COMBINING AHP AND GOAL PROGRAMMING IN THE CONTEXT OF THE ASSESSMENT OF E-LEARNING

Mónica de Castro-Pardo*1
Universidad Rey Juan Carlos
monica.decastro@urjc.es

Concepción de la Fuente-Cabrero
Universidad Rey Juan Carlos
concepcion.delafuente@urjc.es

Pilar Laguna-Sánchez
Universidad Rey Juan Carlos
pilar.laguna@urjc.es

Fernando Pérez-Rodríguez
fora forest technologies SLL
fernando.perez@fora.es

*Corresponding Author

ABSTRACT

The Analytical Hierarchy Process is a very common method used in Multi-Criteria Decision Making (MCDM) to analyze participative assessments. However, due to the qualitative nature of this methodology, a high percentage of inconsistencies need to be addressed when analyzing user preferences. This work analyzes the efficiency of the Goal Programming model in order to reduce inconsistencies with pairwise comparisons when working with inexpert participants and time limitations. A case study has been carried out that assesses online courses in higher education with the Analytical Hierarchy Process in order to understand the usefulness and feasibility of the method. Evaluation of four e-learning tools (collaboration tools, content tools, tutorial sessions and evaluation tools) used in an online business degree were collected from 72 students through a ‘Saaty-type’ survey, and the model was applied to improve the consistency of these results. This model has been able to minimize the inconsistencies of individual preferences while avoiding the loss of primary information.

Keywords: Goal Programming; Analytical Hierarchy Process; inconsistencies; e-learning; participative decision making

1. Introduction

Effective quality measures for e-learning have been described as being “urgently required” (Martínez-Caro et al., 2014). In this sense, it is important to remark on the

This work was supported by the Government of Spain, Department of Economy, Industry and Competitiveness under the Torres Quevedo Contract PTQ-16-08633.

importance of assessment in e-learning environments (Strother, 2002; García-Peñalvo & Seoane-Pardo, 2015). The participative processes oriented to distance learning assessment have been broadly studied (Bozkurt et al., 2016). Nevertheless, participating stakeholders could show inconsistencies, which could be circular or undefined preferences (Brunelli, 2017). In such a situation, special attention must be given to the methodology of these processes in order to maintain objectivity and representation without losing usefulness or efficacy. For some participants, such as students, it is difficult to define individual preferences in the early stages of the decision-making process which makes it necessary to incorporate a high degree of iteration in certain phases of the evaluation process (Owen, 2015). This can be tedious and adds complexity to the process, consuming additional resources (Belton & Steward, 2002). Moreover, reviewing responses or asking for a repetition of responses from the same participant does not guarantee the reduction of inconsistencies. Furthermore, on many occasions it is not possible to do this because of time limitations.

Some studies have utilized pairwise comparisons to assess e-learning systems (Jeong & Yeo, 2014; de Castro et al., 2017). In fact, one of most used methods when designing participative processes is a methodology based on paired comparisons, the Analytical Hierarchy Process (AHP) (Saaty, 1990). The AHP has been applied to educational environments and has been applied to participative processes with users in order to assess e-learning (Ho, 2008; Lin, Ho & Chang, 2014). Recently, studies have been published by Anggrainingsih et al. (2018) that use AHP to evaluate e-learning criteria such as ‘Quality of Design and Material’, and by Mohammed et al. (2018) that apply the same method with more technical criteria. Nevertheless, it is common to obtain a high number of inconsistent primary observations because of the subjective nature of the human mind. To solve this problem, inconsistent responses are generally removed, or the valuations are repeated until they generate results with an acceptable consistency level (Shee & Wang, 2008; Li & Ma, 2007; Lin et al., 2014). The first option, discarding inconsistent results, causes valuable information to be lost and may negatively impact the reliability of the result as it reduces the sample size. The second option, iterating the evaluation process, requires more resources and increases complexity. Some studies evidence positive results for iteration in the participative process as it reduces conflicts and increases consensus. On the other hand, this option is only practical for small groups that are easily managed, with plenty of time, and an intimate knowledge of the evaluation process. However, this iterative process is generally unpractical due to the limited availability of some students.

In this study, we propose the use of weighted goal programming to correct inconsistencies in the primary results and thus avoid information loss without modifying the data collection process (Chen, Kou & Li, 2018). This model allows researchers to obtain consistent results that are as similar as possible to the original results. Therefore, the objective of this work is to evaluate the applicability of using a goal programming model to reduce inconsistencies in participative evaluation systems which collect the preferences of higher education students that study business administration through online courses.

2. Methodology

“AHP is a multi-criteria decision making approach in which factors are arranged in a hierarchic structure” (Saaty, 1990). AHP can measure preferences through pairwise
comparisons to derive priority scales. It is these scales that measure intangibles in relative terms (Saaty, 2001). For these reasons, it has a wide area of applicability and has been successfully used to solve a wide variety of public and private sector decision making problems that require group consensus (Belton & Stewart, 2002). AHP measures individual preferences through judgment evaluations on the relative importance of different paired criteria that are being considered. The decision maker can express the intensity of their preference on a 9-point scale. If two criteria are equally important, they receive a score of 1. A score of 9 indicates that a criterion is extremely preferable over another (Saaty, 2001). Through it, pairwise comparison scores are used to build reciprocal matrices. From these, the relative weights of each attribute are measured. Based on these weights, the different alternatives are ranked.

Pairwise comparison matrices must be reciprocal, homogeneous and consistent (Saaty, 2001). Let $M=(m_{ij})_{ij}$ a pairwise comparison matrix, $M$ verifies the reciprocity condition when $m_{ij}m_{ji} = 1 \forall i \neq j$, and verifies the consistency condition when $m_{ij}m_{jk} = m_{ik} = 0 \forall i \neq j j \neq i$ (González-Pachón & Romero, 2004). Notwithstanding, in the decision making processes the consistency condition is usually not accomplished.

Participants in a decision-making process usually provide inconsistent results because the judgment calls have innate subjectivity. The level of consistency can be measured with the Consistency Index (CI), the cumulative average of matrix inconsistencies. The Consistency Ratio (CR) is the comparison between the CI and the Random Consistency Index (RI). An acceptable CR is equal to or less than 0.10 (Saaty, 2001). AHP inconsistency reduction has been studied in depth using different approaches (Kulakowski, 2018). Khatwani and Kumar (2017) used a stochastic method to define the Cosine Consistency Index. This method is based on a cosine maximization that uses an iterative basis to achieve the most consistent solution. In that case, AHP can be used iteratively until it achieves a consistent ratio.

Some studies have been oriented to deal with inconsistencies to improve the group decision making processes. Fuhua et al. (2010) used two qualitative strategies allocating a weight vector based on the “expert’s experience value”. The main limitations of these methods are twofold: first, the loss of information is important and second, the participants may not feel the decision-making process and final result is their own. Srdjevic et al. (2013) proposed a model to assign the weights to the users in order to obtain consistent results. However, the problem related to the lost information remained. Moreover, the qualitative approach involves subjectivity and bias on the part of the user who is determining the weights.

Ivanco et al. (2017) used sensitivity analysis to improve the consistency of the AHP matrices, taking into account the consensus of the group solution. This method presents more flexibility in order to obtain a consensual solution, however it proposes to address users’ disparities without quantifying them.

Benítez et al. (2014) proposed a linear approach to obtain the closest consistent matrix through a suitable orthogonal projection expressed in terms of a Fourier-like expansion. This method achieves the proposed goal, however, modeling the problem is very complex.
In this paper, we apply a simpler linear approach to optimize the consistency of the AHP matrices based on the Goal Programming method in order to improve the group decision making processes with inexpert participants.

Goal Programming (GP) is a versatile multi-criteria technique used to resolve complex problems. In addition, it has been applied in other management science techniques (Tamiz et al., 1998). GP finds compromise solutions that may not fully satisfy all the goals but do reach certain satisfaction levels set by the decision-maker. For this, an objective function and some constraints are defined. The constraints of the model are formed by the relationship between the objectives of the achievement level for each attribute with these attributes linking themselves through negative and positive deviations. GP can be modelled with different approaches: MinMax GP, Lexicographic GP and Weighted GP. Weighted GP is a linear model that minimizes the weighted sum of the deviations from each goal and provides the most balanced solution. MinMax GP minimizes the maximum deviation between all possible deviations. Lexicographic GP seeks to minimize an achievement function based on a pre-emptive or non-Archimedean priorities approach (Romero, 2014). In this specific case, we applied an Archimedean GP model as laid out by González-Pachón and Romero (2004). With a n×n matrix the model is as follows:

\[
\begin{align*}
\text{Min } & \sum_l (n_l^{(1)} + p_l^{(1)})^p + \sum_s (n_s^{(2)} + p_s^{(2)})^p + \sum_t (n_t^{(2)} + p_t^{(2)})^p \\
\text{s.t. } & w_{ij} - m_{ij} + n_{ij}^{(1)} - p_{ij}^{(1)} = 0, \quad l = 1, 2, \ldots, n(n-1), \\
& w_{ij}w_{ji} + n_{s}^{(2)} - p_{s}^{(2)} = 1, \quad s = 1, 2, \ldots, \frac{n(n-1)}{2}, \\
& w_{ij}w_{jk} - w_{ik} + n_{t}^{(3)} - p_{t}^{(3)} = 0, \quad t = 1, 2, \ldots, n(n-1)(n-2), \\
& 0.11 \leq w_{ij} \leq 9 \quad \forall i, j.
\end{align*}
\]

Where:
- \(n_l^{(1)}\) and \(p_l^{(1)}\) are the negative and positive deviations of the goal, respectively, for constraints that ensure the condition of similarity in the position \(l\).
- \(n_s^{(2)}\) and \(p_s^{(2)}\) are the negative and positive deviations of the goal, respectively, for constraints that ensure the condition of reciprocity in the position \(s\).
- \(n_t^{(3)}\) and \(p_t^{(3)}\) are the negative and positive deviations of the goal, respectively, for constraints that ensure the condition of consistency in the position \(t\).

\(m_{ij}\) are the components of the matrix \(M\) for each pair of criteria.
\(w_{ij}\) are the components of the matrix \(W\), formed by the weights that represent the most similar weights to the components of the original matrix \(M\) for each pair of criteria \(ij\). These are the results of the model.

Let \(M=(m_{ij})_{ij}\) a general matrix given by a student, there exists a set of positive numbers, \((w_1 \ldots w_n)\), such that \(m_{ij} = \frac{w_i}{w_j}\) for every \(i, j = 1, \ldots, n\).

This model uses a distance-based framework approach to inconsistencies in pairwise comparison matrices. The goal is to obtain a matrix that is as similar as possible to the
one generated by the decision maker while meeting Saaty’s conditions of similarity, reciprocity, and consistency (González-Pachón & Romero, 2004).

After correcting inconsistencies in individual pairwise comparison matrices, we aggregated their values by calculating their geometric mean. The resulting matrix represents the collective evaluation of all participants. From this aggregated matrix, we generated the matrix of weights indicating the priorities of each tool in the achievement of the competency under study using the eigenvector method.

3. Application
The process was organized into two steps. First, we collected the evaluations obtained through a survey designed according to Saaty’s (2001) guidelines and corrected the inconsistent matrices. Second, conjoint results were assessed by the students in order to identify the consensus between individual preferences and aggregated results. The high levels of agreement suggest that this method was effective, which improves the entire process by making it more flexible, efficient, and practical.

Students were asked to evaluate four e-learning tools (collaboration tools, content tools, tutorial sessions and evaluation tools) used in an online business degree course, based on how well they helped them acquire a specific competency (Table 1). The competency being evaluated was ‘Ability to work autonomously’, which is especially relevant in online courses.

Table 1
Description of e-learning tools analyzed as criteria in an inquiry

<table>
<thead>
<tr>
<th>E-learning Tools</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaborative</td>
<td>Facilitate the interaction with the professor and among students through chat, messages and a forum.</td>
</tr>
<tr>
<td>Contents</td>
<td>Providing the courses theoretic and practical assignments.</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Allow students to follow the continuous evaluation process through tasks and test type exams.</td>
</tr>
<tr>
<td>Tutorial sessions</td>
<td>To solve course doubts and questions at a one-to-one level. Online or in person.</td>
</tr>
</tbody>
</table>

The GP model was applied in the inconsistent matrices collected, as shown in the example below.

Example: Let \( P=\{c_l, c_l, t, e\} \) be a set of pairwise comparisons that represent the individual preferences of one student about the importance of each type of e-learning tool (collaboratives \( c_l \), contents \( c_t \), tutorials \( t \) and evaluations \( e \)), for the achievement of the competency “Ability to work autonomously” explained in the Table 1. All these preferences were collected using Saaty’s scale.

Also, the matrix \( M \) formed by the cardinal pairwise comparisons over \( P \) is:

\[
M = \begin{pmatrix}
1 & 1/7 & 1 & 7 \\
7 & 1 & 7 & 7 \\
1 & 1/7 & 1 & 1 \\
1/7 & 1 & 1 & 1
\end{pmatrix}
\]
So, the matrix M can be approximated by a reciprocal and consistent model, by using the following GP model:

\[
\begin{align*}
\text{Min} & \sum_{i=1}^{12} \left( n_i \right) + \sum_{s=1}^{6} \left( n_s \right) + \sum_{t=1}^{24} \left( n_t \right), \\
\text{s.t.} & \quad w_{12} - \frac{1}{7} + n_{1} - p_{1} = 0, \quad w_{13} - 1 + n_{2} - p_{2} = 0, \quad w_{14} - 7 + n_{3} - p_{3} = 0, \\
& \quad w_{21} + 7 + n_{4} - p_{1} = 0, \quad w_{23} - 7 + n_{5} - p_{2} = 0, \quad w_{24} + 7 + n_{6} - p_{3} = 0, \\
& \quad w_{31} - 1 + n_{7} - p_{1} = 0, \quad w_{32} - 1/7 + n_{8} - p_{2} = 0, \quad w_{34} - 1 + n_{9} - p_{3} = 0, \\
& \quad w_{41} - 1 + n_{10} - p_{1} = 0, \quad w_{42} - 1/7 + n_{11} - p_{11} = 0, \quad w_{43} - 1 + n_{12} - p_{12} = 0; \\
& \quad w_{12}w_{21} + n_{12} - p_{12} = 1, \quad w_{13}w_{31} + n_{13} - p_{13} = 1, \quad w_{14}w_{41} + n_{14} - p_{14} = 1, \\
& \quad w_{23}w_{32} + n_{23} - p_{23} = 1, \quad w_{24}w_{42} + n_{24} - p_{24} = 1, \quad w_{34}w_{43} + n_{34} - p_{34} = 1; \\
& \quad w_{13}w_{32} - w_{12} + n_{13} - p_{13} = 0, \quad w_{14}w_{42} - w_{12} + n_{14} - p_{14} = 0, \\
& \quad w_{12}w_{23} - w_{13} + n_{13} - p_{13} = 0, \quad w_{14}w_{43} - w_{13} + n_{14} - p_{14} = 0, \\
& \quad w_{12}w_{24} - w_{14} + n_{14} - p_{14} = 0, \quad w_{13}w_{34} - w_{14} + n_{13} - p_{13} = 0, \\
& \quad w_{23}w_{31} - w_{21} + n_{21} - p_{21} = 0, \quad w_{24}w_{41} - w_{21} + n_{24} - p_{24} = 0, \\
& \quad w_{21}w_{31} - w_{23} + n_{23} - p_{23} = 0, \quad w_{24}w_{43} - w_{23} + n_{24} - p_{24} = 0, \\
& \quad w_{21}w_{34} - w_{24} + n_{24} - p_{24} = 0, \quad w_{23}w_{34} - w_{24} + n_{23} - p_{23} = 0, \\
& \quad w_{31}w_{32} - w_{32} + n_{32} - p_{32} = 0, \quad w_{34}w_{41} - w_{32} + n_{34} - p_{34} = 0, \\
& \quad w_{31}w_{34} - w_{34} + n_{34} - p_{34} = 0, \quad w_{32}w_{41} - w_{32} + n_{32} - p_{32} = 0, \\
& \quad w_{31}w_{31} - w_{31} + n_{31} - p_{31} = 0, \quad w_{34}w_{42} - w_{31} + n_{34} - p_{34} = 0, \\
& \quad w_{31}w_{34} - w_{34} + n_{34} - p_{34} = 0, \quad w_{32}w_{41} - w_{32} + n_{32} - p_{32} = 0, \\
& \quad w_{31}w_{31} - w_{31} + n_{31} - p_{31} = 0, \quad w_{34}w_{42} - w_{31} + n_{34} - p_{34} = 0, \\
& \quad w_{31}w_{34} - w_{34} + n_{34} - p_{34} = 0, \quad w_{32}w_{41} - w_{32} + n_{32} - p_{32} = 0; \\
\end{align*}
\]

As a result of the application of the model in the matrix M, we obtained a consistent matrix W:

\[
W= \begin{pmatrix}
1 & 1/7 & 1 & 1 \\
7 & 1 & 7 & 7 \\
1 & 1/7 & 1 & 1 \\
1 & 1/7 & 1 & 1
\end{pmatrix}
\]

was \((m_{14})\). This change only permitted one to obtain a matrix with a CR=0 when the index of the original matrix was CR=0.4192.

This process was applied to improve the consistency of all the matrices that represented the students' preferences, that were obtained using the Saaty survey, with a CR>0.10.

IJAHP Article: de Castro-Pardo, de la Fuente-Cabrero, Laguna-Sánchez, Pérez-Rodriguez/ Combining AHP and goal programming in the context of the assessment of e-learning
4. Results and discussion

Individual opinions were collected from the 72 students through a ‘Saaty-type’ survey with a 1 to 9 scale. In it, they were asked to perform pairwise comparisons of collaboration, content, tutorial sessions, and evaluation tools as a means of acquiring the ‘Ability to work autonomously’ competency. From the resulting 72 pairwise comparison matrices, 7 were excluded for being incomplete or incorrectly completed. Of the remaining 65 valid matrices, 8 had a consistency ratio of less than 0.1. The inconsistent results were corrected by modeling a goal programming function using LINGO 17.0. We obtained improvements in the consistency of 57 matrices. As a result, 65 n=4 matrices were obtained with a consistency ratio under 0.1.

Differences between the weights obtained with corrected inconsistencies and the weights obtained with the original results removing inconsistent answers were not relevant, but these differences changed the final priorities over each e-learning tool (Table 2).

Table 2
Results and differences between results with corrected inconsistencies and results with original results removing inconsistent answers

<table>
<thead>
<tr>
<th>E-learning tools</th>
<th>Consistent results without corrections</th>
<th>Results with corrected inconsistencies (de Castro et al., 2017)</th>
<th>Differences in percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaborative</td>
<td>20.33%</td>
<td>18.91%</td>
<td>1.42%</td>
</tr>
<tr>
<td>Contents</td>
<td>32.23%</td>
<td>32.53%</td>
<td>-0.30%</td>
</tr>
<tr>
<td>Evaluation</td>
<td>16.84%</td>
<td>21.89%</td>
<td>-5.05%</td>
</tr>
<tr>
<td>Tutorial sessions</td>
<td>30.59%</td>
<td>26.66%</td>
<td>3.93%</td>
</tr>
</tbody>
</table>

Table 3
Results and differences between results with corrected inconsistencies and results with original results considering inconsistent answers

<table>
<thead>
<tr>
<th>E-learning tools</th>
<th>Results with inconsistent and consistent matrices</th>
<th>Results with corrected inconsistencies (de Castro et al., 2017)</th>
<th>Differences in percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaborative</td>
<td>19.05%</td>
<td>18.91%</td>
<td>0.14%</td>
</tr>
<tr>
<td>Contents</td>
<td>34.67%</td>
<td>32.53%</td>
<td>2.14%</td>
</tr>
<tr>
<td>Evaluation</td>
<td>25.39%</td>
<td>21.89%</td>
<td>3.5%</td>
</tr>
<tr>
<td>Tutorial sessions</td>
<td>20.87%</td>
<td>26.66%</td>
<td>-5.79%</td>
</tr>
</tbody>
</table>

The results after correcting the inconsistencies give more weight to evaluation tools and less to collaborative tools. Collaborative tools are prioritized in the same order with just 1.42% less relative importance than the original results. Content tools are similarly ranked in the original and corrected versions (Table 2).
Furthermore, when the original inconsistent and consistent results were compared with the results, after correcting the inconsistencies, differences were found. Tutorial sessions showed the highest divergences when the original inconsistencies remained (Table 3). This shows the effect of the inconsistent results over the group solution.

Table 4
Ranking provided by the inconsistent and consistent original results, the original results removing inconsistent answers and the results with corrected inconsistencies

<table>
<thead>
<tr>
<th>E-learning tools</th>
<th>Ranking inconsistent and consistent results</th>
<th>Ranking only consistent results</th>
<th>Ranking with corrected inconsistencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaborative</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Contents</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Evaluation</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Tutorial sessions</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

These differing prioritizations between the results after discarding the inconsistent preferences and the results with corrected inconsistencies, illustrate the effect of information loss on results (Table 4). A sensitivity analysis performed on the aggregated results shows the importance of evaluation tools in acquiring the competency under study, providing the most robust prioritization (Figure 1).

Figure 1 Graphic of sensitivity analysis with variation of one positive point in each pair of aggregated comparison matrices
Furthermore, the differing prioritizations between the results considering inconsistencies, the results with corrected inconsistencies and the results after discarding inconsistent preferences show the effect of inconsistent responses on distorting global results (Table 4).

The goal programming model proposed in this work consolidates the weights of content and evaluation tools, thus reducing the sensitivity of the overall results. The data suggests that information loss can distort evaluation results and diminish the quality of the process.

After this phase, students were asked to complete an online, Likert-scale survey that focused on their level of agreement or disagreement with the results, where 1 represented the minimum agreement and 5 represented the maximum agreement. Here, 87.5% of respondents agreed highly or very highly with the priority ranking generated by the aggregated matrices corrected for inconsistencies. This high level of agreement demonstrated the effectiveness of the proposed model to treat inconsistencies in pairwise comparison matrices of e-learning tools for acquiring competencies. Further validation was provided by high levels of participant satisfaction with the aggregated results. This would seem to suggest that the changes carried out to diminish inconsistencies did not significantly alter the opinion of the group.

Finally, the model allowed researchers to recover 90.47% of the missing information while maintaining the flexibility of the evaluation process; thus, making it more practical. In addition, the high level of agreement from the participants with the results validates the effectiveness of this method.

5. Conclusion

Global results are different when consistency is improved using the proposed GP model. Thus, e-learning tools received different weights when inconsistencies were corrected. Both results, corrected and primary, agree with the assessment that content and tutorial sessions are the most important elements, even though tutorial sessions received a lower weight with the corrected matrices. Notably, the corrected model prioritizes evaluation tools over collaborative tools.

GP is an effective technique when correcting inconsistencies in pairwise comparison matrices as applied to higher education evaluation systems. By correcting the primary results, both the quality and the agility of the evaluation process are improved. The GP model has improved the performance of the AHP method in order to reduce the inconsistencies of the pairwise comparison matrices, solving some of the limitations of previously proposed methods. First, the loss of information has been avoided. Second, the preferences of all the participants have been considered in the decision-making process. Third, the applicability of the process has remained, thus avoiding the use of iterations. Finally, rigor has been maintained throughout the process.

Ultimately, the use of GP in the proposed model efficiently improved the Analytical Hierarchy Process in the context of working with inexpert users such as those who might evaluate an online business degree course.
REFERENCES


Mathematical Modelling, 37(10-11), 6670-6682. Doi: https://doi.org/10.1016/j.apm.2013.01.028

Strother, JB (2002). An assessment of the effectiveness of e-learning in corporate training programs. The International Review of Research in Open and Distributed Learning, 3(1).