Palestine Journal of Mathematics

Vol. 11(Special Issue I)(2022) , 109–124

A FUZZY DATA ENVELOPMENT ANALYSIS MODEL FOR MEASURING EFFICIENCY OF MALAYSIAN PUBLIC RESEARCH UNIVERSITIES

S.A.M. Ahmed, M.A. Talib, N.F.M. Noor and R. Jani

Communicated by Shimpi Singh Jadon

2010 AMS Subject Classification: 00A35, 90C08.

Keywords and phrases: DEA, efficiency, fuzzy data, higher education, decision maker unit.

Abstract The technical efficiency scores of Malaysian public research universities for the next academic year 2021/2022 are estimated based on crisp data gathered from the previous academic years. The objective is achieved by proposing an algorithm established on a triangular fuzzy number (TFN) theorem where the public research universities are regarded as Decision Making Units (DMUs) in this research. The Data Envelopment Analysis (DEA) model is converted to Fuzzy Data Envelopment Analysis (FDEA) based on Charnes Cooper and Rhodes (CCR) model. The efficiency scores for the Malaysian public research universities are successfully obtained and are compared with the previous years' efficiency scores and positions based on the respective local and international ranking systems. Sufficient agreement persists through this process. All five research universities are expected to achieve full efficiency in the next academic year of 2021/2022. In conclusion, the proposed Fuzzy DEA model can be a reliable and promising tool for efficiency estimation of any higher education institution (HEI) in the future.

1 Introduction

Efficiency measurement is an effective method of assessing the performance of organizations. Data Envelopment Analysis (DEA) is a popular technique which has been used to measure the efficiency of a group of presumably homogeneous entities termed as Decision Making Units (DMUs). Universities are such examples where this technique can measure the efficiency of a university in producing as large as possible outputs, for instance, the number of graduates from a given set of inputs such as the number of teaching staff. Changes in the expected efficiency scores can be made based on capacity of certain input and output variables for effective decisions to increase the efficiency for each university, as an individual DMU, or for the whole University by using total average efficiency score [1]. Measuring the efficiency of public universities or any higher education institution (HEI) is so critical as returns on investment in the higher education sector are characterized by time-lags of decades. More often than not, governments are under great pressures to fund higher education at levels deemed necessary [1] so that the identified budget execution, planning and management can be accomplished at an increased transparency level too [2]. In the Malaysian 2021 Budget, a huge amount totalling to more than RM14 billion that was allocated to the Higher Education Ministry includes RM50 million for infrastructure and equipment replacement in the public universities alone [3]. Even more budget was designated for five public universities in Malaysia which were established as research universities (RU) in 2015 to shoulder the responsibility in meeting the national goals.

For these vast arrays of objectives, RUs obviously stand to gain additional government funding for research activities, research management, RU incentive grants and specialised research services like patenting, IPR and repository [3]. At the same time, there are always concerns on issues like public resources not being allocated in a way that promotes efficiency, or to meet the established goals of the higher education sector which have been raised abroad [3]. For that reason, the pressures are all the same for the RUs to reorganise activities and priorities, increasing research output and quality and to increase their international ranking and reputations [4]. For it is a known fact that research has become the main evaluation criteria in determining a university's world ranking, measuring the performance of RUs on efficiency-based ranking is an alternative tool to enhance accountability and transparency of these institutions [4]. The ranking of research universities has some encouraging profits, as not only does it boost their reputation, but it can also be used as an advertising material to entice more new students from local and abroad [5], [6] as well as it is the requirement from working industry for qualified and well-trained human resources to have received quality education in highly ranked universities [7]. University efficiency scores can be used as indicators in allocating resources to universities [7] particularly, when the focus of transforming the Malaysian Higher Education is made based on outcomes and performance.

Evaluating efficiency of organizations is an effective method of assessing performance of organizations, marketplaces and the entire economy system of organizations in a country. As underlined in the Malaysia Education Blueprint (2015-2025), efficiency measurement of the higher education institutions and universities alike is no longer an option [8]. Efficiency measurement can provide detailed information on universities performances which enable the government to make the future framework of long-term plans based on valid data. Data Envelopment Analysis (DEA) technical efficiency-based performance measurement by Charnes, Cooper and Rhodes (CCR) has been widely used to calculate the technical efficiency of universities [9]. Its application in HEIs can also help the universities to be more effective in their operations, hence improving their standards and ranking [7]. DEA technical efficiency approach correlates to how much total of output can be achieved from an offered total input, or a certain mixture of inputs [10]. Thus, this unique approach can handle efficiency estimation and assessment for the coming years based on the expected efficiency scores. Despite the basic DEA method claims that input and output data are precise and accurate, however in the real life situation, the observed values of the input and output data are sometimes imprecise or vague as a result of measurement inaccuracies, unquantifiable and incomplete data [4], [10].

More often there are variable data which are not clear or ambiguous and they are available in the form of linguistic or qualitative data [11]. For instance, the expected quality of graduates produced by the universities going into the labour market are unknown [7]. To deal with imprecise and vague data in DEA models, fuzzy set theory has been proposed to extend the traditional DEA models into a fuzzy framework known as Fuzzy Data Envelopment Analysis (FDEA) [12]. FDEA model measures and estimate the technical efficiency scores by proposing a FDEA model and algorithm. The topic on university efficiency measurement dealing with imprecision and vagueness of input/output data is obviously under researched and to authors' best knowledge, [7] is the only such study that has introduced a fuzzy DEA model for a university system.

There are several kinds of efficiencies as follows; allocative efficiency, productive efficiency, technical efficiency, 'X' efficiency, dynamic efficiency, and social efficiency. Productive efficiency arises when an organization is merging supplies in such a manner to make an offered output at the least mean aggregate cost as possible. Specifically, expenses should be decreased at the least possible level on an organization's short run of mean sum cost curve. However, this is not applied in this research because the variables are not representing any general cost/quality cost. Whereas technical efficiency correlates to how much total of output can be achieved from an offered total input, such as a machine or a worker, or a certain mixture of inputs. Particularly for this study, the focus is technical efficiency of the public research universities in Malaysia (PRUM). In a hierarchical order of university model, the efficiency of PRUM could be one of the benchmarks to get extra grant for research actions, research managing, hence, research quality assurance. Incentive grants, specialized research services such as patenting, international publishing research and repository for the academic year needs [3] are some means of making these universities efficient.

DEA can handle efficiency estimation and assessment for the following years based on the expected efficiency scores. In fact, we cannot find any study on HEIs that applied FDEA to measure and estimate the technical efficiency score by proposing FDEA model and algorithm. Also, up to our best knowledge, there is no any existing study in Malaysia that introduced a fuzzy DEA model for the university system. Therefore, this study resorts to propose the Fuzzy Data Envelopment Analysis (FDEA) approach to deal with the crisp, linguistic or qualitative data variables in measuring the efficiency of public research universities in Malaysia (PRUM). This can be done by employing a suitable algorithm to get the fuzzy numbers by utilizing the Triangular Fuzzy Numbers theory. In a nutshell, FDEA can help to measure the efficiency better with clear results and more importantly, FDEA draws a good relationship between the university performance and its efficiency which can then be put into a suitable rank [9].

The current study introduces some questions on PRUM efficiency-based performance for the latest four academic years (2017/2018, 2018/2019, 2019/2020 and 2020/2021) as follows:

- What are the technical efficiency scores of the public research universities in Malaysia for the latest academic years ?
- What is the algorithm to convert crisp data of input and output variables to fuzzy data in the form of Fuzzy Triangular Number (FTNs)?
- What is the suitable FDEA model to measure the technical efficiency under uncertainty in input and output variables?
- What is the expected efficiency score for PRUM for the next academic year 2020/2022?

As objectives of this study, we are trying to answer these questions. So the main objectives are:

- To measure the technical efficiency scores of the public research universities in Malaysia for the latest academic years 2017/2018, 2018/2019, 2019/2020 and 2020/2021.
- To identify the algorithm which can be used to convert crisp data of input and output variables to fuzzy data in the form of FTNs for a FDEA model of this study.
- To propose a suitable FDEA model to measure the technical efficiency under uncertainty in input and output variables.
- To estimate the efficiency score for PRUM for the next academic year 2021/2022.

2 Data Envelopment Analysis

In discrepancy to parametric techniques that need the preceded description of a cost or a production function, there is Data Envelopment Analysis (DEA), a non-parametric approach that compares convenient outputs and inputs mixture just based on the offered data. The primary DEA model designed by Charnes, Cooper and Rhodes in 1978 assessed the comparative efficiencies of any organization, therefore it is coined as CCR-model [13]. Based on this model, the DEA method utilizes Linear Programming (LP) techniques to perform the relative connection between the outputs and inputs formed by Decision Making Units (DMUs) hence formulate an effective production boundary as an inference in a vision frame for the finest practices and experimental methods. DEA calculates the efficiency score of each DMU by comparing along with all the other DMUs in the case study (institution, association, firm, or any) involving itself. That relative (technical) efficiency is evaluated by getting the relative amount of the weighted outcome after adding together all outputs and the summation of weighted outcome after combining all inputs to achieve the maximum optimality. This DEA approach concerned with technical efficiency in PRUM case as HEI, is in line with the work by Askari et al [14]. DEA is a common method that progressed as the benchmarking approach for formal budgeting, research and planning. Fundamental Data Envelopment Analysis model works from the outputs and inputs variables, that are mostly fixed, selected for the organizations under study which are termed as DMUs in DEA context. However, in practical terms, these ouput and input data variables are dynamic and often fluctuate. Fuzzy Data Envelopment Analysis (FDEA) is the DEA approach to cater for such cases.

3 Fuzzy Data Envelopment Analysis (FDEA)

FDEA is a non-parametric approach to assess the comparative efficiencies of a group of DMUs with input and output data that are frequently vacillated. These are termed as common crisp inputs and outputs. These hesitated figures can be exemplified as linguistic variables described by fuzzy figures [15]. Based on CCR-model, the first work using fuzzy theory in DEA has been developed [15].

3.1 FDEA Model Terms (or DEA Model Variables)

Many fuzzy set-based techniques have been suggested in DEA in the previous three decades. According to Marbini et al. [16], their work develops the fuzziness concept and provides more explanation and interpretations on using the fuzzy set theory in DEA. FDEA models are in general introduced as Fuzzy Linear Programming (FLP) models with fuzzy coefficients (i.e., fuzzy input-output data) and crispy decision variables. For FDEA model, it is so important to know that some of these variables can be in terms of categorical variables. In some studies, these variables are termed to be as membership functions. Because FDEA models take on the form of FLP problems, the various FDEA methods have been updated and created as distinct methods of resolving the equivalent FLP models [16].

3.2 Review on FDEA Models

In typical, the linear programming DEA models are translated into fuzzy linear programming (FLP) models after the input and/or output data are considered by fuzzy figures. Based on this concept, Marbini et al. [9] classify the applications of fuzzy set theory in DEA into four sub-area groups; (1) tolerance approach, (2) α-level based approach, (3) fuzzy ranking approach and (4) possibility approach. Later, another two methods, the fuzzy arithmetic and the fuzzy random type-2 were added [16]. To our recent knowledge, the latest classification of FDEA approach was conducted for the ranking of stocks in the Tehran Stock Exchange [11]. Out of the six approaches, the tolerance approach is the most powerful and commonly used method [16]. FLP dilemmas can be categorized into the following six sets to manage the fuzzy data:

- Both the right-hand-side constraints and the decision variables are described as fuzzy figures.
- The decision variables' coefficients in the objective function are characterized by fuzzy figures.
- The decision variables coefficient in the constraints and the right-hand-side of the constraints are described by fuzzy figures.
- The decision variables' coefficients , the decision variables in the objective function and the right-hand-side of the constraints are described as fuzzy figures.
- The decision variables' coefficients in the objective function, the decision variables coefficients in the constraints and the right-hand-side of the constraints are described as fuzzy figures.
- FLP models when all the above variables and paarameters are described as fuzzy figures

Qin et al. [17] established a DEA model with type-2 fuzzy outputs and inputs befitting to linguistic ambiguities alongside numerical ambiguities with reverence to fuzzy membership functions. The type-2 fuzzy random is the sixth FDEA model [11]. Standing on the likely values of the fuzzy variables, a reducing approach has been employed for the type-2 random fuzzy variables, henceforth the FDEA model was created. Since uncertainty, fuzziness and randomness are synchronized in one assessment system and the fuzzy accidental data are described as well-known probability distributions, a category of type-2 fuzzy random DEA (FRDEA) models with fuzzy accidental outputs-inputs was proposed [18]. A stochastic simulation and hybrid genetic process method to evaluate the objective function of the suggested DEA was also introduced.

3.3 Theoretical Approach with Conceptual Framework of Charnes Cooper and Rhodes DEA (CCR-DEA) model

DEA model with fuzzy data utilization is the major path to this recent study approach as it routing from the CCR-DEA model. This DEA model is used to detect the efficiency scores of the DMUs with unstable or hesitant variables (outputs and inputs) [13]. The current study attempts to offer a model as an expansion of the CCR model to a fuzzy structure that can be utilized for higher educational institution cases. The main formation of a fuzzy interpretation system called Type 2 has been designed [19] and developed for PRUM case in this study as depicted in Figure 1.

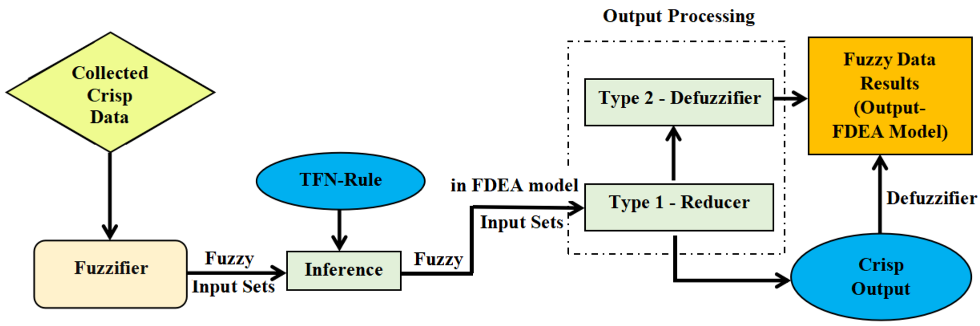

Figure 1. Fuzzy framework for PRUM Case (Source: Author)

In a system of fuzzy, some or all the variables are described as fuzzy figures or linguistic data (numerical values are likely more preferred by decision maker departments), so the variables are defined in Figure 2 as follows:

$$\tilde{x}_i : i\text{-th fuzzy input} \quad \text{and} \quad \tilde{y}_i : i\text{-th fuzzy input},$$
$$\tilde{\mu}_i\text{: weight of the fuzzy input } \tilde{x}_i \quad \text{and} \quad \tilde{v}_i\text{: weight of the fuzzy output } \tilde{y}_j.$$

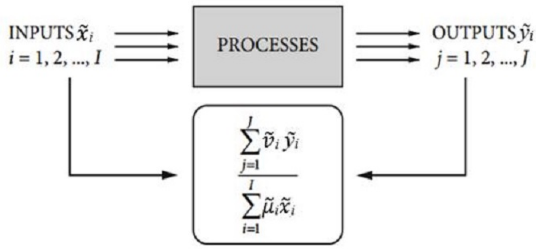

Figure 2. The conceptual framework of DEA using fuzzy variables(Source: Author)

4 DMUs and Variables Selection

Data collected on 5 Research Universities as the DMUs (in Table 1), are panel primary data sought from annual reports and from the Ministry of Higher Education in Malaysia. These universities represent all public research universities in Malaysia (PRUM) under consideration for this research to determine and predict the technical efficiency of each of the university. All DMUs and their initials are catalogued in the following Table 1:

Table 1. DMUs under the study

DMU	University Name	Code
1	Universiti Malaya	UM
2	Universiti Sains Malaysia	USM
3	Universiti Kebangsaan Malaysia	UKM
4	Universiti Putra Malaysia	UPM
5	Universiti Teknologi Malaysia	UTM

Person correlations matrix (PCM) is the statistical measures to ensure reliability and accuracy of the preliminary data obtained. All the data used in this study have been verified with Pearson Correlation tests in the matrix form for each year to ensure their reliability and accuracy. Table 2 shows an instance for PCM for academic year 2017/2018. The inputs and outputs are found to be significant at both levels of 0.01 and 0.05 which confirmed the validity of the collected data. This result is also similar for the other 3 considered academic years.

Table 2. Pearson Correlations Matrix for the variables of academic year 2017/2018

Correlations	Input 1	Input2	Input3	Output1	Output2	Output3
Input1	1	0.846**	0.242*	-0.169*	0.268*	0.425*
Input2	0.846**	1	-0.047	-0.338*	-0.057	-0.093
Input3	0.242*	-0.047	1	-0.625**	.912**	0.588*
Output1	-0.169*	-0.338*	-0.625**	1	-0.401*	0.21*
Output2	0.268*	-0.057	.912**	-0.401*	1	0.75**
Output3	0.425*	-0.093	0.588*	0.21*	0.75**	1

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Next, the working flowchart of this study is illustrated in Figure 3.

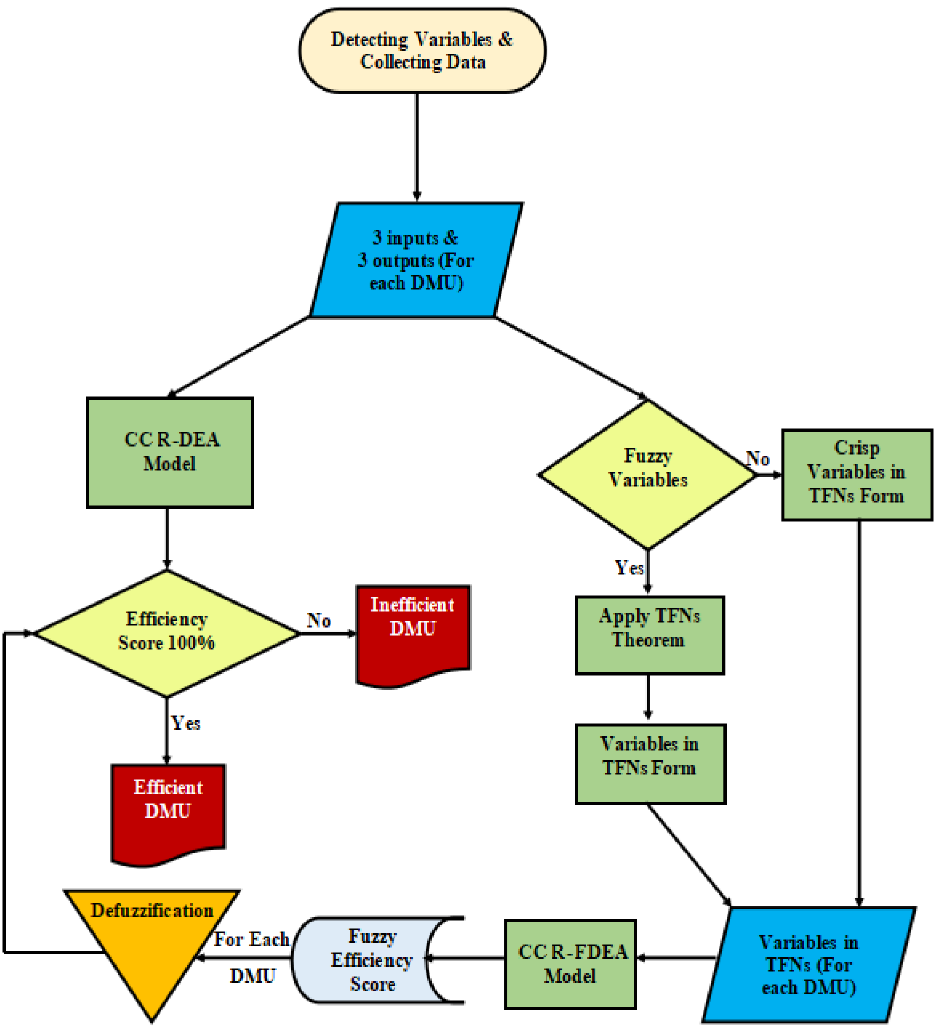

Figure 3. Flowchart for PRUM Case (Source: Author)

4.1 Input Variables

The establishment of PRUM underlines the objectives summarized as: to increase research and development activities and commercialization, to increase the intake of Postgraduate and Post-Doctoral students as well as the number of lecturers with doctoral qualifications, to establish and strengthen Centres of Excellence, to enhance the recruitment of foreign students and thus, the university's ranking at international level [4]. Based on this, the selection of input and output variables are made as follows:

- Input 1: Full Time Equivalent Staff Number (FTE Staff No) inclusive of all number of academic staff, academic staff of international/overseas origin and number of research staff.
- Input 2: Full Time Equivalent Student Number (FTE Stud No) inclusive of all total number of students and students of international/overseas origin.
- Input 3: The percentage % of International Students (Inter Stud %); the percentage ratio of FTE international Student to FTE Student.

Also, important to note that in this PRUM case, both Input 1 and 2 are pre-fixed or can be controlled by each DMU (the university) and as such these are the crisp data.

4.2 Output Variables

In this research, all the output variables are set as fuzzy numbers. The output variables are as follows:

- Output 1: Teaching Reputation in Percentage (%) is the percentage ratio of total number of undergraduate degrees, masters and doctorates awarded to the total students FTE (Tech Reput %).
- Output 2: Research Reputation Percentage (%). This is based on the volume of research reputation ratio inclusive of all the ratios of research income, reputation of university survey and research productivity of the university (Research Reput %).
- Output 3: Citation Percentage (%) which depicts the research influence ratio of the university.

4.3 The Original DEA Model and the Proposed FDEA Model

The original model of DEA is the CCR model (Model 1) [13] for the relative efficiency. This model looks as a non-linear programming model converted to a linear programming model (Model 2). This is because the linear programming (LP) helps and supports to find the finest results from a set of requirements or parameters that have a linear relationship and the DEA must be an LP model with LP parameters [13] as shown in Figure 4:

$$\theta_p^* = \max \frac{\sum_{r=1}^{s} u_r\, y_{rp}}{\sum_{i=1}^{m} v_i\, x_{ip}}$$

s. t.

$$\frac{\sum_{r=1}^{s} u_r\, y_{rj}}{\sum_{i=1}^{s} v_i\, x_{ij}} \leq 1, \quad \forall j$$

$$u_r\ v_i \geq 0 \qquad \forall\, r, i.$$

Model 1: CCR Non-LP

$$\Longrightarrow$$

$$\theta_p^* = \max \sum_{r=1}^{s} u_r\, y_{rp}$$

s. t.

$$\sum_{i=1}^{m} v_r\, x_{ip} = 1$$

$$\sum_{r=1}^{s} u_r\, y_{rj} - \sum_{i=1}^{m} v_i\, x_{ij} \leq 0, \quad \forall j$$

$$u_r\ v_i \geq 0 \qquad \forall\, r, i.$$

Model 2: CCR-LP

Figure 4. Model 1 and Model 2

Each DMU (assume that there are n DMUs) uses m inputs $x_{ij}(i = 1, 2, ..., m)$ to obtain s outputs $y_{rj}(r = 1, 2, ..., s)$. Here $u_r(r = 1, 2, ..., s)$ and $v_i(i = 1, 2, ..., m)$ are the weights of the i^{th} input and r^{th} output. Unlike the basic DEA, FDEA is a powerful approach to assess the efficiency of DMUs with unclear data that used precise data (input and output). This agrees with real life situations, where most of the situations have hesitated or crisp data. So, in the PRUM case, the variables can be defined based on Model 3 of FDEA CCR as introduced as follows:

$$\theta_p = max \sum_{r=1}^{s} u_r y_{rp}$$

such that

$$\sum_{i=1}^{m} v_i x_{ip} = 1$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \forall j$$

$$u_r, v_i \geq 0 \forall r, i.$$

Model 3: Fuzzy DEA - CCR model

where, $\tilde{x}_{ij}(i = 1, 2, ..., m)$ and $\tilde{y}_{rj}(r = 1, 2, ..., s)$ are precise inputs and imprecise (fuzzy) outputs need to be consistent on this. Here in the PRUM case, only the outputs are fuzzy for the j^{th} DMU (DMU_j).

5 Evaluating Technical Efficiency of Public Research Universities in Malaysia

In this research, it is assumed that all five public research universities in Malaysia are homogeneous. We use Solver 365 which is the latest version of Solver application of 2021. Solver is a powerful application for connection and efficiency with Microsoft. As Solver is a Microsoft Excel add-in package, it can be used for "What-If Analysis". Solver can also be used to find an optimum (minimum or maximum) value for a formulation in one cell, which called the objective cell and subject to limits, or constraints, on the values of other formula cells in a worksheet. Solver works with a group of cells, simply called variable cells or decision variables that are used in computing the formulations in the objective and constraint cells. Solver regulates the values in the decision variable cells to satisfy the bounds on constraint cells and to yield the outcome value that we are looking for in the objective cell [20]. In Figure 5, Solver 365 interface for calculation on DMU4 in the academic year 2017/2018 is demonstrated as an example.

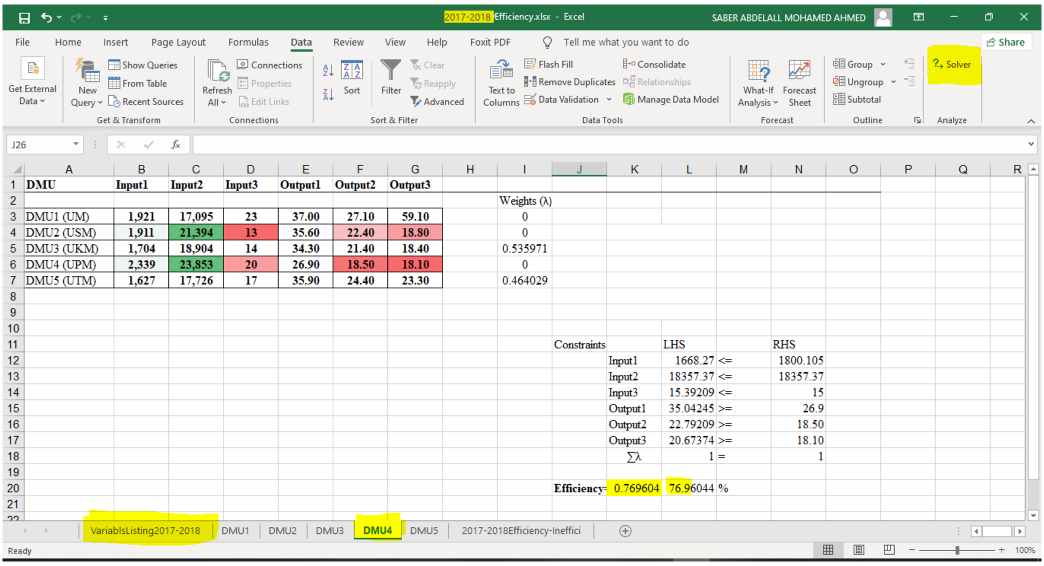

Figure 5. An example of Solver 365 interface for DMU4 calculation (2017/2018)

5.1 DEA Empirical Results

In Model 2 for the PRUM case, each DMU has $s = 3$ inputs and $m = 3$ outputs, i.e., each DMU uses inputs $x_{ij}(i = 1, 2, 3)$ to obtain $m = 3$ outputs of $y_{rj}(r = 1, 2, 3)$. Here $u_r(r = 1, 2, 3)$ and $v_i(i = 1, 2, 3)$, are the weights of the i^{th} input and r^{th} output. This partial program is calculated for every DMU to find out its optimum input and output weights. The efficiency scores calculated by the Solver-365 software are ordered in 4 groups by referring to the four academic years (2017/2018), (2018/2019), (2019/2020) and (2020/2021). Each group in the table has efficiency score for each DMU based on data collected. The results preview the earlier academic year technical efficiency score. The generated results contain 2 sets i.e., column one has the efficiency scores and column two has efficiency status of each DMU for each academic year. Table 3 and Table 4 both show that all DMUs are fully efficient with 100% score for 2 academic years of 2017/2018 and 2018/2019 except for DMU4 which is less efficient with the 77% score in the academic year 2017/2018 and 89% score in the academic year 2018/2019. In Table 5 for the academic year 2019/2020, all the DMUs are fully efficient with 100% scores except for DMU2. Then in Table 6 for the academic year 2020/2021, all DMUs are fully efficient with 100% scores.

Table 3. Efficiency score using DEA for the academic year 2017/2018

DMU	Efficiency	Efficiency Status
DMU1	100%	Efficient
DMU2	100%	Efficient
DMU3	100%	Efficient
DMU4	77%	Inefficient (refers to DMU1, DMU2, DMU3 & DMU5)
DMU5	100%	Efficient

Table 4. Efficiency score using DEA for the academic year 2018/2019

DMU	Efficiency	Efficiency Status
DMU1	100%	Efficient
DMU2	100%	Efficient
DMU3	100%	Efficient
DMU4	89%	Inefficient (refers to DMU1, DMU2, DMU3 & DMU5)
DMU5	100%	Efficient

Table 5. Efficiency score using DEA for the academic year 2019/2020

DMU	Efficiency	Efficiency Status
DMU1	100%	Efficient
DMU2	93%	Inefficient (refers to DMU1, DMU3, DMU4 & DMU5)
DMU3	100%	Efficient
DMU4	100%	Efficient
DMU5	100%	Efficient

Table 6. Efficiency score using DEA for the academic year 2020/2021

DMU	Efficiency Score	Efficiency Status
DMU1	100%	Efficient
DMU2	100%	Efficient
DMU3	100%	Efficient
DMU4	100%	Efficient
DMU5	100%	Efficient

Table 7. Efficiency score for 4 years in a group and average for each university

DMU	2017/18 Efficiency Score	2018/19 Efficiency Score	2019/20 Efficiency Score	2020/21 Efficiency Score	Average	Efficiency Status
DMU1 (UM)	100%	100%	100%	100%	100%	Efficient
DMU2 (USM)	100%	100%	93%	100%	98%	Not Fully Efficient
DMU3 (UKM)	100%	100%	100%	100%	100%	Efficient
DMU4 (UPM)	89%	77%	100%	100%	92%	Not Fully Efficient
DMU5 (UTM)	100%	100%	100%	100%	100%	Efficient

5.2 Average Efficiency Score of Each DMU for All the Academic Years 2017/2018, 2018/2019, 2019/2020 and 2020/2021

Table 7 and Figure 6 show the efficiency scores for all DMUs throughout all the four academic years. DMU1, DMU3 and DMU5, all of them have the full score in the latest four academic years. So, the same set of data was applied using FDEA to estimate the next academic year of 2021/2022. Any result of the efficiency estimation score that is not equal to 100% although it is very close to full score will not be used. Furthermore and based on all above, the DEA empirical results show that DMU1 (UM), DMU3 (UKM) and DMU5 (UTM) are fully efficient for all academic years under the study, but for DMU2 (USM) and DMU4 (UPM), both of them are not fully efficient for the same period of the study as shown in Table 3 until Table 6. This is also illustrated in Figure 6.

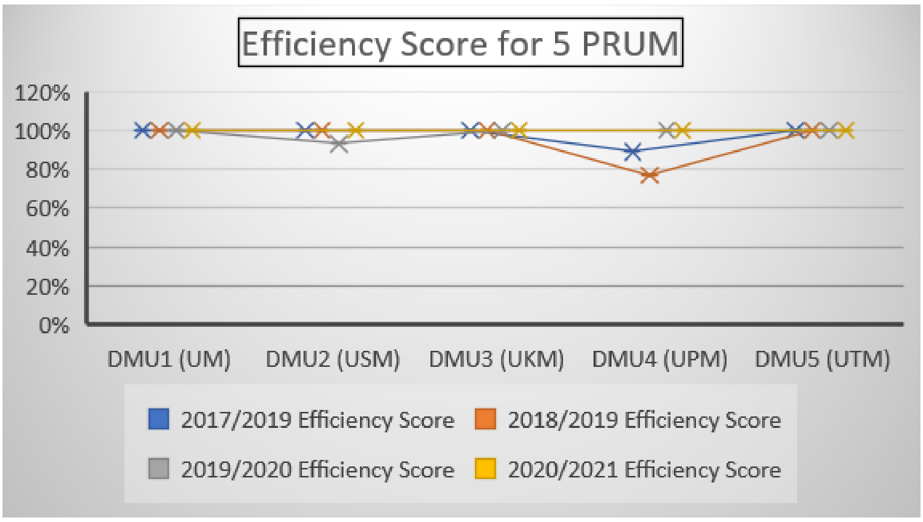

Figure 6. Efficiency Score for 4 years in a group

6 Estimating Technical Efficiency of PRUM by FDEA Model

By using the same data of Public Research Universities in Malaysia (PRUM) for the 4 consecutive academic years (2017/2018, 2018/2019, 2019/2020 and 2020/2021), the second part of this research utilizes the FDEA model. To establish the model for this study, the process starts with deducing the Triangular Fuzzy Numbers (TFNs) for each crisp fuzzy variable (3 inputs), then the technical efficiency score of PRUM for the academic year of 2021/2022 is estimated by applying the proposed expanding model which is derived from Model 3 of linear programming FDEA model for evaluating the technical efficiency score for HEI cases.

6.1 Fuzzifying Crisp Data of the Fuzzy Variables

By using the three output data of four academic years 2017/2018, 2018/2019, 2019/2020 and 2020/2021 which are determined as crisp data, these outputs will be considered as fuzzy variables. Therefore, they must be converted into fuzzy numbers by applying Triangular Fuzzy Numbers Theory [21].

6.2 Algorithm for Finding Fuzzy Numbers using Triangular Fuzzy Numbers Definitions

For explanation on this algorithm in general, assume that there exists a numeric vector x_i where $x_i = (x_1, x_2, ..., x_n; n \in Z)$ are fuzzy variables with collected crisp data. The followings are the steps to covert x_i to a fuzzy variable y_i by using TFNs theorem and its definitions:

(i) Transform exact value numbers of x_i to TFNs by using R-coding. Key in R-soft as: fuzzify (*x*, *y* = NULL, method = "mean", err = 0, dimnames = list("*x*", "*y*"), ...).

(ii) Run R-soft then get the results in the form $(y_L,\ y_M,\ y_U)$, y_L = lower value, y_M = medium value and y_U = upper value for each x_i.

(iii) Transform hesitated (crisp) figures in *x* to TFN. Subsequently, the values in *y* can be utilized as combination aspects and are coerced to a component (factor).

6.3 Details on TFN Format

The followings are the details on TFN format in R-soft :

(i) Mean determines the middle value of a TFN as the arithmetic mean of *x* given *y*, the right and left terminal as standard deviations.

(ii) Median offers the middle value as a median whereas the right and left spread are determined as spaces of the third and first quartiles from the median.

(iii) Zero enclosures zeros to the terminals.

(iv) Error utilizes a user-assuming numeric value or vector for the terminals.

The numeric vector length in argument *err* must be in dimension of (1, length (*x*), 2× length (*x*)). Using any of the above details is up to the type of the collected crisp data and by doing data analysis, we can check which process will be better fitted to get the fuzzy numbers in the R-soft.

6.4 Determining Fuzzy Numbers for PRUM Case

Based on the collected data for PRUM case, it is assumed that the three output variables are fuzzy. Table 8 presents the crisp data that should be converted to TFNs. These TFNs will be applied in the FDEA model. In this approach, where the data values are close to each other and that its range is small, the mean process in R-soft is used to generate the TFNs.

An example of computation for DMU1 and Teaching Reputation (%), data from Table 8 shows 39.3, 41.6, 37, 31.2 as the crisp data. By using TFNs definition and R-soft mean process in Section 6.2, these crisp data values will be TFNs (31.2, 37.28, 41.6) as in Figure 7 such that by applying all the output data (for the three variables), the results of TFNs are shown in Table 9.

Table 8. Data collected for DMUs' latest 4 academic years (Crisp Data)

DMU	Input			Output (Crisp)		
(University)	FTE Staff No	FTE Stud No	Inter Stud %	Teach Reput%	Research Reput %	Citations %
DMU1(UM)	1903	15794	20	(39.3, 41.6, 37.28,31.2)	(30.5,31.5 ,27.1,26.6)	(60,56.6 ,59.1,54.4)
DMU2(USM)	1967	21039	14	(34.7,35.6 ,35.6,32.2)	(23.3,22.7 ,22.4,17.7)	(32.2,26.7 ,18.8,15)
DMU3(UKM)	1709	17601	16	(35.3,34.2 ,34.3,30.5)	(21.4,19.6 ,21.4,21.4)	(42.5,32.3 ,18.4,11)
DMU4(UPM)	1648	19937	25	(33.3,32 ,26.9,26.6)	(29.7,31.4 ,18.5,23.7)	(24.4,19.1 ,18.1,17.2)
DMU5(UTM)	1694	19988	17	(30.4,36.4 ,35.9,31.1)	(24.3,25.2 ,24.4,20.4)	(38.8,29.2 ,23.3,22.7)

Note that, in Table 9, the input data for the TFNs form (L, M, U) are all having the same values L= M = U. The procedure on how to apply the FDEA model for these data (in Table 9) and to run FDEA to get the fuzzy technical efficiency score results will be shown next followed by the last step of defuzzification to predict the technical efficiency score for the following academic year of 2021/2022.

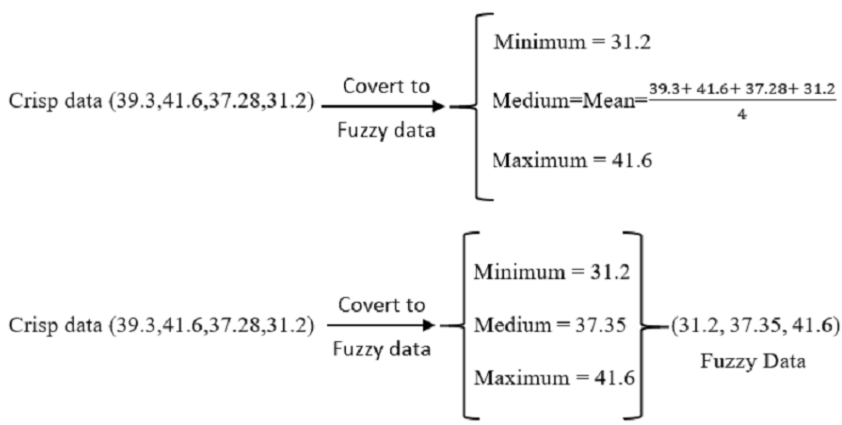

Figure 7. Fuzzy data converter

6.5 Expanding FDEA Model for PRUM case

FDEA-CCR efficiency definition [22] in Model 3:

(i) DMU$_p$ is CCR-efficient if $\theta_p{}^*$ = 1 and there exist at least one optimal $u^* > 0$ and $v^* > 0$.

(ii) Otherwise, DMU$_p$ is CCR-inefficient.

Expanding the CCR transformation Model 3 by using TFNs definitions, the forms $\tilde{x}_{ij}$ and $\tilde{y}_{rj}$ can be defined such that $\tilde{x}_{ij}$ = (x^L_{ij}, x^M_{ij} , x^U_{ij}) and $\tilde{y}_{ij}$ = (y^L_{rj}, y^M_{rj} , y^U_{rj}) whereas the Model 3 can be reformulated as Model 4:

Table 9. Fuzzify: Converting data to TFNs

DMU (University)	Input			Output (Fuzzy)		
	FTE Staff No	FTE Stud No	Inter Stud %	Teaching Reput % (y_1^L, y_1^M, y_1^U)	Research Reput% (y_2^L, y_2^M, y_2^U)	Citations % (y_3^L, y_3^M, y_3^U)
DMU1(UM)	(1903 ,1903 ,1903)	(15794 ,15794 ,15794)	(20 ,20 ,20)	(31.2 ,37.35 ,41.6)	(26.6 ,28.93 ,31.5)	(54.4 ,57.53 ,60)
DMU(USM)	(1967 ,1967 ,1967)	(21039 ,21039 ,21039)	(14 ,14 ,14)	(32.2 ,34.53 ,35.6)	(17.7 ,21.53 ,23.3)	(15 ,23.18 ,32.2)
DMU3(UKM)	(1709 ,1709 ,1709)	(17601 ,17601 ,17601)	(16 ,16 ,16)	(30.5 ,33.58 ,35.3)	(19.6 ,20.95 ,21.4)	(11 ,26.05 ,42.5)
DMU4(UPM)	(1648 ,1648 ,1648)	(19937 ,19937 ,19937)	(25 ,25 ,25)	(26.6 ,29.7 ,33.3)	(18.5 ,25.8 ,31.4)	(17.2 ,19.7 ,24.4)
DMU5(UTM)	(1694 ,1694 ,1694)	(19988 ,19988 ,19988)	(17 ,17 ,17)	(30.4 ,33.45 ,36.4)	(20.4 ,23.58 ,25.2)	(22.7 ,28.5 ,38.8)

$$\theta_p = max \sum_{r=1}^{s} u_r \bigotimes (y_{rp}^L, y_{rp}^M, y_{rp}^U)$$

such that

$$\sum_{i=1}^{m} v_i \bigotimes (x_{ip}^L, x_{ip}^M, x_{ip}^U) = (1, 1, 1)$$

$$\sum_{r=1}^{s} u_r \bigotimes (y_{rj}^L, y_{rj}^M, y_{rj}^U) - \sum_{i=1}^{m} v_i \bigotimes (x_{ij}^L, x_{ij}^M, x_{ij}^U) \leq (0, 0, 0), \forall j$$

$$u_r, v_i \geq 0 \qquad \forall r, i.$$

Model 4: Fuzzy DEA - CCR model

where, $\tilde{x}_{ij}$ (i = 1, 2, . . . , m) and for the PRUM case, m = 3. Also, $\tilde{y}_{rj}$ (r = 1, 2, . . . , s) and for the PRUM case s = 3, the inputs are in fuzzy form only while the real fuzzy outputs are for the jth DMU (DMU$_j$). The FDEA model (Model 4) is developed to be applied in Solver-365 for further technical efficiency estimation for the next academic year of 2021/2022.

6.6 FDEA Empirical Results

By applying Model 4 using data in Table 9 and Solver-365, the following results are produced and showed in Table 10. Table 10 shows the technical efficiency fuzzy scores for the academic year 2021/2022 for 5 DMUs. All levels (lower, medium, upper) are with 100% efficiency score, which make defuzzification so fast and easy to be calculated.

6.7 Defuzzification of FDEA Empirical Results and Final Results

In general, central tendency measurement is employed to defuzzify the fuzzy empirical results. A central tendency value tries to explain a group of data by detecting the middle value within that group of data. Intrinsically, gauges of central tendency are occasionally point to measure the location of the centre. They are also classified as brief statistics. The mean arithmetic (often called the average) is more frequently taken as the central tendency measure than the median or the mode [23].

Table 10. Fuzzy efficiency score of TFNs

(L, M, U) of TFNs Data	
DMU	Efficiency score (L, M, U)
DMU1	(100%, 100%, 100%)
DMU2	(100%, 100%, 100%)
DMU3	(100%, 100%, 100%)
DMU4	(100%, 100%, 100%)
DMU5	(100%, 100%, 100%)

In the PRUM case, the best method is by hiring the mode because all results are 100(%) and all DMUs are fully efficient based on the final estimation of efficiency score and efficiency status in Table 11 the for the next academic year 2021/2022. The data are collected from 4 preceding academic years. For each DMU, it can be said that all public research universities in Malaysia are fully efficient. As observed from the empirical results in Table 7 of Section 5.2, DMU1 (Universiti Malaya), DMU3 (Universiti Kebangsaan Malaysia) and DMU5 (Universiti Teknologi Malaysia) are found to be fully efficient in all 4 previous academic years so in the next academic year of 2021/2022, they are expected to be fully efficient too. On the other hand, the results in Table 7 show that DMU2 (Universiti Sains Malaysia) was not fully efficient in the academic year 2019/2020 with the score of 93% as compared to the other academic years. Table 11 shows the forecast for it to be fully efficient will be in the academic year 2021/2022. Similarly, for DMU4 (Universiti Putra Malaysia), results in Table 7 show that it did not achieve full efficiency for the first two academic years 2017/2018 and 2018/2019. However, DMU4 achieved its full efficiency score of 100% for the next two academic years 2019/2020 and 2020/2021. Therefore, it is very reasonable to stay fully efficient in the next academic year of 2021/2022, so that the expectation results for DMU4 will be consistent too. One more reason that makes the results in Table 11 acceptable and reasonable is that each university (DMU) under the study is being extra careful and concern about the variables used in this study, in coinciding with the Covid-19 circumstances in Malaysia.

Table 11. Estimated efficiency score for academic year 2021/2022

DMU	Efficiency score	Efficiency status
DMU1	100%	Efficient
DMU2	100%	Efficient
DMU3	100%	Efficient
DMU4	100%	Efficient
DMU5	100%	Efficient

7 Conclusion

The proposed Fuzzy DEA (FDEA) model and algorithm to convert crisp output data to fuzzy data by applying the Triangular Fuzzy Numbers (TFNs) Theorem are presented in this paper. The R-soft coding is proposed if the number of variables and DMUs increase to greater numbers. This study has shown how the fuzzy efficiency scores for each DMU can also be used to estimate the efficiency score for the next academic year of 2021/2022 or whenever is required. These results so far confirmed that all DMUs or public research universities in Malaysia (PRUM) are expected to achieve full efficiency score for the next academic year 2021/2022. Results of DEA for the four academic years 2017/2018, 2018/2019, 2019/2020 and 2020/2021, are consistent with the local/international QS website rankings [5]. These 5 PRUM are coming in a good rank for each academic year under the study. For the academic years of 2017/2018 and 2018/2017, the DMU4 (Universiti Putra Malaysia) was ranked at the lowest position among the 5 PRUM while in the academic year of 2019/2020, DMU2 (Universiti Sains Malaysia) received the lowest ranking for PRUM based on TopUniversities 2021 database [5]. Findings on the actual local and international rankings support the results of this study. In general, good ranking is rewarding not only as a strategic marketing for competitive student admission, but it also serves as a benchmark for additional allocation of funding for the research activities and research management. More importantly for PRUM, better research works open up future avenue for income generation, business ventures with the industries and commercialization for their own research products.

References

[1] W. W. Cooper, L. M. Seiford and J. Zhu, *Data Envelopment Analysis: History, Models, and Interpretations*, Springer, Boston (2011).

[2] O. G. Lupke, Performance measurement methods in the public sector, *Pozn. Univ. Econ. Rev.* **9**, 67–88 (2009).

[3] Ministry of Higher of Education in Malaysia, *https://www.mohe.gov.my/en/*

[4] S. Nooraini and A. Noordini, Research universities in Malaysia: What beholds?, *Asian J. Univ. Educ.* **13**, 35–50 (2017).

[5] QS World University Rankings, *https://www.topuniversities.com/qs-world-university-rankings*

[6] J. Nazarko, A. K. Kuzmicz, E. Szubzda and J. Urban, The general concept of benchmarking and its application its higher education in Europe, *Eur. J. High. Educ.* **18**, 497–510 (2009).

[7] U. Mahmudah and M. S. Lola, The efficiency measurement of Indonesian universities based on a fuzzy data envelopment analysis, *Open J. Stat.* **6**, 1050–1066 (2016).

[8] Malaysia Education Development Plan 2015-2025, *https://jpt.mohe.gov.my/portal/index.php/en/corporate/development-plan/16-malaysia-education-development-plan-2015-2025*

[9] A. H. Marbini, A. Ebrahimnejad and M. Tavana, A taxonomy and review of the fuzzy data envelopment analysis literature: Two decades in the making, *Eur. J. Oper. Res.* **214**, 457 – 472 (2011).

[10] J. Nazarko, M. Komuda, A. K. Kuzmicz, E. Szubzda and J. Urban, The DEA method in public sector institutions efficiency analysis on the basis of higher education institutions, *Bad. Oper. Decyz.* **4**, 89–105 (2008).

[11] P. Peykani, E. Mohammadi, M. Rostamy and F.H. Lotfi, Fuzzy data envelopment analysis approach for ranking of stocks with an application to Tehran Stock Exchange, *Adv. Math. Finance Appl.* **4**, 211–240 (2019).

[12] M. Wen and H. Li, Fuzzy data envelopment analysis (DEA): Model and ranking method, *J. Comput. Appl. Math.* **223**, 872—878 (2009).

[13] A. Charnes, W. W. Cooper and E. Rhodes, Measuring the efficiency of decision making units, *Eur. J. Oper. Res.* **6**, 429– 444 (1978).

[14] R. Askari, S. Rafiei, M. Ranjbar, M. Pakdaman and F. Sepase, Estimation of the efficiency of different academic departments using data envelopment analysis: A Study in an Iranian medical university, *J. Med. Educ. Curric. Dev.* **14**, 755–796 (2019).

[15] J. K. Sengupta, A fuzzy systems approach in data envelopment analysis, *Comput. Math. Appl.* **24**, 259–266 (1992).

[16] A. H. Marbini, A. Ebrahimnejad and S. Lozano, Fuzzy efficiency measures in data envelopment analysis using lexicographic multiobjective approach, *Comput. Ind. Eng.* **105**, 362–376 (2017).

[17] R. Qin, Y. Liu, Z. Liu and G. Wang, *Modeling Fuzzy DEA With Type-2 Fuzzy Variable Coefficients*, Lecture Notes in Computer Science, Springer, Berlin (2009).

[18] R. Qin and Y. K. Liu, A new data envelopment analysis model with fuzzy random inputs and outputs, *J. Appl. Math. Comput.* **10**, 99–116 (2009).

[19] N. N. Karnik and J. M. Mendel, Operations on type-2 fuzzy sets, *Fuzzy Sets Syst.* **122**, 327–348 (2001).

[20] Define And Solve A Problem By Using Solver, *https://support.microsoft.com/en-us/office/define-and-solve-a-problem-by-using-solver-5d1a388f-079d-43ac-a7eb-f63e45925040*

[21] H. J. Zimmermann, Fuzzy programming and linear programming with several objective functions, *Fuzzy Sets Syst.* **11**, 45–55 (2001).

[22] E. Mansourirad, M. R. A. Rizam, L. S. Lee and A. Jaafar, Fuzzy weights in data envelopment analysis, *Int. Math. Forum.* **5**, 1871–1886 (2010).

[23] F. J. Gravetter and L. B. Wallnau, *Statistics For The Behavioural Sciences*, Cengage Learning, Boston (2013).

Author information

S.A.M. Ahmed, Department of Applied Statistics, Faculty of Business and Economics, Universiti Malaya 50603 Kuala Lumpur, Malaysia.
E-mail: wana_am@um.edu.my (*Corresponding Author)

M.A. Talib, Department of Applied Statistics, Faculty of Business and Economics, Universiti Malaya 50603 Kuala Lumpur, Malaysia.

N.F.M. Noor, Institute of Mathematical Sciences, Faculty of Science, Universiti Malaya, 50603 Kuala Lumpur, Malaysia.

R. Jani, Ungku Aziz Centre for Development Studies, Universiti Malaya, 50603 Kuala Lumpur, Malaysia.