MEASUREMENT OF DMU-EFFICIENCY BY MODIFIED CROSS EFFICIENCY APPROACH

LAVOSLAV ČAKLOVIĆ AND TIHOMIR HUNJAK

Abstract. The fundamental weakness of a Data Envelopment Analysis (DEA) is its weak discrimination in cases when small number of Decision Making Units (DMU) have been compared. Therefore in such cases the basic DEA model is used in combination with other methods or additional constraints are added to the model. In this paper, the Cross-efficiency method has been combined with the Potential Method (PM).

1. Introduction

DEA is a well know method [1, Cooper at all. 2002] used to evaluate the efficiency of the entities named DMU and it is based on the ratio comparison between outputs produced by a DMU and inputs spent by DMU for the production purpose. The results of the analysis conducted with this method has been successfully applied to banks management, schools, hospitals and similar institutions because, based on the results obtained, activities which increase their operative efficiency can be generated. Numerous variations of this method have been developed depending on following specificities:

(1) if variables under control of a decision-maker are inputs (input oriented model) or outputs (output oriented model),
(2) if there is a constant or varying return on scale,
(3) if there are any limitations regarding weights of inputs and outputs etc.

In the first part of the paper a basic DEA model has been described and the concept of the DMU efficiency has been explained. Then, discrimination problem in case of smaller number of DMU has been commented and elementary approaches on how to solve the problem have been mentioned briefly. The Cross-efficiency method has been described in details. The Cross-efficiency analysis has been demonstrated on a specific example. The concept of input and output flow has been introduced as options for ranking the efficient DMU. The Potential Method has been demonstrated in more details and it has been shown how this method could be used to improve discrimination of DMU by appreciating basic DEA principles.

2. Basic DEA model

Let's assume that a DMU is the entity which transforms a certain number of inputs into outputs through a specific process. Those entities are banks, schools, hospitals etc. It could be assumed that inputs are resources which are used to

---

2000 *Mathematics Subject Classification.* 90B50, 91B06.

Key words and phrases. Multi-criteria decision making, DEA, crossefficiency, Potential Method, selfranking.

achieve certain outputs and it is desirable to create the highest level of outputs for the lowest possible set of inputs. The measure of efficiency for a particular DMU is a ratio between outputs earned and inputs spent. It could be assumed that a quality management would have an effect on inputs spent. As all DMU do not have equal opportunities in accomplishing that, therefore as a measure of efficiency ratio between weighted outputs and inputs is used and weights are given such values which would benefit particular DMU mostly.

Let us denote decision making units by DMU_j, $j = 1, \ldots, n$. A basic concept in this approach is *cross efficiency table*. The the main step in calculating its entries is calculating an output/input ratio for each DMU

$$\text{eff}_o(u, v) = \frac{\sum_{r=1}^{s} u_r y_r^o}{\sum_{i=1}^{m} v_i x_i^o}, \quad o \in \{1, \ldots, n\} \tag{2.1}$$

and it's maximization

$$\text{eff}_o = \max_{u>0, v>0} \frac{\sum_{r=1}^{s} u_r y_r^o}{\sum_{i=1}^{m} v_i x_i^o}, \tag{2.2}$$

where

$$y_r^o = \text{an amaunt of } r\text{-th output for } o\text{-th DMU}$$
$$x_i^o = \text{an amaunt of } i\text{-th input for } o\text{-th DMU}$$

with respect to the following conditions

$$\frac{\sum_{r=1}^{s} u_r y_r^o}{\sum_{i=1}^{m} v_i x_i^o} \leq 1, \quad \text{for each } o \in \{1, \ldots, n\}.$$

The maximal value of goal function eff is a measure of DMU efficiency. If the value equals 1, DMU is considered efficient; therefore there is no other DMU which would produce bigger output with equal inputs (output oriented model) or the same output as the observed DMU with the lesser input expenditure. Instead of this problem of fractional linear programming, the equivalent problem of linear programming has being solved in practice (or his dual), see [1, page 43].

Apart from the strengths of this method, there are also some limitations which could reduce potential of using it in certain situations. That especially applies to the fact that a particular DMU can be efficient thanks to its significant advantage in single input (output) or in its small subset. The second limitation is its weak discrimination potential in case of smaller number of DMU. Namely, in cases when a small number of DMU is compared, it can happen that the number of efficient DMUs is so high that the results of analysis lose their practical value. The rule of thumb is: number of DMUs which are compared should be threefold higher than the sum of inputs and outputs.

2.1. **Concepts for Discrimination Improvement.** In certain cases, when the number of efficient DMUs is too high, in order to efficient DMs, additional assumptions such as specific restrictions on the severity of criteria have to ben introduced, or a basic DMU model has to be combined with other methods. The most renowned concepts for the discrimination improvement are [4, Despotis, 2002]:

- Cross-efficiency approach

- Multi-criteria DEA approach
- Global efficiency approach
- Assurance Region Method (AR)
- Multicriteria Benefit/Cost analysis

As this paper proposes combination between the Cross-efficiency method and the Potential Method, the other mentioned procedures used for improving discriminations between DMUs are described only as concep. In a multi-criteria DEA approach, together with the original DEA criteria, two new criteria (min max criterion and min sum criteria) have been introduced, so associated problem of multiobjective linear programming is being solved. Beside the extreme optimal solutions of the problem which are also DEA efficient, the number of nondominated (efficient) solution is not as high as the number of DEA efficient solutions because all the efficient DEA solutions are nondominated solutions of the added multiobjective programming problem.

Global efficiency approach is also related to the multi-criteria programming. In this approach, a problem of multi-criteria fractional linear programming has been joined to the original DEA model, and its ideal point has been searched for optimization of each objective functions. Components of ideal point have optimal efficiencies of all DMU, and the idea of global scores has been connected with the distance from the ideal point. Some DMUs which are efficient for the basic DEA model and which are successful in maintaining efficiency in relation to the associated problem are considered globally efficient.

In the Assurance Region (Cooper) method the possibility of choosing outputs and inputs weights has been restricted according to certain assumptions based on the Analyst's preferences. That way, certain DMU have been prevented from becoming efficient based on some specificity which could arise as a consequence of some anomalies. Multicriteria Benefit/Cost ratio is an extension of the Benefit/Cost (B/C) analyses. Based on the Decision-maker assessment, the weights of outputs (in the terms of B/C analysis we consider them benefits) and inputs (in the terms of B/C analysis those are costs) has been determined. Therefore, for each DMU the B/C ratio between sum of weighted outputs and sum of weighted inputs is calculated separately, and DMUs are ranked according to those ratios.

(2.3)

	DMU_1	$\cdots$	DMU_i	$\cdots$	DMU_n
DMU_1	c_{11}	$\cdots$	c_{1i}	$\cdots$	c_{1n}
$\vdots$					$\vdots$
DMU_i	c_{i1}	$\cdots$	c_{ii}	$\cdots$	c_{in}
$\vdots$					$\vdots$
DMU_n	c_{n1}	$\cdots$	c_{ni}	$\cdots$	c_{nn}

TABLE 1. Cross efficiency table

2.2. **Cross efficiency approach (C-E).** The idea of the Cross efficiency approach that alleviates the weak discrimination of the basic DEA model could be explained in two steps. Firstly, the basic DEA analysis is carried out and for each DMU optimal weights of inputs and outputs are calculated. Let (u^i, v^i) be the vectors of optimal weights for DMU_i.

In the next step so called cross-efficiency matrix C has to be constructed. The element c_{ij}, in i-th row and j-th column of matrix C is the ratio

$$c_{ij} = \text{eff}_j(u^i, v^i)$$

of outputs and inputs of the DMU_j , weighted by the optimal weights of DMU_i. This means that column j consists of the efficiencies of the DMU_j measured by optimal weights of DMU_i .

The average value of the elements has been calculated for every column and DMUs are ranked according to those values.

2.3. **Income flow, outcome flow and net flow.** Is it possible to utilize the information potential of the cross efficiency analysis in a better way? Namely, in that analysis the standard measure of efficiency is the Average by peers index. According to the intuitive interpretation of the DEA analysis (each particular DMU will be evaluated from the position that suits it best) Average by peers index reflects the influence of the remaining DMU on the one being observed. Roughly speaking, in terms of the graph theory, if a DMU is observed as the graph's knot, Average by peers of the observed DMU could be interpreted as an income flow from the related knots. In fact, data in crossefficiency table are more subtle. Each entry c_{ij} in that table is an amount of strength of DMU_j given by DMU_i. Let us suppose that weights of DMUs are known. Then, the flow obtained as a difference of that strength can be calculated and, using Potential Method, new weights can be obtained. Repeating this proces leads to the fixed point problem. Details are given in section 4.3.

3. An example

Described procedure is illustrated with an example using business data of a community bank's branches. Following performance measures of the branches ere observed: number of employees, operating costs, interest paid per savings, interests earned on loans and non-interest incomes. The data are shown in the Table 2. To rate the branches' efficiency, the DEA input oriented model has been used, where the number of employees and the non-interest income are treated as controlled inputs, and the other performances are treated as outputs. This model has been chosen because in the branches observed it is possible to influence more rational use of the employees and, in certain terms the majority of the costs could be controlled. The cost of interests paid per savings is used as output because they indicate the capital collected that is used for the commercial loans. All collected savings are not stored only within branches which collect them, but they are also stored through the other branches. Therefore, as a specific output, interests' data which has been reimbursed for the credits has been used. As branches also offer other services, the non-interest incomes data has been used as output. In the efficiency analyses *Frontier analyst, Banxia* software is used.

In the Table 3 the result of the DEA method is shown based on the data from the Table 2. It is shown that three branches are 100% efficient: Branch 4, Branch 6

DMU (Branch)	employee	operating costs*	interest paid per savings*	interests earned on loans*	non-interest incomes
1	20	829326	4449202	4786608	1000188
2	7	342554	1020605	1686859	307375
3	7	262008	861443	1516144	426604
4	11	301114	4022446	6491851	1151494
5	9	244918	400783	654434	407243
6	6	326759	3056784	1994946	1055240
7	7	269277	1634220	2291636	1083105
8	6	288521	1232645	1788427	1001151
9	4	165573	445955	904764	462190
10	6	218150	536914	1036494	545877
11	3	132788	229635	387528	160227
12	23	924037	4879496	8471185	470160

* output

Table 2. Input data for eleven bank branches.

and Branch 7. Because of the high level of efficiency (99,65%), Branch 8 has been included in this group of efficient branches as well (the further analysis will show that this branch has been ranked as the weakest between the efficient branches).

DMU (Branch)	CCR score
1	51.22
2	41.27
3	45.59
4	100.00
5	41.34
6	100.00
7	100.00
8	99.65
9	72.86
10	62.21
11	34.24
12	62.41

Table 3. Identification of efficient DMU.

	DMU6	DMU7	DMU4	DMU8
DMU6	100	91.57	100	93.08
DMU7	80.29	100	95.07	86.27
DMU4	100	91.57	100	93.08
DMU8	100	100	74.74	99.65

Table 4. Cross efficiency table of efficient banks

The other results, which are integral part of the standard DEA analysis, will not be commented, as the emphasis of this work is not on them. We are interested how the efficient branches could be ranked based on the data available and by accepting the results of the DEA analysis in the highest possible measures. With that purpose, the Cross efficiency analysis has been made based on data from the Table 2.

Based on the standard data interpretation from the analysis i.e. based on the Average by peers' index, the branches are ranked in the stated order: Branch 4, Branch 6, Branch 7 and Branch 8. However, in the efficiency measuring of all the branches, the order is an application result of the optimal value of their performances. It seems that, in order to compare efficiency of the branches, it makes sense to conduct the Cross-efficiency analysis only between them, so that they could be compared without redundancy which is the result of data describing their relation with inefficient branches. Table 4 provides results of the Cross-efficiency analysis only for the efficient branches.

The order of the efficient branches is significantly different from the structure assigned on all the data; Branch 6 is the best and thereafter Branches 7, 8 and 4 are ranked. Average by peers gives a slightly different ranking.

	DMU6	DMU7	DMU8	DMU4
PM-rank	0.299	0.244	0.235	0.222
Average by peers	95.07	95.79	93.02	92.45

Table 5. Ranking of efficient DMU.

4. Potential method, a brief description

Results described in Table 5 are obtained by Potential Method (PM). We shall briefly describe this method in the sequel. Details can be found in [2, Čaklović, to app.] and [3, Čaklović, to app.].

4.1. **Single criterion case.** For simplicity, let us suppose that one decision maker makes pairwise comparisons on the set of alternatives V, with respect to a single criterion. The results of pairwise comparisons can be organized in a positive reciprocal matrix as it is done in Saaty, [5, Saaty, 1996]. If we want to treat the data using matrix analysis, all elements of this matrix should be given. In practice, it often happens that all comparisons cannot be done i.e. some data is missing.

Potential Method is designed to handle incomplete data as well, organizing them as a weighted oriented graph.

A pair $\alpha = (u, v) \in V \times V$ is declared to be an arc of a directed graph if v is more preferred than u. An un-compared pair is not adjacent in the graph. The set of all arcs is denoted by $\mathcal{A}$. A function $\mathcal{F} : \mathcal{A} \to \mathbb{R}$ which assigns to each arc $\alpha \in \mathcal{A}$ its weight of preference is called a `preference flow`. The flow component $\mathcal{F}(\alpha)$ is usually taken from a given nonnegative interval of real numbers. It is equal to zero if u and v have equal preference. Evidently, preference flow is always non-negative and can be represented as an $m \times 1$ matrix. Oriented graph $(V, \mathcal{A})$ is called a `preference graph`. The preference graph is `complete` if each pair of alternatives is compared. For a given flow $\mathcal{F}$ on the preference graph $(V, \mathcal{A})$ and $\alpha \in \mathcal{A}$ we use a convention $\mathcal{F}(-\alpha) := -\mathcal{F}(\alpha)$.

Let us denote by $n = \#V$ and $m = \#\mathcal{A}$ the cardinality of V and $\mathcal{A}$ respectively. Incidence matrix of the preference graph is denoted by B and it is $m \times n$ matrix defined by

$$B_{\alpha,i} = \begin{cases} -1, & \text{if } \alpha \text{ leaves } i \\ 1, & \text{if } \alpha \text{ enters } i \\ 0, & \text{otherwise.} \end{cases}$$

Let $\mathcal{F}$ be a given preference flow, and B incidence matrix of the graph. Moreover, let us assume that the graph is weakly connected (connected in the rest of the paper). System

$$B^{\tau} B X = B^{\tau} \mathcal{F}, \quad \textstyle\sum_{i=1}^{m} X_i = 0 \tag{4.1}$$

has a unique solution called `normal integral` of $\mathcal{F}$. One can think of the first equation in (4.1) as the normal equation associated to $BX = \mathcal{F}$. The potential difference BX of normal integral is the best approximation of $\mathcal{F}$ by column space of incidence matrix (the least square problem). If the graph is not connected, the normal integral is unique on each connected component (maximal connected subgaph) of the graph. To obtain the ranking from X, the following formula is used

$$w = \frac{a^X}{\|a^X\|_1}, \quad a > 1$$

where exponential function of X is defined componentwise, i.e. $(a^X)_i = a^{X_i}$, and $\| \cdot \|_1$ is 1-norm. Parameter $a > 1$ can be arbitrary. When preference graph is complete then the equation (4.1) has a solution in the following explicit form

$$X_i = \frac{1}{n} \sum_{i=1}^{n} (B^{\tau} \mathcal{F})_i$$

The component X_i of the potential X is then interpreted as a flow difference for i-th node, i.e. the difference between input and output flow for i-th node, divided by the number of nodes in the graph.

`Measure of inconsistency` is defined as

$$\mathrm{Inc}(\mathcal{F}) = \frac{\|\mathcal{F} - BX\|_2}{\|BX\|_2},$$

where $\|\cdot\|_2$ denotes 2-norm and $\beta = \arctan(\mathrm{Inc}(\mathcal{F}))$ is an `angle of inconsistency.` Ranking is considered 'acceptable' if β is less than 12 degrees. The last statement should not be taken for granted, as there is no serious argument, appart from statistical research, to support it due to the freshness of the method.

4.2. **Consensus flow.** If more than one criterion is present then, each criterion C_i generates its own graph $(V, \mathcal{A}_i)$ and its own flow $\mathcal{F}_i$. Let us denote the weight of the i-th criterion by w_i. We are going to describe a procedure of making a `consensus graph` $(V, \mathcal{A})$ and `consensus flow` $\mathcal{F}$ for the group of all criteria.

First, for a given pair $\alpha = (u, v)$ we calculate

$$\mathcal{F}_\alpha := \sum_{\substack{i=1 \\ \pm\alpha\in\mathcal{A}_i}}^{k} w_i \mathcal{F}_i(\alpha) \tag{4.2}$$

where the term $w_i\mathcal{F}_i(\alpha)$ contributes if and only if $\pm\alpha \in \mathcal{A}_i$ i.e. if and only if $\mathcal{F}_i(\alpha)$ or $\mathcal{F}_i(-\alpha)$ is defined. If this sum is non-negative, then we put α in the set of arcs $\mathcal{A}$ and $\mathcal{F}(\alpha) := \mathcal{F}_\alpha$. Otherwise, we define $-\alpha = (v, u)$ as an arc in $\mathcal{A}$ and $\mathcal{F}(-\alpha) := -\mathcal{F}_\alpha$. The flow $\mathcal{F}$ becomes a non-negative flow and is called `consensus flow`. It can happen that the consensus graph has a cycle. Anyway, normal integral of $\mathcal{F}$ exists and it is unique. The presence of cycles can only generate bigger inconsistency $\mathrm{Inc}(\mathcal{F})$.

4.3. **Self-ranking.** A powerful feature of Potential Method is self-ranking. By self-ranking we mean a decision model in which criteria and alternatives are the same. This situation happens in a group self-evaluation when each member gives a preference graph on the set of group members. A member can include himself in a preference graph or not. Moreover each group member can use its own criteria in ranking the others.

Let us denote by $G = \{1, 2, \ldots, n\}$ a group of decision makers and by $\Sigma = \{\xi \mid \xi_i = 1,\, \xi \geq 0\}$ the standard simplex in $\mathbb{R}^n$. Let us suppose that each decision maker $i \in G$ gave the preference flow $\mathcal{F}_i$ on the set G with respect to certain criteria. If $\xi \in \Sigma$ is an a priori given group ranking and $\mathcal{F}_G$ is the consensus flow, then it should satisfy the equation

$$\sum_{i\in G} \xi_i \mathcal{F}_i = \mathcal{F}_G. \tag{4.3}$$

The same relation should take place for the normal integrals X_i and the normal integral X_G of the consensus flow, because of linearity of the equation (4.1). Now we can write

$$\sum_{g\in G} \xi_i X_i = X_G. \tag{4.4}$$

If X denotes the matrix with columns $X_i,\ i = 1, \ldots, n$ then, the left side of the above equation can be written as a product $X\xi$ between matrix X and column ξ. A function

$$\Phi_X : \xi \mapsto \frac{a^{X\xi}}{\|a^{X\xi}\|_1} \tag{4.5}$$

defined on the standard simplex Σ to itself is now well defined and the group ranking derived from the consensus flow should be a fixed point of Φ, i.e. it should satisfy the equation

$$\xi = \Phi_X(\xi). \tag{4.6}$$

Existence of the fixed point is a consequence of Brouwer's fixed point theorem. The following theorem gives uniqueness of the fix point under some restrictions on Φ_X. The proof can be found in [3].

Theorem 4.1. *Let us suppose that*

$$2 \ln a \|X\|_\infty < 1 \tag{4.7}$$

Then, Φ_X is a contraction and for each $\xi \in \Sigma$ the sequence $(\Phi_X)^n(\xi)$ converges to the unique fixed point $\xi_0 \in \Sigma$ of Φ_X.

The condition (4.7) can be satisfied by changing parameter a. Moreover, for complete graphs, the norm $\|X\|_\infty$ is smaller as n becomes bigger because the matrix $B^\tau B$ on the main diagonal $n-1$ and the system (4.1) can be transformed by Gauss method of elimination to the equivalent system with matrix that has n on the main diagonal. Different situation is with sparse graphs but this will not be the case in our application.

5. Consensus flow for cross efficiency table

As the application of Theorem 4.1, we shall calculate the importance (weight) of each DMU from their Cross-efficiency table. As an example we shall take eleven branches of one Croatian bank. Input data are is given in Table 2.

5.1. **Identification of efficient DMUs.** Efficient branches are given in Table 3 and their Cross-efficiency table is Table 4.

5.2. **Ranking of efficient DMUs.** Each (j, i)-entry in Cross-efficiency table, Table 4, is of the form (2.1) for some input and output weights u, v, and taking its logarithm we can suppose that it has the form

$$\log Y_{ji}.$$

Let us denote by $\mathcal{F}^{(i)}$ a preference flow associated to i-th DMU, i.e. to i-th column. As a component of that flow on arc (j, k) we take the difference of (k, i)-entry and (j, i)-entry, i. e.

$$\begin{aligned} \mathcal{F}^{(i)}_{(j,k)} &= \log Y_{ki} - \log Y_{ji} \\ &= \log \frac{Y_{ki}}{Y_{ji}}, \end{aligned}$$

and component of the consensus flow $\mathcal{F}^G$ for that arc equals

$$\mathcal{F}^G_{(j,k)} = \sum_i w_i \log \frac{Y_{ki}}{Y_{ji}} = \log \prod_i \left(\frac{Y_{ki}}{Y_{ji}} \right)^{w_i},$$

where $w \in \Sigma$ denotes the unknown group weight. In other words, the component of consensus flow is the logarithm of the quotient of geometric means of the corresponding rows. By substituting $\mathcal{F}^G$ into equations (4.3) and using (4.4), (4.5) we obtain equation (4.6). For

Fixed point obtained from data in crosseffitiency Table 4 is given in Table 5. It can be seen that Average by peers gives and ranking obtained by Potential method are not the same. The reason can be in introducing the fourth ineffitient DMU in consideration. May be that introducing all DMU in consideration is the right way but this should be carefully investigated.

References

[1] Cooper, W. W. and Seiford, L. M. and K. Tone, *Data Envelopment Analysis*, Kluwer, 2002
[2] Čaklović L. *Decision Making by Potential Method*, to appear in Int. J. Pure Appl. Math.
[3] Čaklović L. *Self–duality in Group Decision*, to appear in Int. J. Pure Appl. Math.
[4] Despotis, D. K., *Improving the discriminating power of DEA: focus on globally efficient units*, Journal of the Operational Research Society, 2002, **53**,314-323
[5] Saaty T. L. (1996) *The Analytic Hierarchy Process*, RWS Publications, Pittsburg

Faculty of Natural Sciences – Department of Mathematics, Bijenička 30, 10 000 Zagreb, Croatia
E-mail address: caklovic@math.hr

Faculty of Organization and Informatics, Pavlinska 2, 42000 Varaždin, Croatia
E-mail address: thunjak@foi.hr