or by the classifier (C) as the crisp class A_i, can be modelled as a k dimensional vector, k being the number of different classes under consideration, in such a way that all coordinates take value 0 except the i-th coordinate, which takes value 1. A crisp classifier or a crisp reference data set can be then considered as a function assigning to each pixel p a vector in

$$\left\{ x \in \{0,1\}^k / \sum_i x_i = 1 \right\}$$

Hence, in case our k classes are fuzzy and we assume that assignments is made in terms of a Ruspini's partition [17] (see also [2]), both classifier C and expert E will be defined as mappings

$$E : P \longrightarrow \left\{ x \in [0,1]^k / \sum_i x_i = 1 \right\}$$

and

$$C : P \longrightarrow \left\{ x \in [0,1]^k / \sum_i x_i = 1 \right\}$$

Hence, disagreement between classifier and expert can be measured by means of a distance in such a k dimensional space, assigning to each pixel a real value

$$D : T \longrightarrow \Re$$

where T is the subset of pixels (or polygons) selected for the accuracy assessment (an alternative valuation set can be given in terms of linguistic terms, see [20]). A standard measure for disagreement between classifier and expert is given below.

Definition 3.1 *Given a remote sensing image P and a family of classes $A_1 \ldots A_k$ under consideration, E an expert function and C a classifier function, then the error of the pixel p given by the classifier C with reference data set E is defined as:*

$$D_f(E(p), C(p), p) = \sum_{i=1}^{k} |E(p)_i - C(p)_i|$$

where $E(p)_i$ is the i-th coordinate of $E(p)$ and $C(p)_i$ is the i-th coordinate of the $C(p)$.

Notice that this definition does not assume a crisp expert neither a crisp classifier. If classes are fuzzy in nature, mathematical models should acknowledge such a situation, and the expert should in terms of fuzzy classes. But in practice we find that most expert classifiers are crisp, perhaps because of the complexity in defining all parameters of a fuzzy classification. In this case, the above distance is unrealistic, giving excessive distance values to transition zones. So, we propose the following modified definition.

Definition 3.2 *Given a remote sensing image P and a family of classes $A_1 \dots A_k$ under consideration, E a crisp expert function and C a fuzzy classifier function, then the error of the pixel p given by the classifier C with reference data set E is defined as:*

$$D_c(E(p), C(p), p) = \sum_{i=1}^{k} w_{ij} |E(p)_i - C(p)_i|$$

where $E(p)_i$ is the i-th coordinate of $E(p)$, $C(p)_i$ is the i-th coordinate of the $C(p)$, j is represents the class to which p is assigned ($E(p)_j = 1$) and each $w_{ij} \in \Re^+$ represents the importance of the error when a unit sampling that belongs to class j is classified as class i.

Notice that whenever both classifier and reference data are crisp, the above error function can be viewed as a weighted error function that takes value w_{ij} if the expert $E(p)$ has classified the pixel as A_j and the classifier has classified the same pixel as A_i. Moreover, if we take $w_{ij} = 1, \forall j$, then the disagreement measure is just the classical one. In a more general approach, these weights can depend on the maximum value E takes, or any other dispersion measure for E.

Notice also that our approach does not impose any particular structure on the classification system (as pointed out in [1, 2], fuzzy partitions in the sense of Ruspini as quite often unrealistic).

4 Accuracy measures

Last definition of disagreement between expert and classifier allows to generalize the concept of producer accuracy, user accuracy and overall accuracy for the case in which not all errors are equal, and for the case in which the classifier is fuzzy but the expert is crisp.

Definition 4.1 *Given a remote sensing image P with $A_1, \dots A_k$ the family of available classes, an accuracy data set $T \subset P$ with cardinality $|T|$, a crisp reference data set $E(p)$ for all $p \in T$, a fuzzy classifier C and an error $D_c(E, C, p)$ for all $p \in T$, the producer's accuracy for a class A_i is defined as*

$$\frac{\sum_{p \in P/E(p)_i = 1} (1 - D_c(E, C, p))}{n_i}$$

where $n_i = card\{p \in P/E(p)_i = 1\}$ is the number of pixels assigned by the expert to class A_i.

Notice that $1 - D_c(E, C, p)$ can be viewed as the agreement between the crisp expert and the fuzzy classifier for the classification of pixel p. Hence, from now on we will denote

$$A_c(E, C, p) = 1 - D_c(E, C, p)$$

the agreement degree between the crisp expert and the fuzzy classifier.

Example 4.1 *Let us continue the previous example, and let us assume the following weight matrix for error:*

$$W = \begin{pmatrix} 0 & 2/3 & 2/3 \\ 1 & 0 & 4/3 \\ 1 & 4/3 & 0 \end{pmatrix}$$

Lets also assume the following reference data set versus classifier matrix:

	Forest	*Wetland*	*Urban*
Forest	*23*	*9*	*6*
Wetland	*3*	*18*	*5*
Urban	*4*	*3*	*29*
Total	*30*	*30*	*40*

Then the producer accuracy for Wetland class can be determined as

$$\frac{9 * (\frac{1}{3}) + 18 + 3 * (\frac{-1}{3})}{30} = \frac{20}{30} = 66\%$$

It should be noticed that under the classical approach, the producer accuracy for Wetland class was $\frac{18}{30}$, obtained when considering equally important all errors (see [4]).

Definition 4.2 *Given a remote sensing image P with classes $A_1, \ldots A_k$, the accuracy data set $T \subset P$, the crisp reference data set E, the fuzzy classifier C and the error $D_c(E, C, p)$ for all $p \in T$, the user's accuracy for a class A_i is defined as:*

$$\frac{\sum_{p \in P/Max\{C(p)\}=C(p)_i} A_c(E, C, p)}{n_i}$$

where $n_i = card\{p \in P/Max\{C(p)\} = C(p)_i\}$.

Definition 4.3 *Given a remote sensing image P with classes $A_1, \ldots A_k$, the accuracy data set $T \subset P$, the reference data set E, the classifier C, and the error $D_c(E, C, p)$ for all $p \in T$, the overall accuracy is defined as:*

$$\hat{p}_o = \frac{\sum_{p \in P} A_c(E, C, p)}{|T|}$$

Notice from the previous examples that the coefficients of the weighted matrix must sum $k^2 - k$ as in the classical way. This condition should be imposed in order to compare the overall accuracy from different classifications.

In case in which the classifier is crisp and all the weights are equal, all three the accuracy measures (producer,user and overall) will be the classical measure.

Proposition 4.1 *Given a digital image P with accuracy training site $T \subset P$, E being a crisp expert and C a fuzzy classifier, if p_0 represents the expected value for a perfect matching between the expert and the classifier, then the overall accuracy of definition 4.3 can be approximated, for a large enough training site and assuming independent behavior between pixels, by a Normal distribution with expectation p_0 and variance*

$$\frac{p_0(1 - p_0)}{|T|}$$

Proof: once C and E are fixed, $D(C, E, .)$ or $A = 1 - D(C, E, .)$ can be viewed as random variables taking a real value for each $p \in P$. If we denote by μ the expected value of $D(C, E, .)$ then the expected value of $A(C, E, .)$ will be $1 - \mu$.

In the classical case, for example, $D(C, E, .) \in \{0, 1\}$ in such a way that $D(C, E, .)$ can be viewed as a Bernoulli distribution and μ represents the probability of disagreement $1 - p_o$. In the general case, $\mu = 1 - p_o$ represents the expected error frequency between the classifier C and the reference data E. Analogously, $1 - \mu = p_o$ will represent the expected agreement frequency.

Hence, if

$$\hat{p}_o = \sum_{p \in T} \frac{A(E, C, p)}{|T|}$$

represents a sample mean of $|T|$ independent observations of the random variable $\{A(E, C, \dot{)}\}$, the Central Limit theorem assures the above result when $|T|$ goes to infinity.

Example 4.2 *Let E be the crisp reference data set given by an expert, let C_1 be a classifier (shown in table 1) and let C_2 be the crisp classifier that assigns to each pixel the class with highest degree of membership according to C_1. Suppose that all errors have the same importance ($w_{ij} = 1$ if $i \neq j$ and 0 otherwise).*

Table 1: Error table for fuzzy classifier C_1

Frequency	Data	Classifier
8	$(1, 0, 0)$	$(0.8, 0.1, 0.1)$
17	$(0, 1, 0)$	$(0.6, 0.4, 0)$
14	$(0, 1, 0)$	$(0.5, 0.4, 0.1)$
21	$(0, 1, 0)$	$(0.1, 0.4, 0.5)$
5	$(0, 0, 1)$	$(0, 0.4, 0.6)$
5	$(0, 1, 0)$	$(0, 1, 0)$
10	$(0, 0, 1)$	$(0.5, 0.5, 0)$
12	$(0, 0, 1)$	$(0, 0.1, 0.9)$
8	$(1, 0, 0)$	$(0.7, 0.3, 0)$
10	$(1, 0, 0)$	$(0.6, 0, 0.4)$

Errors associated to C_1 and C_2 can be then computed, obtaining a final accuracy of 61% for C_1 and 48% for C_2. But notice that under a standard approach accuracy assigned to both classifiers will be the same, despite we realize it is unrealistic. In fact, from a statistical *point of view (signification level $\alpha < 0.0001$) the accuracy of C_1 is greater than the C_2.*

It should be pointed out that the fuzzy error matrix proposed by Binaghi [3] gives the same overall accuracy for the classifier in the above example. In fact, it can be proved that if the reference data is crisp and the fuzzy classification is a Ruspini partition [17], then both approaches are equivalent for the overall and producer accuracies whenever $w_{ij} = 1, \forall i \neq j$ and $w_{ii} = 0$ otherwise. Notice that Binaghi's approach only takes into account one coordinate (the one chosen by the crisp expert), while our approach is based upon the behavior of the remaining coordinates. Future research should explore the possibility of a mixed approach.

Example 4.3 *Let us suppose a 100 pixels image, where a crisp expert proposes* $(1, 0, 0, 0)$ *for all pixels. Let us assume two classifiers, the first one assigning* $(0.6, 0.1, 0.1, 0)$ *to every pixel and the second one assigning* $(0.6, 0.4, 0.1, 0)$ *to every pixel. The approach given by [3] does not make difference between these two classifiers, assigning in both cases* 60% *to the overall accuracy. But it is clear that those situations are not equivalent. In fact, our approach (taking* $w_{ij} = 1, \forall i \neq j$ *and 0 otherwise) assigns* 80% *for the first classifier, and* 50% *for the second classifier, which seems rather more appropriate.*

If the classification is not given in terms of a Ruspini partition both approaches can be very different. For example, the Binaghi approach for the overall measure considers that $E(p) = (1, 0, 0)$ and $C(p) = (1, 0.2, 0)$ represents a perfect agreement, which in our opinion is not appropriate (the agreement for this case gets a 0.8 value under our approach).

5 Obtaining weights

As it can be perceived from the disagreement measure given in definition 3.2, the weights that represent the importance of the different errors play an extremely important role. In the following two subsection we propose two alternative techniques in order to determine the importance of errors. The first one is based on a Multicriteria Decision Making approach, and the second one is based on the distance between fuzzy sets.

5.1 A multicriteria approach

It is a standard assumption in accuracy assessment that all errors are equally important. Introducing weights allows to take into account the opinion of the expert and the main objectives of the study (by means of the relative importance of errors). Therefore, a different weight matrix for each measure (producer, user and overall) is required.

There are different ways to obtain an appropriate weight matrix. If we are building the producer weight matrix, it can be imposed that the all coefficients sum $k(k-1)$, as in the classical approach.

From a multicriteria point of view there are several approaches available in order to determine weights (see, e.g., [13, 16, 18] and [10]). For example, if we want to apply Saaty's approach, first we obtain the Saaty matrix (asking the decision maker to compare each pair of errors in order to define $w_{i,j}$ and $w_{i',j'}$, by means of linguistic values: 1 Same Importance, 3 Moderate Importance, 5 Strong Importance, 7 Show Importance and 9 Extreme Importance, with 2,4,6,8 intermediate values). Once Saaty matrix has been defined, the weights are computed as the eigenvector associated to the maximum eigenvalue (see [18]).

5.2 Fuzzy distances

The classes of the remote sensing problem can be described in fuzzy terms by means of the spectral features of each class. Consequently, for each class A_i, and for each band B_r, we have the functions $\mu_{A_k}^{B_r}$. Taking into account that for each A_j we have $\left(\mu_{A_j}^{B1}, \ldots, \mu_{A_j}^{Bm}\right)$, a distance function (see for example [6, 8, 15]) between fuzzy sets could be applied for each pair of classes, $D(A_i, A_j) = d_{ij}$. On one hand, small distances d_{ij} represent a hight similitude between classes, so error will not be relevant. On the other hand, high values of d_{ij} represent different classes or big errors. Taking into account this information, the weights matrix could be calculated proportionally to distance values.

6 Final remarks

We have emphasized in this paper that we do need to develop accuracy measures for fuzzy classification. Our proposal pursues an evaluation of accuracy of fuzzy classifications, so we can properly deal with those large transition zones so frequently present in remote sensing. This is being done by means of a producer, an user and an overall accuracy measure in which the classifier is fuzzy but the reference data is crisp. Of course the number of classes could be in practice very high, much more than the three classes examples considered in this paper for a pedagogical purpose. But we do not see in principle any suitability difficulty in producing a continuum of different membership degrees resulting in a huge amount of rows having a very small frequency, except for its manageability (developing useful representation techniques is a key problem when dealing with complex information, as pointed out for example in [9]). A real case study is under progress.

This paper mainly addresses the assessment problem when the classifier is fuzzy and the reference data is crisp. In fact, all reference data you can find in remote sensing literature are crisp. But as is pointed in [12] more efforts are needed in order to built fuzzy reference data sets that gather the fuzzy expert's opinions. For this case, is important to note that the disagreement measure here proposed can be easily generalized in order to be able to assess the accuracy when the classifier and the reference data be fuzzy.

Although fuzzy reference sets will be consider in a future stage, it should be pointed out that the producer or the user accuracy have proper sense when dealing with crisp reference data or crisp classifier respectively. In a complete fuzzy framework we have to be extremely careful. For example, if we consider an image in which most of the pixels have been classified as

$$C(p) = (0.8, 0.1, 0.1)$$

by the classifier and $E(p) = (1, 0, 0)$ by the expert, our approach gives an user accuracy of 80%, which seems more adequate than the 100% given in Binaghi [3]. Nevertheless, if we have

$$C(p) = (0.34, 0.33, 0.33)$$

then this pixel p should not be considered in order to measure the user accuracy for the class A_1, since assigning this pixel to the class A_1 is not significatively supported with respect with the other classes. In this situation, user accuracy may not be justified, due to the high degree of fuzziness, which makes unrealistic the association to a unique class. In order to measure the degree of fuzziness, some authors (see, e.g., [7]) consider entropy measures, which can be therefore considered in order to determine which pixels should be taken into account for the user's accuracy. This criticism does not apply to the overall accuracy measure proposed in this paper, which seems to us a more relevant index about the quality of the classification.

Moreover, whenever our fuzzy classifier defines a Ruspini fuzzy partitions, the reference data set is crisp and we take $W_{ij} = 1, \forall i \neq j$ and 0 otherwise, our approach assigns the same overall accuracy as Binaghi approach [3].

Acknowledgements

Research supported by the Government of Spain, grant number MTM2005-08982, and a Del Amo joint programme between Complutense University and the University of California.

References

[1] A. Amo, D. Gomez, J. Montero and G. Biging (2001): Relevance and redundancy in fuzzy classification systems. *Mathware and Soft Computing* 8:203-216.

[2] A. Amo, J. Montero, G. Biging and V. Cutello (2004): Fuzzy classification systems. *European Journal of Operational* 156:459-507.

[3] E. Binaghi, P.A. Brivio, P. Ghezzi and A. Rampini (1999): A fuzzy set based accuracy assessment of soft classification. *Pattern Recognition Letters* 20:935948.

[4] R.G. Congalton and K. Green (1999): *Assessing the Accuracy of Remote Sensed Data: Principles and Practices*. Lewis Publishers, London.

[5] R.G. Congalton and G. Biging (1992): A pilot study evaluating ground reference data collection efforts for use in forestry inventory. *Photogrametic Engineering and Remote Sensing* 58:1669-1671.

[6] D. Dubois and H. Prade (1997): The three semantics of fuzzy sets. *Fuzzy Sets and Systems* 90:141-150.

[7] G.M. Foody (1996): Approaches for the production and evaluation of fuzzy land cover classification from remotely sense data. *International Journal of Remote Sensing* 17:1317-1340.

[8] L. Goldfarb (1992): What is distance and why we do need the metric model for pattern learning. *Pattern Recognition* 25:431-438.

[9] D. Gómez, J. Montero, J. Yáñez, C. Poidomani: *A graph coloring algorithm approach for image segmentation*. Omega (to appear).

[10] J.G. Pachón, D. Gómez, J. Montero and J. Yánez (2003): Soft dimension theory. *Fuzzy set and Systems* 137:137-149.

[11] J.G. Pachón, D. Gómez, J. Montero and J. Yánez (2003): Searching for the dimension of binary valued preference relations. *International Journal of Approximate Reasoning* 33:133-157.

[12] M. Laba, S.K. Gregory, J. Braden, D. Ogurcak, E. Hill, E. Fegraus, J. Fiore and S.D. DeGloria (2002): Conventional and fuzzy accuracy assessment of the New York Gap Analysis Project land cover map. *Remote Sensing of Enviroment* 81:443-455.

[13] C. Macharis and J.P. Brans (1998): The GDSS Promethee procedure. *Journal of Decision Systems* 7:283-307.

[14] C.R. Rao (1965): *Linear Statistical Inference and its Applications*. Wiley, New York.

[15] A. Rosenfeld (1985): Distances between fuzzy sets. *Pattern Recognition Letters* 3:229-233.

[16] B. Roy (1990): Decision aid and decision making. *European Journal of Operational Research* 45:324-331.

[17] E.H. Ruspini (1969): A new approach to clustering. *Information and Control* 15:22-32.

[18] T.L. Saaty (1994): *Fundamentals of Decision Making with the Analytic Hierarchy Process*. RWS Publications, Pittsburgh (Revised in 2000).

[19] S.V. Stehman and R.L. Czaplewski (2003): Introduction to special issues on map accuracy. *Enviromental and Ecological Statistics* 10:301-308.

[20] C.E. Woodcock and S. Gopal (2000): Fuzzy set theory and thematics maps: accuracy assessment and area estimation. *International Journal Geographical Information Science* 14:153-172.

[21] L.A. Zadeh (1965): Fuzzy sets. *Informations Sciences* 8:338-353.