SOBRAPO

A PROBABILISTIC APPROACH APPLIED TO THE CLASSIFICATION OF COURSES BY MULTIPLE EVALUATORS

Annibal Parracho Sant'Anna[1], Helder Gomes Costa[1*],
Lívia Dias de Oliveira Nepomuceno[2] and Valdecy Pereira[1]

ABSTRACT. How to measure the perceived influence of a course on its alumni skills? This paper describes the use of CPP-TRI as a tool to face this problem. The method was applied here in the context of a M.Sc. course evaluation. Levels of impact and importance previously determined for different features provide the framework for the analysis. Classifications by different groups of evaluators are combined. Taking into account the subjectivity in the assessments, CPP-TRI treats them as realizations of random variables. The combination of the evaluations is performed by computing joint probabilities, what avoids the assignment of weights to evaluators. Interval classifications between a hostile and a benevolent limit are provided offerings the educational evaluator a deeper understanding of the results. An additional study is here performed on the classification of the features. A total of sixteen features are sorted.

Keywords: multi-attribute decision making, educational planning, probabilistic preferences composition.

1 INTRODUCTION

This paper studies the problem of combining assessments of a group of evaluators about features of a project or institution, evaluated by their impact and by their importance. By impact is meant the effectiveness of the presence of the feature and by importance the significance of such presence for the objectives of the project or institution.

Levels of impact and importance previously determined for the different features provide a framework for the analysis. The information extracted from the data about which features are more and less present in the institution or project and more and less relevant for its finalities form a basis for the managers dealing with them to decide what to improve and what to explore.

---

*Corresponding author.
[1]Escola de Engenharia, Universidade Federal Fluminense – UFF, Rua Passos da Pátria, 156, São Domingos, 24210-240 Niterói, RJ, Brasil. E-mails: annibal.parracho@gmail.com; helder.hgc@gmail.com; valdecypereira@yahoo.com.br
[2]Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (CEFET - RJ), Av. Maracanã, 229, Maracanã, 20271-110 Rio de Janeiro, RJ, Brasil. E-mail: livia@vm.uff.br

The classification procedure employs an approach to the sorting problem based on randomizing the trichotomic decision of classifying any evaluated alternative as superior, inferior or equivalent to previously determined profiles based on the different levels of impact and importance that they may have for the evaluators. The object of evaluation is finally allocated in the class for which the probabilities of receiving evaluations above and below the class profiles are nearest to each other.

This method is named CPP-Tri by Sant'Anna et al. (2015) to emphasize its use of the composition of probabilistic preferences (Sant'Anna, 2002) in a trichotomic procedure. Its first step consists of obtaining probability distributions for the profiles that define the classes in advance, as well as for the value judgments of the impact and the importance of the features by each evaluator. Probabilistic comparisons between evaluations according to each evaluator and profiles are then performed. A composition of these probabilistic preferences into global scores concludes the classification procedure.

To combine preferences of different evaluators, a method based on previous probability distributions liberates the decision-maker from the burden of assigning weights or determining threshold parameters. Instead, general assumptions on the form of the joint distribution that may be modified to take into account special information about the decision-makers, if judged necessary, are employed.

Here are examined features of a graduate course by two different groups of evaluators. The goal is that the course be classified as of high level for the impact of the features classified as of higher importance.

The next section brings a review of the relevant literature. Section 3 describes the probabilistic framework and the trichotomic classification procedure. Section 4 presents the features evaluated and the data set studied. Section 5 applies the methodology. Section 6 concludes the paper.

2 LITERATURE REVIEW

Classical texts like Arrow (1963), Roy (1968), Fishburn (1973), Saaty (1980), Changkong & Haimes (1983), Zeleny (1982) and Roy & Boyssou (1985) considered that decision involves combining multiple criteria, giving rise to the research field of Multicriteria Decision Analysis.

A basis to deal with the multicriteria decision was constituted by methods nowadays considered as classical: MAUT (Multi-Attribute Utility Theory, Keeney & Raiffa, 1976; Fishburn, 1970), AHP (Analytic Hierarchy Process), Saaty (1980), SMART (Edward & Barron, 1994, Edward, 1974), TOPSIS (Hwang & Yoon, 1981), PROMETHÉE (Brans, 1982) and the ELECTRE family.

The ELECTRE family is composed by ELECTRE I (Roy, 1968), ELECTRE II (Roy & Bertier, 1971), ELECTRE III (Roy, 1978), ELECTRE IV (Roy & Hugonnard, 1981), ELECTRE IS (Roy & Skalka, 1984) and finally, ELECTRE TRI (Mousseau et al., 1999; Yu, 1992), with the more recent variants ELECTRE TRI-C (Almeida-Dias et al., 2010) and ELECTRE TRI-nC (Almeida-Dias et al., 2012).

In recent years, the difficulties associated to uncertainty in the evaluation of the criteria have been highlighted in papers like Greco et al. (2006), Abdellaoui et al. (2007), Desimpleaire & Marchant (2007), Birnbaum (2008), Jain & Khare, 2010, Krokhmal et al., 2011, Durbach (2012), Durbach & Stewart (2009, 2012), Comes et al. (2013) and Mosadeghi et al., 2013.

Besides, developments to deal with uncertainty have been designed to address in new manners the difficulties brought by the presence of uncertainty in the measurements. This is the case of the statistical modeling of the parameter space by Lahdelma et al. (1998) and of the probabilistic transformation of the evaluations vectors by Sant'Anna (2002).

More recent examples of such developments designed without a specific original method in mind are, for instance, in Yang et al. (2006), Ben Amor et al. (2007), Greco et al. (2010), Zhang et al (2010), Boujelben et al (2011) and Bouyssou & Marchand (2013).

Other works were addressed to extend particular classical methods. Different approaches to interval numbers and fuzzy sets (Zadeh, 1965) were applied in this context. Shanian et al. (2008), employing to access the importance of the criteria the weighting method of Simos (1990) extended Electre III, Mustajoki et al. (2005) employed interval numbers in a combination of SMART with SWING (von Winterfeld & Edwards, 1986) criteria weighting method. Mahdavi et al. (2008), Dheena, & Mohanraj (2011) and Eiselt & Marianov (2014) addressed TOPSIS. Wang et al. (2008) and Fan et al. (2010) extended the use of the matrices of pairwise comparison of AHP to stochastic measures. Nefeslioglu et al. (2013) and Feizizadeh et al (2014) also took AHP as starting point. Chang & Wang (2009) and Merigó et. al. (2014) used the fuzzy approach to extend MAUT approaches. Chou & Lin (2009), van der Pas et al. (2010), Liu et al. (2011) and Perez et al. (2014) also developed MAUT models.

Dealing with the sorting problem, it may be found the work of Janssen & Nemery (2013), using interval parameters to take into account uncertainty and extending the models of Nemery & Lamboray (2008), based on PROMETHEE, and of Ishizaka et al. (2012), based on AHP. In another direction, Almeida Dias et al. (2010, 2012) developed ELECTRI-C and ELECTRE-TRI-nC to extend ELECTRE TRI. While ELECTRE-TRI places the alternatives into classes separated by boundaries determined by fictitious alternatives with vectors of performances determined according to the multiple criteria, in ELECTRE-TRI-C a vector of central performances replaces the boundary vectors. ELECTRE-TRI-nC extends ELECTRE-TRI-C, allowing for a larger number of vectors determining each class.

Different applications of multicriteria approaches to deal with the problem of evaluation of educational features are presented in Ishizaka (2012), Silva et al. (2012), Bortoluzzi et al. (2013) and Menezes & Pizzolato (2014).

3   CPP-TRI

The problem here faced is analogous to the sorting of alternatives that are evaluated under multiple criteria into classes determined by a variable number of profiles described by evaluations on the same criteria in Sant'Anna et al. (2014).

The initial problem is that of classifying a course in one of five ordered classes. What makes this problem more complex is that different evaluators of different types: students, coordinators and managers of the students in the companies they work apply different criteria associated to desired qualities of the graduates. The evaluation by a different feature from a different point of view by a different evaluator is treated as the application of a different criterion.

A second problem addressed here is that of exploring the data to classify the features. Five ordered classes are considered again.

The classification procedure taken is similar to that proposed by Almeida Dias et al. (2012). The classes are simply identified by numbers from 1 to 5 and the evaluators are asked to make their assignments according to each feature in terms of grades from 1 to 5. The final classification is made in terms of proximity of the vectors of evaluations to vectors of identifying profiles.

In the present case, instead of identifying the classes by a larger number of profiles, only one profile is used to identify each class, as in ELECTRE-TRI-C (Almeida Dias et al., 2010). This unique profile has for each of its coordinates one of the five values employed to measure the feature represented in this coordinate.

The main difference of the method herein to those of the ELECTRE-TRI family is that it employs a previous probabilistic transformation proposed in Sant'Anna (2002). The probabilistic transformation is based on assuming random disturbances affecting the evaluations.

Following the principles of classical statistical modeling, independence between the evaluations of different alternatives, and identical normal distributions for the random disturbances in the measurements are assumed. If there is knowledge advising other distributions, it can be employed without substantial changes in the procedure. Nevertheless, the influence of the form of the distribution is limited in this technique by the trichotomic classification being made by comparing only sizes of intervals that always include the tails of the distribution.

Employing these assumptions, each numerical evaluation is replaced by a probability distribution centered on it. The comparisons between observed evaluations and representative profiles are then made in probabilistic terms. The whole technique involves the following three stages:

i) Calculus of preference probabilities by each criterion

As above described, the initial numerical assessments by each criterion are treated as means of probability distributions. Assuming normality, mean and variance are enough to determine each distribution. A common variance may be estimated from the observed variance of the vector of profiles provided to represent the different classes.

ii) Calculus of probabilistic distances from evaluations to profiles

Once a probability distribution for the evaluation by each criterion is determined, the probability of each alternative being above or below each profile is computed for each criterion. Assuming

independence, the probabilities of being above and below all the profiles of a class are given by the product of the probabilities of each being above or below each profile.

iii) Allocation in the nearest class

After the probabilities of the alternative being above or below the profiles of a class are computed, it is possible to calculate the difference between these probabilities. A null value for this difference has a null probability, while positive probabilistic distances favor the classification above the class or in it and negative values, on the contrary, indicate that the alternative should not be classified above the class. This difference is calculated for all classes, and the alternative is allocated to that class for which the absolute value of the difference is the lowest.

To formally describe the technique, consider m criteria, n classes and, to simplify, only one profile for each class. For each k from 1 to m, the alternative to be classified receives a numerical evaluation a_k. For each i, from 1 to n, the i-th class is represented by the profile C_i. These representative profiles are vector of m coordinates. Let $(C_{i1}, \ldots, C_{im})$ denote the representative profile of class i. The classes are ordered:

$$\text{if } a < b, \text{ then } C_{ak} < C_{bk}. \tag{1}$$

Each a_k is thought as the mean of a normal distribution. The variances of these distributions will be given by the observed variances V_k in the sample of profile coordinates $\{C_{1k}, C_{2k}, \ldots, C_{nk}\}$. Let X_k represent a random variable with a normal distribution of mean a_k and variance V_k and let A_{ik}^+ and A_{ik}^- denote, respectively, the probability of X_k presenting a value above and below the k-th value of the profile C_i. By independence,

$$A_i^+ = \prod_k A_{ik}^+ = \prod_k P[X_k \geq C_{ik}] \tag{2}$$

and

$$A_i^- = \prod_k A_{ik}^- = \prod_k P[X_k \leq C_{ik}]. \tag{3}$$

The probabilistic score for the i-th class is given by

$$\delta_i = A_i^+ - A_i^-. \tag{4}$$

The alternative is assigned to the class i with the minimum absolute value for δ_i.

A tie may occur is in this computation. The alternative will then be allocated in the two neighbor classes, reflecting the imprecision of the assessments.

From the beginning we deal with rough approximations. The numbers that the evaluators provide are their intuitions about the center of probability distributions. To generate more information on the presence of uncertainty in the classification, the classification derived from replacing the

hypothesis of independence between the evaluations by the different evaluators may be compared to that resulting from a different assumption on correlation. For instance assuming maximal correlation between evaluations by different criteria instead of independence, A_i^+ and A_i^- will be given by

$$A_i^+ = \min\left\{A_{i1}^+, \ldots, A_{im}^+\right\} \tag{5}$$

and

$$A_i^- = \max\left\{A_{i1}^-, \ldots, A_{im}^-\right\} \tag{6}$$

Other classifications may be generated with the same aim of gathering information on the uncertainty in the final classification. Sorting processes based on more and less benevolent rules to classify above and below the representative profiles determined by fixed rates of reduction applied to the probabilities of being above or below the representative profiles may be applied.

A benevolent classification determined by the reduction percent c will place alternative A in the class $C_c(A)^+$ whose index i minimizes the absolute value of the difference $\delta_i^{c+} = A_i^+ - cA_i^-$.

Analogously, a hostile classification for the same reduction percent will place alternative A in the class $C_c(A)^-$ whose index i minimizes the absolute value of the difference $\delta_i^{c-} = cA_i^+ - A_i^-$.

4   COMPOSITION OF FEATURES EVALUATIONS

The probabilistic approach is here applied to replicate the evaluations of the M.Sc. Course of Management Systems of Universidade Federal Fluminense performed by Nepomuceno et al. (2010) and Nepomuceno & Costa (2015).

Table 1 presents the different features of the course, identified by changes in the qualifications of the students. This selection of features has an initial source in Politis & Siskos (2004) and Neves (2005). Politis & Siskos (2004) evaluated a Greek Engineering Department using perceptions from students, graduates and also from companies where they work. (Neves, 2005) adapted the features used in Politis & Siskos (2004) to evaluate a Brazilian M.Sc. Course under a strategic viewpoint. Based on these and other works indicated in Table 1, a draft of a questionnaire was produced and submitted to a group of experts with a large experience in the management of similar courses in Brazil. Features F7 to F16 were included following the analysis in Nepomuceno & Costa (2012) of the suggestions of the group of experts.

The impact of the course by different features and the importance of these features had been evaluated by a set of 21 students who took the course, four managers of the company that employs them, 20 teachers of the course and nine coordinators of courses at other universities evaluated together with the course in a Brazilian evaluation system.

The classifications were on five classes, each represented by a unique profile with all evaluations equal to each of the five values of a Likert scale. For the impact was employed a scale with values $-2, -1, 0, 1$ and 2 and for importance values from 0 to 4. The variance of 2 that characterizes samples of these five values was assumed for all the distributions.

Table 1 – Features evaluated.

	Feature	Quality enhanced	Source
F1	Employability	Suitability to new jobs	Neves (2005); Giannoulis & Ishizaka (2010); Lukman, Krajnc & Glavião (2010)
F2	Market suitability	Adjustment to desired skills	Neves (2005); Politis & Siskos (2004)
F3	Earnings	Worthiness for a high pay	Neves (2005); Politis & Siskos (2004); (Ginevicius & Gineviciene, 2009)
F4	Entrepreneurship	Willingness to take risks	Neves (2005); (Ginevicius & Gineviciene, 2009)
F5	Sociability	Easiness to associate	Neves (2005)
F6	Research ability	Dealing with academic tasks	Neves (2005)
F7	Negotiation ability	Handle convincing skills	Experts suggestion
F8	Self-esteem	Satisfaction with own image	Experts suggestion
F9	Oratory	Persuasive speaking	Experts suggestion
F10	Personal life	Success in private affairs	Experts suggestion
F11	Disinhibition	Freedom to act in public	Celik, Kandakoglu & Er (2009); Experts suggestion
F12	Fluency	Fast oral communication	Experts suggestion
F13	Critical thinking	Master evaluation skills	Experts suggestion
F14	Openness	Readiness to exposition	Experts suggestion
F15	Problem solving	Reasoning skills	Experts suggestion
F16	Special skills	Ability in technical tasks	Experts suggestion

To produce a new global evaluation of the course, the different experts' evaluations of the impact of different features were treated separately. This approach resulted in hundreds of evaluations of the course that were combined in different ways. First, all evaluations were considered as individually final. Then, a combination of the evaluations by the members of each group with a same relation to the course was performed.

All these evaluations confirmed the location of the course obtained in previous analyses by Nepomuceno et al. (2010) and Nepomuceno & Costa (2012) as "good", which corresponds to classification in the fourth of the five classes considered in ascending order. This result follows from the classification of the impact of the course in that class in 21 of 52 evaluations by managers (corresponding to 40% of this subsample), 46 of 134 evaluations by coordinators (33%), 120 of 145 evaluations by teachers (83%) and 146 of 308 evaluations by students (47%).

5   CLASSIFICATION OF THE FEATURES

A sorting of the features was then performed, considering their evaluations in terms of impact and importance. Following Kahneman (2011) preference for the contribution of outer views, this last study was based on the evaluations of the managers of the students and the coordinators of other courses only. The teachers and students of the course were not considered. Thus, 58 alternatives were evaluated, as the coordinators evaluated the 16 features on impact and importance and the managers did not evaluate features 1, 3 and 10 because they were considered unable to access the effect of the course on such personal subjects.

The ordered classes and their unique profiles have the same form of the previous analysis. The evaluations of, respectively, managers of the students and coordinators of other courses are presented in Tables 2 and 3. The letter N in a cell of these tables is used to indicate that the evaluator did not evaluate the feature.

Table 2 – Managers evaluations.

	Impact												
	F_2	F_4	F_5	F_6	F_7	F_8	F_9	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}	F_{16}
M1	2	1	2	2	1	2	1	2	2	1	2	2	2
M2	2	2	1	2	1	2	1	1	1	2	1	2	1
M3	1	1	2	1	1	1	1	1	2	1	1	1	2
M4	2	2	2	2	2	2	2	2	2	2	2	2	2
	Importance												
	F_2	F_4	F_5	F_6	F_7	F_8	F_9	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}	F_{16}
M1	4	4	4	3	3	4	3	3	4	3	3	4	4
M2	4	4	4	4	4	3	3	3	3	4	3	4	4
M3	3	3	4	4	4	4	4	4	4	4	4	4	4
M4	4	4	4	4	3	4	4	4	4	4	4	4	4

Table 4 shows the final classifications for the global comparisons in terms of joint preference. Assuming independence, this joint preference is computed as the product of the probabilities derived from the evaluations by the different individuals.

The classification thus obtained in terms of the impact of the course was that the features of improvement in market suitability, problem-solving abilities, and ability to employ technical skills are classified in a higher class than the others by both coordinators and managers. The first feature, employability, may be included in this group of highest preference as it was not evaluated by the managers and was placed in the highest class by the coordinators. The features of sociability, research ability, self-esteem, and fluency were classified in the highest class with respect to impact only by the managers.

Regarding the importance of the feature, sociability was elevated to the highest level by both groups, replacing the ability to employ technical skills, which was considered less important by

Table 3 – Coordinators' evaluations.

	Impact															
	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}	F_{16}
O1	2	2	2	1	2	0	0	0	0	N	N	1	2	1	2	2
O2	2	2	N	1	2	0	N	2	2	N	1	2	0	2	2	1
O3	2	2	2	1	2	2	1	2	2	2	2	1	2	1	2	2
O4	1	2	N	N	0	1	0	1	0	N	1	0	1	N	1	1
O5	1	1	1	1	1	1	2	1	1	0	1	2	1	1	2	2
O6	1	1	2	0	0	1	0	1	0	N	0	0	1	1	1	1
O7	2	1	2	1	2	1	2	2	1	2	1	1	2	1	2	2
O8	2	2	2	2	2	2	2	1	1	2	2	2	2	2	2	2
O9	2	2	2	2	1	2	1	2	2	2	1	2	2	1	2	2
	Importance															
	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}	F_{16}
O1	4	4	3	2	4	2	2	2	2	N	N	4	4	3	4	4
O2	3	4	N	2	4	2	N	4	4	N	4	4	3	2	3	4
O3	3	4	4	2	4	2	2	N	2	2	4	3	2	4	4	N
O4	3	3	N	N	N	3	2	3	2	N	2	N	2	N	2	2
O5	2	4	3	4	3	2	4	4	2	2	3	4	4	3	4	4
O6	3	4	3	2	3	4	3	2	2	N	2	3	3	3	4	4
O7	3	3	3	4	4	4	4	4	4	3	3	4	4	3	4	4
O8	4	4	2	3	4	3	4	1	0	0	1	4	4	4	4	4
O9	4	4	4	4	4	4	4	4	3	4	3	3	4	4	4	4

the coordinators. Whereas the features of entrepreneurship and critical thinking were elevated to the highest class by the managers, oratory skills and personal life (this last feature not having been evaluated by the managers) were classified by the coordinators as the least important.

Thus, when evaluating importance and impact, the managers of the students working at their companies appear more benevolent than the coordinators of similar courses. This difference may be due to the members of the latter group having a better knowledge of similar projects while being less involved with the participants of the project.

The classification of the features was employed to review the evaluation of the course. In this stage, the impact of only the features of highest importance was considered. Then it could be seen that, by those three features most important for the coordinators, the course would be classified in the highest impact class for market suitability and problem solving and would be classified as good for the impact on sociability. For the managers, the classification by the features in the highest class for importance is in the highest impact class for seven of the nine features of highest importance, the other two features resulting in the classification in the fourth class.

To evaluate the effect of the hypothesis of independence, a second classification was performed assuming that the evaluations from members of the same group were highly correlated. In this

Table 4 – Classification of features by impact and importance under independence.

Feature	Impact for coordinators	Impact for managers	Importance for coordinators	Importance for managers
Employability	5		4	
Market suitability	5	5	5	5
Earnings	4		4	
Entrepreneurship	4	4	4	5
Sociability	4	5	5	5
Research ability	4	5	4	5
Negotiation ability	4	4	4	4
Self-esteem	4	5	4	5
Oratory	4	4	3	4
Personal life	4		3	
Disinhibition	4	4	4	4
Fluency	4	5	4	5
Critical thinking	4	4	4	5
Openness	4	4	4	4
Problem solving	5	5	5	5
Technical skills	5	5	4	5

second approach the probabilities of classification above and below each class are made by equations (5) and (6), respectively, i.e., by the minimum and the maximum of the probabilities of the respective classification by the different members of the group. The results of this second classification are presented in Table 5.

It can be seen in Table 5 that the assumption of high correlation resulted in all features being classified in the fourth class with respect to impact. With respect to importance, there is a slightly higher discrimination, but it is still lower than that found by the evaluation under the hypothesis of independence. This is due to the downgrading of some features. In fact, under this last assumption, even for the managers, only three features are considered in the highest level: sociability, problem-solving, and technical skills. For the coordinators, all features classified under independence in the highest level come now to the fourth level. Between those previously classified in the fourth level, the use of technical skills as well as self-esteem and disinhibition fall also one level and are now added to oratory and personal life in the third level.

In another assessment of the variability of the results, Tables 6 and 7 present the results of taking a benevolent and a hostile approach to obtain class intervals. Percentual changes of 50% in the limits are employed.

Another indication of the lower discriminative power of the second approach is by the amplification of the ranges of the interval classification of the different features, resulting from a larger number of divergences between the hostile and benevolent classifications.

Table 5 – Classification of features assuming maximal correlation.

Feature	Impact for coordinators	Impact for managers	Importance for coordinators	Importance for managers
Employability	4		4	
Market suitability	4	4	4	4
Earnings	4		4	
Entrepreneurship	4	4	4	4
Sociability	4	4	4	5
Research ability	4	4	4	4
Negotiation ability	4	4	4	4
Self-esteem	4	4	3	4
Oratory	4	4	3	4
Personal life	4		3	
Disinhibition	4	4	3	4
Fluency	4	4	4	4
Critical thinking	4	4	4	4
Openness	4	4	4	4
Problem solving	4	4	4	5
Technical skills	4	4	3	5

Table 6 – Hostile and benevolent classifications with respect to impact.

Feature	Independent coordinators		Correlated coordinators		Independent managers		Correlated managers	
	Hostile	Benevolent	Hostile	Benevolent	Hostile	Benevolent	Hostile	Benevolent
Employability	5	5	4	5				
Market suitability	5	5	4	5	5	5	4	5
Earnings	4	5	4	4				
Entrepreneurship	4	4	4	4	4	5	4	5
Sociability	4	4	4	4	5	5	4	5
Research ability	4	4	4	4	5	5	4	5
Negotiation ability	4	4	4	4	4	4	4	5
Self-esteem	4	4	4	4	5	5	4	5
Oratory	4	4	4	4	4	4	4	5
Personal life	4	4	4	4				
Disinhibition	4	4	4	4	4	5	4	5
Fluency	4	4	4	4	5	5	4	5
Critical thinking	4	5	4	4	4	5	4	5
Openness	4	4	4	4	4	5	4	5
Problem solving	5	5	4	5	5	5	4	5
Technical skills	5	5	4	5	5	5	4	5

Table 7 – Hostile and benevolent classifications with respect to importance.

Feature	Independent coordinators		Correlated coordinators		Independent managers		Correlated managers	
	Hostile	Benevolent	Hostile	Benevolent	Hostile	Benevolent	Hostile	Benevolent
Employability	4	4	4	4				
Market suitability	5	5	4	5	5	5	4	5
Earnings	4	4	4	4				
Entrepreneurship	4	4	4	4	5	5	4	5
Sociability	5	5	4	4	5	5	5	5
Research ability	4	4	4	4	5	5	4	5
Negotiation ability	4	4	4	4	4	5	4	5
Self-esteem	4	4	3	4	5	5	4	5
Oratory	3	3	3	3	4	5	4	5
Personal life	3	3	3	3				
Disinhibition	4	4	3	4	4	5	4	5
Fluency	4	5	4	4	5	5	4	5
Critical thinking	4	4	4	4	5	5	4	5
Openness	4	4	4	4	4	5	4	5
Problem solving	5	5	4	4	5	5	5	5
Technical skills	4	4	3	3	5	5	5	5

Considering importance, for instance, when the changes in the limits are applied, three differences are observed in the assessments by the coordinators if correlation is assumed, of which only one is maintained under the independence assumption. In the assessments by the managers, this relation increases to ten against four.

For impact, the total of divergences jumps from two to four for coordinators and from four to 13 for managers, when independence is replaced by maximal correlation. In fact, it can be noticed in Table 6, that all the 13 features evaluated by the managers had a hostile classification lower than the benevolent classification if maximal correlation is assumed. It is interesting to notice also in Tables 6 and 7 that all the divergences when independence is assumed only affect features for which divergence is already observed when correlation is assumed.

6   CONCLUSION

The probabilistic approach here employed to combine the evaluations by a group of evaluators was successfully applied to classify features in terms of importance and impact and to evaluate a course.

The final evaluation as "Good" obtained in a previous analysis was confirmed. Besides, indications that for the features evaluated as more important the impact was evaluated as higher were revealed.

New information was also obtained on the perceived importance of the features. From the 16 features examined, clear agreement was obtained on the classification in the highest importance class of two distinct features: market suitability and problem solving.

Issues related to the mathematical modeling of the classification problem were also addressed. Empirical evidence was obtained favoring the idea that a higher discriminative power is obtained if independence between the disturbances is assumed.

A question that arises from situations where perceptions of several evaluators are combined is how to reduce the stress on the evaluators to reach coherent judgments. The knowledge that they don't need to weight criteria because a probabilistic composition will be able to deal with the assignment of importance may contribute to simplify matters for the evaluators along the decision process.

Another interesting aspect of employing a probabilistic approach to combine preferences like those here considered is that it allows the final results to be provided in terms of interval classifications between a hostile and a benevolent limit. The probabilistic formulation of these limits offers the decision maker a deeper understanding of the results.

Finally, it should be noticed that this approach can be extended to other situations where weighting features or weighting evaluators may be inadequate.

ACKNOWLEDGMENTS

REFERENCES

[1] ABDELLAOUI M, BARRIOS C & WAKKER P. 2007. Reconciling introspective utility with revealed preference: Experimental arguments based on prospect theory. *Journal of Econometrics*, **138**(1): 356–378.

[2] ALMEIDA-DIAS J, FIGUEIRA JR & ROY B. 2010. Electre Tri-C: A multiple criteria sorting method based on characteristic reference actions. *European Journal of Operational Research* **204**: 565–580.

[3] ALMEIDA-DIAS J, FIGUEIRA JR & ROY B. 2012. A multiple criteria sorting method where each category is characterized by several reference actions: The Electre Tri-nC method. *European Journal of Operational Research*, **217**: 567–579.

[4] ARROW KJ. 1963. *Social choice and individual values*. Wiley, London.

[5] BEN AMOR S, JABEUR K & MARTEL J. 2007. Multiple criteria aggregation procedure for mixed evaluations. *European Journal of Operational Research*, **181**(3): 1506–1515.

[6] BIRNBAUM M. 2008. New paradoxes of risky decision making. *Psychological Review*, **115**(2): 463.

[7] BOUJELBEN M, SMET Y DE, FRIKHA A & CHABCHOUB H. 2011. A ranking model in uncertain, imprecise and multi-experts contexts: the application of evidence theory. *International Journal of Approximate Reasoning*, **52**(8): 1171–1194.

[8] BORTOLUZZI SC, SILVA MR, ENSSLIN SR & ENSSLIN L. 2013. Structuring of a Performance Assessment Model for the Management of Accountancy Programs at Universidade Tecnológica Federal do Paraná. *Journal of Education and Research in Accounting*, **7**(1) 36–57.

[9]   BOUYSSOU D & MARCHAND T. 2013. Multiattribute preference models with reference points. *European Journal of Operational Research*, **229**(2): 470–481.

[10] BRANS JP. 1982. *L'ingénierie de la décision: élaboration d'instruments d'aide à la décision. La méthode PROMETHEE*. Presses de l'Université Laval.

[11] CHANG TH & WANG TC. 2009. Using the fuzzy multi-criteria decision making approach for measuring the possibility of successful knowledge management. *Information Sciences*, **179**(4): 355–370.

[12] CHANGKONG Y & HAIMES Y. 1983. *Multiobjective Decision Making*. North Holland, Amsterdam.

[13] CHU TC & LIN YT. 2009. An extension to fuzzy MCDM. *Computers and Mathematics with Applications*, **57**: 445–454.

[14] CELIK M, KANDAKOGLU A & ER ID. 2009. *Structuring fuzzy integrated multi-stages evaluation model on academic personnel recruitment in MET institutions*. Expert Systems with Applications, **36**: 6918–6927.

[15] COMES T, HIETE M & SCHULTMANN F. 2013. An Approach to Multi-Criteria Decision Problems under Severe Uncertainty. *Journal of Multicriteria Decision Analysis*, **20**(1-2): 29–48.

[16] DHEENA P & MOHANRAJ G. 2011. Multicriteria decision-making combining fuzzy Set theory, ideal and anti-ideal points for location site selection. *Information Sciences*, **38**: 13260–13265.

[17] DESIMPELAERE C & MARCHANT T. 2007. An empirical test of some measurement-theoretic axioms for fuzzy sets. *Fuzzy sets and systems*, **158**(12): 1348–1359.

[18] DURBACH I. 2012. An empirical test of the evidential reasoning approach's synthesis axioms. *Expert Systems with Applications*, **39**(12): 11048–11054.

[19] DURBACH I & STEWART T. 2009. Using expected values to simplify decision making under uncertainty. *Omega*, **37**(2): 312–330.

[20] DURBACH I & STEWART T. 2012. A comparison of simplified value function approaches for treating uncertainty in multi-criteria decision analysis. *Omega*, **40**: 456–464.

[21] EDWARDS W. 1977. How to use multiattribute utility measurement for social decisionmaking. *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-7**: 326–340.

[22] EDWARDS W & BARRON FH. 1994. SMARTS and SMARTER: Improved simple methods for multiattribute utility measurements. *Organizational Behavior and Human Decision Processes*, **60**: 306–325.

[23] EISELT HA & MARIANOV V. 2014. Multicriteria decision making under uncertainty: a visual approach. *International Transactions on Operational Research*, **21**(4): 525–540.

[24] FAN Z, LIU Y & FENG B. 2010. A method for stochastic multiple criteria decision making based on pairwise comparisons of alternatives with random evaluations. *European Journal of Operational Research*, **207**(2): 906–915.

[25] FEIZIZADEH B, JANKOWSKI P & BLASCHKE T. 2014. A GIS based spatially-explicit sensitivity and uncertainty analysis approach for multi-criteria decision analysis. *Computers & Geosciences*, **64**: 81–95.

[26] FISHBURN PC. 1970. *Utility theory for decision making*. Wiley (Operations Research Society of America Publications in Operations Research), New York.

[27] FISHBURN PC. 1973. *Les mathématiques de la décision*. Mouton (Mathématiques et sciences de l'homme), Paris.

[28] GINEVICIUS R & GINEVICIENE VB. 2009. The compliance of master's degree studies with the economic needs of the country. *Technological and Economic Development of Economy*, **15**: 136–153.

[29] GIANNOULIS C & ISHIZAKA A. 2010. A Web-based decision support system with ELECTRE III for a personalised ranking of British universities. *Decision Support Systems*, **48**: 488–497.

[30] GRECO S, INUIGUCHI M & SLOWINSKI R. 2006. Fuzzy Rough Sets and Multiple-premise Gradual Decision Rules. *International Journal of Approximate Reasoning*, **41**(2): 179–211.

[31] GRECO S, MATARAZZO B & SLOWINSKI R. 2010. Dominance-based rough set approach to decision under uncertainty and time preference. *Annals of Operations Research*, **176**(1): 41–75.

[32] HWANG CL & YOON K. 1981. *Multiple Attributes Decision Making Methods and Applications*, Springer, Berlin.

[33] ISHIZAKA A. 2012. Clusters and pivots for evaluating a large number of alternatives in AHP. *Pesquisa Operacional*, **32**(1): 87–101.

[34] ISHIZAKA A, PEARLMAN C & NEMERY P. 2012. AHPSort: an AHP-based method for sorting problems. *International Journal of Production Research*, **50**(17): 4767–4784.

[35] JAIN S & KHARE M. 2010. Construction of fuzzy membership functions for urban vehicular exhaust emissions modeling. *Environmental Monitoring and Assessment*, **167**: 691–699.

[36] JANSSEN P & NEMERY P. 2013. An extension of the FlowSort sorting method to deal with imprecision. *4OR*, **11**(2): 171–193.

[37] KAHNEMAN D. 2011. *Thinking, fast and slow*. New York, USA: Farrar, Straus and Giroux.

[38] KEENEY RL & RAIFFA H. 1976. *Decisions with Multiple Objectives: preferences and value trade-offs*. Wiley, New York.

[39] KROKHMAL P, ZABARANKIN M & URYASEV S. 2011. Modeling and optimization of risk. *Surveys in Operations Research and Management Science*, **16**(2): 49–66.

[40] LAHDELMA R, HOKKANEN J & SALMINEN P. 1998. SMAA Stochastic multiobjective acceptability analysis. *European Journal of Operational Research*, **106**(1): 137–143.

[41] LIU Y, FAN Z & ZHANG Y. 2011. A method for stochastic multiple criteria decision making based on dominance degrees. *Information Sciences*, **181**(19): 4139–4153.

[42] LUKMAN R, KRAJNC D & GLAVIÃO P. 2010. University ranking using research, educational and environmental indicators. *Journal of Cleaner Production*, **18**: 619–628.

[43] MAHDAVI I, MAHDAVI NA, HEIDARZADE A & NOURIFAR R. 2008. Designing a model of fuzzy TOPSIS in multiple criteria decision making. *Applied Mathematics and Computation*, **206**: 607–617.

[44] MENEZES RC & PIZZOLATO ND. 2014. Locating public schools in fast expanding areas: application of the capacitated p-median and maximal covering location models. *Pesquisa Operacional*, **34**(2): 301–317.

[45] MERIGÓ JM, GIL-LAFUENTE AM & PALACIOS-MARQUÉS D. 2014. A new method for fuzzy decision making under risk and uncertainty. *International Journal of Business Continuity and Risk Management*, **5**(1): 29–42.

[46] MOSADEGHI R, WARNKEN J, TOMLINSON R & MIRFENDERESK H. 2013. Uncertainty analysis in the application of multi-criteria decision-making methods in Australian strategic environmental decisions. *Journal of Environmental Planning and Management*, **56**(8): 1097–1124.

[47] MOUSSEAU V, SLOWINSKI R & ZIELNIEWICZ P. 1999. *ELECTRE TRI 2.0a. methodological guide and user's manual*. Documents du Lamsade. Université de Paris-Dauphine. Paris.

[48] MUSTAJOKI J, HAMALAINE RP & SAILO A. 2005. Decision Support by Incorporating Uncertainty in the SMART and SWING Methods. *Decision Sciences*, **36**(2): 317–339.

[49] NEFESLIOGLU HA, SEZER EA, GOKCEOGLU C & AYAS Z. 2013. A modified analytical hierarchy process (M-AHP) approach for decision support systems in natural hazard assessments. *Computers & Geosciences*, **59**: 1–8.

[50] NEMERY P & LAMBORAY C. 2008. FlowSort: a flow-based sorting method with limiting or central profiles. *TOP*, **16**(1): 90–113.

[51] NEPOMUCENO LDO & COSTA HG. 2012. Mapeamento de percepções na avaliação dos impactos do mestrado profissional no perfil do seu egresso (Perceptions mapping on the impacts of professional Master's Programs in the profile of their alumni). *Produção*, **22**: 865–873.

[52] NEPOMUCENO LDO & COSTA HG. 2015. Analyzing perceptions about the influence of a master course over the professional skills of its alumni: A multicriteria approach. *Pesquisa Operacional*, **35**(1): 187–211.

[53] NEPOMUCENO LDO, COSTA HG & SHIMODA E. 2010. Impact of professional master's programs on graduate students: intercomparison of perceptions of students, faculty members, coordinators, and companies involved. *Gestão & Produção*, **17**: 817–828.

[54] NEVES RB. 2005. *Integração da análise SWOT com o método ELECTRE TRI na avaliação do desempenho de programas de pós-graduação*. M. Sc. Thesis, Production Engineering Program. Universidade Federal Fluminense, Niterói, Rio de Janeiro, Brasil.

[55] PAS J VAN DER, WALKER W, MARCHAU V, WEE G VAN & AGUSDINATA D. 2010. Exploratory MCDA for handling deep uncertainties: the case of intelligent speed adaptation implementation. *Journal of Multi-Criteria Decision Analysis*, **17**(1/2): 1–23.

[56] PEREZ IJ, CABRERIZO FJ, ALONSO S & HERRERA-VIEDMA E. 2014. A new consensus model for group decision making problems with non-homogeneous experts. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, **44**(4): 494–498.

[57] POLITIS Y & SISKOS Y. 2004. Multicriteria methodology for the evaluation of a Greek engineering department. *European Journal of Operational Research*, **156**: 223–240.

[58] ROY B. 1968. *Classement et choix en presence de points de vue multiples (la methode ELECTRE)*. Presses Polytechniques et Universitaires Romandes, Lausanne.

[59] ROY B. 1978. ELECTRE III: Un alghoritme de méthode de classements fondé sur une representation floue des préférences em présence de critéres multiples. *Cahiers de CERO*, **20**: 3–24.

[60] ROY B & BERTIER PM. 1971. *La methode ELECTRE II: Une methode de classement en presence de criteres multiples*. SEMA (Metra International), Paris, 45.

[61] ROY B & BOYSSOU D. 1985. *Methodologie Multicritère d'Aide à la Decision*. Economica, Paris.

[62] ROY B & HUGONNARD JC. 1981. *Classement des prolongements de lignes de stations en banlieu parisienne*. Cahiers du LAMSADE. Université Dauphine et RATP. Paris.

[63] ROY B & SKALKA JM. 1984. *ELECTRE IS: Aspécts methodologiques et guide d'utilization*. Cahiers du LAMSADE. Université de Paris-Dauphine, Paris.

[64] SAATY TL. 1980. *The Analytic Hierarquic Process*. RWS Publications, Pittsburg.

[65] SANT'ANNA AP. 2002. Randomization and composition of preference measures. *Pesquisa Operacional*, **22**: 87–103.

[66] SANT'ANNA AP. 2005. Probabilistic compositions in courses evaluation. *RBPG*, **2**: 40–54.

[67] SANT'ANNA AP, COSTA HG & PEREIRA V. 2015. CPP-TRI: A Sorting Method Based on the Probabilistic Composition of Preferences. *International Journal of Information and Decision Science*, **7**(3): 194–200.

[68] SHANIAN A, MILANI AS, CARSON C & ABEYARATNE RC. 2008. A new application of ELECTRE III and revised Simos' procedure for group material selection under weighting uncertainty. *Knowledge-Based Systems*, **21**(7): 709–720.

[69] SILVA LE, CUNHA KSG & DIAS EP. 2012. Metodologia para a identificação de fatores que contribuam para a melhoria da qualidade dos cursos de pós-graduação com base nas avaliações trienais da Capes. *RBPG*, **9**(18): 1–16.

[70] SIMOS J. 1990. *Evaluer l'impact sur l'environnement: Une approache originale par l'analyse multicritere et la negociation*. Presses Polytechniques et Universitaires Romandes, Lausanne.

[71] WANG Y, LUO Y & HUA Z. 2008. On the extent analysis method for fuzzy AHP and its applications. *European Journal of Operational Research*, **186**(2): 735–747.

[72] WINTERFELDT D VON & EDWARDS W. 1986. *Decision analysis and behavioral research*. Cambridge University Press, Cambridge, UK.

[73] YANG J, WANG Y, XU D & CHIN K. 2006. The evidential reasoning approach for MADA under both probabilistic and fuzzy uncertainties. *European Journal of Operational Research*, **171**(1): 309–343.

[74] YU W. 1992. ELECTRE TRI – *Aspects Methodologiques et Guide d'Utilisation*. Documents du Lamsade. Université de Paris-Dauphine. Paris.

[75] ZADEH LA 1965. Fuzzy Sets. *Information and Control*, **8**: 338–358.

[76] ZELENY M. 1982. *Multiple Criteria Decision Making*. McGraw-Hill, New York.

[77] ZHANG Y, FAN Z & LIU Y. 2010. A method based on stochastic dominance degrees for stochastic multiple criteria decision making. *Computers & Industrial Engineering*, **58**(4): 544–552.