Carrillo, Marianela; Jorge, Jesús M.

Article

Integrated approach for computing aggregation weights in cross-efficiency evaluation

Operations Research Perspectives

Provided in Cooperation with:
Elsevier

Suggested Citation: Carrillo, Marianela; Jorge, Jesús M. (2018) : Integrated approach for computing aggregation weights in cross-efficiency evaluation, Operations Research Perspectives, ISSN 2214-7160, Elsevier, Amsterdam, Vol. 5, pp. 256-264, https://doi.org/10.1016/j.orp.2018.08.005

This Version is available at:
https://hdl.handle.net/10419/246354

Operations Research Perspectives 5 (2018) 256–264

Contents lists available at ScienceDirect

Operations Research Perspectives

journal homepage: www.elsevier.com/locate/orp

Integrated approach for computing aggregation weights in cross-efficiency evaluation

Marianela Carrillo[a,*], Jesús M. Jorge[b]

[a] Departamento de Economía Aplicada y Métodos Cuantitativos, Universidad de La Laguna (ULL), 38071 La Laguna, Spain
[b] Departamento de Ingeniería Informática y Sistemas, Universidad de La Laguna (ULL), 38071 La Laguna, Spain

ARTICLE INFO

Keywords:
Data envelopment analysis
Cross-efficiency evaluation
Aggregation weights
Ranking

ABSTRACT

Cross-efficiency evaluation is an extension of data envelopment analysis that has been widely used in many different applications aimed at producing a ranking of the set of decision making units. Besides the traditional self-appraisal of units, cross-evaluation methods also take into account peer-appraisals, which are then summarized into an overall performance measure. The standard approach for this aggregation process relies on an equally-weighted average that disregards that some cross-efficiency scores might be considered more relevant or reliable than others. This paper focuses on the aggregation process of cross-efficiency scores and proposes a new approach for deriving meaningful aggregation weights for a more comprehensive evaluation of the units. Our method integrates two complementary perspectives that weights should reflect: the discriminatory ability of the information contained in the cross-efficiency matrix and the relative importance that can be attributed to each of the peer-appraisals. In this sense, the approach presented here provides a more accurate evaluation of the units than previous approaches and therefore it is likely to produce more meaningful rankings. Some numerical examples are provided that validate the approach proposed and examine the results obtained in comparison with previous known methods.

1. Introduction

Data envelopment analysis (DEA), first introduced in [1], is a linear programming technique useful for assessing the relative efficiency of a homogeneous set of decision making units (DMUs) that operate in a production system where multiple inputs are consumed to produce multiple outputs. After four decades of development, there is still intense research activity in the field both at a theoretical as well as empirical level [2], which has proved DEA to be a valuable tool for performance evaluation in many different contexts, with interesting applications in health care, education, banking, manufacturing, etc.

In the traditional DEA model DMUs' performance is assessed using an efficiency score defined as a ratio of a weighted sum of outputs to a weighted sum of inputs. These efficiency scores are obtained through a self-evaluation process, where each DMU is allowed to choose its own set of optimal input and output weights that guarantee a maximum efficiency ratio, as long as the scores of all DMUs calculated from the same weights do not exceed one. According to this evaluation framework, DMUs obtaining a unitary efficiency value are regarded as efficient units whereas DMUs that are unable to attain the maximum efficiency level are considered to perform inefficiently.

This self-assessment scheme with total flexibility in weights selection is especially suitable for the identification of inefficient DMUs, but very commonly too many units are classified as efficient performers not allowing further discrimination or ranking among them, which may result unsatisfactory in several decision contexts aiming at finding the best performer. This lack of discrimination in DEA applications is well documented, particularly when the number of inputs and outputs is too high relative to the number of DMUs being evaluated, and some empirical rules have been suggested to avoid too many units being classified as efficient [3]. Further research has been undertaken with the aim of increasing the discriminative power of DEA. As a result, the traditional model has been extended in different directions (see for example [4,5] and references therein), including for example weight-restriction models, super-efficiency models or common-weight models, although perhaps the methods based on a cross-evaluation approach stand among the most commonly used for ranking DMUs.

Cross-efficiency evaluation, firstly proposed in [6], complements the traditional self-evaluation mode of DEA with a peer-evaluation mode, in such a way that each DMU is also assessed using the most favorable weight set of its peer DMUs. Consequently, in a cross-evaluation framework DMUs act both as evaluated and evaluating units.

* Corresponding author.
E-mail addresses: mcarrif@ull.es (M. Carrillo), jjorge@ull.es (J.M. Jorge).

https://doi.org/10.1016/j.orp.2018.08.005
Received 1 June 2018; Received in revised form 22 July 2018; Accepted 29 August 2018
Available online 31 August 2018

This evaluation scheme, where DMUs are repeatedly assessed using a range of input and output weights instead of a single set of weights, provides a more detailed view of the performance of DMUs, allowing a glimpse of how sensitive the performance assessments of units are to the weight pattern used. Eventually, the scores that a DMU obtains when it is rated by its peers and the self-rated efficiency score are averaged into an overall performance measure that summarizes the different appraisals received by that DMU. As Doyle and Green [7] pose it, cross-evaluation approaches enjoy the same connotations of a democratic process, in the sense that each and every DMU's preferences regarding weights are taking into account within the evaluation procedure, and therefore the results obtained are likely to be considered as a consensual assessment.

Besides this interesting feature, cross-evaluation approaches are found to achieve complete discrimination among DEA-efficient units, which is particularly effective for purposes of ranking. These advantages explain the extensive use of cross-efficiency based approaches in applications involving performance evaluation of DMUs for decision-making within a wide range of fields (see for example, [8–13]). The relevance of this line of research has also been confirmed in the study carried out by Liu et al. [2] who identified cross-efficiency related studies to be one of the more active DEA research subareas in recent years.

Despite the interesting advantages and vast applicability of cross-efficiency evaluation, it has received some critics due to non-uniqueness of cross-efficiency scores. Given that the optimal set of DEA weights selected by each DMU is not necessarily unique, multiple cross-efficiency scores can be obtained depending on the specific optimal solution that the LP solver generates. To overcome this problem, Sexton et al. [6] suggested the use of some secondary objectives to guide the selection of a particular set of optimal weights in a specified direction. Particularly, they proposed the so-called benevolent and aggressive formulations, aimed at finding a set of weights that guarantee the optimal efficiency score of the evaluated DMU while making the others DMUs' cross-efficiencies as large (for benevolent) or small (for aggressive) as possible. Throughout the years, a great deal of research has focused on this topic and many different secondary goal models have been proposed to handle the non-uniqueness issue, either following a benevolent or aggressive strategy [7,14,15] or introducing neutral objectives that avoid taking a position for or against the peers [16,17]. All in all, the introduction of a secondary goal can be seen as an opportunity to intentionally specify a particular strategy for selecting weights that makes the procedure better fit the desired aim of the analysis. Moreover, it has significantly enriched the theoretical development of the methodology.

Less attention has been paid in the literature to the issue of the aggregation of the self and peer evaluated cross-efficiencies into a single cross-efficiency score for each unit. In the standard approach the cross-efficiency score of each DMU is defined as the average of the cross-appraisals received, and although other aggregation measures could also be adopted, they are rarely applied [18]. Particularly, the use of a simple average implicitly assumes that the assessments provided by all the DMUs are equally relevant or reliable. However, this is not necessarily always the case and several arguments can be used to justify that attaching equal aggregation weights to all the cross-efficiency scores may not be completely satisfactory and can fail to reflect the real performance of the evaluated units [19]. In this sense, grounded in the belief that using a weighted average aggregation introduces a higher degree of modeling flexibility that may lead to more accurately assessments of DMUs' performance, some authors have studied the calculation of relative importance weights to be used during the aggregation of cross-efficiency scores, and a few approaches have subsequently been developed that differ in the way the notion of DMUs importance is tackled.

In this work an alternative method is proposed to derive aggregation weights that estimate the relative importance of the cross-efficiencies provided by different DMUs. The rationale of our approach develops from known strategies for weight elicitation in multicriteria decision making, given that the scores that need to be aggregated in a cross-evaluation scheme bear a notable resemblance with the ratings of an alternative set across a number of criteria. Particularly, we suggest considering two components in the definition of the importance weights: an intrinsic component, which is intended to reflect how much information can be inferred from the DMUs appraisals for a discrimination purpose, and a contextual component, which is intended to reflect how relevant or valuable the DMUs appraisals can be considered within the background conditions where the evaluation process takes place.

The rest of the paper is organized as follows. Section 2 describes the traditional cross-evaluation approach and how the aggregation of cross-efficiency scores is tackled in the literature. In Section 3 an alternative approach that computes relative importance weights for cross-efficiency aggregation is presented. Then, some numerical examples from the DEA literature are examined in Section 4 to illustrate the application of our approach in comparison with the traditional average cross-efficiency and other weighted cross-efficiency approaches and finally, some concluding remarks are provided in Section 5.

2. Average and weighted cross-efficiency evaluation

Let us consider n production units or DMUs, each of them being evaluated in terms of r inputs and s outputs. Using the standard notation, let x_{ij} and y_{kj} be nonnegative values denoting respectively the amount of input i consumed and the amount of output k produced by the jth DMU ($i = 1, \ldots, r$, $k = 1, \ldots, s$, $j = 1, \ldots, n$). In this setting, the original DEA ratio model developed in [1] allows each DMU to choose an optimal set of input and output weights to achieve the maximum efficiency score, defined as the ratio of the weighted sum of outputs to the weighted sum of inputs, constrained to no other DMU scoring more than one. Therefore, for each DMU q under evaluation the following non-linear program is formulated:

$$\begin{aligned} E_{qq} = \max \quad & \frac{\sum_{k=1}^{s} u_{kq} y_{kq}}{\sum_{i=1}^{r} v_{iq} x_{iq}} \\ \text{s. t.} \quad & \\ & \frac{\sum_{k=1}^{s} u_{kq} y_{kj}}{\sum_{i=1}^{r} v_{iq} x_{ij}} \leq 1 \quad j = 1, \ldots n \\ & u_{kq},\ v_{iq} \geq 0 \end{aligned} \tag{1}$$

which can be suitably transformed into an equivalent linear program that is usually known as CCR model, named after its authors Charnes, Cooper and Rhodes:

$$\begin{aligned} E_{qq} = \max \quad & \sum_{k=1}^{s} u_{kq} y_{kq} \\ \text{s. t.} \quad & \sum_{i=1}^{r} v_{iq} x_{iq} = 1 \\ & \sum_{k=1}^{s} u_{kq} y_{kj} - \sum_{i=1}^{r} v_{iq} x_{ij} \leq 0 \quad j = 1, \ldots n \\ & u_{kq},\ v_{iq} \geq 0 \end{aligned} \tag{2}$$

Problem (2) must be solved n times in order to obtain a set of optimal weights u_{kq}^*, v_{iq}^* ($i = 1, \ldots, r$, $k = 1, \ldots, s$) and the efficiency scores E_{qq} for all the units analyzed $q \in \{1, \ldots, n\}$, allowing a classification of the DMU set into efficient ($E_{qq} = 1$) and non-efficient ($E_{qq} < 1$) units. When the most preferred weights for a given DMU are used to compute an efficiency score for the other DMUs we obtain the so-called cross-efficiency values, $E_{jq} = \frac{\sum_{k=1}^{s} u_{kq}^* y_{kj}}{\sum_{i=1}^{r} v_{iq}^* x_{ij}}$ ($j = 1, \ldots, n$) which represent the evaluation of DMU j under the perspective of DMU q. The resulting $n \times n$ values can be gathered into a matrix that will be referred to as cross-efficiency matrix (CEM), which is represented in Table 1. As seen,

Table 1
Cross-efficiency matrix (CEM).

Rated DMU	Rating DMU				
	1	…	q	…	n
1	E_{11}	…	E_{1q}	…	E_{1n}
⋮	⋮		⋮		⋮
j	E_{j1}	…	E_{jq}	…	E_{jn}
⋮	⋮		⋮		⋮
n	E_{n1}	…	E_{nq}	…	E_{nn}

the entries in each column of the CEM contains the evaluations provided by the corresponding DMU (rating DMU) while each row compiles the efficiency scores attained by the corresponding DMU (rating DMU) with respect to the optimal weighs of the peers. Being this so, the self-evaluated efficiency values lie on the leading diagonal positions.

A row-wise exploration of the CEM gives an image of the performance of each unit across the different perspectives determined by all the units. Then, the cross-evaluation method proceeds by aggregating the self and peer cross-efficiency scores in each row into a single value, called cross-efficiency value, that will be used for a ranking performance of the DMU set.

Two main issues arise in relation to the practical application of cross-efficiency evaluation methods. First, given that program (2) may have multiple optimal solutions, multiple values of cross-efficiencies E_{jq} for $q \neq j$ can therefore be obtained, depending on the particular optimal weight pattern used by DMU q for the assessments. To solve this ambiguity, Sexton et al. [6] proposed the use of a secondary objective function for the choice among the alternative optimal solutions to problem (2). They suggested that a DMU could assume a benevolent or aggressive attitude towards the other DMUs, in the sense of selecting weights that maintain its own self-evaluated score while being as advantageous (for benevolent) or harmful (for aggressive) as possible to the others DMUs. Following this principle, the formulations proposed by Doyle and Green [7] are perhaps the most commonly-used in applications. Their secondary-goal models are expressed as follows:

$$
\begin{aligned}
\max/\min \quad & \sum_{k=1}^{s}\left(u_{kq}\sum_{j=1,j\neq q}^{n} y_{kj}\right)\\
\text{s. t.}\quad & \\
& \sum_{i=1}^{r}\left(v_{iq}\sum_{j=1,j\neq q}^{n} x_{ij}\right)=1\\
& \sum_{k=1}^{s} u_{kq}y_{kj}-\sum_{i=1}^{r} v_{iq}x_{ij}\leq 0 \qquad j=1,\ldots n;\ j\neq q\\
& \sum_{k=1}^{s} u_{kq}y_{kq}-E_{qq}\sum_{i=1}^{r} v_{iq}x_{iq}=0\\
& u_{kq},\ v_{iq}\geq 0 \qquad k=1,\ldots s;\ i=1,\ldots r
\end{aligned}
\tag{3}
$$

representing linear surrogates for the maximization (for a benevolent approach) or minimization (for an aggressive formulation) of the average of the peer DMUs' scores when assessed with the weights of the qth DMU. Several alternative secondary goals have been formulated throughout the years that approach benevolent or aggressive strategies from different perspectives [14,20]. Regarding the selection between both schemes, some authors argue that benevolent formulations can be considered more consistent with the classical DEA spirits based on being as favorable as possible to each DMU, while aggressive models are deemed more beneficial as regards the discrimination problem [15]. Alternatively, to avoid taking a position for or against the peers, further methodological options have emerged that determine an optimal weight set for each unit without being particularly benevolent or aggressive to the others. The so-called neutral models are only concerned with obtaining the most favorable result for the unit under evaluation without paying attention to the effect of the selected weights on the peer DMUs [21,22].

Having obtained the cross-efficiency matrix, a second issue that has to be addressed to complete the cross-evaluation procedure concerns the aggregation of the cross-efficiency scores in CEM. The simple average is by far the most commonly-used aggregation measure, with little attention being paid in the literature to alternative aggregation schemes in comparison with the many pages devoted to alternative secondary goals. Therefore, the traditional approaches compute the average cross efficiency (ACE) of DMU j as the arithmetic mean of the n cross-efficiency scores that it receives when systematically evaluated by all the units, which are found in the jth row of the CEM:

$$\mathrm{ACE}_j = \frac{1}{n}\sum_{q=1}^{n} E_{jq} \tag{4}$$

However, a reasonable alternative to the simple average that is likely to produce more comprehensive assessments is a weighted average, with weights determined to approximate the importance of each individual cross-efficiency score instead of assuming the same importance for all of them. Then, for some given importance weights adding up to 1, the weighted cross efficiency (WCE) of DMU j is computed as:

$$\mathrm{WCE}_j = \sum_{q=1}^{n} w_q E_{jq} \tag{5}$$

Not many studies have considered the use of relative importance weights for cross-efficiency aggregation. One of the pioneering contributions in this regard was proposed by Wu, Liang and Yang [23], who used game theory concepts with significant computational burdens to derive aggregation weights based on the Shapley value in the cooperative game. Wang and Chin [24] later suggested that self-rated efficiencies should play a more important role in the aggregation than peer-rated efficiencies, for all the evaluated DMUs. Working on this idea, they used an ordered weighted averaging (OWA) operator for the aggregation process, in such a way that each DMU would always attach the greater weights to the more favorable assessments received. The use of OWA operators allows additional flexibility into the modeling process by adjusting the orness degree, which measures to what extent the aggregation emulates the maximum operator. In that way, as the selected orness degree gets closer to 1, more weight is put on the highest assessment and more cross-scores are discarded from the aggregation. However, an important drawback of this approach is that the weights attached to the cross-efficiencies provided by a given DMU will be different across the evaluated units, which causes that the overall cross-efficiency values are hardly comparable and the ranking derived can be put under question [25,26]. This shortcoming is addressed in [27] by using an induced ordered weighted averaging (IOWA) approach, where the order of importance of the DMUs is determined by the order-inducing variable selected, that should be related to the evaluations provided by the DMUs, while the intensity of importance (i.e., the value of the aggregation weights) is determined by the orness degree and the particular method used for deriving the weights (the minimax disparity approach, in the above case), irrespective of the magnitudes of the cross-efficiency scores.

From a different perspective, Wang and Wang [28] proposed some interesting approaches to compute importance weights for the DMUs based on the analysis of the assessments they provide. Their least-square dissimilarity approach considers that those evaluating units providing assessments that deviate too much from the other DMUs' assessments are more unreliable and therefore they should be assigned the lower weights. Alternatively, their least-square deviation approach attaches the bigger weights to the DMUs providing scores with minimum deviation from the optimal DEA efficiency scores. A similar approach was also used in [29] who aimed at reducing the inconsistencies in the evaluation process by minimizing the deviation

between the entropy values of self and peer-evaluated scores.

Admittedly, choosing aggregation weights for summarizing the information contained in the CEM remind us of the problem of selecting criteria weights for evaluating alternatives in multicriteria decision making (MCDM), if we identify criteria with CEM columns. In that context, a well-known principle states that a criterion that rates all alternatives very similarly does not offer any additional information and therefore it becomes useless for the decision-making process, since it does not help discriminating among the alternatives [30,31]. Recently, Song and Liu [19] noticed that, according to this principle, it is reasonable to attach the lower importance weights to the evaluating DMUs that provide similar cross-efficiency scores across all the evaluated units, and methods that fail to comply with this rule may eventually lead to unsatisfactory cross-evaluation. Based on that conception, they propose to use the variation coefficient method based on the Shannon entropy of the cross-efficiency scores as an alternative to deviation approaches like the used in [29] and also in [28].

Following this line of reasoning, next section presents a proposal that takes ideas for weight elicitation from the field of MCDM and extends the work in [19] by complementing the importance weight based on discriminatory ability with a second component reflecting the relevance that can be attributed to the evaluating DMUs in their peers' view.

3. Integrated aggregation weights for cross-efficiency evaluation

The approach presented here is theoretically supported by well-known techniques for measuring attribute importance in MCDM. A close relationship between DEA and MCDM has been highlighted on different occasions [32], and that connection comes into sight once again when facing the selection of aggregation weights for computing cross-efficiency scores, a task that bears a close resemblance to the problem of measuring criteria importance for the assessment and ranking of a given set of alternatives. To see this, the CEM should be interpreted as a decision matrix where n alternatives (rated DMUs) are evaluated under n different criteria (represented by the different DEA-weights used by rating DMUs).

The problem of measuring importance weights for criteria has received considerable attention in MCDM. Undoubtedly, considering equal weights can be deemed a practical, acceptable option in many decision contexts, but it seems much more reasonable to admit that not all the criteria are equally relevant for the decision making process, which has led to many different approaches for the computation of weights. In that regard, Zeleny's view is that in a given decision situation attribute weights should be related to the concept of intrinsic information contained in the decision matrix and, in parallel, to a concept of a priori importance that reflects the evaluation background [30]. In order to simultaneously account for both dimensions, an overall importance weight can be formulated as a multiplicative combination of them. As a result, the weight values obtained integrate the merits of the two notions included.

According to this general scheme, a new approach for cross-efficiency evaluation is proposed as follows:

Step 1. Determination of the intrinsic weight component. As is widely accepted in MCDM, attributes that rate all available alternatives very similarly are not helpful in making a decision. In this sense, the relative importance of a given criterion should be proportional to its ability to discriminate among the alternatives. In our context this means that the cross-efficiency scores obtained from a given evaluating DMU that generates similar scores across all DMUs should be assigned a smaller weight when computing the aggregate score. Consequently, we will define intrinsic information weights in direct proportion to the dispersion degree of the data as given by the coefficient of variation. Therefore, for each column q in CEM we compute

$$\mathrm{CV}_q = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(E_{iq}-\overline{E}_q)^2}}{\overline{E}_q}, \qquad q = 1, \ldots, n \tag{6}$$

where $\overline{E}_q = \frac{1}{n}\sum_{i=1}^{n} E_{iq}$, and then the intrinsic information weights will be determined by

$$\lambda_q^1 = \frac{\mathrm{CV}_q}{\sum_{q=1}^{n}\mathrm{CV}_q} \tag{7}$$

Step 2. Determination of the contextual weight component. To account for the evaluation background we will define importance weights that represent the relevance of each evaluating DMU in the view of the peer DMUs. Inspired by the way that peer-citations that an author receives are used in academic contexts to judge his/her relevance in the field, we propose to estimate the importance of a DMU in terms of how highly it ranks when its performance is compared to others by a group of peers.

To take this idea into practice, let us consider the matrix $R = (R_{zt})$ where R_{zt} represents the ranking order of unit z according to the evaluations provided by unit t. We assume than when two or more units tie for a certain position, all of them receive the same (minimum) ranking order and then subsequent ranks are skipped in such a way that the position of the units ranked below them is unaffected. In this way, the columns of R highlight the preference order that each evaluating unit expresses with regard to the performance efficiency of its peers.

Our goal is to aggregate the preferential information in R to obtain a measure of the overall standing of each unit in terms of the quality and frequency of the ranking places earned. The problem of aggregating individual preferences expressed as ordinal rankings is rather common in studies of voting and elections, as well as in sporting championship where the final result of a competition must combine the results obtained in a series of games or trials. The simplest models for consensus formation are based on positional scoring rules that assign p_j points to each jth-ranked position and then compute the total number of points that the units get in terms of their ranking positions. Any strictly decreasing sequence of points could be used in this fashion to obtain a measure that summarizes the intensity of preference of each candidate over the others. In particular, Borda method is a well-known example of a positional scoring rule, which is commonly used in Social Choice theory. In Borda method, each candidate gets n points for each first place vote received, n-1 points for each second place vote received, and consecutively this way down to 1 point for each last place vote received.

Then, using the marks defined according to Borda's rule by $p_j = n + 1 - j$, and considering the matrix $M = (m_{zt})$ with m_{zt} being the number of times that DMU z has earned the tth ranking position, the score S_q computed as $S_q = \sum_{j=1}^{n} p_j m_{qj}$ can be taken as a measure of the relevance attributed to the qth DMU by the group of peers. Then, the weights reflecting contextual information about the importance of DMUs will be taken in direct proportion to this score as follows:

$$\lambda_q^2 = \frac{S_q}{\sum_{q=1}^{n} S_q} \tag{8}$$

Step 3. Computation of integrated aggregation weights. In order to obtain overall importance weights for the aggregation of cross-efficiencies, we perform a multiplicative combination of both previous components, that after normalization leaves us with the following:

$$w_q = \frac{\lambda_q^1 \lambda_q^2}{\sum_{q=1}^{n} \lambda_q^1 \lambda_q^2} \tag{9}$$

Step 4. Computation of the weighted cross-efficiency score and ranking. Finally, the peer cross-efficiency scores are weighted and aggregated using (5) and the DMUs are ranked accordingly.

From the procedure described, the following properties are derived:

1. For all q $w_q \geq 0$ holds, with $w_q = 0$ if and only if all assessments

provided by DMU q are identical, in which case the consideration of those evaluations in the aggregate score is irrelevant for the final ranking of the units. Since that is a rather unusual occurrence, $w_q > 0\ \forall q$ in practice and therefore all the peer-evaluations are effectively taken into account during cross-evaluation.
2. The integrated aggregation weights approach the estimation of the relative importance of the cross-efficiency scores through two complementing perspectives that simultaneously account for the discriminatory ability of the DMUs as well as the relevance attributed to DMUs according to the peers.
3. The higher weights are attained to DMUs that provide reliable assessments and at the same time show a high discriminatory ability whereas DMUs that provide unreliable assessments or show an extremely low discriminatory ability are assigned low weights.

As a consequence, the proposed cross-evaluation procedure is expected to produce fair and convincing assessments and ranking results.

4. Numerical examples

In this section different datasets from the DEA literature are used to illustrate the application of the proposed method for weighted cross-efficiency aggregation and compare the results obtained with other approaches. The cross-efficiency matrix obtained by the aggressive formulation of Doyle and Green [7] will be taken as our starting point for each of the examples considered, since we are assuming that discrimination of the given alternatives is a key objective of the analysis.

Example 1. Table 2 contains a dataset used in [33] for the efficiency evaluation of fourteen international passenger airlines that are evaluated in the light of three inputs ($x1$: aircraft capacity, $x2$: operating cost, $x3$: non-flight assets) and two outputs ($y1$: passenger kilometers, $y2$: non-passenger revenue). The last column of Table 2 shows the CCR efficiency of the fourteen airlines or DMUs, showing that as far as half of them are efficient and therefore a cross-efficiency analysis can be particularly convenient for achieving a better discrimination and ranking of the analyzed units.

For a comprehensive aggregation of the self and peer assessments we will apply the proposed procedure to compute aggregation weights that simultaneously reflect the intrinsic information contained in the CEM and the contextual information about the evaluating criteria (DMUs), as reported in Table 3. The intrinsic component (λ_q^1) is derived from the coefficient of variation criterion, and therefore it takes into consideration the discriminatory ability of the assessments provided by different DMUs, while the contextual component (λ_q^2) accounts for the importance that can be attributed to each evaluating unit by the whole DMU set in terms of the ranking orders they earn. The integrated aggregation weights, defined in a multiplicative fashion, successfully combine both ideas. Table 4 shows the aggregate scores (with the corresponding ranking order in parentheses) obtained with the described cross-evaluation process in comparison with the traditional average approach (ACE), Song and Liu's method [19] and the weighted least square deviation method by Wang and Wang [28]. Certain similarities are clearly observed, particularly between the proposed approach and Song and Liu's approach.

To get further insights into the ranking results obtained, the Spearman rank correlation coefficients are used to see how well the ordinal ranks obtained with different approaches correlate to each other. The figures reported in Table 5 provide evidence of a statistically significant (at least at a 0.01 level) association of ranks, which validates the consistency of the ranking method proposed here. Particularly, the highest correlation is observed between Song and Liu's method and the proposed approach as expected, given that both are based on the multicriteria principle of intrinsic information, while the Wang and Wang's approach shows the lowest correlation figures denoting a different rationale in its conception.

Example 2. Consider the case study described in [34] for the evaluation of 15 Chinese thermoelectric enterprises in terms of four inputs ($x1$:

Table 2
Input and Output data for 14 passenger airlines.

DMUs	$x1$	$x2$	$x3$	$y1$	$y2$	CCR efficiency
1	5723	3239	2003	26,677	697	0.868
2	5895	4225	4557	3081	539	0.338
3	24,099	9560	6267	124,055	1266	0.948
4	13,565	7499	3213	64,734	1563	0.958
5	5183	1880	783	23,604	513	1
6	19,080	8032	3272	95,011	572	0.977
7	4603	3457	2360	22,112	969	1
8	12,097	6779	6474	52,363	2001	0.859
9	6587	3341	3581	26,504	1297	0.948
10	5654	1878	1916	19,277	972	1
11	12,559	8098	3310	41,925	3398	1
12	5728	2481	2254	27,754	982	1
13	4715	1792	2485	31,332	543	1
14	22,793	9874	4145	122,528	1404	1

Table 3
Intrinsic weights (λ_q^1), contextual weights (λ_q^2) and integrated aggregation weights (w_q) for Example 1.

Evaluating DMU	CV_q	λ_q^1	S_q	λ_q^2	w_q
1	0.248	0.044	97	0.065	0.039
2	0.503	0.089	18	0.012	0.015
3	0.409	0.072	87	0.058	0.058
4	0.248	0.044	113	0.076	0.046
5	0.591	0.104	139	0.093	0.134
6	0.563	0.099	81	0.054	0.074
7	0.251	0.044	110	0.074	0.045
8	0.245	0.043	79	0.053	0.032
9	0.278	0.049	94	0.063	0.043
10	0.480	0.085	116	0.078	0.091
11	0.579	0.102	151	0.101	0.143
12	0.277	0.049	133	0.089	0.060
13	0.427	0.075	153	0.102	0.107
14	0.563	0.099	123	0.082	0.113

Table 4
Results of different cross-evaluation approaches for Example 1.

DMU	ACE		Song and Liu		Wang and Wang		Integrated approach	
1	0.599	(12)	0.553	(12)	0.789	(8)	0.533	(10)
2	0.165	(14)	0.142	(14)	0.193	(14)	0.123	(14)
3	0.623	(11)	0.603	(9)	0.786	(9)	0.596	(9)
4	0.673	(7)	0.646	(7)	0.853	(7)	0.638	(7)
5	0.798	(1)	0.804	(1)	0.886	(4)	0.818	(1)
6	0.638	(9)	0.650	(6)	0.767	(12)	0.652	(5)
7	0.648	(8)	0.575	(10)	0.867	(6)	0.531	(11)
8	0.585	(13)	0.521	(13)	0.756	(13)	0.488	(13)
9	0.631	(10)	0.560	(11)	0.785	(10)	0.523	(12)
10	0.681	(6)	0.627	(8)	0.785	(11)	0.612	(8)
11	0.774	(2)	0.713	(3)	0.947	(2)	0.692	(3)
12	0.731	(5)	0.668	(5)	0.909	(3)	0.643	(6)
13	0.750	(3)	0.681	(4)	0.957	(1)	0.659	(4)
14	0.732	(4)	0.733	(2)	0.881	(5)	0.735	(2)

Table 5
Spearman's rank correlation coefficients for Example 1.

	Equally weighted	Song and Liu	Wang and Wang	Integrated approach
Equally weighted	1.000	0.938	0.833	0.895
Song and Liu		1.000	0.758	0.982
Wang and Wang			1.000	0.719
Integrated approach				1.000

Table 6
Input and output data for 15 Chinese thermoelectric enterprises.

DMUs	$x1$	$x2$	$x3$	$x4$	$y1$	$y2$	CCR efficiency
1	8092	819.8	4647.3	180	45,526	155,355	0.965
2	5000	1067.2	2200	910	78,878	209	1.000
3	8760	1879.3	75,920.5	324.8	13,523	3782	0.085
4	4498	2618.6	17,503.9	24,357	179,815	53,979	1.000
5	8760	3253.6	33,493.3	219	276,299.4	716,472	1.000
6	8440	3057.6	782,956	440	231,083	705,561	1.000
7	8760	781.6	10,691.6	19,082.6	47,065	1482	0.709
8	7140	3013.2	583,762	203.1	218,698.9	712	0.971
9	6404	95.6	610.6	1144.7	3222.7	7053	0.435
10	8760	940.1	493,131	28,324.1	150	387,546.5	1.000
11	8760	208.5	272	1000	5378	14,685	0.480
12	8760	863.3	1344	2.13	56,807	151,143	1.000
13	8760	98.7	3700	765.65	7862	3210	0.932
14	5139	38	613.2	1200	2995	6167	0.927
15	4224	62.4	9160	1320.46	5456	7274.4	1.000

Table 7
Intrinsic weights (λ_q^1), contextual weights (λ_q^2) and integrated aggregation weights (w_q) for Example 2.

Evaluating DMU	CV_q	λ_q^1	S_q	λ_q^2	w_q
1	1.144	0.063	141	0.078	0.066
2	1.391	0.077	121	0.067	0.069
3	0.483	0.027	54	0.030	0.011
4	1.305	0.072	126	0.069	0.068
5	1.918	0.106	212	0.117	0.167
6	1.533	0.085	147	0.081	0.093
7	0.481	0.027	81	0.045	0.016
8	1.218	0.067	89	0.049	0.045
9	0.690	0.038	94	0.052	0.027
10	1.038	0.057	68	0.037	0.029
11	1.847	0.102	106	0.058	0.080
12	3.564	0.197	171	0.094	0.250
13	0.481	0.027	127	0.070	0.025
14	0.481	0.027	136	0.075	0.027
15	0.483	0.027	141	0.078	0.028

annual production time (h), $x2$: total coal consumption (1000 tons), $x3$: total water consumption (1000 tons), $x4$: total electric power consumption (10 kWh)) and two outputs ($y1$: total industrial output value (10,000 RMB), $y2$: generating capacity (10 kWh)). Table 6 reports the evaluation data as well as the CCR efficiency scores which identify seven DMUs as efficient performers. In order to achieve a full ranking of the DMUs according to their performance, a cross-evaluation analysis is conducted as explained in Section 3. Table 7 shows the importance weights used for the aggregation of the cross-efficiency scores, that were obtained following the procedure that integrates the intrinsic information (λ_q^1) and background information (λ_q^2). The resulting weighted cross-efficiency is compared in Table 8 with the traditional average cross-efficiency and other two weighted-average approaches in the literature. Once again, we see how the use of different aggregation weights has a non-negligible effect on the overall cross-efficiency value and the resulting ranking order, which emphasizes the importance of selecting an appropriate set of weights for the aggregation.

The Spearman's rank correlation coefficients listed in Table 9 reveal statistically significant positive association of ranks at a 0.01 level between all the methods considered, despite relying on different conceptions about DMU importance. Noteworthy, the new approach and Song and Liu's approach seem to provide the most highly correlated rankings, which can be explained by the fact that both methods are built on a similar basis that computes aggregation weights according to the dispersion degree of the data. However, the approach proposed here additionally takes into account the importance of each evaluating unit as perceived by the peers, in such a way that a low contextual weight

Table 8
Results of different cross-evaluation approaches for Example 2.

DMU	ACE		Song and Liu		Wang and Wang		Integrated approach	
1	0.506	(5)	0.422	(4)	0.604	(10)	0.360	(3)
2	0.522	(3)	0.358	(5)	0.743	(4)	0.278	(5)
3	0.037	(15)	0.023	(15)	0.065	(14)	0.016	(15)
4	0.489	(6)	0.343	(6)	0.692	(8)	0.253	(6)
5	0.786	(1)	0.695	(2)	0.924	(1)	0.616	(2)
6	0.508	(4)	0.423	(3)	0.709	(6)	0.353	(4)
7	0.310	(11)	0.176	(12)	0.549	(11)	0.124	(12)
8	0.355	(10)	0.214	(8)	0.634	(9)	0.14	(9)
9	0.217	(13)	0.142	(13)	0.329	(12)	0.106	(13)
10	0.085	(14)	0.082	(14)	0.039	(15)	0.049	(14)
11	0.224	(12)	0.177	(11)	0.275	(13)	0.142	(8)
12	0.786	(2)	0.816	(1)	0.775	(2)	0.872	(1)
13	0.366	(9)	0.194	(10)	0.701	(7)	0.133	(11)
14	0.447	(7)	0.271	(7)	0.735	(5)	0.195	(7)
15	0.385	(8)	0.207	(9)	0.744	(3)	0.140	(10)

would reduce the overall importance of that unit even if it presented a high intrinsic weight, that is, a high discriminatory ability. In that sense, since DMUs have been given an opportunity to judge the reliability of the peer-evaluation scores, the cross-evaluation results obtained with this method will be likely considered as fair and reliable.

Example 3. Table 10 shows a real-world dataset presented in [35] concerning 14 Chinese city commercial banks. Each bank is described using three inputs ($x1$: fixed assets, $x2$: employees´ pay, $x3$: general expenses) and two outputs ($y1$: profit, $y2$: loans), all of them measured in thousand Yuan. CCR efficiency scores, reported in the last column of Table 10, find that six banks are performing efficiently.

The results of the proposed weighted cross-evaluation approach in comparison with León et al.'s approach [27] are presented in Table 11. For both methods, the aggregation weights, the weighted cross-efficiency score and the ranking derived are reported. For the computation of IOWA weights, we have used an inducing variable based on the entropy of the evaluations provided by each DMU, as proposed in [27]. In order to obtain positive aggregation weights for all the units, we have

Table 9
Spearman's rank correlations for Example 2.

	Equally weighted	Song and Liu	Wang and Wang	Integrated approach
Equally weighted	1.000	0.973	0.879	0.943
Song and Liu		1.000	0.820	0.977
Wang and Wang			1.000	0.746
Integrated approach				1.000

Table 10
Input and output data for 14 Chinese banks.

DMUs	$x1$	$x2$	$x3$	$y1$	$y2$	CCR efficiency
1	775,310	283,893	644,280	2,193,085	26,392,924	1.000
2	1,120,795	679,087	1,674,877	3,262,998	91,799,733	0.854
3	5,086,018	420,404	2,867,703	2,845,104	57,320,923	0.442
4	1,114,333	1,239,207	3,187,168	4,474,572	148,341,428	0.974
5	1,178,797	1,155,569	2,433,518	5,680,038	159,941,475	1.000
6	5,810,688	332,044	2,466,586	3,789,997	85,298,079	0.715
7	3,203,353	2,864,359	6,617,403	9,625,877	342,827,271	0.818
8	150,570	207,802	877,996	636,046	21,152,061	1.000
9	1,267,597	296,940	1,820,934	3,298,603	118,767,291	1.000
10	1,015,688	142,024	1,577,106	2,518,447	75,256,873	1.000
11	1,154,635	619,940	2,327,875	2,253,330	98,880,747	0.773
12	1,088,335	95,372	1,221,129	2,414,775	57,757,208	1.000
13	2,172,992	635,475	2,721,211	4,980,404	121,962,186	0.827
14	2,397,820	624,031	3,529,395	5,098,041	142,564,629	0.769

Table 11
Results of different cross-evaluation approaches for Example 3.

	Doyle and Green (1994)			León et al.			Integrated approach		
DMU	w_q	ACE	Ranking	w_q	WCE	Ranking	w_q	WCE	Ranking
1	0.071	0.608	(7)	0.059	0.465	(12)	0.065	0.482	(11)
2	0.071	0.615	(6)	0.016	0.550	(8)	0.043	0.531	(6)
3	0.071	0.241	(14)	0.050	0.197	(14)	0.015	0.210	(14)
4	0.071	0.653	(5)	0.110	0.652	(5)	0.082	0.591	(5)
5	0.071	0.804	(2)	0.093	0.757	(2)	0.110	0.707	(4)
6	0.071	0.360	(13)	0.024	0.293	(13)	0.026	0.315	(13)
7	0.071	0.589	(8)	0.076	0.566	(7)	0.057	0.523	(8)
8	0.071	0.554	(11)	0.119	0.603	(6)	0.065	0.529	(7)
9	0.071	0.821	(1)	0.067	0.758	(1)	0.113	0.762	(2)
10	0.071	0.777	(3)	0.101	0.736	(3)	0.130	0.765	(1)
11	0.071	0.509	(12)	0.084	0.493	(11)	0.047	0.463	(12)
12	0.071	0.770	(4)	0.127	0.712	(4)	0.153	0.761	(3)
13	0.071	0.585	(9)	0.033	0.507	(10)	0.048	0.511	(10)
14	0.071	0.565	(10)	0.041	0.510	(9)	0.045	0.513	(9)

Table 12
Input and output data for six nursing homes.

DMUs	$x1$	$x2$	$y1$	$y2$	CCR efficiency
1	1.50	0.2	1.40	0.35	1.000
2	4.00	0.7	1.40	2.10	1.000
3	3.20	1.2	4.20	1.05	1.000
4	5.20	2.0	2.80	4.20	1.000
5	3.50	1.2	1.90	2.50	0.977
6	3.20	0.7	1.40	1.50	0.867

selected an orness degree equal to 0.65.[1] Note that these two methods are not completely opposed in their rationale. In fact, in the IOWA approach the order of importance of the evaluating units is determined by the entropy value of the CEM columns, which is an indicator of the amount of information contained in the CEM. However, the different modeling assumptions used for the computation of weight values naturally lead to different outcomes, and although the ranking results obtained with both approaches are fairly consistent, with a Spearman's rank correlation coefficient $\rho = 0.96$, some differences can be observed, particularly in the top 5 ranking places. Since rankings can eventually have consequences for decision making, evaluated units may well be interested in monitoring the procedures used to elaborate the ranked lists. Using aggregation weights that account for the relevance that DMUs attribute to their peers is an interesting option in order to prevent final results to be affected by possibly unfair assessments.

Example 4. We finally consider a classical example studied by Sexton [36] and frequently revisited in the literature. The data provided in Table 12 describes six nursing homes that are evaluated in terms of two inputs (x_1: staff hours per day, x_2: supplies per day measured in thousands of dollars) and two outputs (y_1: total Medicare-plus Medicaid-reimbursed patient days (10^4), y_2: total privately paid patient days (10^4)) and their CCR efficiency scores. The results for different cross-evaluation approaches are summarized in Table 13, where once more the differences in the modeling assumptions become explicit, particularly for Wang and Wang's method which is based on a rationale that do not relate to the amount of information contained in the CEM. Different rankings are almost inevitable when different methods are used, and it is inappropriate to say which is the most rational ranking or which is the best strategy, given that it depends to a large extent on the evaluation context and the preferences of decision makers. For that reason, it is imperative for managers and investors to be aware of the modeling assumptions underlying the procedures used to derive the ranking results. In that sense, the cross-evaluation procedure proposed here fits particularly well for highly competitive situations, where countries, firms or institutions are being evaluated and certain economic or financial decisions are being taken on the ranked lists derived. In those cases, DMUs will be interested in avoid unfair outcomes so they will likely be more willing to accept a set of aggregation weights that accounts for the reliability of the assessments received.

[1] Recall that an orness degree of 0.5 would assign equal weights to all DMUs and when the orness degree equals 1.0 only one DMU would be assigned a nonzero weight.

Table 13
Weights and ranking results obtained with different cross-evaluation approaches for Example 4.

DMU	ACE		Wang and Wang		Song and Liu		León et al.		Integrated approach	
	w_q	Ranking	w_q	Ranking	w_q	Ranking	w_q	Ranking	w_q	Ranking
1	0.167	1	0.006	3	0.261	1	0.274	1	0.372	1
2	0.167	3	0.016	2	0.199	5	0.145	5	0.182	4
3	0.167	5	0.009	6	0.230	3	0.231	2	0.167	3
4	0.167	2	0.013	1	0.228	2	0.188	3	0.218	2
5	0.167	4	0.478	4	0.041	4	0.060	4	0.032	5
6	0.167	6	0.478	5	0.041	6	0.102	6	0.029	6

5. Conclusions

Cross-efficiency evaluation is an extension of data envelopment analysis that has been widely used within different fields for ranking decision making units according to their performance. The main idea of cross-efficiency evaluation relies on the combination of the traditional DEA self-evaluation mode with a peer-evaluation mode where the optimal DEA weights selected by DMUs are also used to assess their peers. Since all the evaluations that a DMU receives are eventually averaged, the overall cross-efficiency score represents a quite complete assessment of the performance of units when appraised under a range of preferences (defined by the input and output weights), which brings an appealing democratic connotation into the whole evaluation process. A vast number of applications have proved the ability of the approach for effectively differentiating performance among all units, thus allowing a meaningful ranking of the DMU set.

Most cross-evaluation approaches are based on a simple average of the cross-efficiency scores. However, it has been pointed out how the assessments provided by different DMUs may not be equally useful for discrimination purposes, and not all the units should necessarily be equally relied on in their evaluating role due to possibly extremely specialized selection of weights. In those circumstances the cross-evaluation results might be affected by potentially unfair assessments, and the use of a weighted instead of a simple average becomes a plausible strategy to account for the relative importance of the DMUs, therefore introducing a greater degree of flexibility in the model that allows for more realistic evaluations.

Working along this line, in this paper the computation of aggregation weights for cross-efficiency evaluation is further investigated on the basis of two basic assumptions. First, if the aggregate score is aimed at discriminating the units, the aggregation weights should take into account the discriminatory ability of the peer-assessed cross-efficiencies. Second, if the aggregate score is expected to be fair, the reliability of DMUs as evaluating units should also be acknowledged. These assumptions are consistent with the MCDM conception that considers attribute weights as measures that simultaneously comprise the intrinsic information provided by attributes as well as their contextual importance. In order to put this idea into practice, the aggregation weights are then defined as the multiplicative combination of two components. The intrinsic weight component reflects the discrimination power of the DMUs, which is here defined in direct proportion to the coefficient of variation of the vector of cross-efficiencies provided by each DMU. The contextual weight component reflects the relevance of each DMU within the background where the evaluation takes place, and is obtained through an original quantification of the reputation that the DMU enjoys in the view of its peers.

The numerical examples presented illustrate the practical application of the proposed approach and demonstrate that it represents a valid methodological option that contributes to DEA cross-efficiency evaluation and complements previous existing methods. The approach presented uses ideas from the field of multicriteria decision making to define aggregation weights that reflect the importance of the evaluating units from two complementing perspectives and, as a consequence, this method can be said to involve a rather comprehensive mechanism for deriving aggregation weights that is expected to provide more realistic assessments and rankings than previous weighted approaches that rely on a single aspect. Besides, since DMUs perceptions about their peers are taken into consideration, the democratic character of the cross-evaluation approach is enhanced and all the units are likely to accept the results obtained. All in all, it has become clear that including relative importance weights in the cross-efficiency aggregation process may have a critical impact on the ranking order of the studied units and therefore, the customarily application of the average aggregation in cross-efficiency evaluations should be carefully reconsidered. Particularly, the alternative aggregation presented here can be specially useful in extremely competitive situations where the ranking positions determine any kind of investment or economic decision.

Lastly, the procedure presented can be extended by changing the way the two components of the derived weights are handled. For example, the use of variation coefficient method for the intrinsic component can be substituted by other multicriteria approaches such as the standard deviation method, entropy method or CRITIC method. Likewise, different scoring systems other than Borda's rule could also be used to estimate the importance of DMUs in terms of the ranking positions.

References

[1] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. Eur J Oper Res 1978;2:429–44.
[2] Liu JS, Lu LYY, Lu WM. Research fronts in data envelopment analysis. Omega 2016;58:33–45.
[3] Dyson RG, Allen R, Camanho AS, Podinovski VV, Sarrico CS, Shale EA. Pitfalls and protocols in DEA. Eur J Oper Res 2001;132:245–59.
[4] Adler N, Friedman L, Sinuany-Stern Z. Review of ranking methods in the data envelopment analysis context. Eur J Oper Res 2002;140:249–65.
[5] Angulo-Meza L, Lins MPE. Review of methods for increasing discrimination in data envelopment analysis. Ann Oper Res 2002;11:225–42.
[6] Sexton TR, Silkman RH, Hogan AJ. Data envelopment analysis: critique and extensions. New Dir Program Eval 1986;32:73–105. https://doi.org/10.1002/ev.1441.
[7] Doyle J, Green R. Efficiency and cross-efficiency in DEA: derivations, meaning and uses. J Oper Res Soc 1994;45:567–78.
[8] Baker RC, Talluri S. A closer look at the use of data envelopment analysis for technology selection. Comput Ind Eng 1997;32:101–8.
[9] Falagario M, Sciancalepore F, Costantino N, Pietroforte R. Using a DEA-cross efficiency approach in public procurement tenders. Eur J Oper Res 2012;218:523–9.
[10] Lim S, Oh KW, Zhu J. Use of DEA cross-efficiency evaluation in portfolio selection: an application to Korean stock market. Eur J Oper Res 2014;236:361–8.
[11] Wu J, Liang L, Chen Y. DEA game cross-efficiency approach to Olympic rankings. Omega 2009;37:909–18.
[12] Yu MM, Ting SC, Chen MC. Evaluating the cross-efficiency of information sharing in supply chains. Expert Syst Appl 2010;37:2891–7.
[13] Tan Y, Zhang Y, Khodavedi R. Service performance evaluation using data envelopment analysis and balance scoreboard approach: an application to automotive industry. Ann Oper Res 2017;248:449–70.
[14] Liang L, Wu J, Cook WD, Zhu J. Alternative secondary goals in DEA cross efficiency evaluation. Int J Prod Econ 2008;113(2):1025–30.
[15] Örkcü HH, Bal H. Goal programming approaches for data envelopment analysis cross efficiency evaluation. Appl Math Comput 2011;218:346–56.
[16] Wu J, Sun J, Liang L. Cross efficiency evaluation method based on weight-balanced data envelopment analysis model. Comput Ind Eng 2012;63:513–9.
[17] Carrillo J, Jorge JM. An alternative neutral approach for cross-efficiency evaluation. Comput Ind Eng 2018;120:137–45.
[18] Ding T, Liang L, Yang M, Wu H. Multiple attribute decision making based on cross-

evaluation with uncertain decision parameters. Math Probl Eng 2016:4313247. 10 pages.
[19] Song L, Liu F. An improvement in DEA cross-efficiency aggregation based on the Shannon entropy. Int Trans Oper Res 2018;25:705–14.
[20] Lim S. Minimax and maximin formulations of cross-efficiency in DEA. Comput Ind Eng 2012;62(3):726–31.
[21] Wang YM, Chin KS, Jiang P. Weight determination in the cross-efficiency evaluation. Comput Ind Eng 2011;61(3):497–502.
[22] Ramón N, Ruiz JL, Sirvent I. On the choice of weights profiles in cross-efficiency evaluations. Eur J Oper Res 2010;207(3):1564–72.
[23] Wu J, Liang L, Yang F. Determination of the weights for the ultimate cross-efficiency using Shapley value in cooperative game. Expert Syst Appl 2009;36(1):872–6.
[24] Wang YM, Chin KS. The use of OWA operator weights for cross-efficiency aggregation. Omega 2011;39:493–503.
[25] Despotis DK. Improving the discriminating power of DEA: focus on globally efficient units. J Oper Res Soc 2002;53:314–23.
[26] Ruiz JL, Sirvent I. On the DEA total weight flexibility and the aggregation in cross efficiency evaluations. Eur J Oper Res 2012;223(3):732–8.
[27] León T, Ramón N, Ruiz JL, Sirvent I. Using induced ordered weighted averaging (IOWA) operators for aggregation in cross-efficiency evaluations. Int J Intell Syst 2014;29:1100–16.
[28] Wang YM, Wang S. Approaches to determining the relative importance weights for cross-efficiency aggregation in data envelopment analysis. J Oper Res Soc 2013;64(1):60–9.
[29] Wu J, Sun J, Liang L. DEA cross-efficiency aggregation method based upon Shannon entropy. Int J Prod Res 2012;50(23):6726–36.
[30] Zeleny M. Multiple criteria decision making. New York: McGraw Hill; 1982.
[31] Diakoulaki D, Mavrotas G, Papayannakis L. Determining objective weights in multiple criteria problems: the CRITIC method. Comput Oper Res 1995;22(7):763–70.
[32] Sarkis J. A comparative analysis of DEA as a discrete alternative multiple criteria decision tool. Eur J Oper Res 2000;123:543–57.
[33] Tofallis C. Input efficiency profiling: an application to airlines. Comput Oper Res 1997;24(3):253–8.
[34] Song M, Zhu Q, Peng J, Santibanez EDR. Improving the evaluation of cross efficiencies: a method based on Shannon entropy weight. Comput Ind Eng 2017;112:99–106.
[35] Li Y, Xie J, Wang M, Liang L. Super efficiency evaluation using a common platform on a cooperative game. Eur J Oper Res 2016;255:884–92.
[36] Sexton TR. The methodology of data envelopment analysis. In: Silkman RH, editor. Measuring efficiency: an assessment of data envelopment analysis, 1986. San Francisco: Jossey-Bass; 1986. p. 7–29.