# Performance Evaluation of Faculty Departments by a Delphi Method Based on 2-Tuple fuzzy Linguistic Representation Model and TOPSIS

Oğuz TORAĞAY[a], Murat ARIKAN[b],
[a, b] Engineering Faculty, Gazi University, 06570, Ankara, TURKEY
Corresponding author, Tel:+903125823808; e-mail adress: marikan@gazi.edu.tr

***Abstract*--** **The development of and competition in educational facilities gradually increase the service quality's importance. To accommodate this rapid process, educational organisations attempt to measure their performance and to enhance their standards. In general, an organisation's performance does not depend solely on one criterion; instead, it should be evaluated based on multiple criteria. In this study, the academic performances of the departments within the Engineering Faculty of one of the largest universities in Turkey, Gazi University, have been compared using a multi-attribute decision making (MADM) method, TOPSIS. The criteria weights for the TOPSIS method are determined dependent upon linguistically expressed expert opinions. Therefore, a Delphi method based on the 2-tuple fuzzy linguistic representation model for computing with words is proposed. A sensitivity analysis study is also performed to determine the most critical criterion.**

***Index Term*--** **Multi Attribute Decision Making, TOPSIS, Delphi method, performance evaluation, education.**

## 1. INTRODUCTİON

To maintain pace with today's innovation speed, it is very important to conserve the limited resources and use them in the most efficient manner without waste. As a result of economic growth and in the light of science and technology, economic power and independence have an intimate relation with the efficient utilization of a country's industrial potential. Therefore, universities, which are the various branches' sources of science production, become important. The higher education system in every country is considered a factor that affects sustainable development. This is because universities and cenres of higher education provide the education to produce skilled and experienced manpower, which is the basis of a country's comprehensive development (Ahmadi et al., 2011). Therefore, it is necessary to evaluate higher education institutions and control resources to increase the general success of academic programs and to meet their missions. Consequently, performance evaluation of universities or units such as faculties and departments and the comparison of these will guide researchers to achieve the goals of improving educational quality and productivity in higher education centres.

In many studies, the literature shows comparisons between universities and/or various departments and units in a particular university. In certain studies, the data envelopment analysis (DEA) method has been used to focus on the efficiency of the units which have more than one input and output (Tomkins and Green, 1988; Sinuany-Stern et al., 1994; Beasley, 1995; Johnes and Johnes, 1995; Leitner et al., 2007; Rayeni et al., 2010). However, this method is not able to fully rank the alternatives, and it classifies the units into two categories, efficient and inefficient; moreover, all efficient units are equally suitable (Rouyendegh and Erol, 2009). To compensate for this gap, there are hybrid approaches in which the DEA was combined with other methods such as ANP (Rouyendegh and Erol, 2009) and Fuzzy TOPSIS (Rouyendegh, 2011). In addition, in certain studies, universities have been evaluated by multi criteria decision making methods using a set of performance indicators. Wu et al. (2011) and Zolfani and Ghadikolaei (2013) selected critical performance indicators based on the balanced score card perspectives by a survey study to evaluate units. The researchers determined the cause and effect relations between those perspectives via DEMATEL and then applied ANP to calculate the relative weights of the performance indicators. Finally, VIKOR is used for ranking and comparing the considered alternatives. Wu et al. (2012), after indicating evaluation indices via a survey, used a hybrid MCDM model of AHP and VIKOR methods to rank 12 private universities in Taiwan.

In this paper, the objective is the evaluation and ranking of various departments in the engineering faculty of a large university in Turkey. Therefore, one MADM method, TOPSIS, is used. To maintain continuous improvement via performance measurement, the determination of performance indicators and their weights is critical. Here, performance indicators are selected from the related literature and to determine their associated weights; the Delphi method is employed to ensure experts' opinions can be considered. Expert opinions are expressed by linguistic terms and to adress linguistic variables without loss of information; the 2-tuple fuzzy linguistic representation model, which is proposed by Herrera and Martinez (2000), is used for computing with words. To our knowledge there is no research study that intends to determine the crisp normalized numeric weights of the performance indicators with the Delphi method based on the 2-tuple fuzzy linguistic representation model. After the consensus is achieved using the Delphi method, the numerical equivalences of 2-tuple representation of expert opinions are used to calculate normalized crisp weights, which will form an input to TOPSIS. The proposed Delphi method can easily be implemented by practitioners whose objective is to determine criteria weights when the expert opinions regarding the importance level are expressed linguistically. Additionally, a sensitivity analysis study is performed to specify the most critical criterion. In the next section of this paper, the proposed Delphi method, which is based on the 2-tuple fuzzy linguistic representation model and the TOPSIS method, are explained. The third section involves implementation of the proposed approach and the sensitivity analysis performed. In the last section, the results are summarized.

## 2. METHODOLOGY

### 2.1. Fuzzy Delphi Method

The Delphi method which was developed by Norman Delkay and Olaf Helmer in 1963, is a repetitive procedure employed to gather and distil expert opinions using a series of questionnaires interspersed with feedback. The questionnaires are designed to focus on problems, opportunities, solutions or forecasts. Each subsequent questionnaire is developed based on the result of the previous questionnaires. The procedure is stopped when the research question is answered, for example, when consensus is reached, theoretical saturation is achieved, or when sufficient information has been exchanged. The Delphi method has its origins in the American business community and has since been widely accepted throughout the world in many industry sectors, including health care, defence, business, education, information technology, transportation and engineering (Skulmoski etal., 2007, p. 2).

Consensus of the experts is considered as the main goal of a Delphi study. Researchers have used many different measures to determine when a sufficient level of agreement among the expert panel has been attained (von der Gracht, 2008, p. 51-52). Certain consensus measures are based on measures of statistical dispersion, such as the range, interquartile range, standard deviation, and coefficient of (relative) variation. The range is the simplest measure of dispersion because it is easily calculated as the difference between the lowest and the highest score in a distribution. The range changes with the values of the extreme scores. The interquartile range (IQR) is equal to the difference between the upper and lower quartiles; in other words, the IQR is the 1st quartile subtracted from the 3rd quartile. Thus, an IQR of less than 1 means that more than 50% of all opinions fall within 1 point on the scale. This is a frequently used measure in Delphi studies, and it is generally accepted as an objective and rigorous means of determining consensus. The coefficient of variation is a dimensionless number defined as the ratio of the standard deviation to the arithmetic mean. It is accepted that a coefficient of variation at or below 0.5 indicates reasonable internal agreement. Another measure is associated with mean and standard deviation. Standard deviation is a measure of dispersion for the mean and is usually considered together with the mean. In certain studies, a range of mean +/- 1.64 standard deviation is used as the consensus criterion (von der Gracht, 2008, p. 57). A comprehensive investigation of consensus measures can be found in von der Gracht (2008).

Although the Delphi method has been widely applied in many different areas, the traditional Delphi method has been criticized for the low convergence in generating results, the long process of interrogation, and the loss of valuable information from expert opinions (Chang et al., 2011). Many researchers adapted the method for use in a fuzzy environment**.** Murray et al. (1985) proposed an application of Fuzzy Theory to the Delphi Method. These researchers used the membership degree to determine the membership function of each expert. Ishikawa et al. (1993) used the advantage of the Maximum-Minimum Method in combination with cumulative frequency distribution and fuzzy scoring to compile expert opinions into fuzzy numbers. The expert prediction interval value was then used to derive the fuzzy numbers, resulting in the Fuzzy Delphi Method (FDM). Chang et al. (1995) estimated project activity times using FDM. These researchers employed triangular fuzzy numbers to represent the pessimistic, moderate and optimistic expert opinions associated with the times of each activity. Hsu and Chen (1996) proposed the fuzzy similarity aggregation method to combine individual subjective estimates, which are represented by positive trapezoidal fuzzy numbers. The Delphi method is utilized to gather the positive trapezoidal fuzzy numbers of each expert's estimate. Using the similarity function, similarities between experts were collated and fuzzy numbers were assigned directly to each expert to determine the agreement degree between them. The consensus coefficient was then used to aggregate all experts' fuzzy evaluation values. Kuo and Chen (2008) used a Fuzzy Delphi Method to select performance appraisal indicators for service industries' mobility. In this study, expert opinions are collected from questionnaires, and triangular fuzzy numbers are created. For each indicator, lower, upper and mid-levels of expert opinions are established as the minimum, maximum and geometric mean of all experts' appraisal values. The geometric mean of each indicator's triangular fuzzy number was used to denote the expert group's consensus regarding the indicator's appraisal value. In many real-world cases, experts' judgments cannot be properly reflected in quantitative terms. When experts are not able to provide exact numerical values to express their opinions, a more realistic alternative option is to use linguistic assessments instead of numerical values (Fasanghari and Montazer, 2008). Certain researchers have translated those linguistic variables that are used to gather experts' opinions into triangular (Büyüközkan and Ruan, 2008; Hsu et al., 2010) or into trapezoid fuzzy numbers (Cheng and Lin, 2002). It is unavoidable that some information has been lost during this transfer. In this paper to avoid information loss, a different representation model proposed by Herrera and Martinez (2000), a 2-tuple model, is used in the Delphi method. In this model, every linguistic variable demonstrated by a representation composed of two parts ($s$, $\alpha$); the first part, $s$ includes the linguistic variable and the second part, $\alpha$ denotes a number belongs to the [-0.5, 0.5) set. The main advantage of this representation is a continuous domain. Therefore, this representation can express any counting of information in the discourse's universe (Wang, 2010). The Delphi method steps, which are employed to determine the weights of the evaluation criteria in this study, are provided as follows.

*Step 1:* Construct a committee of $K$ experts and determine the evaluation criteria set. Let the evaluation criteria set contain $M$ constituents $(m = 1, \dots, M)$.

*Step 2:* Identify the linguistic terms $s_i \in S = \{s_0, \dots, s_g\}$ for the importance weight of criteria; $g+1$ is the cardinality of set $S$. For each evaluation criterion, collect the opinions of $K$ experts based on the linguistic term set.

*Step 3:* Let $\gamma_{km} \in (0, 1, \dots, g)$ be the numerical equivalence of the $k^{th}$ expert's linguistically expressed opinion on the $m^{th}$ evaluation criterion. By considering $\gamma_{km}$, specify the level of consensus with respect to a certain dispersion measure. If a consensus is attained among the participants, STOP; otherwise calculate the average importance weight of the $m^{th}$ criterion regarding the opinions collected from the experts ( $\beta_m = \frac{1}{K}\sum_{k=1}^{K}\gamma_{km}$, $(\beta_m \in [0,\ g])$). Obtain the 2-tuple that represents the equivalent information for $\beta_m$ as follows:

$$\Delta: [0, g] \rightarrow S \times [-0.5, 0.5)$$

$$\Delta(\beta_m) = \left\{ \begin{matrix} s_i & i = round(\beta_m) \\ \alpha = \beta_m - i & \alpha = [-.5, .5) \end{matrix} \right\}$$

*Step 4:* Return the 2-tuple information associated with each criterion to the committee of $K$ experts for re-examination.

*Step 5:* To evaluate the answers of the new round return to Step 3.

In the Delphi method above, the linguistic variables concerning the expert opinions on the criteria set's importance are transformed into numerical values $\in(0, 1,..,g)$. If the collected opinions do not reflect a degree of consensus, the numerical average of linguistically expressed opinions of the expert panel related to each criterion ($\beta_m$) is calculated as an arithmetic mean. These average numerical equivalences are translated into the 2-tuple information composed of the linguistic variable ($s_i$) and the distance from the closest linguistic variable ($\alpha_i \in[-0,5,\ 0.5)$), and the obtained 2-tuple information regarding the average importance level of criteria set are returned to the expert panel. The 2-tuples are very understandable, enabling interpretation without loss of information; by considering the results of the previous Delphi round, the experts can modify or retain their opinions. The main advantage of the proposed Delphi method is the convenience of its application by practitioners, although the participants are able to express their opinions linguistically.

### 2.2. TOPSIS method

Hwang and Yoon (1981) developed the *Technique for Order Preference by Similarity to Ideal Solution* (TOPSIS) method based on the concept that the chosen alternatives should have the shortest distance from the positive-ideal solution and the longest distance from the negative-ideal solution. Formally, for a Multi-Attribute Decision Making (MADM) problem with $m$ alternatives that are evaluated by $n$ attributes (criteria), the positive-ideal solution is denoted as:

$A^* = (x_1^*, \dots, x_j^*, \dots, x_n^*)$

where $x$ is the best value for the $j^{th}$ attribute among all available alternatives. Then, the negative-level solution is given as:

$A^- = (x_1^-, \dots, x_j^-, \dots, x_n^-)$

where $x_j^-$ is the worst value for the $j^{th}$ attribute among all available alternatives.

The method's steps are presented as follows:

<u>Step 1:</u> Calculate the normalized ratings.

The objective of this process is to transform the various attribute dimensions into nondimensional attributes, which allows comparisons among the attributes. Each element of the normalized decision matrix ($R$) is calculated as

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}} \qquad (1)$$

In this equation, $r_{ij}$ and $x_{ij}$ show the quantity of normalized decision matrices and the numerical outcome of the $i^{th}$ alternative with respect to $j^{th}$ criterion, respectively.

<u>Step 2:</u> Calculate the weighted normalized ratings.

These values are calculated as

$$v_{ij} = w_j r_{ij} \qquad (2)$$

where $w_j$ is the weight of $j^{th}$ attribute.

<u>Step 3:</u> Identify positive-ideal and negative-ideal solutions.

$A^*$ and $A^-$ are defined in terms of weighted normalized values:

$A^* = \{ v_1^*, v_2^*, \dots, v_j^*, \dots, v_n^* \} = \{(max\ v_{ij} \mid j \in j_1), (min\ v_{ij} \mid j \in j_2) \mid i=1, \dots, m \}$

$A^- = \{ v_1^-, v_2^-, \dots, v_j^-, \dots, v_n^- \} = \{(min\ v_{ij} \mid j \in j_1), (max\ v_{ij} \mid j \in j_2) \mid i=1, \dots, m \}$

In the aforementioned sets, $J_1$ shows the benefit criteria set, and $J_2$ shows the cost criteria set.

<u>Step 4:</u> Calculate the separation measures.

The separation of each alternative from the positive-ideal solution, $A^*$, is given by

$$S_i^* = \sqrt{\sum_{j=1}^{n}\left(v_{ij} - v_j^*\right)^2}, \quad i=1,2,\dots,m \qquad (3)$$

Similarly, the separation from the negative-ideal solution, $A^-$, is given by

$$S_i^- = \sqrt{\sum_{j=1}^{n}\left(v_{ij} - v_j^-\right)^2}, \quad i=1, 2\dots m \qquad (4)$$

<u>Step 5:</u> Calculate similarities to the positive-ideal solution.

$$C_i^* = \frac{s_i^-}{(s_i^* + s_i^-)}, \qquad i=1, 2 \ldots m \qquad (5)$$

*Step 6*: Rank the preference order

Choose an alternative with the maximum $C_i^*$ or rank alternatives according to $C_i^*$ in descending order.

## 3. IMPLEMENTATİON

### 3.1. The implementation institution

In this study, the academic performance of the departments within the Engineering Faculty of Gazi University is evaluated. Gazi University is one of the largest universities in Turkey; it is located in the capital city, Ankara. The faculty is built on a 70000 square meter area and composed of three separate but interrelated blocks, including administrative, educational and laboratory buildings. There are 12 amps, 25 classrooms, 44 laboratories, 10 seminar rooms and a convention centre with two conference halls. The faculty library was built on a 1400 square meter area and continues to serve in cooperation with the University's Central Library. The Engineering Faculty consists of six departments, including industrial engineering, chemical engineering, mechanical engineering, electrical and electronics engineering, civil engineering and computer engineering. Each department executes a specific program. In the electric and electronics department, after the third year, the students proceed with their education either in the electronics-communication or electric subsections. However, the graduates obtain the same diploma. Computer engineering is a relatively new department and began accepting students in the 2003-2004 academic year.

### 3.2. Performance Criteria Set and determination of their weights

In the literature on academic efficiency field of study, many different criteria have been used in various research (Rouyendegh, 2009; Rouyendegh and Erol, 2009; Ateş et al., 2006). From these criteria, 12 criteria appropriate for this study have been determined.

The list of selected criteria is as follows; 1) number of students per academic staff; 2) educational area per student; 3) number of SCI papers per academic staff; 4) annual budget per student; 5) teaching hours per academic staff; 6) number of graduated students from the graduate programs in any department; 7) undergraduate grade point average (GPA); 8) number of annual projects (as an indicator of the relation with industry); 9) number of organized international or national conferences, symposiums, panels or meetings per department, 10) percentage of the resources allocated to academic research or R&D in annual budget; 11) percentage of the students who are not able to graduate in their supposed semester; and 12) average duration of the education for undergraduate students.

To specify the weights of the foregoing criteria set to be used in TOPSIS, a Delphi method is implemented that allows expert opinions to be considered. After two rounds of the Delphi method, consensus is achieved. In the first round, a survey of the selected criteria is conducted among 10 experts who are chosen from the department chairs and experienced academicians. In addition to the criteria listed above, the participants have the opportunity to note their own important criterion. The results of the first round are analysed, and their summary is returned to the expert panel in addition to the survey's second round to determine whether there is a change in their opinion. After evaluating the second round results, it is determined that the consensus level is satisfactory; hence, the procedure is halted. The first and second round results are shown in table 1. During the implementation of the Delphi method, the participants are asked to note their opinion on the importance of the performance criteria by choosing the most appropriate linguistic variable from the following 7-point scale.

Linguistic variable set = {$s_0$: none, $s_1$: very low, $s_2$: low, $s_3$: medium, $s_4$: high, $s_5$: very high, $s_6$: precise}

The linguistic terms, none, very low, low, medium, high, very high, and precise correspond to numerical values 0, 1, 2, 3, 4, 5, and 6, respectively.

Table I
Results of the first and second Delphi rounds.

| | First Round | | Second Round | |
|---|---|---|---|---|
| Criteria | Average ($\beta_m$) | Linguistic label and the distance | Average ($\beta_m$) | Linguistic label and the distance |
| 1 | 5,2 | (VERY HIGH, 0,2) | 4,9 | (VERY HIGH, -0,1) |
| 2 | 3,5 | (HIGH, -0,5) | 3,4 | (MEDIUM, 0,4) |
| 3 | 4,9 | (VERY HIGH, -0,1) | 4,8 | (VERY HIGH, -0,2) |
| 4 | 4,3 | (HIGH, 0,3) | 4,2 | (HIGH, 0,2) |
| 5 | 4,6 | (VERY HIGH, -0,4) | 4,4 | (HIGH, 0,4) |
| 6 | 4,4 | (HIGH, 0,4) | 4,4 | (HIGH, 0,4) |
| 7 | 3,6 | (HIGH, -0,4) | 4 | (HIGH, 0) |
| 8 | 4,7 | (VERY HIGH, -0,3) | 4,7 | (VERY HIGH, -0,3) |
| 9 | 3,7 | (HIGH, -0,3) | 3,9 | (HIGH, -0,1) |
| 10 | 5 | (VERY HIGH, 0) | 5 | (VERY HIGH, 0) |
| 11 | 2,6 | (MEDIUM, -0,4) | 2,7 | (MEDIUM, -0,3) |
| 12 | 2,7 | (MEDIUM, -0,3) | 3 | (MEDIUM, 0) |

The distance score in Table I represents the distance between the average expert opinion and the nearest linguistic variable in the 7-point scale by a numeric value in the [-0.5, +0.5) range. A distance score near 0 indicates that the average group opinion is not far from the associated linguistic variable. In addition, the distance value implies the closeness of the average expert opinion to the associated variable. For example, let the average group opinion value on two different criteria be (high, -0,33) and (very high, +0,49); then, these values can be illustrated on a seven point scala as shown in figure 1.

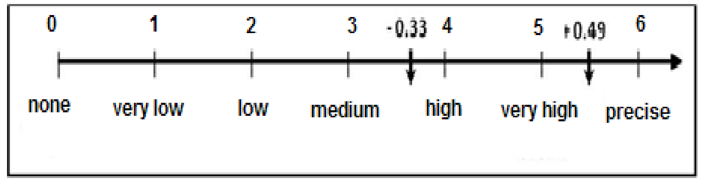


Fig. 1.Example for symbolic translation of the values

Although the linguistic variables are naturally vague, the numerical equivalences of the linguistically expressed expert opinions are used as crisp numbers. To observe whether consensus has occurred, one measure noted in Section 2.1 can be employed on the numerical values of linguistic variables collected from each expert ($\gamma_{km}$). The results with respect to the Coefficient of Variation consensus measure are shown in Table II.

Table II
The results with respect to the Coefficient of Variation consensus measure for two rounds of Delphi

| | First round | | | Second round | | |
|---|---|---|---|---|---|---|
| Criterion | Standard deviation | Mean | Coefficient of variation | Standard deviation | Mean | Coefficient of variation |
| 1 | 1,032796 | 5,2 | 0,198615 | 0,875595 | 4,9 | 0,178693 |
| 2 | 1,354006 | 3,5 | 0,386859 | 0,699206 | 3,4 | 0,205649 |
| 3 | 0,737865 | 4,9 | 0,150585 | 0,632456 | 4,8 | 0,131762 |
| 4 | 0,948683 | 4,3 | 0,220624 | 0,632456 | 4,2 | 0,150585 |
| 5 | 0,966092 | 4,6 | 0,21002 | 0,699206 | 4,4 | 0,15891 |
| 6 | 1,505545 | 4,4 | 0,342169 | 1,264911 | 4,4 | 0,28748 |
| 7 | 1,837873 | 3,6 | 0,51052 | 1,247219 | 4 | 0,311805 |
| 8 | 0,948683 | 4,7 | 0,201848 | 0,674949 | 4,7 | 0,143606 |
| 9 | 1,337494 | 3,7 | 0,361485 | 0,994429 | 3,9 | 0,254982 |
| 10 | 0,942809 | 5 | 0,188562 | 0,666667 | 5 | 0,133333 |
| 11 | 1,429841 | 2,6 | 0,549939 | 1,05935 | 2,7 | 0,392352 |
| 12 | 1,418136 | 2,7 | 0,525236 | 1,154701 | 3 | 0,3849 |

For instance, when criterion 1 is considered in the first Delphi method round, the coefficient of variation (CoV) is 0.198615 (=1.32796/5.2), although it decreases to 0.178693 in the second round. Moreover, it is observed that the coefficient of variation values are improved for all criteria in the second round. If all of the CoV values are ≤0.5, a suitable degree of consensus is achieved, and for Delphi research, no additional round is required (von der Gracht, 2008, p. 58). After implementing of the Delphi method, the numerical equivalences of linguistic importance levels ($\beta_m$) of the criteria are transformed into normalized weights ($w_m$) using equation (6); these are provided in Table 3. Although the TOPSIS method can be implemented without normalizing the weights, the normalization procedure allows convenient comparison of the relative importance of the criteria.

$$w_m = \frac{\beta_m}{\sum_{m=1}^{M} \beta_m} \quad (6)$$

### 3.3. Gathered data about the institution

Raw data, which are gathered to quantify the determined criteria set, are provided in Table 4. The pertinent explanations are as follows:

- All of the data belong to the same academic year.
- The number of students is the sum of the registered students for each department.
- Faculty members consist of all full-time and part-time professors, associate professors, asistant professors and instructors.
- Because the annual budget of each faculty is determined collectively, the allocated budget for each department is calculated as the sum of relevant salaries and allowances.
- GPA is the arithmetical mean of all lectured course grades in the given academic year for the associated department.
- The number of projects includes the projects that have been started in the related academic year.
- Symposiums include the number of national and international meetings, symposiums, and conferences organized by the department in the given academic year.
- The percentage of the students who could not graduate on time column represents the ratio of the students that completed eight semesters but still had not graduated.

- The average graduation duration represents the average number of semesters of the students who graduated from the related department at the end of the given academic year.
- Weekly lecture time data demonstrate the weekly teaching hours per academic staff.
- Educational area represents the total departmental space used for education, such as classrooms and laboratories, in terms of square meters.

Table III
Weights of criteria and their normalized weights obtained from Delphi method

| Criterion | Delphi weight | Normalized weight |
|---|---|---|
| 1 | 4,9 | 0,0992 |
| 2 | 3,4 | 0,0688 |
| 3 | 4,8 | 0,0972 |
| 4 | 4,2 | 0,0850 |
| 5 | 4,4 | 0,0891 |
| 6 | 4,4 | 0,0891 |
| 7 | 4 | 0,0810 |
| 8 | 4,7 | 0,0951 |
| 9 | 3,9 | 0,0789 |
| 10 | 5 | 0,1012 |
| 11 | 2,7 | 0,0547 |
| 12 | 3 | 0,0607 |
| | Total | 1,0000000 |

Table IV
Raw data concerning the departments

| | Number of students | Number of faculty members | SCI papers | Annual budget | GPA | Number of projects | Symposiums | Percentage of students who could not graduate on time | Average graduation duration | Graduate students numbers | Weekly lecture time | Educational area |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mechanical | 996 | 29 | 44 | 908942.64 | 57.3875 | 22 | 1 | 0.293 | 10.8973 | 7 | 17 | 2700 |
| Industrial | 746 | 21 | 41 | 598391.4 | 65.5027 | 10 | 1 | 0.895 | 9.6475 | 25 | 19.55 | 1030 |
| Computer | 271 | 7 | 11 | 164644 | 68.1294 | 3 | 2 | 0.676 | 9.3637 | 3 | 15.2 | 350 |
| Chemistry | 824 | 30 | 51 | 944902.44 | 58.0131 | 17 | 0 | 0.674 | 10.7114 | 2 | 18 | 2311 |
| Civil | 579 | 24 | 27 | 720885.12 | 60.2903 | 32 | 3 | 0.786 | 11.8688 | 21 | 17 | 2816 |
| Electric-Electronic | 431 | 12 | 21 | 339450 | 66.1068 | 10 | 0 | 0.733 | 10.3731 | 19 | 10.6 | 1030 |

### 3.4. Implementation of TOPSIS method

By using the gathered information regarding the departments in the Engineering Faculty, the criteria set is quantified as provided in Table 5. The data related to criterion 10 (percentage of the resources allocated to academic research or R&D in the annual budget) could not be obtained; hence, the associated values for all departments are considered equally. Therefore, it is guaranteed that the following steps and the sequence of the departments are not affected.

Table V
Decision matrix

| Criteria | | Alternatives(Departments) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Mechanical | Industrial | Computer | Chemical | Civil | Electronic |
| | 1 | 34.345 | 35.524 | 38.714 | 27.467 | 24.125 | 35.917 |
| | 2 | 2.711 | 1.348 | 1.292 | 2.805 | 4.864 | 2.390 |
| | 3 | 1.517 | 1.952 | 1.571 | 1.70 | 1.125 | 1.750 |
| | 4 | 912.6 | 802.1 | 607.5 | 1146.7 | 1245.0 | 787.6 |
| | 5 | 17 | 19.55 | 15.2 | 18 | 17 | 10.6 |
| | 6 | 7 | 25 | 3 | 2 | 21 | 19 |
| | 7 | 57.3875 | 65.5027 | 68.1294 | 58.0131 | 60.2903 | 66.1068 |
| | 8 | 22 | 10 | 3 | 17 | 32 | 10 |
| | 9 | 1 | 1 | 2 | 0 | 3 | 0 |
| | 10 | 5000 | 5000 | 5000 | 5000 | 5000 | 5000 |
| | 11 | 29.3 | 89.5 | 67.6 | 67.4 | 78.6 | 73.3 |
| | 12 | 10.8973 | 9.6475 | 9.3637 | 10.7114 | 11.8688 | 10.3731 |

The TOPSIS method steps are as follows:

Step 1: Construct the normalized decision matrix. By utilizing the normalization procedure, various criteria dimensions are converted into non-dimensional criteria (Table 6).

Step 2: The weights in Table 3 are used with the decision matrix in Table 6 to generate the weighted normalized matrix (Table 7).

Step 3: The positive and negative ideal solution are determined.

- The positive ideal solution consists of the largest values among all alternatives for any of the benefit criteria and the smallest values among all of the alternatives for any of the cost criteria,
- The negative ideal solution consists of the smallest values among all alternatives for any of the benefit criteria and the largest values among all of the alternatives for any of the cost criteria,

For the problem under consideration, benefit criteria set={2, 3, 6, 7, 8, 9, 10} and the cost criteria set={1, 4, 5, 11, 12}. The positive ideal solution and the negative ideal solution are shown in Table VIII.

Table VI
Normalized decision matrix

| Criteria | | Alternatives(Departments) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Mechanical | Industrial | Computer | Chemical | Civil | Electronic |
| | 1 | 0,42382 | 0,43837 | 0,47775 | 0,33895 | 0,29771 | 0,44322 |
| | 2 | 0,39065 | 0,19897 | 0,18612 | 0,40416 | 0,70087 | 0,34438 |
| | 3 | 0,38169 | 0,49115 | 0,39532 | 0,42766 | 0,28301 | 0,44024 |
| | 4 | 0,39526 | 0,34742 | 0,26314 | 0,49667 | 0,53925 | 0,34112 |
| | 5 | 0,42138 | 0,48459 | 0,37676 | 0,44617 | 0,42138 | 0,26274 |
| | 6 | 0,18141 | 0,64788 | 0,07775 | 0,05183 | 0,54422 | 0,49239 |
| | 7 | 0,37359 | 0,42642 | 0,44352 | 0,37766 | 0,39249 | 0,43035 |
| | 8 | 0,49120 | 0,22327 | 0,06698 | 0,37956 | 0,71447 | 0,22327 |
| | 9 | 0,25820 | 0,25820 | 0,51640 | 0,00000 | 0,77460 | 0,00000 |
| | 10 | 0,40825 | 0,40825 | 0,40825 | 0,40825 | 0,40825 | 0,40825 |
| | 11 | 0,17050 | 0,52081 | 0,39337 | 0,39220 | 0,45738 | 0,42654 |
| | 12 | 0,42331 | 0,37476 | 0,36374 | 0,41609 | 0,46105 | 0,40295 |

Table VII
Weighted normalized decision matrix

| | | Alternatives(departments) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Mechanical | Industrial | Computer | Chemical | Civil | Electronic |
| Criteria | 1 | 0,04204 | 0,04348 | 0,04739 | 0,03362 | 0,02953 | 0,04396 |
| | 2 | 0,02689 | 0,01369 | 0,01281 | 0,02782 | 0,04824 | 0,02370 |
| | 3 | 0,03709 | 0,04772 | 0,03841 | 0,04155 | 0,02750 | 0,04278 |
| | 4 | 0,03361 | 0,02954 | 0,02237 | 0,04223 | 0,04585 | 0,02900 |
| | 5 | 0,03753 | 0,04316 | 0,03356 | 0,03974 | 0,03753 | 0,02340 |
| | 6 | 0,01616 | 0,05771 | 0,00692 | 0,00462 | 0,04847 | 0,04386 |
| | 7 | 0,03025 | 0,03453 | 0,03591 | 0,03058 | 0,03178 | 0,03485 |
| | 8 | 0,04673 | 0,02124 | 0,00637 | 0,03611 | 0,06798 | 0,02124 |
| | 9 | 0,02038 | 0,02038 | 0,04077 | 0,00000 | 0,06115 | 0,00000 |
| | 10 | 0,04132 | 0,04132 | 0,04132 | 0,04132 | 0,04132 | 0,04132 |
| | 11 | 0,00932 | 0,02847 | 0,02150 | 0,02144 | 0,02500 | 0,02331 |
| | 12 | 0,02571 | 0,02276 | 0,02209 | 0,02527 | 0,02800 | 0,02447 |

Table VIII
The positive and negative ideal solutions

| | | Positive ideal solution | Negative ideal solution |
|---|---|---|---|
| Criteria | 1 | 0,02953 | 0,04739 |
| | 2 | 0,04824 | 0,01281 |
| | 3 | 0,04772 | 0,02750 |
| | 4 | 0,02237 | 0,04585 |
| | 5 | 0,02340 | 0,04316 |
| | 6 | 0,05771 | 0,00462 |
| | 7 | 0,03591 | 0,03025 |
| | 8 | 0,06798 | 0,00637 |
| | 9 | 0,06115 | 0,00000 |
| | 10 | 0,04132 | 0,04132 |
| | 11 | 0,00932 | 0,02847 |
| | 12 | 0,02209 | 0,02800 |

Step 4: The separation measures of each alternative, the Euclidean distances from the positive ideal solution ($A^{*}$) and from the negative ideal solution ($A^{-}$) are calculated using equations (3) and (4) (Table 9).

Step 5: The relative closeness of the alternatives with respect to the ideal solution is calculated using Equation (5). $C_i^*$ values and the orderings among alternatives are provided in Table 10.

Step 6: According to the calculations above, civil engineering is in first place; it is followed by industrial engineering, mechanical engineering, electric-electronic engineering, computer engineering and chemical engineering.

Table IX
Separation measures of each department

| Department | $S_i^*$ | $S_i^-$ |
|---|---|---|
| Mechanical | 0,07026 | 0,05523 |
| Industrial | 0,07775 | 0,06474 |
| Computer | 0,09328 | 0,05045 |
| Chemical | 0,09430 | 0,03973 |
| Civil | 0,03928 | 0,10525 |
| Electronic | 0,08483 | 0,05347 |

Table X
$C_i^*$ values and the orderings among alternatives

| Department | $C_i^*$ | Rank |
|---|---|---|
| Mechanical | 0,44010 | 3 |
| Industrial | 0,45437 | 2 |
| Computer | 0,35100 | 5 |
| Chemical | 0,29644 | 6 |
| Civil | 0,72823 | 1 |
| Electronic | 0,38661 | 4 |

### 3.5. Sensitivity analysis

In this study, sensitivity analysis is performed to determine the most critical criterion. The intuitive belief is that the criterion with the highest weight is the most critical criterion. This is not always true, and occasionally, the criterion with the lowest weight may be the most critical. Furthermore, if the decision maker can determine how sensitive the actual ranking of alternatives is to changes in the current weights of the criteria, more judicious decisions can be made (Triantaphyllou, 2000, p. 132). In the sensitivity analysis conducted, first, for each criterion, the range of numerical equivalences of linguistic weights, in which the current ranking of alternatives remains unchanged, is determined by means of an Excel macro program. Those numerical equivalences cannot be larger than the highest importance level of 6. With respect to their upper limits, increasing the weights of criteria 1, 3, 4, 5, 7, 10, and 12 does not affect the current ranking. Additionally, decreasing the weights associated with criteria 1, 2, 3, 4, 5, 7, 10, 11, and 12 does not lead to a change in the actual order of the departments (Table 11). After the ranges for the numerical equivalences of linguistic weights are determined, those ranges are normalized to be compatible with the weights used in the TOPSIS implementation. The critical criteria are determined by considering the normalized lower/upper limits. The most critical criterion is defined as the one that alters the existing ranking of the alternatives with the smallest change in its weight. The smallest change notion can also be depicted as two alternatives, absolute and relative. Although absolute change represents the difference between the current weight and its lower/upper limits, the relative change is calculated by dividing absolute change into current weight for the related criterion.

In table 11, the lower and upper limits both for the numerical equivalences of fuzzy linguistic weights and normalized weights are indicated. In addition, the smallest absolute and relative changes between the lower/upper limits of the normalized weights are provided. The range between the lower and upper limits represents the weight interval in which the current ranking does not change. The absolute and relative deviations from the lower limits for criteria 1, 2, 3, 4, 5, 7, 10, 11, and 12 and from the upper limits for criteria 1, 3, 4, 5, 7, 10, and 12 are not calculated because of their insensitivity to changes in the rankings.

Thus, it is shown that criterion 6 is the performance indicator that alters the current ranking of alternatives with the smallest decrease in its weight regarding both absolute and relative changes. Moreover, criterion 8 is another critical indicator because it modifies the actual ordering with a minimum increase in its weight with respect to absolute and relative changes.

Table XI
Lower and upper limits for the criteria weights and their absolute and relative changes

| Criteria | Numerical equivalences of fuzzy linguistic weights | | | Normalized weights | | | Absolute deviation | | Relative deviation | |
|---|---|---|---|---|---|---|---|---|---|---|
| | current | lower limit | upper limit | current | lower limit | upper limit | from lower limit | from upper limit | from lower limit | from upper limit |
| 1 | 4,9 | 0,00 | 6,00 | 0,0992 | 0,0000 | 0,1188 | - | - | - | - |
| 2 | 3,4 | 0,00 | 4,47 | 0,0688 | 0,0000 | 0,0886 | - | 0,0197 | - | 0,2868 |
| 3 | 4,8 | 0,00 | 6,00 | 0,0972 | 0,0000 | 0,1186 | - | - | - | - |
| 4 | 4,2 | 0,00 | 6,00 | 0,0850 | 0,0000 | 0,1172 | - | - | - | - |
| 5 | 4,4 | 0,00 | 6,00 | 0,0891 | 0,0000 | 0,1176 | - | - | - | - |
| 6 | 4,4 | 4,15 | 5,60 | 0,0891 | 0,0844 | 0,1107 | **0,0046** | 0,0216 | **0,0520** | 0,2425 |
| 7 | 4 | 0,00 | 6,00 | 0,0810 | 0,0000 | 0,1167 | - | - | - | - |
| 8 | 4,7 | 3,13 | 5,06 | 0,0951 | 0,0654 | 0,1017 | 0,0297 | **0,0065** | 0,3122 | **0,0688** |
| 9 | 3,9 | 2,97 | 4,44 | 0,0789 | 0,0613 | 0,0889 | 0,0177 | 0,0100 | 0,2238 | 0,1262 |
| 10 | 5 | 0,00 | 6,00 | 0,1012 | 0,0000 | 0,1190 | - | - | - | - |
| 11 | 2,7 | 0,00 | 3,47 | 0,0547 | 0,0000 | 0,0692 | - | 0,0145 | - | 0,2655 |
| 12 | 3 | 0,00 | 6,00 | 0,0607 | 0,0000 | 0,1145 | - | - | - | - |

## 4. CONCLUSİON

To achieve sustainable improvement in higher education institutions, conducting periodic performance evaluations is crucial. In this study, the objective is that the performances of the departments in the Engineering Faculty of Gazi University are evaluated to establish a ranking among them. Thus, a multi-criteria decision making technique, TOPSIS, which enables a ranking based on the ideal solutions, is utilized. Accordingly, in addition to the ranking among alternatives, their distance to each other and to the ideal situation can be determined.

In this context, determining the performance indicators and their relative weights is important. Here, although performance indicators are selected by reviewing the related literature, their weights are obtained through a Fuzzy Delphi method that allows collection of expert opinions in linguistic terms. The Fuzzy Delphi method employed in this study incorporates the 2-tuple linguistic representation model developed by Herrera and Martinez (2000). The proposed Delphi method, which has the objective of finding the criteria weights, is straightforward to implement by practitioners when expert opinions are expressed using linguistic variables. The determined criteria weights are used in TOPSIS implementation, and the general ordering of the departments is obtained. Finally, a sensitivity analysis is performed to identify the most critical criteria that affect the performance of the academic departments.

Concerning the outcomes of the sensitivity analysis, the 6$^{th}$ and 8$^{th}$ criteria are found to be more effectual in the ranking of departments; to improve the performance and achieve an enhancement in the rankings, departments must place emphasis on these criteria. Thus, two recommendations to departments to improve their performance are to increase the number of students graduating from their M.Sc. and Ph.D. programs and to increase the number of projects conducted.

## REFERENCES


[1] Ahmadi, A., Bazargan, A., Beigi, & F. H. (2011). Relationship between organizational characteristics and implementation of internal evaluation in universities educational departments, case: University of Tehran. *Procedia Social and Behavioral Sciences*, 15, 221–228.

[2] Ateş, N.Y., Çevik, S., Kahraman, C., Gülbay, M., & Erdoğan, S.A. (2006). Multi Attribute Performance Evaluation Using a Hierarchical Fuzzy TOPSIS Method. In Kahraman, C. (Eds.), Fuzzy Applications in Industrial Engineering (pp. 537-572). New York: Springer-Verlag.

[3] Beasley, J. E. (1995). Determining teaching and research efficiencies. *Journal of the Operational Research Society*, 46(4), 441–452.

[4] Büyüközkan, G., & Ruan, D. (2008). Evaluation of software development projects using a fuzzy multi-criteria decision approach. *Mathematics and Computers in Simulation*, 77, 464-475.

[5] Chang, P. L., Hsu, C. W., & Chang, P. C. (2011). Fuzzy Delphi Method for evaluating hydrogen production Technologies. *International Journal of Hydrogen Energy*, 36, 14172-14179.

[6] Chang, S. I., Tsujimura, Y., Gen, M., & Tozawa, T. (1995). An efficient approach for large scale project planning based on fuzzy Delphi method. *Fuzzy Sets and Systems*, 76, 277-288.

[7] Cheng, C. H., & Lin, Y. (2002). Evaluating the best main battle tank using fuzzy decision theory with linguistic criteria evaluation. *European Journal of Operational Research*, 142, 174-186.

[8] Fasanghari, M., Montazer, G. A. (2008). A Stock Portfolio Selection Method through Fuzzy Delphi. In Fuchun Sun, Jianwei Zhang, Ying Tan, Jinde Cao, & WenYu (Eds.), Proceedings of Advances in Neural Networks-ISNN 2008 (pp. 615-623). New York: Springer.

[9] Herrera, F., Martinez, L. (2000). A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 8(6), 746-752.

[10] Hsu, H. M., & Chen, C.T. (1996). Aggregation of fuzzy opinions under group Decision making. *Fuzzy Sets and Systems*, 79, 279-285.

[11] Hsu, Y. L., Lee, C. H., & Kreng, V. B. (2010). The application of Fuzzy Delphi Method and Fuzzy AHP in lubricant regenerative technology selection. *Expert Systems with Applications*, 37, 419-425.

[12] Hwang C. L., & Yoon, K. (1981). *Multiple Attribute Decision Making: Methods and Applications: A State of the Art Survey*. New York: Springer-Verlag, (Chapter 2).

[13] Ishikawa, A., Amagasa, M., Shiga, T., Tomizawa, G., Tatsuta, R., & Mieno, H. (1993). The Max-Min Delphi method and Fuzzy Delphi method via fuzzy integration. *Fuzzy Sets and Systems*, 55(3), 241-253.

[14] Johnes, J., & Johnes, G. (1995). Research funding and performance in U.K. University Departments of Economics: a frontier analysis. *Economics of Education Review*, 14(3), 301–314.

[15] Kuo, Y. F., & Chen, P. C. (2008). Constructing performance appraisal indicators for the mobility of service industries using Fuzzy Delphi Method. *Expert Systems with Applications*, 35, 1930-1939.

[16] Leitner, K. H., Prikoszovits, J., Schaffenhauser-Linzatti, M., Stowasser, R., & Wagner, K. (2007). The impact of size and specialisation on universities' department performance: A DEA analysis applied to Austrian universities. *Higher Education*, 53, 517–538.

[17] Murray, T. J., Pipino, L. L., & Gigch, J. P. (1985). A Pilot Study of Fuzzy Set Modification of Delphi. *Human Systems Management*, 5, 76-80.

[18] Rayeni, M. M., Vardanyan, G., & Saljooghi, F. H. (2010). The Measurement of Productivity Growth in the Academic Departments using Malmquist Productivity Index. *Journal of Applied Science*, 10(22), 2875-2880.

[19] Rouyendegh, B. D. (2009). DEA-ANP sequential hybrid algorithm for multiple-criteria decision-making process, a long with an application. Ph.D. Thesis, *Gazi University Institute of Science and Technology*, Ankara.

[20] Rouyendegh, B. D. (2011). The DEA and Intuitionistic Fuzzy TOPSIS Approach to Departments' Performances: A Pilot Study. *Journal of Applied Mathematics*, 2011, 1-16.

[21] Rouyendegh, B. D., & Erol, S. (2009). A DEA–ANP hybrid Algorithm Approach to Evaluate a University's Performance. *International Journal of Basic & Applied Sciences IJBAS*, 9(10), 76-86.

[22] Sinuany-Stern, Z., Mehrez, A., & Barboy, A. (1994). Academic departments efficiency via DEA. *Computers and Operations Research*, 21(5), 543–556.

[23] Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *Journal of Information Technology Education*, 6, 1–21.

[24] Tomkins, C., & Green, R. (1988). An experiment in the use of Data Envelopment Analysis for evaluating the efficiency of UK university departments of accounting. *Financial Accountability and Management*, 4(2), 147-164.

[25] Triantaphyllou, E. (2000). *Multi-Criteria Decision Making Methods: A Comparative Study*. Dordrecht: Kluwer Academic Publishers, (Chapter 8).

[26] Von der Gracht, H. A. (2008). *The Future of Logistics, Scenarios for 2025*. (1st ed.). Wiesbaden: GWV Fachverlage GmbH, (Chapter 3).

[27] Wang, W. P. (2010). A fuzzy linguistic computing approach to supplier evaluation. *Applied Mathematical Modelling*, 34, 3130-3141.

[28] Wu, H. Y., Chen, J. K., Chen, I. S., & Zhuo, H. H. (2012). Ranking universities based on performance evaluation by a hybrid MCDM model. *Measurement*, 45, 856-880.

[29] Wu, H. Y., Lin, Y. K., & Chang, C. H. (2011). Performance evaluation of extension education centers in universities based on the balanced scorecard. *Evaluation and Program Planning*, 34, 37–50.

[30] Zolfani, S. H., & Ghadikolaei, A. S. (2013). Performance Evaluation of Private Universities Based on Balanced Scorecard: Emprical Study Based on Iran. *Journal of Business Economics and Management*, 14(4), 696–714.