

A Web-based Decision Support System with ELECTRE III for a Personalised Ranking of British Universities

Christos Giannoulis

University of Portsmouth, Faculty of Technology, Department of Electronic and Computer Engineering, Anglesea Building, Anglesea Road, Portsmouth PO1 3DJ, United Kingdom
Email: cgiannoulis@hotmail.com

Alessio Ishizaka

University of Portsmouth, Portsmouth Business School, Richmond Building, Portland Street, Portsmouth PO1 3DE, United Kingdom
Email: Alessio.Ishizaka@port.ac.uk

Abstract: Reliance upon multi-criteria decision methods, like ELECTRE III, has increased many folds in the past few years. However, ELECTRE III has not yet been applied in ranking universities. League tables are important because they may have an impact on the number and quality of the students. The tables serve an indication of prestige. This paper describes a three-tier Web-system, which produces a customised ranking of British Universities with ELECTRE III reflecting personal preferences, where information is uncertain and vague. Using this case study, the benefits of ELECTRE III in the ranking process are illustrated.

Keywords: Multi-criteria decision method, ELECTRE III, University rankings

1. Introduction

Professor William Cooper is particularly known through his work on DEA (Data Envelopment Analysis)[26]. His paper [17] has been elected as one of the most influential papers published in the *European Journal of Operational Research*. Professor Cooper has applied DEA widely to the performance analysis in the public and private sectors, especially in education. He was the first (founding) Dean at Carnegie Mellon University's School of Urban and Public Affairs (now the H.J. Heinz III School of Public Policy and Management, USA) and a founding member of the Graduate School of Industrial Administration at Carnegie Mellon. He always strives for the improvement of the quality in education as it can be seen in his papers [3; 5; 10; 15; 18; 19].

The evaluation of education with ranking lists of universities has become, over the past few years, increasingly popular. Some examples in United Kingdom are the *Times Higher Education*, *The Complete University Guide*, *The Guardian University Guide* and the *Sunday Times University Guide* all of which produce leagues tables based on statistical data from the Higher Education Statistical Agency (HESA) and the National Student Survey (NSS). These rankings have a sizeable impact on universities as they may have some indication of prestige and a direct influence on the number and

quality of applicants. However the ranking of universities does not use rigorous methodologies like ones used in Professor Cooper's work. The methodology used to rank universities is a simple weighted sum, which has several limitations. First, the weights are predetermined with very little, if any, justification of their value. Therefore, it is assumed that the criteria have the same importance (i.e. weight) for everybody. This is clearly not true as each person is different and has different preferences. Moreover, commercial league tables use a simple aggregation, which is compensatory and does not differentiate between universities having strengths in different areas.

This paper has been prepared to celebrate the 95th birthday of Professor Cooper and his motivation to evaluate education with new methods. We have thus developed a new interactive online way to rank universities with the multi-criteria decision method ELECTRE III [42] (<http://www.pbs.port.ac.uk/IshizakaA/>). As ELECTRE III may be complicated for new users, a simple and an advanced version has been developed. These two versions are user-friendly, free, Web accessible and have tailored functionalities, which is not the case for the old commercial off-the-shelf software supporting the ELECTRE III (<http://www.lamsade.dauphine.fr/english/software.html>). However, the commercial software was used to validate the results of our Web decision support tool.

Hereinafter, we will review methods used for rankings universities. In section three, the ELECTRE III algorithm is described. Section four describes the design and implementation of the decision support tool, and section five evaluates the implemented system. Finally, the concluding section summarises the main points arising from this project.

2. Rankings systems

2.1 Commercial rankings

Several commercial universities ranking schemes are annually published. Alongside, criticisms of these rankings have also increased [13; 34; 37; 51; 53; 55]. These leagues tables are based on a weighted sum of performances, which has some methodological problems. As each criterion is measured in a different unit, they need to be transformed to commensurate units in order to be summed together. The problem is that numerous ways of standardising exist (commercial rankings generally uses z-transformation) and they often lead to a different final ranking. An example can be found in [39], where the authors emphasises that "prior normalization of data is not a neutral operation, and the final result of aggregation may well depend on the normalization method used". The same normalisation problem is also observed in the Analytic Hierarchy Process (AHP), where different normalisations may lead to a rank reversal [7; 30]. Moreover, AHP is difficult to use with a large volume of data, due to the high number of pairwise comparisons required [29].

2.2 DEA

Data Envelopment Analysis (DEA) is an often used ranking technique [2; 33; 44; 46; 47], which does not require any normalisation. The global score of each Decision Making Unit (DMU) is defined as the ratio of the sum of its weighted output levels to the sum of its weighted input levels. The analogy with multi-criteria methods is striking if we replace the name "DMU" with "alternatives", "outputs" with "criteria to be maximised" and "inputs" with "criteria to be minimised". The particularity of this method is that weights are not allocated by users or experts; moreover it does not employ a common set of weights for all alternatives. Instead, for each alternative a different set of weights is calculated with a linear optimisation procedure. The aim of the optimisation is to select weights in order to highlight their particular strength. Some constraints are added in order to ensure that when these weights are applied to all other candidates, none of the scores exceed 100%, the perfect efficiency. DEA has been widely used to rank universities or schools [1; 4; 5; 6; 9; 10; 11; 14; 16; 18;

23; 25; 31; 32] and in many other sectors as compiled in [24]. However, there are some limitations to DEA, which are highlighted below:

- “DEA is not designed to select a single winner” [21; 52]. DEA identifies all alternatives that are located on the efficient frontier as the best alternatives without distinction. When the list of alternatives is large, the number of efficient alternatives may also be large. Further analysis must then be applied to select the best alternative. We know that multiplier restriction method (e.g., cone ratio) has been developed to reduce the number of efficient DMUs. It is possible to identify a single best alternative, using DEA [26]. See also [46, 47].
- “The ranking of inefficient alternatives depends upon which DEA model is used for performance evaluation” [12; 45]. See [48].
- “A conventional use of DEA does not consider the weakness of some candidates”[45; 52]. Any alternative which has the highest score on one criterion is often regarded as efficient, irrespectively of how low it scores on all other criteria. This issue is due to the flexibility in allocating weights in its conventional use, which allows DEA to focus on a few criteria, not putting importance on the others. Note that a new type of DEA [46, 47] does not have such a problem. See also [48].
- “DEA becomes less discriminating as more information is provided”[52]. This problem derives from the critic above. The likelihood that one alternative scores well on one criterion increases with the number of criteria. Thus, unlikely other decision supports methods, the more criteria you have, the less discriminating the method becomes.
- Alternatives that are not on the efficient frontier are not considered as candidates for the final selection [12]. The conventional use of DEA does not recognise optimal non supported alternatives as efficient. See Figure 1.

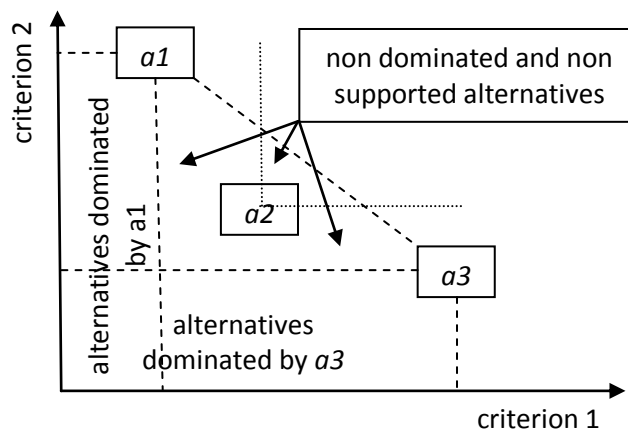


Figure 1: Non-Dominance in Interpretation of DEA Radial Models

Note: Alternatives $a1$ and $a3$ are supported optimal solutions; $a2$ is an optimal non supported solution, the DEA does not consider $a2$ as an efficient solution because it is not on the efficient frontier. This type of problem occurs only in DEA radial models. Non-radial DEA models do not have such a problem. Therefore, $a1$ becomes efficient in the non-radial DEA models.

- All alternatives on the efficient frontier serve as a ranking basis for all other alternatives even if some non-efficient alternatives may be more attractive than efficient alternatives [12]. See Figure 2.

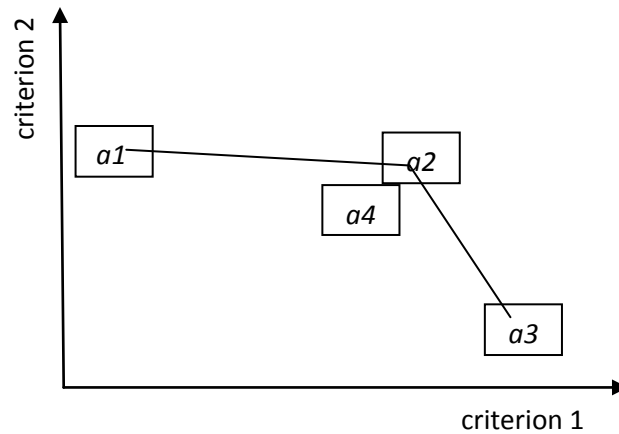


Figure 2: Importance among Alternatives in DEA

Note: Alternatives $a1$, $a2$ and $a3$ are on the efficient frontier serve a ranking basis for $a4$ in the DEA radial model with the assumption of convexity on efficiency frontier. The assumption excludes the alternative $a4$. The alternative is on a higher linear or convex indifference utility curve. The type of problem does not occur in DEA non-radial models. It is true that DEA needs to incorporate information for consensus building among decision makers.

There is an extensive literature which describes techniques to improve the DEA. They generally require more information from the user. The most used techniques use value judgements to constrain weight (multiplier) flexibility [54]. However the exercise of bounding the weights is not trivial as restrictions are subjective and depends on the measurement units of the different inputs and outputs [45]. In order to help the user, visual methods have been developed [8; 23]. These methods are time-consuming and difficult to use with a large amount of inputs and outputs. Of course, this study is fully aware of the recent study [48] that restricts weight (multiplier) by strong complementary slackness condition. Hence, the approach does need any subjective information for weight restriction.

2.3 Ranking with pseudo-criteria

The multi-criteria ranking methods described above, alongside the shortcomings described, are not adapted for uncertain, indeterminate and imprecise data, as explained below:

- Imprecise criteria, because of the difficulty of determining them: students evaluate some criteria (e.g. “*Student satisfaction*”, “*Graduate prospects*”) for the university, where they are studying but judgements are made without a common reference with the other universities [13].
- Indeterminate criteria, because the method for evaluating the criteria is selected relatively arbitrarily between several possible definitions. For example, does the “*Staff/student ratio*” incorporate part-time lecturers and part-time students? How is spending divided between the criteria “*Academic Service Spend*” and “*Facilities Spend*”?
- Uncertain criteria, because the measured values refer only to a point in time and some values vary over time. For example: the “*Employability*” of a university’s graduates depends on the economic situation. The “*Investments in facilities*” may not be uniformly distributed over time.

The ELECTRE III allows imprecise, indeterminate and uncertain criteria inherent to complex human decision processes by relying on the use of pseudo-criteria and indifference and preference thresholds. See Section 3. Furthermore, very bad performance on one criterion may not be

compensated by good scores on the other criteria, depending on the veto threshold. ELECTRE III has been widely used in ranking problems, for instance in ranking the stocks for investment selection [28], for choosing a sustainable demolition waste management strategy [41], for the selection of energy systems [38], for ranking urban stormwater drainage [35] or for housing evaluation [36] but it has not yet been applied for ranking universities.

3. ELECTRE III

3.1 Introduction

ELECTRE III relies upon the construction and the exploitation of the outranking relations. The two distinct phases are depicted in Figure 3:

- a) *Construction of the outranking relation*: Alternatives are pairwise compared (A, B). Each pairwise comparison is characterised by an outranking relation. To say that “alternative A outranks alternative B ” means that “ A is at least as good as B ”. Therefore three outranking relations exists: A is “indifferent”, “weakly preferred” or “strictly preferred” to B depending on the difference between the performance of the alternatives and the thresholds given by the user. See Section 3.2.
- b) *Exploitation of the outranking relation*: Two pre-rankings are then constructed with two antagonist procedures (ascending and descending distillation). The combination of the two pre-ranking gives the final ranking. See Section 3.3.

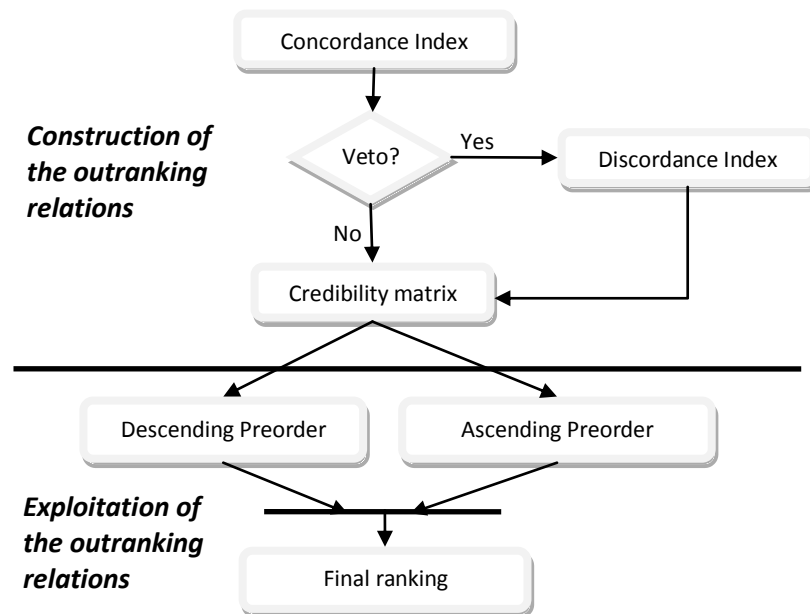


Figure 3: ELECTRE III process flow

3.2 Building the outranking relations

3.2.1 Pseudo-criteria

True criteria, which are the simplest and traditional form of criterion, do not have thresholds. Only the difference between the scores on the criteria is used to determine which option is preferred. In order to take into account imprecision, uncertainty and indetermination in complex decision problems, pseudo-criteria are used. The indifference q and preference p thresholds allow the construction of a pseudo-criterion. Thus, three relations between alternatives A and B can be considered:

- a) A and B are **indifferent** if the difference between the performance of the two alternatives is below the indifference threshold:

$$A \text{ I } B \leftrightarrow z(A) - z(B) \leq q \quad (1)$$

where $z(X)$: performance of the alternative X
 q : indifference threshold

- b) A is **weakly preferred** to B if the difference between the performance of the two alternatives is in between the indifference and the preference threshold:

$$A \text{ Q } B \leftrightarrow q < z(A) - z(B) \leq p \quad (2)$$

where $z(X)$: performance of the alternative X
 q : indifference threshold
 p : preference threshold of the alternative

- c) A is **strictly preferred** to B if the difference between the performance of the two alternatives is higher than the preference threshold:

$$A \text{ P } B \leftrightarrow z(A) - z(B) \geq p \quad (3)$$

where $z(X)$: performance of the alternative X
 p : preference threshold of the alternative

3.2.2 Concordance index

The concordance index (4) indicates the truthfulness of the assertion “ A outranks B ” ($A \text{ S } B$)¹. $C = 1$ indicates the full truthfulness of the assertion and $C = 0$ indicates that the assertion is false. The graphical representation is given in Figure 4.

$$C(A, B) = \frac{1}{W} \sum_{i=1}^n w_i c_i(A, B) \quad (4)$$

where $W = \sum_{i=1}^n w_i$

$$c_i(A, B) = \begin{cases} 1 & \text{if } z_i(B) - z_i(A) \leq q_i \\ \frac{p_i(z_i(A)) + z_i(A) - z_i(B)}{p_i(z_i(A)) - q_i(z_i(A))} & \text{if } q_i < z_i(B) - z_i(A) < p_i \\ 0 & \text{if } z_i(B) - z_i(A) \geq p_i \end{cases} \quad (5)$$

Here,

w_i : weight of the criterion i
 n : number of criteria
 $z_i(X)$: performance of the alternative X as regards to the criterion i
 q_i : indifference threshold for the criterion i
 p_i : preference threshold of the alternative on the criterion i

¹ S is the abbreviation of the French word « Surclasse », as defined in [42].

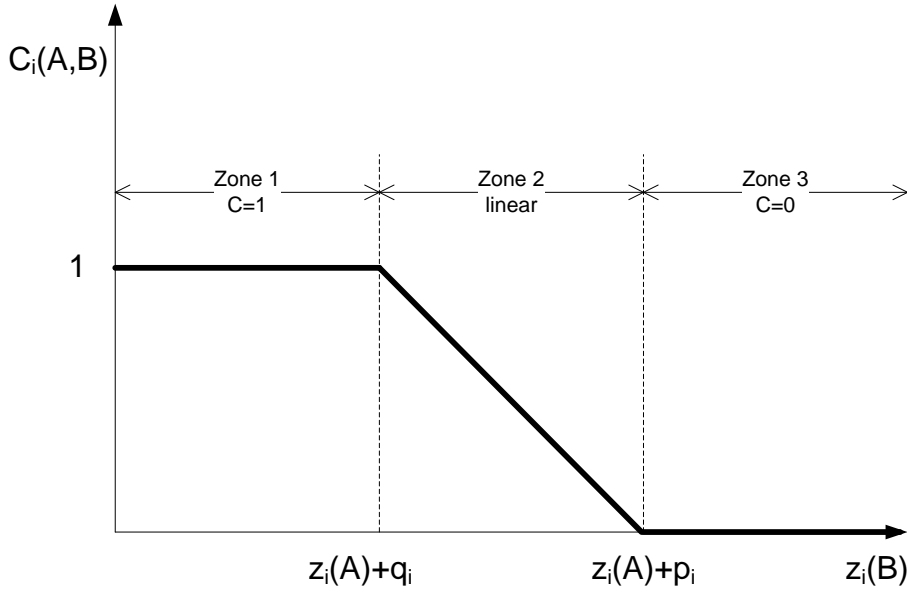


Figure 4: Concordance Index between alternatives A and B

Note:

Zone 1: $z_i(B) - z_i(A) \leq q_i$, the alternatives A and B are indifferent, which means a concordance on the assertion "A outranks B"

Zone 2: $q_i < z_i(B) - z_i(A) < p_i$, the alternative B is weakly preferred to A, which means a partial concordance on the assertion "A outranks B"

Zone 3: $z_i(B) - z_i(A) \geq p_i$, the alternative B is strictly preferred to A, which means a null-concordance on the assertion "A outranks B"

Discordance index: If the difference of performances between the alternative A and B, on a criterion i , is higher than the veto threshold v_i , it is cautious to refuse the assertion "A outranks B". The discordance index for each criterion i is given in (6). Figure 5 shows the graphical representation.

$$D_i(A, B) = \begin{cases} 0 & \text{if } z_i(B) - z_i(A) \leq p_i \\ \frac{z_i(B) - [z_i(A) + p_i]}{v_i - p_i} & \text{if } p_i < z_i(B) - z_i(A) \leq v_i \\ 1 & \text{if } z_i(B) - z_i(A) \geq v_i \end{cases} \quad (6)$$

Here, $z_i(X)$: performance of the alternative X as regards to the criterion i
 p_i : preference threshold of the alternative on the criterion i
 v_i : veto threshold for the criterion i

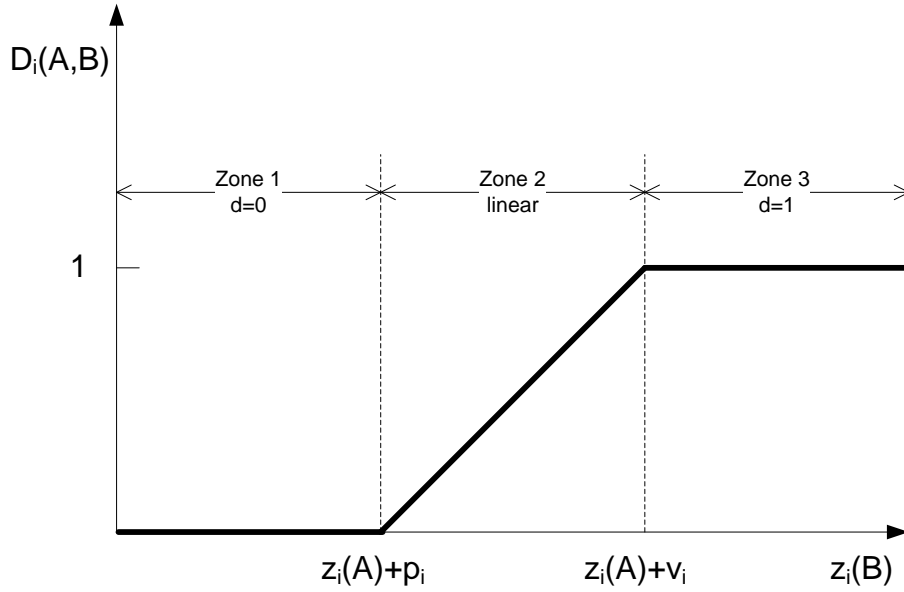


Figure 5: Discordance Index between alternatives A and B

Note:

Zone 1: $z_i(B) - z_i(A) \leq p_i$, the alternatives B is weakly preferred to A, which means no-discordance on the assertion "A outranks B"

Zone 2: $p_i < z_i(B) - z_i(A) < v_i$, the alternative B is strictly preferred to A, which means a weak discordance on the assertion "A outranks B"

Zone 3: $z_i(B) - z_i(A) \geq p_i$, the difference between A and B exceed the veto threshold, which means a total discordance on the assertion "A outranks B"

3.2.3 Degree of credibility

Considering the concordance (4) and discordance indices (6), the degree of credibility (7) indicates if the outranking hypothesis is true or not. If the concordance index (4) is higher or equal to the discordance index of all criteria (6), then the degree of credibility (7) is equal to the concordance index (4). If the concordance index (4) is strictly below the discordance index (6), then the degree of credibility (7) is equal to the concordance index (4) lowered in direct relation to the importance of those discordances.

$$S(A, B) = \begin{cases} C(A, B) & \text{if } D_i(A, B) \leq C(A, B) \forall i \\ C(A, B) \cdot \prod_{i \in J(A, B)} \frac{(1 - D_i(A, B))}{(1 - C(A, B))} & \text{otherwise} \end{cases} \quad (7)$$

where $J(A, B)$ is the set of criteria for which $D_i(A, B) > C(A, B)$.

Then, the degrees of credibility are gathered in a credibility matrix.

Example 1: In order to illustrate the ranking process of ELECTRE III, we will use in following example with six universities and five criteria. See Table 1.

Table 1: Performance Matrix of Universities (U1-U6)

Alternatives	Academic Services Spend	Completion	Entry Standards	Facilities Spend	Good Honours
U1	947	79	400	228	69.7
U2	1406	64	350	204	47.6
U3	677	90	300	349	61.8
U4	561	65	247	188	52.3
U5	1006	88	352	437	65.8
U6	765	77	280	198	55.6

For each criterion of Table 1, thresholds and criteria weights are determined by the user. See Table 2.

Table 2: Thresholds and criteria weights defined by the user

Criterion	Academic Services Spend	Completion	Entry Standards	Facilities Spend	Good Honours
Indifference (q)	0.1	0.1	0.1	0.1	0.05
Preference (p)	0.2	0.2	0.2	0.2	0.2
Veto (v)	0.4	0.5	0.4	0.3	0.4
Weight (w)	0.1	0.3	0.3	0.2	0.1

After calculation of the concordance and discordance indexes, the degrees of credibility are constructed and gathered in the credibility matrix. See Table 3. It can be seen that the two degrees of credibility, attached at each pair of alternatives (one in each way) does not produce a symmetric credibility matrix. The next step is to exploit this matrix. See Section 3.3.

Table 3: Credibility matrix

	U1	U2	U3	U4	U5	U6
U1	1	0	0	1	0	1
U2	0	1	0	0.97	0	0.62
U3	0.0053	0	1	1	0	0.97
U4	0	0	0	1	0	0.22
U5	0.88	0.11	1	1	1	1
U6	0	0	0	1	0	1

3.3 Distillation procedures

From the credibility matrix, a graph can be drawn. Each alternative is linked with each other alternative with two arrows, one each way, indicating the credibility index. For a large number of alternatives, the graph is highly complex. An automated procedure, named distillation, must be used to rank the alternatives. The name distillation has been chosen for the analogy with alchemists, who distil mixtures of liquid to extract a magic ingredient. The algorithm for ranking all alternatives yields two pre-orders.

The first pre-order is obtained with a descending distillation, selecting the best-rated alternatives initially and finishing with the worst. The best alternatives are extracted from the whole set by applying very stringent rules (8). In this sub-set, the best alternatives are selected by applying less restrictive rules (10) (same previously used rules would not bring a different result). The procedure

continues with incrementally lesser restrictive rules and incrementally smaller sub-sets of alternatives. The procedure terminates when only one alternative remains or a group of alternatives that cannot be separated. The second distillation uses the same process but on the original set of alternatives amputated from the best alternative(s) resulting from the first distillation. Thus, a new sub-set is obtained at each distillation, which contains the best alternative(s) of the remaining set. At each distillation, the extracted alternative(s) will be ranked on a lower position.

As each alternative is linked with each other by two arrows, one each way, but not necessarily with a symmetric credibility index, a second pre-order is constructed with an ascending distillation. In this case, the worst rated alternatives are selected first and the distillation terminates with the assignment of the best alternative(s).

For the distillation, the condition needed to state that an alternative A is preferred to B is defined as follow: an alternative A is preferred to B if the degree of credibility of “ A outranks B ” is higher than a threshold λ_2 and significantly higher than the degree of credibility “ B outranks A ” (8).

$$S(A,B) > \lambda_2 \text{ AND } S(A,B) - S(B,A) > s(\lambda_0) \quad (8)$$

where λ_2 is the largest credibility index, which is just below the cut-off level λ_1 as follows:

$$\lambda_2 = \max_{\{S(A,B) \leq \lambda_1\}} S(A,B) \quad \forall \{A,B\} \in G \quad (9)$$

where G is the set of alternatives. λ_1 is the following cut-off level:

$$\lambda_1 = \lambda_0 - s(\lambda_0) \quad (10)$$

where λ_0 is the highest degree of credibility in the following credibility matrix:

$$\lambda_0 = \max_{A,B \in G} S(A,B) \quad (11)$$

and $s(\lambda_0)$ is the following discrimination threshold:

$$s(\lambda_0) = \alpha + \beta \cdot \lambda_0 \quad (12)$$

We use $\alpha = 0.3$ and $\beta = -0.15$ because the two values are recommended values from [43].

With successive distillations, the cut-off level λ_1 is progressively reduced, which makes the condition weaker and it is much easier for A to be preferred than B . However the discrimination threshold contains some arbitrariness as the recommended values α and β are empirical values [50]. Other values could be used, which may slightly change the ranking.

3.4 Extraction

When A outranks B , A is given the score +1 (strength) and B is given -1 (weakness). For each alternative, the individual strengths and weakness are added together to give the final qualification score. Within the descending distillation, the alternative with the highest qualification score is assigned to a rank and removed from the credibility matrix. The process is repeated with the remaining alternatives until all alternatives are ranked.

In the case of several alternatives with the same qualification score, the process is repeated within this subset until either an alternative has a higher qualification score or the highest degree of credibility λ_0 is equal to 0, which means that it is not possible to decide between the remaining options in the subset and therefore they are declared indifferent.

The ascending distillation procedure works in a similar way to the descending distillation with the exception that the procedure assigns the alternative having the lowest qualification score.

Example 2: From the credibility matrix (table 3), the highest credibility degree is $\lambda_0 = 1$ and we can calculate $\lambda_1 = 1 - (0.3 - 0.15 \cdot 1) = 0.85$ and therefore $\lambda_2 = 0.66$. The first qualified university is U5. See Table 4.

Table 4: First round of the qualification

	U1	U2	U3	U4	U5	U6
Outranks	U4, U6	U4	U4, U6	-	U1, U3, U6	U4
Strength	2	1	2	0	3	1
Weakness	1	0	1	4	0	3
Qualification	1	1	1	-4	3	-2

The distillation is repeated with the five remaining universities. Then, the same process is applied for the ascending distillation. See Figure 6.

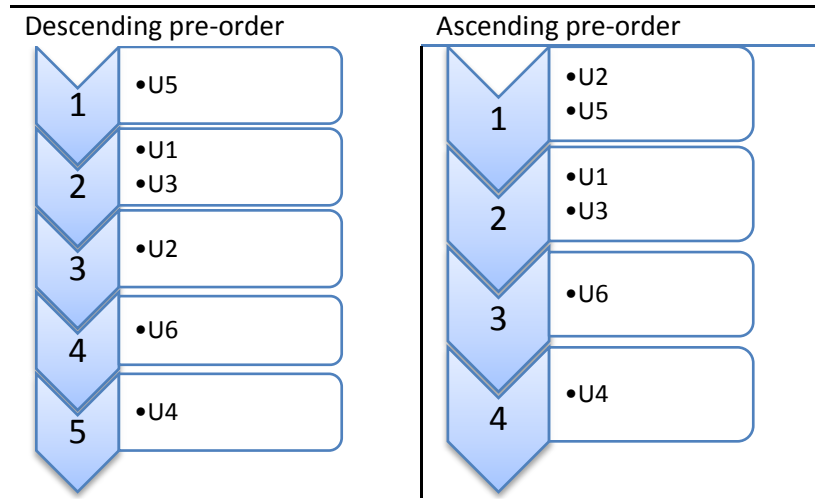


Figure 6: Descending and ascending distillation pre-orders of 6 Universities (U1-U6)

3.5 Final ranking

The final ranking is obtained through the combination of the two pre-orders. See Section 3.4. The results from the partial pre-orders are aggregated into the ranking matrix. We have four possible cases:

- A is higher ranked than B in both distillations or A is better than B in one distillation and has the same ranking in the other distillation
then A is **better** than B : $A \mathbf{P}^+ B$
- A is higher ranked than B in one distillation but B is better ranked than A the other distillation
then A is **incomparable** to B : $A \mathbf{R} B$
- A has the same ranking than B in both distillations
then A is **indifferent** to B : $A \mathbf{I} B$
- A is lower ranked than B in both distillations or A is lower ranked than B in one distillation and has the same rank in the other distillation
then A is **worst** than B : $A \mathbf{P}^- B$

The final ranking is obtained by adding the number of P^+ . In case of tie, the comparison between the two alternatives with the same score decides between an indifferent or incomparable relation See Example 3.

Example 3: If we consider the two pre-orders of Figure 6, the resulting ranking matrix is given in Table 5:

Table 5: Ranking Matrix

	U1	U2	U3	U4	U5	U6	Sum P^+
U1	-	R	I	P^+	P^-	P^+	2
U2	R	-	R	P^+	P^-	P^+	2
U3	I	R	-	P^+	P^-	P^+	2
U4	P^-	P^-	P^-	-	P^-	P^-	0
U5	P^+	P^+	P^+	P^+	-	P^+	5
U6	P^-	P^-	P^-	P^+	P^-	-	1

The final ranking is given in Figure 7, where it can be seen that U1, U2 and U3 have the same scores but U1 and U3 are indifferent and U2 is incomparable to the two other alternatives.

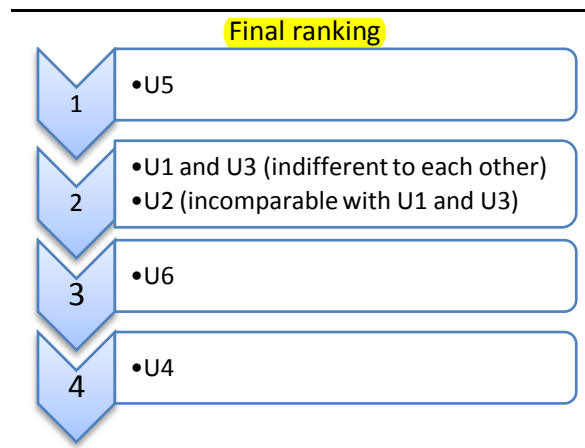


Figure 7: Final ranking

4. Overview of the decision support system

4.1 Introduction

The system architecture is three-tier as shown in Figure 8. It divides the functionality into independent logical layers, each one responsible for different operations of the application and opaque to the other layers.

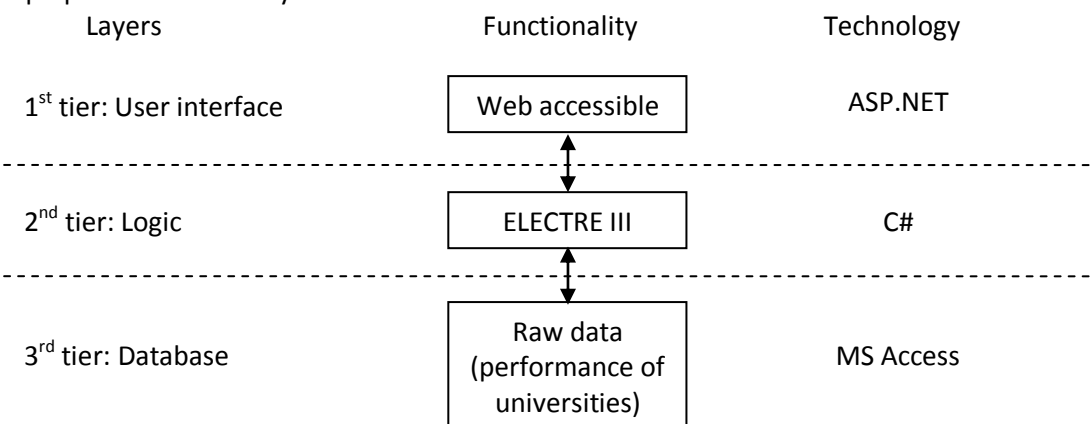


Figure 8: Three-tier architecture

The first layer runs on standard Web browsers. In this tier, the users enter their weighted criteria and related thresholds for the ranking of universities and receive back a personalised ranking. The middle tier contains the Web server, where the ELECTRE III method runs as described in Section 3. Its implementation uses an oriented object design with C#. This layer is independent from the other and therefore can be reusable or upgradable for other decision problems. The bottom layer stores the performance of the universities, which are those used by *The Complete University Guide*. See Table 6. The *Times Higher Education* uses the same criteria with the exception that it merges the criteria “Academic Services Spend” and “Facilities Spend” in one criterion: “Services & Facilities Spend”. The selection of these criteria may be considered controversial. They have been retained in our study because it allows a comparative evaluation of our system with the existing ones (Section 5). However, users are not obliged to select all criteria. In the next section, we discuss the user interface.

Table 6: Criteria used for the ranking of universities [49]

Criterion	Description
Student Satisfaction	Evaluation of students on the teaching quality
Research Assessment	Average quality of the research undertaken
Entry Standards	Average UCAS tariff score of new students under the age of 21
Staff/Student Ratio	Average staffing level
Academic Services Spend	Expenditures per student in all academic services
Facilities Spend	Expenditures per student on staff and student facilities
Good Honours	Percentage of graduates achieving a first or upper second class honours degree
Graduate Prospects	Employability of a university's graduates
Completion	Completion rate of those studying at the university

4.2 User interface

4.2.1 Introduction

The user interface is very important, as it is the link between a person and the system. Because users of this support decision system are unlikely to know ELECTRE III, we have implemented a simple version for them and an advanced version for more experienced users. The goal is to attract users with the simple version and then upgrade them to the advance version. Both versions are based on a wizard style with five steps easily to follow (start page, criteria selection, weights settings, threshold settings and ranking display). The algorithm used is the same for the two versions (see Section 3), only the user interface and the required values differs.

4.2.2 Simple version

The simple version is created for unfamiliar users of the ELECTRE III method or for users wishing to see the universities rankings quickly. As a verbal scale is intuitively appealing, user-friendly and more common in our everyday lives, it has been preferred in this version for the criteria weights and thresholds. The drawback of the user-friendliness is that some arbitrary choice must be made, in this case how the verbal scale is converted to numbers. Table 7 shows the conversion for the weights (scale 2-10) and the indifference thresholds (multiplicative factor 0.2-1).

Table 7: Conversion verbal to numerical scale

Verbal scale	Weights	Indifference thresholds
very low	2	0.2
low	4	0.4
medium	6	0.6
high	8	0.8
very high	10	1

In order to minimise the number of inputs required, we have used the double threshold model, where only the indifference threshold is required. The value of the veto preference threshold is the double of the value of the preference threshold, which is the double of the value of the indifference threshold. See Figure 9. It is an arbitrary choice used in other applications [40], which allows a gain of time and an increase of the usability.

Multiple Criteria UK Universities Rankings

You are in: [Home](#) : [Rankings](#)

1. Criteria : **2. Weights** : 3. Thresholds : 4. My Rankings

Step 2: Choose importance weight for each of the chosen criteria

Student Satisfaction

Medium ▾

Research Assessment

Medium ▾

Entry Standards

Medium ▾

Back

Next

Notes:
For each of the criteria, an importance factor must be set. This factor is a percentage value between 0% to 100%.
The exact meaning of the values are:

Very Low:	20%
Low:	40%
Medium:	60%
High:	80%
Very High	100%

If you prefer to set other values than those proposed here, [Switch to Advanced Version](#)

Figure 9: Simple version with verbal inputs

4.2.3 Advanced version

The advanced version is aimed at users who are or who become familiar with the ELECTRE III. In contrast with the simple version, where there is no fixed scale nor for weights neither for thresholds. Users can enter weights on the numerical scale of their choice (e.g. 1-10 or 1-100). The thresholds are defined with a multiplicative parameter a and an additive parameter b . See Figure 10. For example, suppose that the user select $b=15$ and $a=0.1$ and the performance of alternative X is 100. All alternatives with a performance 100 ± 25 ($25=15+100 \cdot 0.1$) are indifferent to the alternative X.

Multiple Criteria UK Universities Rankings

You are in: [Home](#) : Rankings

1. Criteria : 2. Weights : **3. Thresholds** : 4. My Rankings

Step 3: Define by how much universities should be different

Adjust thresholds for each criterion:

	Student Satisfaction		Research Assessment		Entry Standards	
Indifference:	a= <input type="text" value="0"/>	b= <input type="text" value="0"/>	a= <input type="text" value="0"/>	b= <input type="text" value="0"/>	a= <input type="text" value="0"/>	b= <input type="text" value="0"/>
Preference:	a= <input type="text" value="0"/>	b= <input type="text" value="0"/>	a= <input type="text" value="0"/>	b= <input type="text" value="0"/>	a= <input type="text" value="0"/>	b= <input type="text" value="0"/>
<input type="checkbox"/> Veto:	a= <input type="text" value="0"/>	b= <input type="text" value="0"/>	a= <input type="text" value="0"/>	b= <input type="text" value="0"/>	a= <input type="text" value="0"/>	b= <input type="text" value="0"/>
Min	0		0		0	
Max	4.2		6.6		530	

[Back](#)
[Next](#)

[Contact Help](#)
*We need your feedback.
Please, spend a minute to
participate to our [survey](#).*

Figure 10: Advanced version with numerical inputs

4.2.4 Final ranking

The commercial software for ELECTRE III uses a graph for representing the results. It allows the distinction between indifferent and incomparable. However, this representation is not possible with a large number of alternatives, as it is very difficult to read the results. Our solution uses a table to display the ranking of the 113 universities. It can be seen that rankings depend highly from the criteria and weights selected by the user. See Figures 11 and 12. The distinction between indifferent and incomparable universities is made with colours. See Table 8.

Table 8: Status of ranked universities

Status	Colour	Description
Normal	No colour	One university per rank
Indifferent	<div style="background-color: grey; width: 40px; height: 20px; display: flex; align-items: center; justify-content: center;">grey</div>	Two or more universities have the same rank and are indifferent
Incomparable	<div style="background-color: yellow; width: 40px; height: 20px; display: flex; align-items: center; justify-content: center;">yellow</div>	Two or more universities have the same rank and are incomparable



Figure 11: Final customised ranking

Note: In this case, the universities of York, Bath, Bristol, Manchester and Dundee are all ranked 10. The first four are indifferent (in grey) and the last one is incomparable (in yellow)

Multiple Criteria UK Universities Rankings

You are in: [Home](#) : Rankings

[Contact Help](#)
*We need your feedback.
Please, spend a minute to
participate to our [survey](#).*

1. Criteria : 2. Weights : 3. Thresholds : **4. My Rankings**

Your university rankings are:

Rank	Institution	Student Satisfaction	Student-Staff Ratio	Completion
1	Lampeter	3.8	26.3	77
2	Leeds Metropolitan	3.5	24.2	83
3	Lincoln	3.7	23.7	84
3	London South Bank	3.7	25.7	72
5	Greenwich	3.6	23.7	80
6	York St. John	3.7	22.7	85
7	Bath Spa	3.9	21.7	86
8	Edge Hill	3.8	22.1	78
9	Roehampton	3.7	21.5	80
9	Bedfordshire	3.8	21.7	75
11	University of the Arts - London	3.4	21	87
12	Sheffield Hallam	3.6	20.6	85
12	Cumbria	3.5	21.2	76
14	Middlesex	3.5	22.9	72

Figure 12: Additional Comparison in Final customised ranking

Note: Compared with Figure 11, other criteria and weights can lead to a very different ranking

5. Evaluation

There were two types of evaluation that were conducted: a questionnaire and an observation of users. We sent a questionnaire by email to 800 students of our university, as the application has been conceived mainly to help students in the selection of a university. They were asked to visit the Web-system and answer an anonymous questionnaire. This approach was selected in order that students can complete the questionnaire where, when and only if they want it. The disadvantage of this unpressured voluntary exercise is a low participation: 20 participants (2.5% response rate). However, the collected data gives us some significant indications.

80% of the respondents claimed that the university ranking was of interest to them but only 55% have already visited a Website providing this type of ranking. This observation is in line with past researches [22; 27] indicating that students do not rely on commercial rankings in choosing their university in UK. Only one participant knew ELECTRE III before. 50% of the participants used only the simple version, 20% of the participants used only the advanced version and 30% used both versions. Users, even those unfamiliar with ELECTRE III, were able to understand and appreciate the working of the system. When they were asked to assess the Web-system (Figures 13 and 14), the results clearly indicate that the system was helpful and better than the other current solutions.

For the second evaluation, a group of 10 masters' students were observed. All students first used the simple version, which is easier and selected by default on the welcome page. At the beginning, some students asked the purpose of the thresholds. Then after, they found the system easy and straightforward to use. All students rated the system far superior to the current commercial rankings as the user can select the criteria, their allocated weight and it returns more information (e.g.

indifferent and incomparable universities). Finally, they encouraged the developers to implement a similar system ranking universities by specialities (like business schools or engineering faculties).

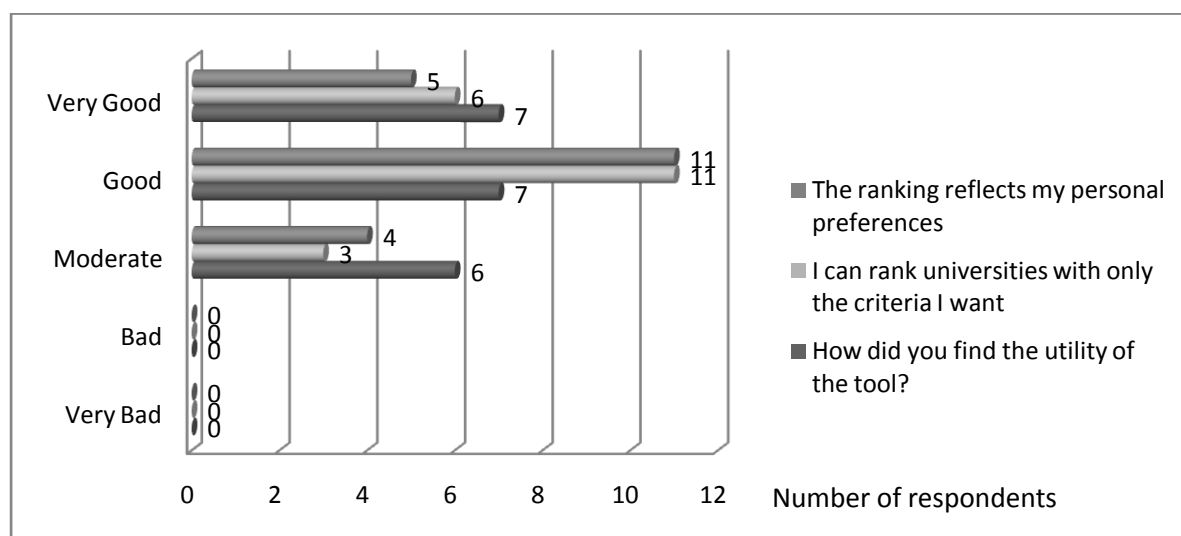


Figure 13: Bar chart evaluating the tool

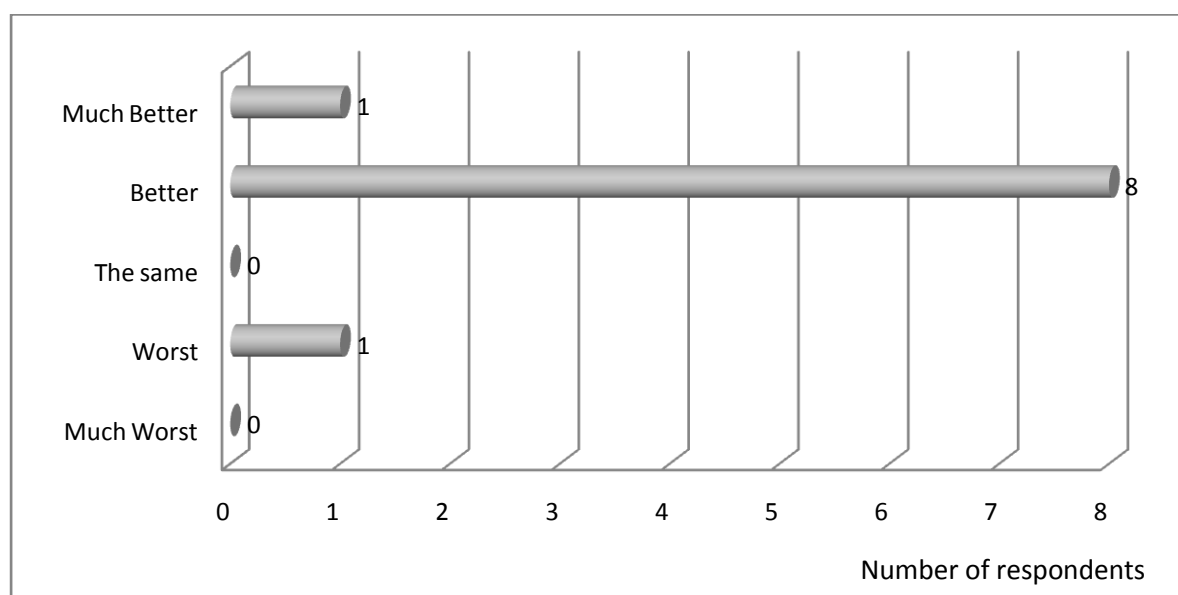


Figure 14: Bar chart of comparison with other Websites

6. Conclusion

Professor Cooper was one of the first researchers to evaluate the performance of education institutions, especially with DEA. Acknowledging his contribution in performance evaluation, we have seen that DEA has a problem when used as a ranking method [45]. It has even been called “Multiattribute for lazy decision maker” [20] as no input is required from the user. Today, several commercial rankings of Universities are periodically published. These ranking have been severely criticised and may not be as useful as students do not rely on them for selecting their university. Oswald [37] concludes that “Britain needs a wider range of rankings” in order to help students. Our web-system respond to this need as it provides a personalised ranking of the 113 official British Universities. It cannot be ascertained that a university is always better than another as the ranking depends on the criteria and weights selected by the user. Ranking universities rely on imprecise,

indeterminate and uncertain criteria; therefore ELECTRE III with its pseudo-criteria was appropriate for this problem. Furthermore it allows:

- to bypass the problem of the full aggregation of incommensurate performances,
- to reveal any disastrous criterion with the veto threshold,
- to distinguish between indifferent and incomparable alternatives, and
- to compare a very large number of universities (the limitation is given by the physical storage of data and not from ELECTRE III).

However, ELECTRE III suffers from some issues in the exploitation process of the outranking relations. As the graph of the relations may be complicated, especially for a large number of alternatives, an automatic process must be used to generate the final ranking. For this purpose, the distillation has been developed. The drawback of the distillation is that an arbitrary threshold has to be selected, which may have an impact on the final ranking. In our future research, we will investigate the impact of this threshold on the final ranking. We will also research other methods to exploit the outranking relations.

Acknowledgements

The authors are indebted to their colleagues Vijay Pereira, David Whitmarsh, Michael Wood and two anonymous referees for their valuable comments and suggestions, which have greatly improved the earlier version of the paper.

7. References

- [1] M. Abbott and C. Doucouliagos, The efficiency of Australian universities: a data envelopment analysis, *Economics of Education Review* 22(1) (2003) 89-97.
- [2] N. Adler, L. Friedman and Z. Sinuany-Stern, Review of ranking methods in the data envelopment analysis context, *European Journal of Operational Research* 140(2) (2002) 249-265.
- [3] T. Ahn, V. Arnold, A. Charnes and W. Cooper, DEA and ratio efficiency analyses for public institutions of higher learning in Texas, *Research in Governmental and Nonprofit Accounting* 5 (1989) 165-185.
- [4] F. Arcelus, An efficiency review of university departments, *International Journal of Systems Science* 28(7) (1997) 721-729.
- [5] V. Arnold, I. Bardhan, W. Cooper and S. Kumbhakar, New uses of DEA and statistical regressions for efficiency evaluation and estimation — with an illustrative application to public secondary schools in Texas, *Annals of Operations Research* 66(4) (1996) 255-277.
- [6] A. Athanassopoulos and E. Shale, Assessing the comparative efficiency of Higher Education Institutions in the UK by the means of Data Envelopment Analysis, *Education Economics* 5(2) (1997) 117-134.
- [7] V. Belton and A. Gear, On a Shortcoming of Saaty's Method of Analytical Hierarchies, *Omega* 11(3) (1983) 228-230.
- [8] E. Bernroider and V. Stix, Profile distance method - a multi-attribute decision making approach for information system investments, *Decision Support Systems* 42(2) (2006) 988-998.
- [9] A. Bessent and W. Bessent, Determining the comparative efficiency of schools through Data Envelopment Analysis, *Educational Administration Quarterly* 16(2) (1980) 57-75.
- [10] A. Bessent, W. Bessent, A. Charnes, W. Cooper and N. Thorogood, Evaluation of educational program proposals by means of DEA, *Educational Administration Quarterly* 2 (1983) 82-107.
- [11] A. Bessent, W. Bessent, J. Kennington and B. Reagan, An application of mathematical programming to assess productivity in the Houston independent school district, *Management Science* 28(12) (1982) 1355-1367.
- [12] D. Bouyssou, Using DEA as a tool for MCDM: some remarks, *Journal of the Operational Research Society* 50(9) (1999) 974-978.

- [13] R. Bowden, Fantasy Higher Education: university and college league tables, *Quality in Higher Education* 6(1) (2000) 41-60.
- [14] T. Breu and R. Raab, Efficiency and perceived quality of the nation's "top 25" National Universities and National Liberal Arts Colleges: An application of data envelopment analysis to higher education, *Socio-Economic Planning Sciences* 28(1) (1994) 33-45.
- [15] P. Brockett, W. Cooper, L. Lasdon and B. Parker, A note extending Grosskopf, Hayes, Taylor and Weber, "Anticipating the consequences of school reform: A new use of DEA", *Socio-Economic Planning Sciences* 39(4) (2005) 351-359.
- [16] H. Cengiz and M. Yuki, Measuring Value in MBA Programmes, *Education Economics* 6(1) (1998) 11-25.
- [17] A. Charnes, W. Cooper and E. Rhodes, Measuring the efficiency of decision making units, *European Journal of Operational Research* 2(6) (1978) 429-444.
- [18] A. Charnes, W. Cooper and E. Rhodes, Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through, *Management Science* 27(6) (1981) 668-697.
- [19] W. Cooper and L. McAlister, Can research be basic and applied? You bet. It better be for B-schools!, *Socio-Economic Planning Sciences* 33(4) (1999) 257-276.
- [20] J. Doyle, Multiattribute choice for the lazy decision maker: let the alternatives decide!, *Organizational Behavior and Human Decision Process* 62(1) (1995) 87-100.
- [21] J. Doyle and R. Green, Data envelopment analysis and multiple criteria decision making, *Omega* 21(6) (1993) 713-715.
- [22] C. Eccles, The use of university rankings in the United Kingdom, *Higher Education in Europe* 27(4) (2002) 423-432.
- [23] S. El-Mahgary and R. Lahdelma, Data envelopment analysis: Visualizing the results, *European Journal of Operational Research* 83(2) (1995) 700-710.
- [24] A. Emrouznejad, B. Parker and G. Tavares, Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA, *Socio-Economic Planning Sciences* 42(3) (2008) 151-157.
- [25] L. Friedman and Z. Sinuany-Stern, Scaling units via the canonical correlation analysis in the DEA context, *European Journal of Operational Research* 100(3) (1997) 629-637.
- [26] F. Glover and T. Sueyoshi, Contributions of Professor William W. Cooper in Operations Research and Management Science, *European Journal of Operational Research* 197(1) (2009) 1-16.
- [27] R. Gunn and S. Hill, The impact of league tables on university application rates, *Higher Education Quarterly* 62(3) (2008) 273-296.
- [28] N. Huck, Pairs selection and outranking: An application to the S&P 100 index, *European Journal of Operational Research* 196(2) (2009) 819-825.
- [29] A. Ishizaka. (2008). *A Multicriteria Approach with AHP and Clusters for Supplier Selection*. Paper presented at the 15th International Annual EurOMA Conference, Groningen.
- [30] A. Ishizaka and A.W. Labib. (2008). *Towards fifty years of the Analytic Hierarchy Process*. Paper presented at the 50th Operational Research Society Conference, York.
- [31] J. Johnes, Data envelopment analysis and its application to the measurement of efficiency in higher education, *Economics of Education Review* 25(3) (2006) 273-288.
- [32] J. Johnes and L. Yu, Measuring the research performance of Chinese higher education institutions using data envelopment analysis, *China Economic Review* 19(4) (2008) 679-696.
- [33] M. Mannino, S.N. Hong and I.J. Choi, Efficiency evaluation of data warehouse operations, *Decision Support Systems* 44(4) (2008) 883-898.
- [34] S. Marginson, Global university rankings: implications in general and for Australia, *Journal of Higher Education Policy and Management* 29(2) (2007) 131-142.
- [35] C. Martin, Y. Ruperd and M. Legret, Urban stormwater drainage management: The development of a multicriteria decision aid approach for best management practices, *European Journal of Operational Research* 181(1) (2007) 338-349.

- [36] E. Natividade-Jesus, J. Coutinho-Rodrigues and C.H. Antunes, A multicriteria decision support system for housing evaluation, *Decision Support Systems* 43(3) (2007) 779-790.
- [37] A. Oswald, An economist's view of university league tables, *Public Money & Management* 21(3) (2001) 5-6.
- [38] A. Papadopoulos and A. Karagiannidis, Application of the multi-criteria analysis method Electre III for the optimisation of decentralised energy systems, *Omega* 36(5) (2008) 766-776.
- [39] J.C. Pomerol and S. Barba-Romero, *Multicriterion Decision in Management: Principles and Practice*, (Kluwer Academic Publishers, 2000).
- [40] M. Rogers, Using ELECTRE III to aid the choice of housing construction process within structural engineering, *Construction Management and Economics* 18(3) (2000) 333-342.
- [41] N. Roussat, C. Dujet and J. Méhu, Choosing a sustainable demolition waste management strategy using multicriteria decision analysis, *Waste Management* 29(1) (2009) 12-20.
- [42] B. Roy, ELECTRE III: algorithme de classement base sur une présentation floue des préférences en présence de critères multiples, *Cahiers du CERO* 20(1) (1978) 3-24.
- [43] B. Roy and D. Bouyssou, *Aide Multicritère d'Aide à la Décision: Méthodes et Cas*, (Economica, Paris, 1993).
- [44] C. Serrano-Cinca, Y. Fuertes-Callén and C. Mar-Molinero, Measuring DEA efficiency in Internet companies, *Decision Support Systems* 38(4) (2005) 557-573.
- [45] T. Stewart, Relationships between data envelopment analysis and multicriteria decision analysis, *Journal of the Operational Research Society* 47(5) (1996) 654-665.
- [46] T. Sueyoshi and M. Goto, DEA–DA for bankruptcy-based performance assessment: Misclassification analysis of Japanese construction industry, *European Journal of Operational Research advance online publication* (2009) doi:10.1016/j.ejor.2008.1011.1039.
- [47] T. Sueyoshi and M. Goto, Can R&D expenditure avoid corporate bankruptcy? Comparison between Japanese machinery and electric equipment industries using DEA–discriminant analysis, *European Journal of Operational Research* 196(1) (2009) 289-311.
- [48] T. Sueyoshi and K. Sekitani, An occurrence of multiple projections in DEA-based measurement of technical efficiency: Theoretical comparison among DEA models from desirable properties. *European Journal of Operational Research* 196(2) (2009) 764-794.
- [49] The league table of UK universities 2009 - The Complete University Guide, <http://www.thecompleteuniversityguide.co.uk/single.htm?ipg=6310#How%20the%20League%20Table%20works> (2008).
- [50] E. Takeda, A method for multiple pseudo-criteria decision problems, *Computers & Operations Research* 28(14) (2001) 1427-1439.
- [51] P. Taylor and R. Braddock, International university ranking systems and the idea of university excellence, *Journal of Higher Education Policy and Management* 29(3) (2007) 245-260.
- [52] C. Tofallis, Selecting the best statistical distribution using multiple criteria, *Computers & Industrial Engineering* 54(3) (2008) 690-694.
- [53] J. Vaughn, Accreditation, commercial rankings, and new approaches to assessing the quality of university research and education programmes in the United states, *Higher Education in Europe* 27(4) (2002) 433-441.
- [54] Y.-M. Wang, K.-S. Chin and G.K.K. Poon, A data envelopment analysis method with assurance region for weight generation in the analytic hierarchy process, *Decision Support Systems* 45(4) (2008) 913-921.
- [55] M. Yorke, A good league table guide, *Quality Assurance in Education* 5(2) (1997) 61-72.

Bibliography

Alessio Ishizaka is Senior Lecturer at the Portsmouth Business School of the University of Portsmouth. He received his PhD from the University of Basel (Switzerland). He worked successively for the University of Exeter (UK), University of York (UK) and Audencia Grande Ecole de Management Nantes (France). His research is in the area of decision-making. He was the co-organiser of a special AHP stream on the 50th Operational Research Society Conference, York, 2008 and he is a founding member of the Decision Deck Association (<http://www.decision-deck.org/>).

Christos Giannoulis obtained his MSc in Software Engineering at the School of Computing of the University of Portsmouth in 2008. His research interests are on the fields of software development methods, project management and decision analysis.