Robust compromise multicriteria decision model to evaluate the level of achievement of graduation profile

Modelo robusto de decisão multicritério de compromisso para avaliar o nível de realização do perfil de graduação

DOI: 10.54022/shsv4n4-026
Recebimento dos originais: 23/10/2023
Aceitação para publicação: 28/11/2023

Fernando Rojas Zúñiga
PhD in Management
Institution: School of Nutrition and Dietetics, Faculty of Pharmacy, Universidad de Valparaíso
Address: Gran Bretaña 1093, Playa Ancha, Valparaíso, Chile
E-mail: fernando.rojas@uv.cl

Pamela Estay Castillo
Graduate in Nutrition
Institution: School of Nutrition and Dietetics, Faculty of Pharmacy, Universidad de Valparaíso
Address: Gran Bretaña 1093, Playa Ancha, Valparaíso, Chile
E-mail: pamela.estay@uv.cl

Kathleen Priken Figueroa
Master in Nutrition
Institution: School of Nutrition and Dietetics, Faculty of Pharmacy, Universidad de Valparaíso
Address: Gran Bretaña 1093, Playa Ancha, Valparaíso, Chile
E-mail: kathleen.priken@uv.cl

Ximena Ceballos Sánchez
Master in Nutrition
Institution: School of Nutrition and Dietetics, Faculty of Pharmacy, Universidad de Valparaíso
Address: Gran Bretaña 1093, Playa Ancha, Valparaíso, Chile
E-mail: ximena.ceballos@uv.cl

Héctor Vargas Oyarzún
Doctor in Systems Engineering and Automation
Institution: Pontificia Universidad Católica de Valparaíso
Address: Av. Brasil 2147, Valparaíso, Chile
E-mail: hector.vargas@pucv.cl
ABSTRACT
We propose an approach of a robust compromise multi-criteria decision-making model that evaluates the level of competency achievement in graduation profiles. Quantitative information on qualifications of specific and generic competencies of a graduation profile was collected from professors, graduates, employers, community partners, and students' self-perception of achievement. This information was categorized using ordinal Likert scales generated through Monte Carlo simulation methods and copula modeling. The proposed model is validated using optimization methods. It provides a single score for the level of competency achievement, and the analysis is conducted using descriptive statistics and multivariate analysis of clusters. Simulated scenarios revealed that higher weights given by stakeholders in the construction of the robust compromise score were associated with higher evaluation scores of competencies from the same stakeholders. Scenarios with negative correlations between stakeholder qualifications about level of achievement of graduation profile yielded lower scores in the robust compromise multi-criteria decision-making model. Our proposal for Robust Compromise Multi-Criteria Decision Making Model offers more precise single scores, reducing the epistemological uncertainty associated with the input information.

Keywords: achievement, copula, graduation profile, multi-criteria decision making.

RESUMO
Propomos uma abordagem de um modelo de tomada de decisão multicritério robusto comprometido que avalia o nível de conquista de competências em perfis de graduação. Foram coletadas informações quantitativas sobre qualificações de competências específicas e genéricas de um perfiel de graduação, fornecidas por professores, graduados, empregadores, parceiros comunitários e autopercepção de conquista dos estudantes. Essas informações foram categorizadas usando escalas ordinais de Likert geradas por métodos de simulação Monte Carlo e modelagem de cópulas. O modelo proposto é validado por meio de métodos de otimização. Ele fornece uma pontuação única para o nível de conquista de competências, e a análise é realizada usando estatísticas descritivas e análise multivariada de conglomerados. Cenários simulados revelaram que pesos mais altos atribuídos pelas partes interessadas na construção da pontuação de...
compromisso robusto estavam associados a pontuações de avaliação mais altas das competências pelas mesmas partes interessadas. Cenários com correlações negativas entre as avaliações das partes interessadas sobre o nível de conquista do perfil de graduação resultaram em pontuações mais baixas no modelo de tomada de decisão multicritério de compromisso robusto. Nossa proposta de modelo de tomada de decisão multicritério de compromisso robusto oferece pontuações únicas mais precisas, reduzindo a incerteza epistemológica associada às informações de entrada.

Palavras-chave: conquista, cópula, perfil de graduação, tomada de decisão multicritério.

1 INTRODUCTION

Universities from different latitudes are currently working based on curricular models of professional training by specific and generic competencies, which make up the graduation profile of their students (Brauer, 2021).

The conformation of these competencies, as well as the evaluation of their level of achievement in the graduating cohorts, should consider more than the classic qualification carried out by teachers in the different subjects that make up the study plans (Jorre de St Jorre et al., 2021). Thus, a comprehensive assessment of the achievement of competencies attained by graduates of an undergraduate degree should also be validated by stakeholders in the learning process and performance of these competencies, such as: graduates (Lavi et al., 2021), employers (Hill et al., 2019), society (Candy, 2000), and self-perception of the achievement obtained by the students themselves (Carless, 2017). In this regard, a question that emerges is: How has the evaluation of levels of achievement of specific and generic competencies of graduation profiles been approached in university education? Without intending to do an exhaustive review of the subject, we will now show some approaches used up to now.

The educational quality of producing graduates along with community expectations such as personal quality, morality, knowledge, and job competency is an absolute requirement for sustainable global community adjustment. To ensure this happens, Mukhaiyar et al. (2019) has indicated that it is required to design a qualified program focused on satisfying all components of educational activities. The authors used three criteria for mapping competencies: basic, vocational, and elective courses. They carried out interviews and focus groups
aimed at: 1) Alumni in the category of “experts” who have worked in the industry, 2) Professors, 3) Managers, 4) Students and 5) Graduates and Recent Graduates who have worked in the industry.

In a more quantitative approach, Vargas et al. (2019) developed an automated algorithm that is used to calculate the achievement of student competencies. The devised method is called “automated assessment and monitoring support for competence-based courses”, and it is an efficient and easily extensible tool to incorporate new ways of adding achievement information (both numerical and linguistic). However, the instrument lacks the important evaluation perspectives that other interest groups could contribute in the graduation profile of a program, which go beyond the input of qualifications that come from the formal qualifications of the learning results of subjects that make up a study plan.

Al-Qatawneh and Altweissi (2019) have made progress in identifying the adequacy of general and specific competencies for the 10th grade English language curriculum, and in identifying any statistically significant differences in teachers’ perspectives related to variables of gender, experience, and the interaction between them. In pursuit of this goal, the authors have developed an item questionnaire based on the learning outcome items approved and officially adopted by the Jordanian Ministry of Education, which assesses the learning outcomes included in the 10th grade English language of the Curriculum in Jordan from the perspective of English teachers.

Based on the theory of Mental Self-Government and using three data collection methods: document analysis, interviews and questionnaires, Huincahue et al. (2019) show an approach that reflects the professional competences expected at the end of the study plan, established both by the institution as well as by the administration of the program.

Universities that have followed the Education 4.0 concept postulate a flexible combination of digital literacy, critical thinking and problem solving in educational settings linked to real world scenarios, see Wasilah et al. (2021). In this context, teachers have faced the challenge of developing new methods and resources to integrate into their planning in order to help students develop these desirable and necessary skills; hence the question that Ramírez-Montoya et al. (2021) helps us to solve: What are the characteristics of a teacher to consider
within the framework of Education 4.0? These authors used descriptive and exploratory approaches, where quantitative and qualitative instruments were applied to undergraduate students in education programs and graduates. Interviews with experts in the educational field and focus groups with rectors, school directors, university professors and specialists in the educational area were also included. The data was triangulated, and the results were organized into the categories of (a) processes as facilitators (b), soft skills, (c) human sense and (d) the use of technologies.

As evidenced by the exploration of the scientific literature, we found a gap of quantitative methods useful to assess the level of achievement of the competencies expressed in the graduation profiles of university graduates, capable of ideally compiling in a single score, the perspectives that could emanate from various actors involved in the learning process and performance of these competencies, such as: graduates, employers, society, and self-perception of the achievement achieved by the students themselves. Multi-criteria decision-making approaches allow the construction of evaluation scores, which can provide useful solutions to this problem, since they reduce the epistemic uncertainty of underrepresented groups and linguistics that qualitative work entails, see Amin et al. (2011). Robust compromise multi-criteria decision model is a continuously growing research field with different approaches continuously being developed and combined to explore the underlying epistemic uncertainty (eg underrepresented minority groups in a data set or the presence of "rare" words in a linguistic modeling context), and useful weighting criteria for measuring performance derived from the data (scores), as indicated in Chai et al. (2013), Chang et al. (2011), Feng et al. (2011), Rajesh and Ravi (2015), Shaw et al. (2012), and Stewart and Durbach (2016).

To our knowledge, none of the previous studies have addressed the issue of defining a score that summarizes a simple way to measure these factors, particularly by applying new stochastic and robust compromise multi-criteria decision model, which we address in this article.

The aim of this paper is to propose a robust compromise multi-criteria decision model to evaluate the level of achievement of competencies of a
graduation profile, taking as an example the competencies of the graduation profile of an anonymous Chilean career in the health area.

2 METHODOLOGY

Quantitative information on qualifications related to specific and generic competencies of the graduation profile of an anonymous Chilean career in the health area, granted by professors, graduates, employers, community partners as representatives of society and the self-perception of achievement achieved by the students themselves, was categorized in ordinal Likert scales generated by Monte Carlo simulation methods. This approach serves to validate the proposed robust compromise multi-criteria decision model based on optimization methods. The application of this model establishes a single score for the level of achievement of competencies of the graduation profile. The analysis of this score was performed by using descriptive statistics and multivariate analysis of clusters.

For the formulation of the robust compromise multi-criteria decision model, we use different pieces of information or items on an ordinal scale (Likert), useful to evaluate their performance and obtain a single final qualification score. The method is based on:

An objective function that establishes obtaining the minimum covariance between normalized, weighted multi-criteria decision model scores by considering the sign (+) or (-) of the best decisions.

Decision variables: the weights of the multi-criteria decision model and the order of relative importance of each item, to finally obtain a single score that leads to a robust compromise solution.

Therefore, we ensure that any unaccounted-for epistemic uncertainty would not affect the performance of the calculated estimates, under the key concept of a trade-off between both disparate methods and their underlying assumptions.

To evaluate the alternatives, the following robust compromise multi-criteria decision models will be used approach of Rojas et al. (2022):

(i) the Complex Proportional Assessment of alternatives (COPRAS) method that evaluates the different alternatives through basic arithmetic operations, such as the utility function used in Zheng et al. (2018);
(ii) the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) that considers the distance to both the ideal solution and the non-ideal solution presented in Wanke et al. (2015); and

(iii) the Vlsekriterijumska Optimizacija I Kompromisno Resenje (VIKOR) method that considers consensus and regret functions described in Opricovic and Tzeng (2004).

The alternatives are compared in pairs and their relative importance is established. Based on:

(iv) the Analytic Hierarchy Process (AHP) method to perform a pairwise comparison between the criteria Partovi (1994), and

(v) the Stepwise Weight Assessment Ratio Analysis (SWARA) method to consider the relative efficiency of each criterion Kerˇsuliene et al. (2010)

2.1 ROBUST COMPROMISE MULTI-CRITERIA DECISION MODEL FORMULATION

Let $F \equiv \{f_k\}_{k \in \{1, \ldots, K\}}$ be a set formed by $K$ multi-criteria decision models used for ranking, such as TOPSIS, VIKOR, and COPRAS. For all $k \in \{1, \ldots, K\}$, each function $f_k \in F$ associated with such multi-criteria decision models returns a vector of performance scores $p_k$ containing values for each of the $m$ alternatives. Each function uses as inputs a weight vector $w$ with values for the $n$ criteria, an $m \times n$ matrix $R$ of normalized criterion values for each alternative, and a vector $s$ with the sign description for each of the $n$ criteria. Hence, we obtain $f_k(w, R, s) = p_k$.

Further, let be a set formed by $L$ distinct multi-criteria decision models employed for criteria weighting, such as AHP and SWARA. For all $l \in \{1, \ldots, L\}$, each function $g_l \in G$ associated with such multi-criteria decision models returns a vector of weights $w_l$ containing values for each of the $n$ criteria; as an input, an $m \times n$ matrix $X$ of original criterion values for each alternative, which can be eventually normalized into an $R$ matrix depending on the method. Therefore, we have $g_l(X) = w_l$ or $g_l(R) = w_l$.

Let $W$ be a vector of length $K$ denoting the weights for each function Heuristic genetic algorithms such as the shown in can be used to solve the robust compromise non-linear program in terms of the optimal values of $W = (W_1, \ldots, W_K)$ and $w = (w_1, \ldots, w_n)$, as follows:
\[
\begin{align*}
\min \{W^T \text{Cov}(f_1(w, R, s), \ldots, f_K(w, R, s))\}, \\
\text{subject to:}
\end{align*}
\]

\[
\begin{align*}
w_{\text{min}} \leq w \leq w_{\text{max}}, \sum_{k=1}^{K} W_k = 1, \sum_{i=1}^{n} w_i = 1,
\end{align*}
\]

where

the objective function stated in (1) represents the weighted covariance matrix of the performance computed using each multi-criteria decision model in F. Additionally, \(w_{\text{min}}\) and \(w_{\text{max}}\) are vectors of length \(n\) computed, respectively, by \(\min\{g_1(X) \text{ or } g_1(R), \ldots, g_L(X) \text{ or } g_L(R)\}\) and \(\max\{g_1(X) \text{ or } g_1(R), \ldots, g_L(X) \text{ or } g_L(R)\}\), that is, these represent the minimal and maximal weights obtained using the multi-criteria decision models in for each criterion. Constraints expressed in (2) indicate that all multi-criteria decision models and criterion weights sum up to one.

2.2 MONTECARLO SIMULATION

To carry out a numerical experiment via Montecarlo simulation, the generation of a simulated data vector of numerical grades ordered on a Likert scale was needed. These data represent the level of achievement of competencies of the graduation profile of the nutrition and dietetics career of the University of Valparaiso, Chile, valued by 5 stakeholders in the learning process and performance of these competencies: teachers, graduates, employers, society and self-perception of students. With this objective, we generated a joint distribution for this data frame, by using the copula package implemented in the open-source software so called R version 4.1.2. This statistical skill can be formulated for a bivariate case as follows:

Let \(Y_1\) and \(Y_2\) be independent and identically distributed random variables, as valuation ordered on a Likert scale, with their unknown joint cumulative distribution function \(F_{Y_1, Y_2}\). Then, we have the relationship:

\[
C(a, b; \rho) = F_{Y_1, Y_2}(F_{Y_1}^{-1}(a), F_{Y_2}^{-1}(b)), \quad a, b \in [0,1],
\]

Where

\(\rho\) is a parameter of dependence between \(Y_1\) and \(Y_2\), which is denoted as \(\rho = \rho_{Y_1, Y_2}\). The copula defined in (3) is assumed to be continuous and twice differentiable. Thus, from (3) we got to that the probability density function associated with \(C(u, v, \rho)\), which can be expressed as
\[ c(F_{Y_2}(y_2), F_{Y_1}(y_1), \rho_{Y_1,Y_2}) = \frac{\partial^2 c}{\partial F_{Y_2} \partial F_{Y_1}}. \]  

(4)

For \( Y \) random variable, the probability density function \( f_Y \) defined on \([0, \infty)\) (non-negative support), is a cumulative density function

\[ F_Y(y) = \int_0^y f_Y(v) \, dv, \]  

(5)

and a quantile function \( y(q) = F_Y^{-1}(q) \), for \( 0 < q < 1 \). Note that \( F_{Y_1}^{-1}(a) \) and \( F_{Y_2}^{-1}(b) \), which can be obtained when marginal statistical distributions are known or adjusted by a statistical model from actual data. This model can be extended for a multi-dimension vector.

Readers interested in using copulas to describe correlated random variables can refer to Rojas et al. (2021). This approach (copula) allowed us to generate 1000 random scenarios with uniform statistical distributions between an interval of 0.5 to 4.5 (rounded to 0 decimal), to obtain a vector with 5 dimensions of simulated scores on a Likert scale ordered from 1 to 4, to establish the level of achievement of the graduation profile of a cohort of 50 individuals (length of sample simulated), where 1 = acceptable level of achievement, 2 = good level of achievement, 3 = very good level of achievement, and 4 = level of achievement of excellence, qualified by each of the stakeholders evaluating the training process: 1) Teachers, 2) Graduates, 3) Employers, 4) Society and 5) Students (self-perception of achievement). For each of these 1000 scenarios generated through the copula method shown in 3, a random generation for the parameter \( \rho \) was considered through a uniform statistical distribution in the interval \([ 0.25, 1]\), in which it is possible to generate the multidimensional joint distribution of size 5. In each of the 1000 scenarios generated, we apply the Robust compromise multi-criteria decision model formulation, with which we obtain the scores that summarize, in a measure between 0-1, the scores given by the stakeholders in the training process. These scores are obtained for each decision-making method (COPRAS, VIKOR and TOPSIS), in addition to our robust compromise approach (Roco). Interested readers can request the programmed codes for R software, both for the Roco function and for the illustrative simulation of this paper.
3 RESULTS

In Table 1 it is possible to appreciate a descriptive statistical summary of the parameters obtained in random way for the Monte Carlo simulation. These results allow us to corroborate that there are no relevant differences in the simulated data of the average Likert scores, their standard deviations and interquartile ranges, for the several stakeholders considered in the numerical experiment. Although we do not show this result, we confirm that there are no statistically significant differences in the medians of these data distributions, by applying the Kruskal-Wallis test.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>mean</th>
<th>sd</th>
<th>IQR</th>
<th>0%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>0.38</td>
<td>0.37</td>
<td>0.65</td>
<td>-0.25</td>
<td>0.05</td>
<td>0.39</td>
<td>0.71</td>
<td>1.00</td>
</tr>
<tr>
<td>Average Likert simulated evaluation</td>
<td>2.50</td>
<td>0.15</td>
<td>0.20</td>
<td>2.06</td>
<td>2.40</td>
<td>2.50</td>
<td>2.60</td>
<td>3.02</td>
</tr>
<tr>
<td>Professors</td>
<td>2.49</td>
<td>0.15</td>
<td>0.22</td>
<td>1.98</td>
<td>2.38</td>
<td>2.48</td>
<td>2.60</td>
<td>2.92</td>
</tr>
<tr>
<td>Graduated</td>
<td>2.50</td>
<td>0.16</td>
<td>0.22</td>
<td>1.98</td>
<td>2.38</td>
<td>2.50</td>
<td>2.60</td>
<td>2.96</td>
</tr>
<tr>
<td>Employers</td>
<td>2.50</td>
<td>0.15</td>
<td>0.20</td>
<td>2.00</td>
<td>2.40</td>
<td>2.50</td>
<td>2.60</td>
<td>2.92</td>
</tr>
<tr>
<td>Society</td>
<td>2.50</td>
<td>0.16</td>
<td>0.23</td>
<td>2.04</td>
<td>2.38</td>
<td>2.50</td>
<td>2.61</td>
<td>2.98</td>
</tr>
<tr>
<td>Students</td>
<td>1.11</td>
<td>0.07</td>
<td>0.09</td>
<td>0.90</td>
<td>1.07</td>
<td>1.11</td>
<td>1.16</td>
<td>1.27</td>
</tr>
<tr>
<td>Standard deviation Likert simulated evaluation Professors</td>
<td>1.11</td>
<td>0.06</td>
<td>0.08</td>
<td>0.91</td>
<td>1.07</td>
<td>1.11</td>
<td>1.15</td>
<td>1.28</td>
</tr>
<tr>
<td>Standard deviation Likert simulated evaluation Graduated</td>
<td>1.11</td>
<td>0.06</td>
<td>0.09</td>
<td>0.93</td>
<td>1.07</td>
<td>1.11</td>
<td>1.16</td>
<td>1.30</td>
</tr>
<tr>
<td>Standard deviation Likert simulated evaluation Employers</td>
<td>1.11</td>
<td>0.06</td>
<td>0.09</td>
<td>0.93</td>
<td>1.07</td>
<td>1.12</td>
<td>1.16</td>
<td>1.28</td>
</tr>
<tr>
<td>Standard deviation Likert simulated evaluation Society</td>
<td>1.11</td>
<td>0.06</td>
<td>0.09</td>
<td>0.89</td>
<td>1.07</td>
<td>1.11</td>
<td>1.16</td>
<td>1.29</td>
</tr>
<tr>
<td>Standard deviation Likert simulated evaluation Students</td>
<td>1.89</td>
<td>0.58</td>
<td>0.25</td>
<td>1.00</td>
<td>1.75</td>
<td>2.00</td>
<td>2.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Interquartile Range Likert simulated evaluation Professors</td>
<td>1.87</td>
<td>0.59</td>
<td>0.25</td>
<td>1.00</td>
<td>1.75</td>
<td>2.00</td>
<td>2.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Interquartile Range Likert simulated evaluation Graduated</td>
<td>1.89</td>
<td>0.57</td>
<td>0.25</td>
<td>1.00</td>
<td>1.75</td>
<td>2.00</td>
<td>2.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Interquartile Range Likert simulated evaluation Employers</td>
<td>1.89</td>
<td>0.57</td>
<td>0.25</td>
<td>1.00</td>
<td>1.75</td>
<td>2.00</td>
<td>2.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Interquartile Range Likert simulated evaluation Society</td>
<td>1.89</td>
<td>0.57</td>
<td>0.25</td>
<td>1.00</td>
<td>1.75</td>
<td>2.00</td>
<td>2.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Interquartile Range Likert simulated evaluation Students</td>
<td>1.91</td>
<td>0.58</td>
<td>0.25</td>
<td>1.00</td>
<td>1.75</td>
<td>2.00</td>
<td>2.00</td>
<td>3.00</td>
</tr>
</tbody>
</table>

Source: own elaboration.

In Table 2 it is possible to appreciate a descriptive statistical summary of the results by criteria weights of each stakeholder that evaluated profile of graduated, and weights of method for decision making for the Monte Carlo simulation.
simulation. Note that the weights of the 5th criterion (evaluating key actor: Students), are slightly greater than the rest of the evaluations to obtain a single robust compromise score, as a decision-making method. This result was corroborated without statistical significance of the differences in medians of criteria weights, using the Kruskal-Wallis test. On the other hand, the weighting of the COPRAS, VIKOR and COPRAS decision-making methods was quite similar, although slightly higher for the COPRAS method in the construction of the robust compromise score. There were also no statistically significant differences when comparing the medians of weights of decision-making methods.

Table 2: Descriptive statistical of results of Montecarlo simulation.

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>sd</th>
<th>IQR</th>
<th>0%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria Weights Professors</td>
<td>0.16</td>
<td>0.05</td>
<td>0.07</td>
<td>0.06</td>
<td>0.12</td>
<td>0.18</td>
<td>0.19</td>
<td>0.28</td>
</tr>
<tr>
<td>Criteria Weights Graduated</td>
<td>0.19</td>
<td>0.13</td>
<td>0.15</td>
<td>0.02</td>
<td>0.09</td>
<td>0.15</td>
<td>0.24</td>
<td>0.55</td>
</tr>
<tr>
<td>Criteria Weights Employers</td>
<td>0.17</td>
<td>0.11</td>
<td>0.07</td>
<td>0.01</td>
<td>0.11</td>
<td>0.15</td>
<td>0.18</td>
<td>0.48</td>
</tr>
<tr>
<td>Criteria Weights Society</td>
<td>0.19</td>
<td>0.14</td>
<td>0.16</td>
<td>0.01</td>
<td>0.08</td>
<td>0.16</td>
<td>0.24</td>
<td>0.56</td>
</tr>
<tr>
<td>Criteria Weights Students</td>
<td>0.28</td>
<td>0.17</td>
<td>0.28</td>
<td>0.07</td>
<td>0.15</td>
<td>0.19</td>
<td>0.43</td>
<td>0.60</td>
</tr>
<tr>
<td>SCORE ROCO</td>
<td>0.55</td>
<td>0.07</td>
<td>0.09</td>
<td>0.40</td>
<td>0.51</td>
<td>0.57</td>
<td>0.60</td>
<td>0.68</td>
</tr>
<tr>
<td>SCORE TOPSIS</td>
<td>0.47</td>
<td>0.24</td>
<td>0.29</td>
<td>0.00</td>
<td>0.33</td>
<td>0.48</td>
<td>0.62</td>
<td>1.00</td>
</tr>
<tr>
<td>SCORE VIKOR</td>
<td>0.63</td>
<td>0.28</td>
<td>0.40</td>
<td>0.00</td>
<td>0.47</td>
<td>0.70</td>
<td>0.87</td>
<td>1.00</td>
</tr>
<tr>
<td>SCORE COPRAS</td>
<td>0.66</td>
<td>0.23</td>
<td>0.35</td>
<td>0.25</td>
<td>0.49</td>
<td>0.70</td>
<td>0.8</td>
<td>1.00</td>
</tr>
<tr>
<td>Weights method TOPSIS</td>
<td>0.34</td>
<td>0.02</td>
<td>0.03</td>
<td>0.26</td>
<td>0.32</td>
<td>0.33</td>
<td>0.35</td>
<td>0.40</td>
</tr>
<tr>
<td>Weights method VIKOR</td>
<td>0.32</td>
<td>0.03</td>
<td>0.03</td>
<td>0.24</td>
<td>0.31</td>
<td>0.33</td>
<td>0.34</td>
<td>0.38</td>
</tr>
<tr>
<td>Weights method COPRAS</td>
<td>0.34</td>
<td>0.03</td>
<td>0.02</td>
<td>0.26</td>
<td>0.33</td>
<td>0.34</td>
<td>0.35</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Source: own elaboration.

In Figure 1 shows a box plot for the Scores built based on each of Multi Criteria Decision Making methods. Note that score of our robust compromise solution proposal deliver more precise scores, indicating a reduction in the epistemological decision of ranking the information used as input.
Figure 1: Box plot for the Scores built based on each of Multi Criteria Decision Making methods.

By means of multivariate hierarchical cluster analysis techniques, we constitute groupings of scenarios, which will help us to characterize the conditions in which it is more likely to obtain criteria results weights of each stakeholder that evaluated profile of graduated, and weights of method of robust compromise for solutions of decision making. A simple examination of the Figure 2 shows us that we have 4 large groups of scenarios, which are heterogeneous among themselves and have homogeneous characteristics within the scenarios that are members of the same group.
Figure 2: Grouping of scenarios constructed through the use of multivariate cluster analysis.

In Table 3 it is possible to review the centroids or means of each of the variables used to generate the groups of scenarios by k-means clusters analysis, whose examination allows us to associate the parameters of mean scores, interquartile ranges and correlation coefficients of the Likert scales of evaluation of the training process of each one of the stakeholders. The results of robust compromise scores, weightings of decision-making ranking methods, and weighting of criteria of each key evaluator actor are showed in Table 4 for each group of scenarios obtained by k-means clusters analysis.

Table 3: Centroids or means of each of the variables used to generate the groups of scenarios.

<table>
<thead>
<tr>
<th>Group</th>
<th>IQR Professors</th>
<th>IQR Graduate Students</th>
<th>IQR Employees</th>
<th>IQR Societies</th>
<th>Mean Professors</th>
<th>Mean Graduate Students</th>
<th>Mean Employees</th>
<th>Mean Societies</th>
<th>Mean Students</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.75</td>
<td>2.00</td>
<td>1.75</td>
<td>2.00</td>
<td>2.56</td>
<td>2.80</td>
<td>2.38</td>
<td>2.26</td>
<td>2.28</td>
<td>-0.25</td>
</tr>
<tr>
<td>2</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>1.00</td>
<td>2.32</td>
<td>2.70</td>
<td>2.22</td>
<td>2.30</td>
<td>2.62</td>
<td>0.05</td>
</tr>
<tr>
<td>3</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>1.88</td>
<td>2.29</td>
<td>2.26</td>
<td>2.44</td>
<td>2.34</td>
<td>2.29</td>
<td>0.87</td>
</tr>
<tr>
<td>4</td>
<td>2.00</td>
<td>3.00</td>
<td>2.75</td>
<td>1.75</td>
<td>2.42</td>
<td>2.34</td>
<td>2.24</td>
<td>2.50</td>
<td>2.50</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Source: own elaboration.
Table 4: Centroids or means of each of the results in the groups of scenarios.

<table>
<thead>
<tr>
<th>Group</th>
<th>Criteria Graduated</th>
<th>Weights</th>
<th>Criteria Employers</th>
<th>Weights</th>
<th>Criteria Students</th>
<th>Weights</th>
<th>Criteria Professors</th>
<th>Weights</th>
<th>COPRAS Weights</th>
<th>VIKOR Weights</th>
<th>TOPSIS Weights</th>
<th>Robust Compromise Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.42</td>
<td>0.16</td>
<td>0.11</td>
<td>0.21</td>
<td>0.11</td>
<td>0.35</td>
<td>0.32</td>
<td>0.33</td>
<td>0.35</td>
<td>0.33</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.23</td>
<td>0.12</td>
<td>0.43</td>
<td>0.13</td>
<td>0.09</td>
<td>0.34</td>
<td>0.32</td>
<td>0.34</td>
<td>0.34</td>
<td>0.34</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.11</td>
<td>0.22</td>
<td>0.34</td>
<td>0.17</td>
<td>0.16</td>
<td>0.33</td>
<td>0.31</td>
<td>0.35</td>
<td>0.35</td>
<td>0.35</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.09</td>
<td>0.15</td>
<td>0.19</td>
<td>0.13</td>
<td>0.43</td>
<td>0.33</td>
<td>0.33</td>
<td>0.34</td>
<td>0.34</td>
<td>0.34</td>
<td>0.55</td>
<td></td>
</tr>
</tbody>
</table>

Source: own elaboration.

We corroborate that all the medians of variables and results, shown in Table 3 and Table 4, respectively, have significant statistical differences when analyzed through the Kruskal-Wallis test (results not shown), which is why we can affirm that these variables and results they are associated with the displayed scenario groups.

4 DISCUSSION

In the realm of university education, it is essential to have a tool that can aggregate the assessment of competency levels from various stakeholders involved in the educational process. This tool would enable a comprehensive evaluation of a student’s attainment of competencies outlined in the graduate profile. It is crucial to involve different stakeholders, such as faculty members, employers, and alumni, as they provide diverse perspectives on the graduates’ abilities. While there is limited literature specifically addressing this topic, the work by Wait and Govender (2016) emphasizes the importance of multi-stakeholder involvement in competency assessment in higher education. In this context, the great contribution to the scientific literature of this work is the use of the robust compromise multi-criteria decision model turns out to be a useful tool to evaluate the level of achievement of competencies of the graduation profile, reducing the epistemic uncertainty of this process.

Like our approach, Monte Carlo simulation studies are widely used in scientific research due to their versatility and robustness. However, the accuracy and reliability of these simulations heavily depend on the selection of unbiased input parameters. Biased parameters can introduce systematic errors, leading to distorted results and misleading conclusions. Several studies have emphasized the significance of unbiased parameter selection (e.g., Arend and Schäfer, 2019).
It is therefore essential to ensure that input parameters in Monte Carlo simulations are free from bias, as our work shows in the results of the table 1. By doing so, researchers can enhance the credibility and validity of their findings, making the simulation outcomes more accurate and dependable. The results obtained from assigning criteria weights to each stakeholder who evaluated the profile of graduates, as well as the weights assigned to the decision-making method in the Monte Carlo simulation, did not exhibit significant biases. This finding suggests that the assessment process and the methodology employed in the simulation were fair and unbiased. The absence of significant biases ensures the integrity and reliability of the evaluation, allowing for a more accurate and comprehensive analysis of the graduates’ competency levels. It is important to ensure transparency and fairness in stakeholder involvement and decision-making methods for robust evaluations.

Reducing epistemic uncertainty is crucial when constructing an indicator based on decision-making models. Epistemic uncertainty refers to the lack of knowledge or information about the true values or probabilities involved in the decision-making process. By minimizing epistemic uncertainty, decision-makers can enhance the reliability and validity of the indicator, leading to more informed and effective decision-making. Studies such as the work by Brüggemann et al. (2020), on the concept of "post-normal science" highlight the importance of addressing epistemic uncertainty in decision-making processes, emphasizing the need for robust approaches to model-based indicators. As illustrated in Figure 1, this study has successfully developed a robust compromise model that effectively reduces the variability in the performance of assessing the attainment level of competencies in the graduate profile. This model incorporates the opinions of key actors in the educational process and decision-ranking methods, ensuring a comprehensive and reliable evaluation. By integrating multiple perspectives and decision-making approaches, the model achieves a more holistic and accurate assessment of competency levels. The results highlight the effectiveness of this approach in enhancing the evaluation process and providing valuable insights for program improvement.

Cluster analysis has emerged as a valuable technique for grouping Monte Carlo simulation scenarios to explore empirical associations or relationships more
effectively. By applying cluster analysis to simulation outputs, researchers can identify similar patterns or clusters of results, revealing underlying structures within the data. This approach allows for a comprehensive exploration of scenario relationships and facilitates a deeper understanding of complex systems. Studies such as the work by Colella et al. (2019) demonstrate the utility of cluster analysis in analyzing Monte Carlo simulation outcomes to uncover empirical associations and gain valuable insights into system dynamics.

On the one hand, as extracted from the results shown in Tables 3 and 4, higher average score of evaluations of competencies of profile of stakeholders used in the construction of the robust compromise score, were associated with higher criteria weights from these same stakeholders. In multicriteria decision-making models, it is commonly observed that when an item receives a higher score, it leads to a higher weighting for that particular item. This phenomenon is attributed to the inherent nature of subjective judgments and the decision-maker's perception of the item's importance. Numerous studies provide explanations and support for this relationship. According to the research by Odu (2019), decision-makers tend to give higher weights to criteria that are perceived as more significant or influential in achieving their goals. This aligns with the notion of "importance weighting" in decision-making, where higher-scoring items are considered more important and thus receive higher weights. Similarly, Bulley and Schacter (2020) discuss the concept of "value-focused thinking" in decision-making, emphasizing that higher-scoring items are typically seen as more valuable and deserving of greater consideration and weight.

While on the other hand, scenarios generated from negative correlations of qualifications of stakeholders in the learning process related to the graduation profile are associated with dissimilar scores of robust compromise multicriteria decision-making. This is explained because the generation of scenarios based on multidimensional copulas with negative correlations can lead to greater differences in the weightings of evaluated item criteria.

Copulas enable the modeling of complex dependencies between variables, including negative correlations, which can result in diverse and contrasting scenarios. Studies such as the work by Safari-Katesari et al. (2020) emphasize the role of copulas in capturing and generating correlated scenarios. Additionally, the
research by Yang et al. (2020) demonstrates that negative correlations in copulas can amplify differences in criteria weightings, highlighting the impact of scenario generation on the evaluation process.

As we have showed, our proposal is based on creating a single evaluation score to assess the level of competency attainment in the graduation profile. This score is derived from ratings obtained on the Likert scale from multiple stakeholders involved in the training process. However, the authors acknowledge limitations associated with Likert scales used for indicator construction. Response style biases, limited response options, interpretation discrepancies, and the lack of interval measurement can affect the validity and reliability of the indicators (Carifio and Perla, 2007; Krosnick, 1999; Nunnally, 1978; and Podsakoff et al., 2003). Addressing these limitations is essential to ensure accurate and comprehensive evaluations.

5 CONCLUSION

This study presents a comprehensive assessment using a robust compromise multi-criteria decision model to evaluate competency levels within graduation profiles. Through descriptive statistics, statistical tests, and multivariate analysis, we affirm the consistency and reliability of our approach in analyzing simulated data and decision-making weights. Our work contributes significantly to the realm of university education by offering a tool that aggregates multiple stakeholder assessments, ensuring a holistic evaluation of graduates' competencies outlined in their profiles.

The Monte Carlo simulation results, as displayed in Table 1, underscore the importance of unbiased input parameters in ensuring the accuracy and reliability of simulation outcomes. Our findings emphasize the fairness and transparency in both stakeholder assessments and decision-making methodologies, ensuring a comprehensive and reliable evaluation process.

Reducing epistemic uncertainty is paramount in constructing indicators based on decision-making models. Our study successfully minimizes this uncertainty, providing a robust compromise model that effectively evaluates competency levels within graduate profiles. By integrating diverse perspectives
and decision-making methodologies, our model achieves a more accurate and comprehensive assessment, as depicted in Figure 1.

Cluster analysis proves instrumental in identifying patterns and relationships within the simulation scenarios, shedding light on the dynamics of complex systems. The associations between stakeholder evaluations and criteria weights align with known phenomena in multicriteria decision-making models, where higher scores often result in higher weightings due to perceived importance.

Negative correlations among stakeholders’ qualifications generate dissimilar scores in the robust compromise multi-criteria decision-making, showcasing the impact of copulas in scenario generation. Copulas enable the modeling of intricate dependencies and highlight the diverse outcomes arising from negative correlations, as noted in previous studies.

However, limitations associated with Likert scales used for indicator construction must be acknowledged. Response biases and scale limitations could affect the validity and reliability of our evaluation process, necessitating further attention in future assessments.

In summary, our work underscores the significance of a robust compromise multi-criteria decision model in evaluating competency levels within graduation profiles. While demonstrating the effectiveness of our approach, future research should focus on addressing the limitations associated with evaluation scales to ensure more accurate and reliable assessments in educational contexts.
REFERENCES


