http://dx.doi.org/10.1016/j.omega.2014.12.003

Benchmarking of maintenance and outage repair in an electricity distribution company using the Value-Based DEA method

MC Gouveia
ISCAC, Quinta Agrícola, Bencanta, 3040-316 Coimbra
INESC Coimbra, Rua Antero de Quental, 199; 3000-033 Coimbra
mgouveia@iscac.pt

LC Dias
Faculdade de Economia, Universidade de Coimbra, Av. Dias da Silva 165, 3004-512 Coimbra
INESC Coimbra, Rua Antero de Quental, 199; 3000-033 Coimbra
lmcdias@fe.uc.pt

CH Antunes
DEEC-FCT Universidade de Coimbra, Pólo II, 3030 Coimbra
INESC Coimbra, Rua Antero de Quental, 199; 3000-033 Coimbra
cantunes@inescc.pt

J Boucinha
EDP Serviço Universal, S.A., Rua Camilo Castelo Branco, nº 43; 1050-044 Lisboa
julia.boucinha@edp.pt

CF Inácio
EDP Serviço Universal, S.A., Rua Camilo Castelo Branco, nº 43; 1050-044 Lisboa
catarinafeteira.inacio@edp.pt

ABSTRACT. Benchmarking of electricity distribution utilities has been widely used as a means to contribute for the adoption or reinforcement of enhanced competitiveness and innovation practices to optimize costs, increase customer satisfaction, improve corporate image and maximize profits. The purpose of this paper is to present a benchmarking study for the maintenance and outage repair activity carried out by a Portuguese electricity distribution company, EDP Distribuição (EDP-D), using the Value-Based DEA method, which builds on links between Data Envelopment Analysis (DEA) and Multiple Criteria Decision Analysis (MCDA). This study illustrates the impact of the incorporation of managerial preferences in the classification and ranking of 40 network areas served by EDP-D, confronting the results with a previous study based on a BCC DEA model. In order to deal with the underlying uncertainty, the Value-Based DEA method for performance evaluation is adapted to include

the concept of super-efficiency. Besides identifying best practices, sources of inefficiency, gaps relatively to best practices and opportunities for improvement, this analysis supports the introduction of corrective measures and informs decisions about future goals.

KEYWORDS. Data Envelopment Analysis; Multi-Criteria Analysis; Electricity Distribution; Super-efficiency.

1. Introduction

Electrical energy is at the heart of modern society, as an essential component of lifestyle and a determining factor in the competitiveness of the economy. In the current context of increasing competition and regulatory pressure on electric utilities, companies must become increasingly efficient. In this framework, benchmarking is a very helpful instrument to identify the most efficient utilities in the sector, providing measures to evaluate the relative performance of the different utilities analyzed. Through benchmarking it is possible to quantify differences in performance, identify the reasons for such differences and also the improvements needed to achieve the targets set by the organization.

Several approaches have been proposed to measure the relative efficiency of utilities with respect to an empirical efficient frontier defined by a set of units. More recently several benchmarking studies use Data Envelopment Analysis (DEA) for the identification of sources of inefficiency in some of the most profitable companies (see, e.g. [1-8]).

DEA [9] is a nonparametric approach based on linear programming for measuring the efficiency of a set of entities called Decision Making Units (DMUs). A DMU is any entity under evaluation in terms of its abilities to convert inputs into outputs, engaged in the same activity.

Since the early 1980's, after a first DEA study conducted by Färe *et al.* [10] to measure the relative efficiency of electricity utilities in Illinois, several other studies to evaluate the efficiency of electricity distribution companies have appeared in the literature. Jamasb and Pollit [11] reported and discussed the main benchmarking methods used by the electricity regulators of the electricity distribution activity in the OECD and a few other countries. Later, the same authors [12] presented an international benchmarking study of 63 regional electricity distribution utilities in six European countries, illustrating the methodological and data difficulties encountered in the use of international benchmarking for utility regulation. Haney and Pollitt [13] presented the results of an international survey of energy regulators in 40 countries. There are many other studies in this area using DEA, some of which evaluate the relative efficiency of distribution utilities in a single country, while others have an inter-country focus (for a comprehensive review see [14]).

Although various studies on the use of DEA have been published, only a few incorporate managerial preferences in the analysis (see e.g. [15-21]). Thanassoulis *et al.* [22] reported a number of reasons for the inclusion of the Decision Maker's (DM's) preferences in DEA. In a work relating possible areas of interaction between DEA and Multiple Criteria Decision Making (MCDM), Bouyssou [23] concluded that both the choice and the ranking of alternatives can be achieved only by introducing a preference structure. In this line of thought Köksalan and Tuncer [24] proposed a DEA-based approach to ranking multi-criteria alternatives, including weight restrictions to incorporate DM's preferences into the analysis. Cook *et al.* [25] offered some clarification and direction on how DEA can be viewed as a tool for multiple-criteria evaluation problems. The present study combines DEA with Multiple Criteria Decision Analysis (MCDA) including relevant preferential information elicited from the DM using Value-Based DEA [26-27]. This method is based on the additive DEA model with oriented projections [28], making use of concepts developed in the field of MCDA under imprecise information [29-30].

The present paper presents a benchmarking study for one of the activities performed by EDP *Distribuição* (EDP-D) — maintenance and outage repair. An internal benchmarking study had been previously undertaken using several DEA models to examine data on this activity in the Portuguese electricity distribution system during 2004–2005 [31]. More recently, for this particular activity, a study that uses DEA with the involvement of DMs was carried out by Amado *et al*. [21], which compares the cost efficiency of medium-voltage power lines belonging to the same regional distribution networks area operated by EDP-D. These authors analyze the impact of different design systems and different maintenance programs, contributing to the reduction of costs and improving service delivery quality.

In both aforementioned studies the experience of the company has been used to draw some lessons about how performance measurement can be implemented within a company. In this context, the Value-Based DEA method can bring useful insights to the company managers, since their preferences and judgments are incorporated into the model. The present work intends to assess whether that is indeed the case, by applying Value-Based DEA to a context the DMs already knew well. The study can be perceived as a learning activity both for the method proponents (analysts) and for the company's DMs involved. This is in line with Dyson and Shale [32], when they state that a greater collaboration between academics and practitioners or a greater involvement of academics with practice is necessary to obtain credible results and to foster the confidence of the DMs in new developments of methods and techniques. Although this study was developed for a particular activity, maintenance and outage repair, the same type of methodology can be applied to other activities, as well as to the whole set of activities developed within EDP-D.

In order to be able to compare results, this study uses the same data previously used to evaluate the efficiency of the 40 network areas then operating in the Portuguese mainland. This approach is useful not only for comparison with results obtained with the classical DEA models, but also to understand the impact of the incorporation of managerial preferences in the classification of the units. With this analysis, besides identifying best practices, sources of inefficiency, gaps relatively to best practices and opportunities for improvement, it is also possible to support the introduction of corrective measures and inform decisions about future goals, as well as to improve the knowledge of the company.

The remainder of this paper is organized as follows. Section 2 introduces the Value-Based DEA method with the modifications to include the super-efficiency concept [27]. Section 3 gives a brief background on EDP-D. In section 4 the input and output factors are presented as well as the protocols used to elicit the DM's preferences. The analysis of results is carried out in section 5. Section 6 highlights the many prospects for improvement in some of the inefficient units, given their specificity. Concluding remarks are presented in section 7.

2. The Value-Based DEA method

The main idea underlying DEA is that by comparing a set of similar DMUs, it is possible to identify best practices and find the efficient frontier formed by DMUs operating efficiently. Hence, the different models for DEA seek to determine the DMUs which form the efficient frontier (or envelopment surface) in the Pareto-Koopmans sense. Since DEA is based on best practices, it is appealing for the manager who prefers to think in terms of benchmarks instead of comparisons with the mean, for example. DEA identifies benchmarks against which the inefficient units can be compared, i.e., it provides measures for the relative efficiency of the non-frontier units. In other words, for the inefficient units, which do not represent the "best practice" from the combination of inputs and outputs, it is possible to identify the units belonging to the efficient frontier, with which they should compare and, consequently, the reductions in inputs and/or increases in outputs necessary for those units to become efficient (this may be interpreted as a mechanism of projection on the efficient frontier). Hence, this technique is widely used for benchmarking because it is very effective in determining the units with best practices. Classic DEA models may consider both constant returns to scale (CRS), as the CCR model [9], and variable returns to scale (VRS), as the BCC model [33]. In the first case, a proportional change in outputs is expected from a given change in inputs, at all levels of scale. In the second case, an increasing or decreasing change in outputs may occur due to a given change in inputs. Charnes *et al*. [34] proposed the additive DEA model as an alternative to the BCC model; the additive model also considers VRS but does not need a choice between input-orientation and output-orientation. In oriented models, firstly all factors are reduced or increased at the same rate towards the envelopment surface, and the second stage yields an optimal set of slack values. The additive DEA model uses the second stage only and measures the excess of inputs and the deficit of outputs for a DMU under evaluation, when confronted with the DMUs operating on the efficient frontier.

The method developed by Gouveia *et al.* [26] builds on Multi-Attribute Utility Theory (MAUT) [35] since the input and output factors are converted into utility functions according to the preference information provided by DMs. In accordance with von Winterfeldt and Edwards [36], some protocols were used in the process of eliciting preferences in order to build the marginal utility functions, as well as constraining the weights for the aggregation of marginal utilities into an additive overall utility function, instead of letting each DMU choose freely the weights associated with these functions.

In the context of additive aggregation with imprecise weights, the min-max regret rule [37] is a meaningful tool to compare the alternatives [30, 38, 39]. In the Value-Based DEA method used in this study, one must find the scale coefficients (weights) that, for each alternative, minimize the utility difference to the best alternative, according to the min-max regret rule, which gives an intuitive meaning (the loss of utility) to the efficiency measure assigned to each DMU.

Let us consider n DMUs to be evaluated, each of them consuming m different inputs to produce p different outputs. The DMU j consumes the quantity $x_{ij} > 0$ of input i and produces the quantity $y_{rj} > 0$ of output r.

Considering that the DMUs are evaluated according to q (with $q = m+p$) criteria, q utility functions $u_1, \ldots, u_q$ must be defined such that the worst level has a 0 value and the best level has value 1. Hence, after being converted into utility values, all factors are treated as outputs to be maximized. For each alternative (DMU), according to the additive MAUT model, the utility obtained is $U(DMU_j) = \sum_{c=1}^{q} w_c u_c(DMU_j)$, where $w_c \geq 0$, $\forall c = 1, \ldots, q$ and $\sum_{c=1}^{q} w_c = 1$ (by convention). The scale coefficients $w_1, \ldots, w_q$ are the weights of the utility functions and reflect the DM's utility trade-offs, since one unit in one marginal utility function is not necessarily valued as much as one unit in the marginal utility corresponding to a different factor.

The Value-Based DEA method [26] is extended to consider the super-efficiency concept introduced in DEA by Andersen and Petersen [40], in the sense that a complete ranking of all DMUs can be obtained [27]. For that purpose, the following linear program is solved (Phase 1):

$$
\begin{aligned}
&\min_{d_k, w} \; d_k \\
&s.t. \; \sum_{c=1}^{q} w_c u_c\left(DMU_j\right) - \sum_{c=1}^{q} w_c u_c\left(DMU_k\right) \leq d_k, \; j = 1, \ldots, n, \; j \neq k \\
&\qquad \sum_{c=1}^{q} w_c = 1, \\
&\qquad w_c \geq 0, \; \forall c = 1, \ldots, q
\end{aligned}
\tag{1}
$$

The optimal value d_k^* denotes the distance defined by the utility difference to the best of all alternatives (excluding the one under evaluation). If $d_k^* < 0$, then DMU k is efficient. This measure gives the extent to which an efficient DMU may worsen its utility while remaining efficient.

The purpose of the method is to calculate the vector w of weights which minimizes the distance (the utility difference) of DMU k to the best one (note that the best alternative will also depend on w), excluding itself from the reference set. Then, an efficient target is determined in case the DMU is inefficient (Phase 2). The details of this process are as follows:

Phase 1: Convert inputs and outputs into utility scales. Compute the efficiency measure, d_k^*, for each DMU, $k = 1, \ldots, n$, and the corresponding weighting vector.

Phase 2: If $d_k^* \geq 0$ then solve the "weighted additive" model (2), using the optimal weighting vector resulting from Phase 1, w^*, and determine the corresponding projected point of the DMU under evaluation.

$$
\begin{aligned}
\min_{\lambda, s} \quad & z_k = -\sum_{c=1}^{q} w_c^* s_c \\
\text{s.t.} \quad & \sum_{j=1, j \neq k}^{n} \lambda_j u_c(DMU_j) - s_c = u_c(DMU_k), \ c = 1, \dots, q \\
& \sum_{j=1}^{n} \lambda_j = 1, \\
& \lambda_j, \ s_c \geq 0, \ \ j = 1, \dots, k-1, k+1, \dots, n, \ \ c = 1, \dots, q
\end{aligned}
\tag{2}
$$

Variables λ_j, $j = 1, \dots, k-1, k+1, \dots, n$, define a convex combination of the n-1 DMUs. The set of efficient DMUs (possibly only one) defining the convex combination (those DMU j such that λ_j>0) are the "peers" of DMU k under evaluation, i.e. the DMUs with which it should compare in terms of overall utility to achieve efficiency. The convex combination corresponds to a point on the efficient frontier which is better than DMU k by a difference given by s_c (slack) on each criterion c. This target point, considering these weights optimal for DMU k, is better than DMU k by a difference of d_k^* in terms of global utility.

3. EDP *Distribuição* - Background for the Case Study

According to Portuguese law, local authorities, at Municipal level, are entitled to perform all the activities related with low voltage electricity distribution. After the creation of EDP – *Electricidade de Portugal*, in 1976, as an integrated company in charge of electricity generation and transmission across the whole Portuguese mainland, a process of integration of the distribution activities into EDP has started to develop.

Under 20 year contract agreements with each of the municipalities, EDP has progressively taken over the distribution activity. By mid 1990s, the company was in charge of all the investment and maintenance activities required in the distribution network. According to these contracts, although the assets' ownership remained within the local authorities, EDP was in charge of all the operations, in exchange for the payment of a concession fee to each municipality. The distribution activity was then organized into four companies within EDP, according to the four main regions of the Portuguese mainland — North, Center, Tagus Valley and South.

In 2000, after deregulation and the creation of the Regulator (ERSE), these four companies were merged and converted into a single company, named EDP *Distribuição*, unbundled from the other activities, with a completely separate management.

Given the very different network background, in terms of assets and organization, benchmarking activities are highly relevant for management purposes, in order to determine best performances which can be used as benchmarks for areas whose practices need to improve.

4. A Model for Maintenance and Outage Repair in Electricity Distribution including the DM's Preferences

In a previous internal benchmarking study, a variety of DEA models (CCR, BCC and SBM) were used to evaluate the efficiency of maintenance and outage repair expenditures in the electricity distribution networks operated by EDP-D [31]. The objective of that study was to identify best practices, taking into account all the relevant explanatory variables for this activity.

The Value-Based DEA method was suggested to EDP-D as an approach that might bring useful insights to the company, since managerial preferences would be incorporated into the model. This work relies on data for 2004-2005, allowing the comparison with results previously obtained, for the same period, in which the company was organized into 40 different DMUs. After that period the company was re-organized into a considerably smaller number of units, which are not comparable with the previous ones. Therefore, the value to the company is not only the comparison with results obtained through other DEA models, already used by Weyman-Jones *et al.* [31], but also to understand the impact of the incorporation of those preferences on the classification of the units and to assess the interest of using Value-Based DEA in benchmarking other activities of the company.

4.1. Factors

Weyman-Jones *et al.* [31], in the study undertaken for EDP-D, constructed a model in which the explanatory factors were selected in an interactive process with the company engineers, with the purpose of evaluating the efficiency of 40 network areas in the years 2004 and 2005. The efficiency analysis was applied to the particular activity of maintenance and outage repairing. In that study, an input orientation was adopted. The analysis included the comparison of a variety of DEA models to examine the relationships between oriented and non-oriented models, and radial and non-radial analysis.

Inputs and outputs, presented in Table 1, have been discussed with the engineering branches of the company related to these activities. The rationale was to include in the models the main variables that reflect the costs and performance of the different units analyzed.

Table 1. Factors**.**

Inputs	**Outputs**
x_{OPEX}: maintenance and outage repairing costs	y_{CLI}: clients (LV+MV)
x_{MLL}: supply interruptions (minutes of lost load)	y_{NLL}: network lines length (LV+MV)
x_{CC}: complaints per customer	
x_{NI}: number of incidents (LV and clients' installations)	

x_{OPEX} represents the resources used, in terms of costs of that particular activity. Inputs x_{MLL}, x_{CC} and x_{NI} are indicators for quality of supply and reflect undesirable outputs. Supply interruptions, measured in minutes of lost load, represent the number of minutes customers are without electricity supply which, ideally, should be zero. The number of complaints per customer also reflects the performance of the area, as a higher number of complaints indicates poorer customer service. A higher number of incidents on the low voltage network or in customer installations also reflects poor service and, hence, must be minimized.

Outputs y_{CLI} and y_{NLL} reflect the activity level of each area and apply to both low voltage (LV) and medium voltage (MV) networks - more clients and a network with higher length will lead to higher costs. Output y_{CLI} is a proxy for the number of customer services provided. Network line length (y_{NLL}) is considered as an exogenous operating characteristic reflecting the maintenance and repair load in the network, treated as an additional output in the input orientation. Regulators have used y_{NLL} as a measure of the difficulty in delivering electricity [31]. These inputs and outputs are typical in the evaluation of this activity. For instance, Amado *et al*. [21] took into account most of these variables (namely x_{OPEX}, x_{MLL}, x_{NI} and y_{NLL}) to assess this activity, although in a different context, with the purpose of evaluating the impact of alternative policies of design and maintenance on the efficiency of lines.

The study uses the experience of the company to draw some lessons about how performance measurement can be implemented within a company, in contrast to the usual objective of regulatory benchmarking procedures. In fact, there is a significant difference in purpose and implementation between public regulatory benchmarking and internal company benchmarking, which is related to the nature of incentives and rewards.

The summary statistics of the inputs and outputs of all DMUs for both years is depicted in Table 2.

Table 2. Summary statistics of factors for 2004-2005**.**

	Inputs				**Outputs**	
Year 2004	x_{OPEX}	x_{MLL}	x_{CC}	x_{NI}	y_{CLI}	y_{NLL}
Average	2541523. 09	253.95	0.92	5333.53	145593.30	4832.00
Std. Dev.	1506629.27	136.82	0.39	6141.89	154080.24	2322.78
Max	7 802 302.10	606.45	1.75	33280.00	859831.00	14337.85
Min	1 129 899.13	86.00	0.39	1578.00	56730.00	2651.81
Year 2005						
Average	2514733.13	221.00	1.17	4576.93	147685.40	4931.90
Std. Dev.	1501549.64	110.80	0.44	5141.97	155802.68	2354.68
Max	8 204 735.48	665.86	1.99	28880.00	868566.00	14717.25
Min	1 129 899.13	98.80	0.29	1479.00	57453.00	2725.99

4.2. Elicitation of factors utility functions

The use of the Value-Based DEA method allows tailoring the analysis according to the DM's preferences. von Winterfelt and Edwards [36] make a detailed presentation of various techniques to question the DM in order to build utility functions consistent with the DM's answers, but these questions must be framed for each particular context. The elicitation of the DM's preferences is a crucial step of a multiple criteria decision aiding process. The purpose of factors conversion into a utility scale in the Value-Based DEA method developed by Gouveia *et al.* [26] is to reflect the DM's preferences. The utility functions have been constructed using a precise protocol (described by Almeida and Dias [41]) to elicit the difference in the DMU's relative merit corresponding to decreases in inputs or increases in outputs, rather than the absolute utility of having these inputs available or outputs produced.

The elicitation protocol is based on comparing the merit of increasing an output (or decreasing an input) from *a* to *b* versus increasing the same output (or decreasing the same input) from *a'* to *b'*, all other performance levels being equal, and asking the DM to adjust one of these four values such that the increase of merit would be approximately equal. This conversion is done assuming the continuity of functions and because utility functions are unique up to positive affine transformations it is usually assumed that both the global utility functions and marginal utility functions are scaled between 0 and 1, as referred to in section 2.

For example, considering the variable x_{MLL} a question raised to the DM was: "Is it more meritorious to decrease supply interruptions (minutes) from 800 to 300 or from 300 to 60, all the other performances being equal?" The answer was that it is more meritorious to decrease from 300 to 60. Then an adjustment has been made and the question was reformulated as: "Is it more meritorious to decrease the number of supply interruptions from 800 to 200 or from 200 to 60, all the other performances being equal?" The answer was that the merit is the same. This means that $u_{MLL}(200)-u_{MLL}(800) = u_{MLL}(60)-u_{MLL}(200)$, i.e., $u_{MLL}(200)=(u_{MLL}(60)+u_{MLL}(800))/2$. The same procedure was used to dichotomize the intervals of merit [60,200] and [200,800].

The DM answered questions about the differences of merit between the performance levels on each factor. A piecewise linear approximation was defined to represent the utility functions for most factors, and known functions (namely logarithmic functions) were used when the DM's answer could be adjusted to predefined curves.

The elicited ranges were chosen to include the observed performance ranges plus or minus the highest tolerance value considered (in this case $\delta = 20\%$).

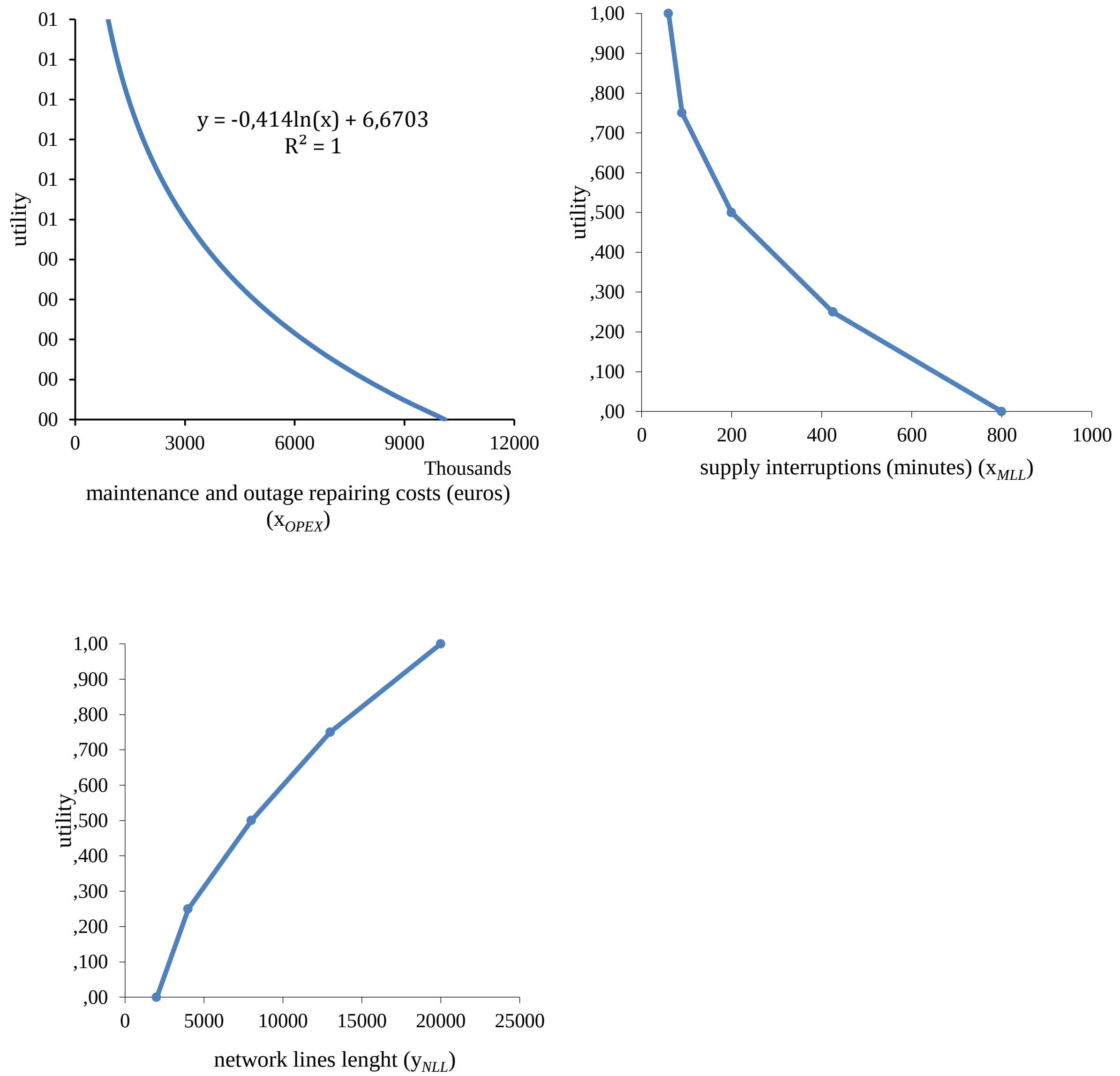

Figure 1. Three of the utility functions elicited for factors.

Figure 1 displays the piece-wise linear utility functions for the inputs x_{OPEX} and x_{MLL} and for the output y_{NLL}. For example, for the supply interruptions: $u_{MLL}(60) - u_{MLL}(90) = u_{MLL}(90) - u_{MLL}(200) = u_{MLL}(200) - u_{MLL}(425) = u_{MLL}(425) - u_{MLL}(800)$, all other performance levels being equal. The input factors x_{CC}, x_{NI} have utility functions similar to the x_{MLL} utility function and the output utility function y_{CLI} is identical to the output utility function y_{NLL}. The x_{OPEX} utility function was obtained by making the corresponding adjustment of a known function to the DM's preferences. Note that the transformation of the original input/output data from original scales to a utility scale, on the basis of preference information provided by the DM, allows dealing with undesirable outputs in a natural way by setting a decreasing utility function.

Table 3 indicates utilities, for the 40 DMUs.

Table 3. Performances converted into utility scales for 2004-2005.

	Factors in utility scales (2004)						**Factors in utility scales (2005)**					
DMUs	u_{OPEX}	u_{MLL}	u_{CC}	u_{NI}	u_{CLI}	u_{NLL}	u_{OPEX}	u_{MLL}	u_{CC}	u_{NI}	u_{CLI}	u_{NLL}
1	0.518	0.426	0.221	0.668	0.063	0.326	0.523	0.466	0.166	0.696	0.064	0.334
2	0.668	0.582	0.464	0.709	0.080	0.094	0.687	0.695	0.347	0.730	0.082	0.103
3	0.672	0.644	0.205	0.687	0.068	0.081	0.619	0.730	0.280	0.715	0.070	0.091
4	0.477	0.462	0.287	0.596	0.139	0.393	0.511	0.693	0.233	0.645	0.142	0.406
5	0.100	0.697	0.438	0.169	0.669	0.724	0.069	0.702	0.291	0.243	0.674	0.725
6	0.841	0.217	0.282	0.777	0.017	0.104	0.882	0.453	0.109	0.880	0.018	0.114
7	0.590	0.478	0.169	0.686	0.073	0.336	0.632	0.591	0.097	0.717	0.075	0.343
8	0.593	0.595	0.322	0.607	0.138	0.178	0.581	0.658	0.243	0.659	0.142	0.189
9	0.677	0.783	0.330	0.632	0.091	0.156	0.683	0.655	0.174	0.668	0.094	0.164
10	0.811	0.700	0.516	0.792	0.029	0.249	0.801	0.560	0.345	0.857	0.030	0.253
11	0.887	0.432	0.317	0.741	0.019	0.117	0.831	0.401	0.315	0.786	0.020	0.123
12	0.900	0.490	0.470	0.749	0.021	0.166	0.901	0.481	0.384	0.856	0.022	0.171
13	0.588	0.417	0.226	0.689	0.088	0.278	0.628	0.610	0.213	0.711	0.089	0.282
14	0.836	0.669	0.571	0.728	0.061	0.326	0.827	0.338	0.489	0.747	0.062	0.332
15	0.801	0.601	0.533	0.856	0.031	0.162	0.804	0.469	0.330	0.865	0.032	0.175
16	0.858	0.668	0.481	0.792	0.036	0.153	0.888	0.659	0.409	0.812	0.037	0.175
17	0.878	0.402	0.527	0.791	0.023	0.171	0.897	0.682	0.632	0.836	0.024	0.181
18	0.397	0.528	0.144	0.563	0.183	0.419	0.462	0.485	0.118	0.602	0.186	0.428
19	0.438	0.363	0.236	0.612	0.150	0.338	0.428	0.358	0.118	0.636	0.154	0.343
20	0.606	0.547	0.419	0.682	0.084	0.144	0.634	0.525	0.228	0.700	0.085	0.155
21	0.659	0.725	0.255	0.662	0.127	0.236	0.618	0.738	0.318	0.689	0.129	0.251
22	0.710	0.505	0.472	0.777	0.032	0.175	0.688	0.428	0.390	0.805	0.033	0.187
23	0.754	0.376	0.360	0.714	0.030	0.226	0.770	0.384	0.138	0.716	0.031	0.239
24	0.529	0.486	0.301	0.632	0.124	0.303	0.497	0.489	0.281	0.680	0.127	0.310
25	0.536	0.177	0.183	0.606	0.109	0.337	0.550	0.089	0.122	0.647	0.112	0.350
26	0.607	0.355	0.300	0.655	0.120	0.200	0.594	0.464	0.215	0.682	0.123	0.208
27	0.646	0.331	0.198	0.671	0.073	0.242	0.614	0.283	0.190	0.697	0.074	0.252
28	0.102	0.593	0.426	0.084	0.828	0.798	0.125	0.649	0.363	0.139	0.835	0.811
29	0.221	0.498	0.306	0.384	0.475	0.459	0.195	0.556	0.205	0.416	0.485	0.464
30	0.564	0.215	0.230	0.680	0.093	0.307	0.530	0.382	0.335	0.695	0.095	0.310
31	0.776	0.131	0.390	0.643	0.031	0.265	0.758	0.316	0.270	0.656	0.031	0.274
32	0.492	0.129	0.180	0.726	0.105	0.379	0.516	0.322	0.118	0.733	0.107	0.381
33	0.439	0.611	0.436	0.544	0.247	0.289	0.470	0.780	0.370	0.579	0.253	0.288
34	0.593	0.651	0.536	0.530	0.246	0.191	0.592	0.732	0.433	0.570	0.249	0.200
35	0.597	0.479	0.406	0.662	0.075	0.365	0.621	0.472	0.438	0.688	0.077	0.375
36	0.635	0.259	0.432	0.714	0.055	0.381	0.675	0.418	0.461	0.730	0.057	0.395
37	0.750	0.266	0.462	0.743	0.025	0.287	0.773	0.442	0.342	0.794	0.026	0.293
38	0.538	0.468	0.482	0.625	0.128	0.304	0.583	0.461	0.414	0.614	0.135	0.307
39	0.554	0.575	0.447	0.592	0.145	0.311	0.591	0.540	0.339	0.589	0.150	0.315
40	0.699	0.450	0.508	0.704	0.055	0.203	0.702	0.436	0.394	0.701	0.058	0.211

4.3. Elicitation of weight restrictions

In DEA, DMUs choose their best conceivable weights; the fact that those may be in contradiction with *a priori* knowledge leads to the introduction of managerial preferences on the relative importance of the inputs and outputs used in the assessment [42]. The introduction of weight restrictions in the model helps to reflect the organization's objectives, ensuring meaningful results which are closer to what the DM considers as best practices.

The weights employed in the additive aggregation MAUT model used in the Value-Based DEA method [26] are the scale coefficients of the utility functions, which allow for utility trade-offs between different factors (see [35]). The weights of the additive aggregation model can be assessed, using again the DM's judgments. There are various techniques available which may help to obtain weights, such as the direct rating, the swings method and the indifference equations [36, 43]. The swings method was considered to be the most appropriate for this case because it is simpler and clearer to the DM.

The swings method begins by constructing two extreme hypotheses, B and G, the first one displaying the worst performance on all criteria and the second one the corresponding best performance. The preference elicitation protocol consists in querying the DM to observe the potential gains from moving from B to G on each criterion and then to decide which of the criteria he/she prefers to shift to hypothesis G. Suppose that the transition from hypothesis B to hypothesis G on a specified criterion is worth 100 units in a hypothetical scale. Then, the DM is asked to give a value (<100) to the second criterion moved to G, then to the third criterion and so on, until the last criterion is moved to G.

The procedure used in this work was to start with a ranking of weights, via the swings method, and then to establish a limit to the ratio between the weights ranked first and last, by means of a trade-off question, to avoid null weights. Let W denote the set of weighting vectors compatible with the elicited ranking and ratio limit.

After the elicitation of weight restrictions, formulation (1) (see section 2) is modified to include the weight restrictions $(w_1,...,w_q) \in W$. With this change in Phase 1, it is necessary to change formulation (2) allowing the slacks to have negative values; otherwise it might not be possible to keep the optimal utility difference d_k^* derived from Phase 1 with the weight vectors incorporated (for details, see [41]).

5. Results

5.1. Comparison of results for standard DEA models and the Value-Based DEA method

In this subsection, we compare the results of the BCC DEA model, obtained by Weyman-Jones *et al*. [31] and the Value-Based DEA method, without considering weight restrictions. Although Weyman-Jones *et al*.'s study presents results for the CCR, BCC and SBM models, for the sake of comparison with this approach only BCC results are referred to, as Value-Based DEA builds on the additive model with a VRS frontier and the results obtained with BCC and SBM models do not vary much, maintaining the number of efficient units.

Figure 2 exhibits a comparison between the results obtained with the Value-Based DEA method and the standard BCC model (input oriented), for the year 2005. As stated in section 2, if d^* is negative then the DMU under analysis is efficient and if $d^* > 0$ then the DMU is inefficient. The number of efficient units decreases in Value-Based DEA method, since the incorporation of preferences during the construction of the utility functions changes the shape of the efficient frontier. In the BCC model 19 DMUs are efficient (the ones with efficiency 1 on the y-axis), but three of these DMUs lose efficiency in Value-Based DEA method: the ones with efficiency 1 (y-axis) but $d^* > 0$ (x-axis), that is in the first quadrant in Figure 2. DMUs that are classified as efficient in the Value-Based DEA method are also efficient in the BCC model. There are no DMUs with $d^* < 0$ and with efficiency score in the BCC model less than 1. For example, DMU 2 (in the 1st quadrant) is efficient in the BCC model, but inefficient considering Value-Based DEA method. This is mainly due to the DM's preferences reflected in the utility functions. DMUs 28 and 33 (2nd quadrant) are efficient in both models (Phase 1 of Value-Based DEA method and BCC model); DMU 1 (1st quadrant) is classified as inefficient in both.

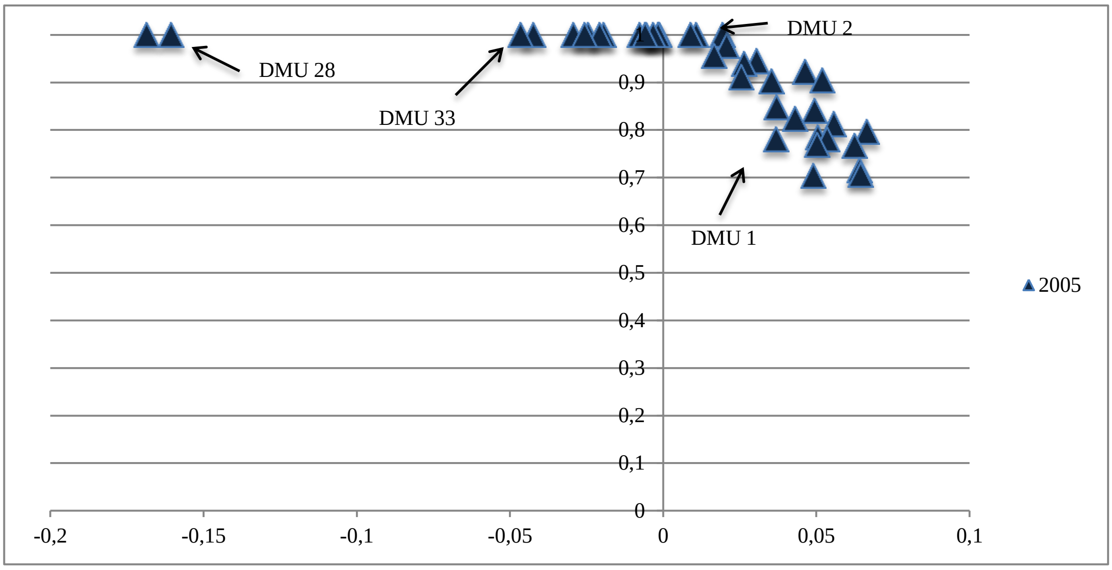

Figure 2. Comparison of the BCC model and Value-Based DEA method (without weight restrictions), for 2005 data (d^* in the x-axis and BCC score in the y-axis).

A study considering the combined set of 2004 and 2005 observations, with 40 DMUs, was also carried out (Figure 3), in order to identify when variation of efficiency has occurred. In the BCC model, the number of efficient units increases from 6 (2004) to 18 (2005), and in the Value-Based DEA method the number of efficient DMUs increases from 7 to 14. Only one unit is classified as inefficient in the BCC and as efficient in the Value-Based DEA method (DMU 12, in 2004), all the other efficient units for the BCC model also having the same classification in the Value-Based DEA method. On the other hand, four DMUs lose the efficiency status with the Value-Based DEA method in the year 2005 (DMUs 2, 3, 21, 37).

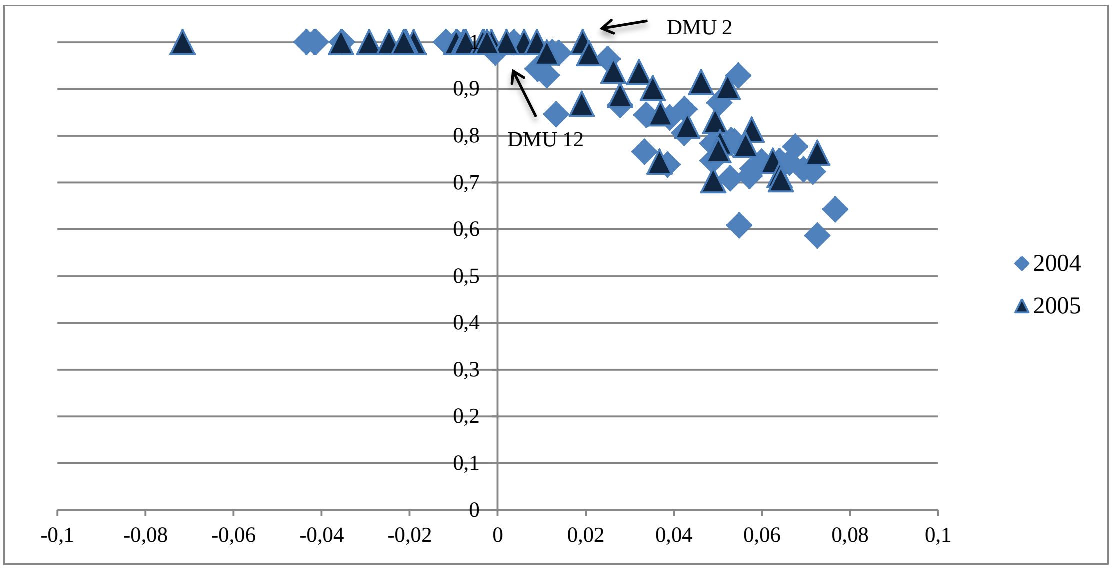

Figure 3. Comparison of the BCC model and Value-Based DEA method (without weight restrictions), for 2004 and 2005 data (d^* in the x-axis and BCC score in the y-axis).

This analysis leads to the conclusion that the radial DEA model is more "generous" in the classification of units in most cases, but there is a reasonable agreement between both approaches.

The DM has learned what was the impact of his / her responses in terms of the overall strength of each DMU, taking into account the diverse factors according to his / her managerial preferences. Additionally, the DM was offered information providing further discrimination of the efficient DMUs, in comparison with standard DEA models.

5.2. Further results of Value-Based DEA method

In the previous subsection the number of efficient units was determined considering the year 2005 and combining the observations of the years 2004 and 2005, when comparing the BCC model with the Value-Based DEA method. In this subsection results for the same years will be presented, but focusing on the Value-Based DEA method to observe performance changes from 2004 to 2005, since this method yields more discriminating results. The inclusion of weight restrictions in the Value-Based DEA method is also discussed.

Table 4 shows the evaluation of DMUs' efficiency across the two years without weights restrictions. The efficiency measure d^* decreased for DMUs 5, 9, 11, 14, 15, 28, 31, 38, 39 and 40. DMU 9 is the only one that lost the efficiency status from 2004 to 2005. In fact, DMUs that change from inefficient to efficient, considering both years, namely DMUs 16 and 17, have better utility values in almost all factors (except y_{CLI}) than DMU 9 (in 2005) and DMU 33. DMU 33, despite having a worse x_{OPEX} utility value than DMU 9, improved from 2004 to 2005 in x_{OPEX}, x_{MLL} (the best unit in this factor), x_{NI} (worse than DMU 9), and y_{CLI} (much better than DMU 9). The factors in which DMU 33 worsened from 2004 to 2005 (x_{CC} and y_{NLL}) have a better utility value when compared with the utilities of the same factors in DMU 9 in 2005.

Table 4. $d*$ for the 40 DMUs (2004-2005) and the difference between d^*considering both years (by increasing order of d^*(2005)).

DMUs	d^* (2004)	d^* (2005)	d^* (2005) − d^* (2004)	DMUs	d^* (2004)	d^* (2005)	d^* (2005) − d^* (2004)
17	0.0112	**-0.0715**	**-0.0827**	2	0.0504	0.0193	**-0.0311**
28	**-0.0416**	**-0.0356**	0.0060	29	0.0425	0.0208	**-0.0217**
10	**-0.0117**	**-0.0291**	**-0.0174**	18	0.0424	0.0263	**-0.0161**
33	0.0546	**-0.0246**	**-0.0793**	22	0.0391	0.0278	**-0.0113**
36	0.0133	**-0.0213**	**-0.0346**	13	0.0600	0.0321	**-0.0279**
6	0.0338	**-0.0208**	**-0.0546**	8	0.0716	0.0353	**-0.0363**
5	**-0.0353**	**-0.0191**	0.0162	1	0.0549	0.0368	**-0.0181**
4	0.0489	**-0.0094**	**-0.0582**	11	0.0091	0.0370	0.0279
12	**-0.0005**	**-0.0077**	**-0.0072**	30	0.0529	0.0431	**-0.0098**
14	**-0.0434**	**-0.0072**	0.0363	24	0.0695	0.0463	**-0.0232**
32	0.0037	**-0.0033**	**-0.0070**	25	0.0726	0.0490	**-0.0236**
15	**-0.0092**	**-0.0031**	0.0061	9	**-0.0413**	0.0494	0.0907
34	0.0250	**-0.0024**	**-0.0275**	23	0.0572	0.0501	**-0.0071**
16	0.0139	**-0.0014**	**-0.0153**	19	0.0643	0.0505	**-0.0138**
21	0.0124	0.0020	**-0.0104**	26	0.0676	0.0523	**-0.0153**
3	0.0538	0.0061	**-0.0478**	31	0.0488	0.0563	0.0075
37	0.0279	0.0090	**-0.0189**	20	0.0663	0.0577	**-0.0086**
7	0.0386	0.0112	**-0.0274**	38	0.0531	0.0625	0.0094
35	0.0334	0.0191	**-0.0142**	40	0.0580	0.0641	0.0061
27	0.0767	0.0643	**-0.0124**	39	0.0641	0.0727	0.0086

Table 5 displays the results from Phase 1, only for efficient units, in the year 2005. When there are no restrictions on the weights, some factors may be disregarded from the assessment because DMUs can assign zero weights to some factors (namely those presenting low levels of outputs and high levels of inputs). For example, DMU 28 is efficient but all weights are null except one, and there are several factors disregarded in other units. Moreover, DMUs can assign weights to their factors ignoring recognized opinions about the value of those factors [44]. The incorporation of weight restrictions expressed by the DM on the efficiency assessment of DMUs is a way to overcome the problem of having less credible efficiency scores.

Table 5. Results of Phase 1 for efficient units without weight restrictions, for 2005 data.

DMUs	d^*	w^*_{OPEX}	w^*_{MLL}	w^*_{CC}	w^*_{NI}	w^*_{CLI}	w^*_{NLL}
3	**-0.001**	0.000	0.724	0.000	0.276	0.000	0.000
4	**-0.020**	0.000	0.278	0.000	0.350	0.000	0.372
5	**-0.025**	0.000	0.800	0.000	0.000	0.000	0.200
6	**-0.021**	0.091	0.000	0.000	0.909	0.000	0.000
10	**-0.030**	0.000	0.069	0.000	0.570	0.000	0.361
12	**-0.008**	0.286	0.000	0.000	0.468	0.000	0.246
14	**-0.042**	0.490	0.000	0.000	0.000	0.000	0.510
15	**-0.006**	0.000	0.000	0.055	0.690	0.254	0.000
16	**-0.001**	0.510	0.003	0.000	0.000	0.486	0.000
17	**-0.169**	0.099	0.165	0.736	0.000	0.000	0.000
21	**-0.002**	0.000	0.504	0.000	0.281	0.000	0.214
28	**-0.161**	0.000	0.000	0.000	0.000	1.000	0.000
32	**-0.003**	0.000	0.000	0.000	0.466	0.146	0.389
33	**-0.047**	0.000	0.937	0.000	0.000	0.052	0.011
34	**-0.006**	0.251	0.491	0.000	0.000	0.259	0.000
36	**-0.026**	0.000	0.017	0.052	0.409	0.000	0.521

Following the procedure explained in subsection 4.3, weight restrictions were elicited by asking the DM to compare the “swings” of utility from 0 to 1 as depicted in Table 6.

The DM was asked to consider one hypothetical unit with the performance level 0 for all factors and the question was: "if you could improve one and only one factor to the maximum utility level (1), which would you choose?". The DM’s answer was: x_{OPEX}. This allows the inference that w_{OPEX} is the highest scaling constant. By repeating this question successively for the remaining factors, the following ranking of the scale coefficients was attained: $w_{OPEX} \geq w_{MLL} \geq w_{NI} \geq w_{CC} \geq w_{NLL} \geq w_{CLI}$.

Table 6. Extreme performances associated with utility levels 0 and 1.

Utility level	x_{OPEX}	x_{MLL}	x_{CC}	x_{NI}	y_{CLI}	y_{NLL}
$u(.) = 0$	10500000	800	2.5	40000	45000	2000
$u(.) = 1$	900000	60	0	1000	1100000	20000

After the DM had established the ranking of the scale coefficients and in order to avoid zero-value weights, an indifference judgment question was asked to limit the ratio of the weights ranked in the first and last position. The answer to the question “What would be the lowest amount *h* that would allow a unit with of 1.1 million clients and with maintenance and outage repairing costs of 10.5 million euros to be considered as having more merit than a unit with 45 000 clients and with maintenance and outage repairing costs of *h*?” was $h = 1$ million euros. This answer is translated into: $w_{CLI}\ u_{CLI}(1\ 100\ 000) + w_{OPEX}\ u_{OPEX}\ (10\ 500\ 000) \geq w_{CLI}\ u_{CLI}\ (45\ 000) + w_{OPEX}\ u_{OPEX}\ (h)$.

Substituting *h* in the previous expression yields: $w_{OPEX} \leq 1.05\ w_{CLI}$.

Table 7 portrays the results from Phase 1 and Phase 2 under weight restrictions and free slacks, only for efficient units, in the year 2005.

Table 7. Results of Phase 1 and Phase 2 under weight restrictions and free slacks for 2005 data.

	Phase 1							Phase 2					
DMUs	d^*	w^*_{OPEX}	w^*_{MLL}	w^*_{CC}	w^*_{NI}	w^*_{CLI}	w^*_{NLL}	s^*_{OPEX}	s^*_{MLL}	s^*_{CC}	s^*_{NI}	s^*_{CLI}	s^*_{NLL}
3	0.125	0.171	0.171	0.163	0.171	0.163	0.163	0.28	**-0.05**	0.35	0.12	**-0.05**	0.09
4	0.104	0.167	0.167	0.167	0.167	0.167	0.167	0.39	**-0.01**	0.40	0.19	**-0.12**	**-0.22**
5	0.091	0.167	0.167	0.167	0.167	0.167	0.167	0.83	**-0.02**	0.34	0.59	**-0.65**	**-0.54**
6	0.131	0.171	0.171	0.163	0.171	0.163	0.163	0.02	0.23	0.52	**-0.05**	0.01	0.07
10	0.068	0.171	0.171	0.163	0.171	0.163	0.163	0.10	0.12	0.29	**-0.02**	**-0.01**	**-0.07**
12	0.072	0.174	0.165	0.165	0.165	0.165	0.165	**-0.00**	0.20	0.25	**-0.02**	0.00	0.01
14	0.076	0.174	0.165	0.165	0.165	0.165	0.165	0.07	0.34	0.14	0.09	**-0.04**	**-0.15**
15	0.096	0.171	0.171	0.163	0.171	0.163	0.163	0.10	0.21	0.30	**-0.03**	**-0.01**	0.01
16	0.045	0.171	0.171	0.163	0.171	0.163	0.163	0.01	0.02	0.22	0.02	**-0.01**	0.01
17	**-0.046**	0.169	0.169	0.169	0.169	0.161	0.161						
21	0.085	0.167	0.167	0.167	0.167	0.167	0.167	0.28	**-0.06**	0.31	0.15	**-0.11**	**-0.07**
28	0.055	0.167	0.167	0.167	0.167	0.167	0.167	0.77	0.03	0.27	0.70	**-0.81**	**-0.63**
32	0.175	0.167	0.167	0.167	0.167	0.167	0.167	0.38	0.36	0.51	0.10	**-0.08**	**-0.20**
33	0.086	0.167	0.167	0.167	0.167	0.167	0.167	0.43	**-0.10**	0.26	0.26	**-0.23**	**-0.11**
34	0.080	0.167	0.167	0.167	0.167	0.167	0.167	0.31	**-0.05**	0.20	0.27	**-0.23**	**-0.02**
36	0.086	0.167	0.167	0.167	0.167	0.167	0.167	0.22	0.26	0.17	0.11	**-0.03**	**-0.21**

When comparing results without weight restrictions and with weight restrictions, as expected d^* is worse (i.e., higher) for all units when the weight restrictions are incorporated into the model (Figure 4). Only DMU 17 (d^* = -0.046) is efficient when the weight restrictions are considered in the Value-Based DEA method. The best unit (DMU 17) without considering weight restrictions is still the best DMU considering the weight restrictions previously stated. This means that all units should aim at achieving DMU 17's utility, but not necessarily trying to imitate the mix of inputs and outputs of that DMU (this is analyzed in section 6).

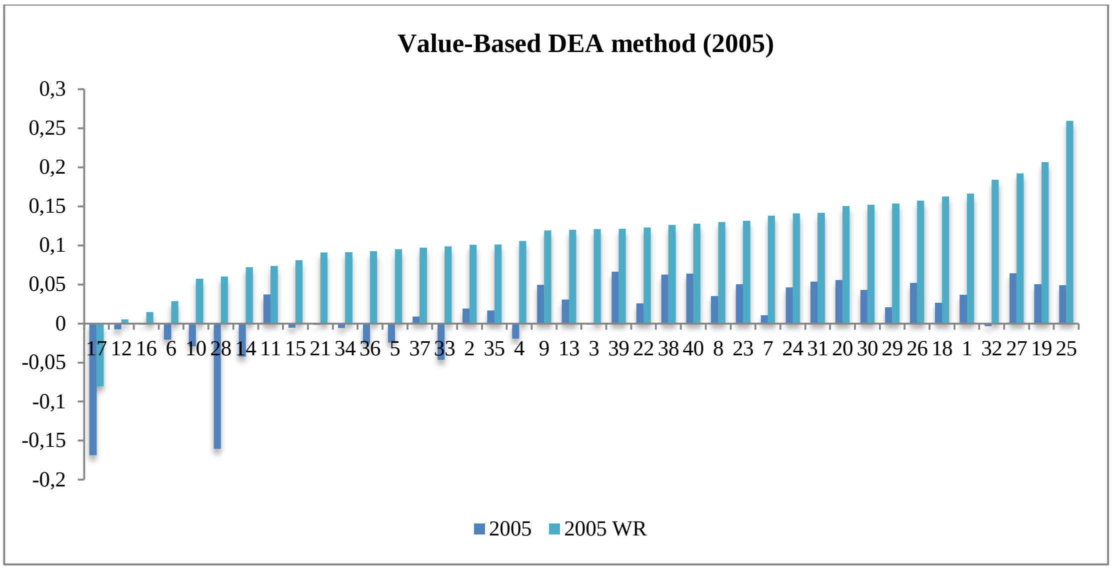

Figure 4. Comparison of Value-Based DEA method results without and with weight restrictions, ranked by 2005 efficiency measure with weight restrictions.

In conclusion, it can be stated that only one unit is classified as efficient in the analysis when managerial preferences are incorporated into the model. This is due to the fact that managerial preferences did not give much freedom to DMUs for choosing the weights. In further experiments we used the 2005 data and tried a less stringent trade-off limit $w_{OPEX} \leq 3.52\ w_{CLI}$, (if the answer to the indifference judgment question was h = 5 million euros, accommodating the response of DM) but the results were very similar. With this trade-off limit ($w_{OPEX} \leq 3.52\ w_{CLI}$) and comparing the results in terms of units ranking with the one established by the DM ($w_{OPEX} \leq 1.05\ w_{CLI}$), we concluded that there was another efficient DMU (DMU 16) besides DMU 17. The ranking changes are small and have to do with a few exchanges of DMUs in consecutive positions. Another experiment was performed, in this case for the ranking of weights. It consisted in exchanging the positions of the factors placed in second and third place (x_{MLL} or x_{NI}), because the first and last position in the ranking of the weights were undoubted. The ranking of scale coefficients for all factors became: $w_{OPEX} \geq w_{MLL} \geq w_{CC} \geq w_{NLL} \geq w_{CLI} \wedge w_{OPEX} \geq w_{NI} \geq w_{CC}$ maintaining the less stringent trade-off limit $w_{OPEX} \leq 3.52\ w_{CLI}$. In fact, there were no changes in the number of efficient units, some exchanges were detected only in the ranking of inefficient DMUs. The results are thus quite insensitive with regards to the respective answers provided by the DM.

6. Prospects for Improvement in Inefficient Units

The main objective of this study was to illustrate the results of a different possibility for making an internal benchmarking exercise with the introduction of managerial preferences. Hence, the results are expected to be closer to what the DM judges to be the best practice for a specific activity of the company. The use of the 2004-2005 data allows a comparison of results from different approaches, the one performed by Weyman-Jones *et al*. [31] and Value-Based DEA.

To illustrate the various possibilities for improvement in inefficient units, only the ten units with higher x_{OPEX} in 2005 (the lower u_{OPEX} value) were considered. The proposed solution of an efficiency target (projection) for each inefficient DMU is displayed in Table 7 in the previous section (Phase 2). For that year all units should match the overall utility of DMU 17, the only efficient unit. Note that to attain the efficiency status these inefficient DMUs must change their utility in each factor by the amount indicated by s^*, which does not necessarily correspond to an improvement since some of the inefficient units may have negative slacks. In fact, an inefficient DMU may be able to match its peers on the efficient frontier having a negative slack corresponding to an input (meaning it would increase) or a negative slack corresponding to an output factor (meaning it would decrease) if this is compensated by enough improvement in other factors. For the DM this revealed to be difficult to understand and thus there was a need to propose another solution.

Let u_k^* denote the utility of the best DMU using the optimal weighting vector resulting from (1) with the weight restrictions added, i.e., the utility value that DMU k ought to achieve:

$$u_k^* = \sum_{c=1}^{q} w_k^* u_c(DMU_k) + d_k^* \tag{3}$$

The problem solved in Phase 2 admits alternative optimal targets, corresponding to different ways of closing the gap d_k^*. These targets correspond to different projections on the efficient frontier. The purpose is to constrain the proposed efficiency targets to achieve u_k^*, not only to avoid those targets that imply an increase of inputs or a decrease of outputs, but also to choose which factors can be changed, given the characteristics of each unit. Hence, this requires that the utility value cannot decrease in any factor and targets are forced to maintain or improve the performance of all factors [41].

To accomplish this it is necessary to define two sets to develop new model formulations (4)-(5). Let $S_< = \left\{ c \in \{1,\dots,q\} : s_c^* < 0 \text{ in Phase 2} \right\}$ denote the negative slacks in the optimal solution obtained for (2) with free slacks; these slacks will now become null constants. Let $S_\geq = \left\{ c \in \{1,\dots,q\} : s_c^* \geq 0 \text{ in Phase 2} \right\}$ denote the remaining slacks, which will be considered as non-negative variables. Therefore, a formulation which yields an alternative target can be obtained by solving the linear problem (4) in which the maximum slack (in terms of value) shall be minimized to achieve the global utility target. No negative slacks are allowed, but the target will no longer be a convex combination of the observed DMUs.

$$\begin{aligned}
\min_{\rho_k, s} \quad & \rho_k \\
s.t. \quad & \sum_{c \in S_\geq} w_c^* s_c = d_k^* \\
& u_c(DMU_k) + s_c \leq 1, c = 1,\dots,q \\
& s_c - \sigma_c \leq 0, c \in S_\geq \\
& \rho_k \geq s_c \geq 0, c = 1,\dots,q
\end{aligned} \tag{4}$$

A parameter σ_c is introduced to bound the value a slack may have. The purpose is to avoid setting unrealistic improvement targets. The restriction $s_c - \sigma_c \leq 0$ in (4) implies that only the factors with negative slacks in Phase 2 have $\sigma_c > 0$ and can be changed in order to overcome the gap with the peer. Targets will never exceed the value 1 in any factor due to the restriction $u_c(DMU_k)+s_c \leq 1$. This ensures that the utility function does not spill over outside the ranges elicited for the performances.

A proposal to improve performance of all inefficient units is to block the changes in the factors with negative slacks in Phase 2 (see proposal 1 in Table 8). However, for some of these ten inefficient units, the x_{OPEX} decrease is attainable, but for others, given their specific characteristics, the proposed value would be impossible to achieve. For those units the DM does not consider attainable a x_{OPEX} reduction greater than 40%. This requires a new formulation (5) according to which the gap to the peer is distributed by all the factors (inputs and outputs), in a balanced way (proposal 2 in Table 8):

$$\begin{aligned} &\min_{\rho_k, s} \quad \rho_k \\ &s.t. \quad \sum_{c=1}^{q} w_c^* s_c = d_k^* \\ &\qquad u_c(DMU_k) + s_c \leq 1, c = 1,\ldots,q \\ &\qquad s_c - \sigma_c \leq 0, c = 1,\ldots,q \\ &\qquad \rho_k \geq s_c \geq 0, c = 1,\ldots,q \end{aligned} \tag{5}$$

However, the result still did not satisfy the DM in some cases. With this proposal all units reduced the x_{OPEX} value by a percentage below 40%; however, for some units it was impossible to increase the y_{CLI} value recommended and the DM did not accept increases in y_{NLL}. Hence, the gap between the unit under evaluation and DMU 17 is closed by decreasing input factors x_{MLL}, x_{CC} and x_{NI} and increasing the output factor y_{CLI} in some cases (DMU 29) (see proposal 3 in Table 8). Factors that have a limit to reduction (which is the case of x_{OPEX}) or a limit to increase (the case of y_{CLI}) are displayed in bold typeface in Table 8 (proposal 3).

The DM approved that the DMUs 4, 5, 28 and 33 maintain the decreases suggested in proposal 1. For the remaining ones there are proposals 2 and 3 to choose from, according to the specific characteristics of each DMU.

It is relevant to point out that the DM was pleased with the possibility of having the entire range of targets available for making the choice.

Table 8. Improvement proposals for the ten inefficient units with higher x_{OPEX} in 2005: required variation in the original units.

	Proposal 1					
DMUs	Δx_{OPEX}	Δx_{MLL}	Δx_{CC}	Δx_{NI}	Δy_{CLI}	Δy_{NLL}
1	-45.48%	-54.98%	-55.17%	-63.16%		
4	-39.42%		-48.29%	-64.91%		
5	-35.68%		-47.14%	-43.69%		
18	-44.65%	-53.41%	-52.76%	-70.80%		
19	-52.04%	-60.79%	-61.90%	-73.81%		
24	-40.87%	-48.00%	-54.68%	-61.66%		
28	-18.23%	-26.87%	-25.90%	-22.96%		
29	-42.95%	-51.26%	-52.97%	-63.26%		
32	-47.87%	-55.56%	-56.73%	-58.51%		
33	-33.97%		-49.65%	-67.26%		
	Proposal 2					
DMUs	Δx_{OPEX}	Δx_{MLL}	Δx_{CC}	Δx_{NI}	Δy_{CLI}	Δy_{NLL}
1	-33.12%	-38.75%	-40.80%	-52.89%	130.90%	49.92%
4	-22.05%	-73.40%	-25.70%	-55.15%	50.68%	26.06%
5	-19.63%	-23.60%	-23.62%	-22.62%	19.67%	18.94%
18	-32.38%	-36.78%	-40.76%	-64.94%	64.92%	42.94%
19	-38.77%	-47.13%	-47.01%	-65.23%	94.44%	62.27%
24	-29.51%	-32.86%	-36.33%	-53.89%	76.17%	46.69%
28	-12.37%	-17.91%	-17.42%	-15.17%	8.92%	10.42%
29	-31.29%	-38.86%	-37.82%	-47.71%	53.07%	39.78%
32	-35.19%	-44.58%	-43.33%	-43.94%	105.25%	51.02%
33	-18.67%	-11.51%	-27.78%	-33.37%	28.18%	29.83%
	Proposal 3					
DMUs	Δx_{OPEX}	Δx_{MLL}	Δx_{CC}	Δx_{NI}	Δy_{CLI}	Δy_{NLL}
1	**-40.00%**	-50.30%	-51.68%	-60.64%	**80.00%**	
4						
5						
18	**-40.00%**	-46.96%	-48.25%	-68.57%	**50.00%**	
19	**-40.00%**	-61.22%	-62.48%	-74.10%	**60.00%**	
24	**-40.00%**	-42.14%	-47.39%	-58.53%	**60.00%**	
28						
29	**-40.00%**	-46.63%	-43.89%	-53.78%	64.11%	
32	**-40.00%**	-53.06%	-53.65%	-55.31%	**70.00%**	
33						

7. Concluding Remarks

In order to enhance certain core activities, EDP *Distribuição* conducted an internal benchmarking study using several DEA models [31]. The current study exploits the results of the application of the Value-Based DEA method to the same data. This approach allows for the incorporation of managerial preferences to identify best practices. The incorporation of preferences was carried out by converting input and output factor performances into utility functions, which required interpreting these utility functions as devices to compare the change in merit corresponding to performance differences. With this approach it was possible to compare the results using Value-Based DEA with the results obtained using the classical DEA models. This study highlights the impact of the incorporation of DM’s preferences (in the construction of utility functions) in the classification of units. In order to bring more managerial information to the units’ assessment and to make this benchmarking study as complete and useful as possible for the company, weight restrictions were introduced and those were elicited using the swings technique to draw a ranking of weights. A trade-off question was used to limit the ratio between the weights ranked first and last, to avoid null weights.

Only one DM was interviewed in the process of elicitation of the utility functions and the introduction of weight restrictions, although later the answers were validated by senior managers of the company. The DM already had used DEA in evaluating efficiency of the company's activities, which greatly facilitated the process. The DM's interest increased in the course of the interaction process and he always gave assertive answers to all questions; however, the answer to the trade-off question was the one that led to longer reflection (which motivated experimenting with different trade-off values).

A ranking of all the units and benchmarks were identified using Value-Based DEA with and without weight restrictions for the year of 2005 and also combining the 80 observations using the 2004-2005 data for the 40 DMUs. The main conclusion is that when managerial preferences are incorporated into Value-Based DEA only one unit is classified as efficient, the one exhibiting best practices taking those preferences into account. Some conclusions were drawn for the utility trends across the two years.

We find three proposals for improvements of the ten inefficient units with the highest x_{OPEX} in 2005, which were then discussed and refined considering the DM´s requests. We leave to the DM the possibility to choose the most suitable proposal for each DMU, given its specific characteristics.

According to the opinion expressed by the DM, the Value-Based DEA is more suitable for management purposes than standard DEA models, as the new approach allows for the inclusion of managerial preferences in the analysis - the introduction of utility functions allows for the consideration of an order of merit for the relative input and output changes, and the DM may define a certain threshold beyond which the values of the variables should be considered as unacceptable. Therefore, the Value-Based approach revealed more adequate for supporting management decisions in shaping future company policies. The results presented illustrate how these differences are reflected in practice, in this particular case. A more clear-cut result is obtained, in the form of a reduction in the number of efficient units, when such managerial preferences are incorporated into the analysis. On the other hand, standard DEA is easier to use as it does not require such an active effort from the DMs. Each approach has its strengths and weaknesses and the choice of the methodology depends on the specific needs of managers, at a given moment.

The main purpose of this study was to show a different possibility for an internal benchmarking exercise incorporating managerial preferences. Hence, the results obtained are closer to what the DM judges to be the best practice for the company. With this analysis, in addition to identifying best practices, sources of inefficiency, gaps relatively to best practices and opportunities for improvement, it is also possible to support the introduction of corrective measures and to inform decisions about future goals, as well as to improve the knowledge about the company.

Acknowledgements

The authors are grateful to Miguel Stilwell d'Andrade, member of the Board of Directors of EDP, and João Gomes Aguiar, president of EDP Serviço Universal, for their support to this study. This work has been partially supported by Fundação para a Ciência e a Tecnologia (FCT) under project grants PEst-OE/EEI/UI0308/2014 and EMSURE - Energy and Mobility for Sustainable Regions (CENTRO-07-0224-FEDER-002004).

References

[1] N. Bagdadioglu, C. M. W. Price, T. G. Weyman-Jones, Efficiency and Ownership in Electricity Distribution: A Non Parametric Model of the Turkish Experience, Energy Economics 18 (1996) 1-23.

[2] N. Bagdadioglu, C. M. W. Price, T. G. Weyman-Jones, Measuring Potential Gains from Mergers among Electricity Distribution Companies in Turkey using a Non-parametric Model, The Energy Journal 28 (2) (2007) 83-110.

[3] W.W. Cooper, L. M. Seiford, K. Tone, Data Envelopment Analysis – A comprehensive text with models, applications, references and DEA-Solver software (Kluwer Academic Publishers, Boston, 2000).

[4] E. Thanassoulis, Introduction to the Theory and Application of Data Envelopment Analysis: A Foundation Text with Integrated Software (Kluwer Academic Publishers, 2001).

[5] J. Zhu, Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets and DEA Excel Solver (Kluwer Academic Publishers, 2003).

[6] Z. Huang, S. X. Li, J. Zhu, A Special Issue on "Data Envelopment Analysis: Theories and Applications" In honor of William W. Cooper, International Journal of Information Technology & Decision Making 4 (3) (2005) 311-316.

[7] T.G. Weyman-Jones, J. Boucinha, C. F. Inácio, Efficiency analysis of maintenance and outage repair in electricity distribution, Portuguese Journal of Management Studies XIII (2008) 89–110.

[8] X. Dai, T. Kuosmanen, Best-practice benchmarking using clustering methods: Application to energy regulation, Omega 42 (1) (2014) 179-188.

[9] A. Charnes, W. W. Cooper, E. Rhodes, Measuring the efficiency of decision making units, European Journal of Operational Research 2 (6) (1978) 429-444.

[10] R. Färe, S. Grosskopf, J. Logan, The relative efficiency of Illinois electric utilities, Resources and Energy 5 (1983) 349–367.

[11] T. Jamasb and M. Pollitt, Benchmarking and regulation: international electricity experience, Utilities Policy 9 (3) (2000) 107-130.

[12] T. Jamasb and M. Pollitt,, International benchmarking and regulation: An application to European electricity distribution utilities, Energy Policy 31 (2003) 1609–1622.

[13] A.B. Haney and M.G. Pollitt, Efficiency analysis of energy networks: An international survey of regulators, Energy Policy 37 (2009) 5814–5830.

[14] S.P. Santos, C.A. Amado, J.R. Rosado, Formative evaluation of electricity distribution utilities using data envelopment analysis, Journal of the Operational Research Society 62 (2011) 1298-1319.

[15] E. Thanassoulis, and R.G. Dyson, Estimating preferred input-output levels using data envelopment analysis, European Journal of Operational Research. 56 (1992) 80-97.

[16] J. Zhu, Data Envelopment Analysis with preference structure, Journal of the Operational Research Society 47 (1996) 136-50.

[17] M. Halme, T. Joro, P. Korhonen, S.E. Salo, J. Wallenius, A value efficiency approach to incorporating preference information in Data Envelopment Analysis, Management Science 45 (1999) 103-15.

[18] Y. Chen, On preference structure in Data Envelopment Analysis, International Journal of Information Technology & Decision Making 4 (3) (2005) 411-431.

[19] N.K. Avkiran, H. Morita, Benchmarking firm performance from a multiple stakeholder perspective with an application to Chinese banking, Omega 38 (2010) 501–508.

[20] C.A. Amado, S.P. Santos, P.M. Marques. Integrating the Data Envelopment Analysis and the Balanced Scorecard approaches for enhanced performance assessment, Omega 40 (2012) 390–403.

[21] C.A.F. Amado, S.P. Santos, J. Sequeira. Using Data Envelopment Analysis to support the design of process improvement interventions in electricity distribution, European Journal of Operational Research 228 (2013) 226-235.

[22] E. Thanassoulis, M.C. Portela, R. Allen, Incorporating value judgments in DEA, Handbook on Data Envelopment Analysis, eds. W.W. Cooper, L. M. Seiford, J. Zhu (Kluwer, Boston, 2004), pp. 99-137.

[23] D. Bouyssou, Using DEA as a tool for MCDM: some remarks, Journal of the Operational Research Society 50 (1999) 974-978.

[24] M. Köksalan and C. Tuncer, A DEA-based approach to ranking Multi-Criteria alternatives, International Journal of Information Technology & Decision Making 8 (1) (2009) 29-54.

[25] W. D. Cook, K. Tone, J. Zhu, Data envelopment analysis: Prior to choosing a model, Omega 44 (2014) 1-4.

[26] M.C. Gouveia, L.C. Dias, C.H. Antunes, Additive DEA based on MCDA with imprecise information, Journal of the Operational Research Society 59 (2008) 54-63.

[27] M.C. Gouveia, L.C. Dias, C.H. Antunes, Super-efficiency and stability intervals in additive DEA, Journal of the Operational Research Society.64 (2013) 86-96.

[28] A.I. Ali, C.S. Lerme, L. Seiford, Components of efficiency evaluation in Data Envelopment Analysis, European Journal of Operational Research 80 (1995) 462-473.

[29] A.D. Athanassopoulos and V.V. Podinovsky, Dominance and potential optimality multiple criteria models in decision analysis with imprecise information, Journal of the Operational Research Society 48 (1) (1997) 142-150.

[30] L.C. Dias and J.N. Clímaco, Additive Aggregation with Variable Interdependent Parameters: the VIP Analysis Software, Journal of the Operational Research Society 51 (9) (2000) 1070-1082.

[31] T.G. Weyman-Jones, J. Boucinha, C.F. Inácio, Quality of Supply and Efficiency: An Analysis of Portuguese Electricity Distribution Networks, SosyoEkonomi 6 (2010) 177-208.

[32] R.G. Dyson and E. Shale, Data envelopment analysis, operational research and uncertainty. Journal of the Operational Research Society 61 (2010) 25-34.

[33] R.D. Banker, A. Charnes, W.W. Cooper, Some models estimating technical and scale inefficiencies in Data Envelopment Analysis, Management Science 3 (9) (1984) 1078-1092.

[34] A. Charnes, W.W. Cooper, B. Golany, L. Seiford, J. Stutz, Foundations of Data Envelopment Analysis for Pareto-Koopmans efficient empirical production functions, Journal of Econometrics 30 (1985) 91-107.

[35] R.L. Keeney and H. Raiffa, Decisions with multiple objectives: preferences and value tradeoff (Wiley, New York, 1976).

[36] D. von Winterfeldt, W. Edwards, Decision analysis behavioural research (Cambridge University Press, New York, 1986).

[37] D.E. Bell, Regret in Decision Making under Uncertainty, Operations Research 30 (1982) 961-981.

[38] A.A. Salo and R.P. Hämäläinen, Preference Ratios in Multiattribute Evaluation (PRIME) - elicitation and decision procedures under incomplete information, IEEE Transactions on Systems Man and Cybernetics part A 31 (2001) 533-545.

[39] Sarabando P. and Dias L.C., Multiattribute Choice With Ordinal Information: A Comparison of Different Decision Rules, Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans 39 (2009) 545 – 554.

[40] P. Andersen and N.C. Petersen, A procedure for ranking efficient units in data envelopment analysis, Management Science 39 (1993) 1261-1264.

[41] P.N. Almeida and L.C. Dias, Value-based DEA models: application-driven developments, Journal of the Operational Research Society. 63 (2012) 16-27.

[42] R.G. Dyson and E. Thanassoulis, Reducing weight flexibility in DEA, Journal of the Operations Research Society 39 (1988) 563–576.

[43] P. Goodwin and G. Wright, Decision analysis for management judgment (Wiley, Chichester, 1998).

[44] R. Allen, A. Athanassopoulos, R.G. Dyson, E. Thanassoulis, Weights restrictions and value judgments in data envelopment analysis: evolution, development and future directions, Annals of Operations Research 73 (1997) 13–34.