Iranian Journal of Operations Research
Vol. 5, No. 1, 2014, pp. 29-46

An Integrated Model with Conservative Levels to Evaluate the DMUs Efficiencies for Uncertain Data

A. H. Shokouhi[1,*], H. Shahriari[2]

In traditional data envelopment analysis (DEA) the uncertainty of inputs and outputs is not considered when evaluating the performance of a unit. In other words, effects of uncertainty on optimality and feasibility of models are ignored. This paper introduces a new model for measuring the efficiency of decision making units (DMUs) having interval inputs and outputs. The proposed model is based on interval DEA (IDEA) in which the inputs and outputs are limited to be within uncertainty bounds. In this model, the inputs and outputs take fixed values for each DMU such that the sum of efficiencies is maximized. The DMUs are evaluated by the same production possibility set (PPS). The efficiency is measured based on the proposed conservatism level for each input and output. Indeed, the inputs and outputs are defined by the presented conservatism level. The proposed model is integrated measuring all the DMUs efficiencies simultaneously. These efficiency scores lie between the optimistic and pessimistic cases introduced by Despotis and Similar (2002) [11].

***Keywords**: Data envelopment analysis, Efficiency, Integrated model, Uncertainty, Conservatism level.*

Manuscript was received on 11/10/2013, revised on 09/03/2014 and accepted for publication on 09/09/2014.

1. Introduction

Data envelopment analysis (DEA) technique, first introduced by Charnes et al. [6], is now widely exploited for the measurement of efficiency of many entities in public and private sectors. An important methodological feature of DEA is its capability to determine the performance of a decision making unit (DMU) in comparison with all other DMUs. Moreover, it is widely known that DEA is developed to measure the relative efficiency of DMUs with multiple inputs and outputs using a linear programming (LP) model (Banker et al. [2]; Charnes et al. [6]). The main purpose of DEA models is to classify DMUs into two classes: efficient and inefficient. The original CCR[3] model is only applicable to technologies characterized by global constant returns to scale (CRS). This model is modified by Banker et al. [2], assuming variable returns to scale (VRS) technologies. Applying these models, the efficiency score of each DMU is obtained by assuming data certainty, not to evaluate DMUs with uncertain data such as imprecision, vagueness, inconsistency, etc..

Since uncertainty is present in real situations, it is observed in DEA. The data uncertainty is dealt with in different ways, such as fuzzy, stochastic, and interval approaches. Knowing membership functions and probability distributions are respectively necessary in fuzzy and stochastic approaches. Interval analysis is developed to model uncertainty in DEA, in which only the bounds of the uncertain

[*] Corresponding Author.
[1] Department of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran.
Email: shokouhi@dena.kntu.ac.ir
[2] Department of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran.
Email: hshahriari@kntu.ac.ir
[3] Charnes, Cooper and Rhodes.

data are required, not the membership functions or the probability distributions. Sengupta [26] initially introduced DEA models under uncertainty. Cooper et al. ([7], [8] and [9]) introduced an interval approach to deal with interval data in DEA. Interval data, strong and weak ordinal data and ratio interval data modeling were proposed by Kim et al. [20]. Despotis and Smirlis [11] proposed two models with interval data in DEA to obtain the upper and lower bounds of efficiency scores for DMUs as the optimistic and pessimistic models, respectively. The DMUs were classified into three groups according to the intervals obtained for the DMUs. Zhu [34] simplified the Cooper et al.'s ([7], [8], [9]) model. Wang et al. [32] proposed DEA models considering intervals to get a fixed production frontier for measuring the efficiencies of DMUs. Their models obtained the lower and upper bounds of the efficiencies for each DMU. Amirteimoori and Kordrostami [1] extended the Zhu's [34] model to multi-component efficiency. Jahanshahloo et al. [16] estimated a radius of stability for all DMUs with interval data and showed that the original classifications remained unchanged under perturbations. Jahanshahloo et al. [17] also introduced a method for measuring the efficiency of DMUs in the free disposal Hull (FDH) model with interval data. Kao [19] formulated the problem as a bi-level mathematical programming model to deal with uncertainty in data and converted the model into a pair of ordinary one-level linear programing one to assess the interval efficiency of DMUs. The use of Despotis and Smirlis's [11] approach in interval DEA was presented by Smirlis et al. [29]. In the proposed method, the upper and the lower bounds for the DMUs' efficiencies were computed while having missed observations. In fact, they proposed a new method based on interval DEA in which the units were evaluated with missing values along with the other units having available crisp data. They replaced missing values with approximations in the form of intervals such that the unknown missing values were likely to belong to the intervals. The bounds could be achieved by using statistical or experimential techniques. Consequently, they achieved upper and lower bounds for the efficiency score of each DMU. Toloo et al. [31] proposed an imprecise DEA model to measure the overall profit efficiency of DMU while the input and output values varied over certain ranges. This model calculated the upper and lower bounds of the overall profit efficiency for each DMU. Then, the DMUs were classified into three groups with respect to their efficiency bounds. Jahanshahloo et al. [18] modified interval generalized DEA (IGDEA) model to treat the above-mentioned basic DEA models with interval data. Park [21] applied duality theory to investigate the relationship between the primal and dual models in IDEA. Emrouznejad et al. [14] proposed two novel approaches based on the traditional profit Malmquist productivity index to measure the overall profit Malmquist productivity index when the inputs, outputs, and price vectors were fuzzy or varied in intervals. Emrouznejad et al. [15] also presented two IDEA models including general non-parametric corporate performance model and multiplicative non-parametric corporate performance with interval data.

An alternative approach proposed to address the data uncertainty is robust optimization. In this approach, the nature of data is assumed to be bounded, not necessarily stochastic. Indeed, robust optimization constructs a model solution that is optimal for any realization of uncertainty in a given set. Soyster [30] investigated explicit approaches to robust optimization and proposed a linear optimization model to obtain a solution that was feasible for all data belonging to a convex set. To generate robust optimization models, some alternative approaches were proposed by Ben-Tal and Nemirovski ([3], [4], [5]), El-Ghaoui et al. [13], and El-Ghaoui and Lebret [12]. Sadjadi and Omrani [22] proposed a robust DEA model assuming uncertainty for output parameters. Sadjadi and Omrani [23] applied the bootstrap techniques to present a robust DEA model with an application in telecommunication. On the basis of a robust optimization model, Shokouhi et al. [27] proposed a robust data envelopment analysis (RDEA) model in which the input and output parameters varied only in some ranges. Wang and Wei [33] developed four different DEA models for CRS technologies based on robust optimization techniques including various discrete combinations of precise and imprecise sub-datasets. Sadjadi et al. [24] proposed a super-efficiency DEA model by utilizing the robust optimization approach of Ben-Tal and Nemirovski [5]. Sadjadi et al. [25] proposed an imprecise interactive DEA to identify the input and output targets. Shokouhi et al. [28] proposed a modified RDEA (MRDEA) model to prevent the problem of incommensurability in the Despotis and Smirlis [11] formulation. The model applied a robust optimization approach to produce an empirical distribution for the interval efficiency where the parameters values were smooth at their extreme values.

Here, a new model for measuring the efficiency of DMUs, when the inputs and outputs vary in an interval, is proposed. The DMUs efficiencies are evaluated using the same production possibility set (PPS). In addition, conservative levels for the inputs and outputs are defined in advance and the inputs and outputs are controlled by the assigned levels. The proposed model is integrated to evaluate all DMUs simultaneously and maximize the sum of the DMUs efficiencies concurrently.

In Section 2, preliminary models for measuring the efficiency scores of DMUs are represented. The integrated model with data uncertainty is introduced in Section 3. The proposed non-linear model for measuring the efficiency score of DMUs with uncertain data, is presented in Section 4. Section 5 consists of two numerical examples. Finally, the discussions and conclusions are provided in Section 6.

2. Preliminaries

Throughout our work, measuring efficiency scores of DMUs for the CCR and integrated models without uncertainty are presented.

Suppose that there are n DMUs to be evaluated, indexed by $j \in \{1, \dots, n\}$, and each DMU is assumed to produce s outputs from m inputs. So, in DEA, each observed DMU is represented by the pair of non-negative input and output vectors $(x_j, y_j) \in R_{+}^{m+s}, j = 1, \dots, n.$ The technology T or production possibility set (PPS) is defined by:

$$T = \{(x, y) | x \text{ can produce } y\}. \quad (1)$$

Since a benchmark technology is constructed by the observed inputs and outputs of the DMUs, the following general assumptions about production technology without specifying any functional form are made. T satisfies the following standard axioms of production. Thus, the PPS of CRS model due to Charnes et al. [6] is the minimal set that satisfies the following axioms:

(A_1) Feasibility of observed data: $(x_j, y_j) \in T$, for $j = 1, \dots, n.$

(A_2) Free disposability: $(x, y) \in T,\ y \ge \bar{y} \ge 0,\ x \le \bar{x} \Rightarrow (\bar{x}, \bar{y}) \in T.$

(A_3) Constant returns to scale: $(x, y) \in T \Rightarrow (\lambda x, \lambda y) \in T,\ \forall \lambda \in R.$

(A_4) Convexity:

$$(x, y), (\bar{x}, \bar{y}) \in T,\ (\tilde{x}, \tilde{y}) = \lambda(x, y) + (1-\lambda)(\bar{x}, \bar{y}),\ 0 \le \lambda \le 1 \Rightarrow (\tilde{x}, \tilde{y}) \in T.$$

Under the axioms (A_1) to (A_4), the minimal PPS for T can be stated as:

$$T_c = \{(x, y) \mid x \ge \sum_{j=1}^{n} \lambda_j x_j,\ y \le \sum_{j=1}^{n} \lambda_j y_j,\ \lambda_j \ge 0,\ j = 1, \dots, n\}. \quad (2)$$

Based on the relation in (2), the envelopment form for the input-oriented model measuring the efficiency of a DMU is defined to be

$$\begin{aligned} \theta^{*} = \min \quad & \theta_p \\ \text{s.t.} \quad & \sum_{j=1}^{n} \lambda_j x_j \le \theta_p x_p, \\ & \sum_{j=1}^{n} \lambda_j y_j \ge y_p, \\ & \lambda_j \ge 0, \quad \forall j. \end{aligned} \tag{3}$$

The dual of (3) is expressed as follows:

$$\begin{aligned} \max \quad & \theta_p = \sum_{r=1}^{s} u_r y_{rp} \\ \text{s.t.} \quad & \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \ \ \forall j, && (4a) \\ & \sum_{i=1}^{m} v_i x_{ip} = 1, && (4b) \\ & u_r \ge 0, \ v_i \ge 0, \qquad \forall r, i, && (4c) \end{aligned} \tag{4}$$

where the v_i and the u_r are the multipliers (weights) respectively assigned to the ith input and the rth output. Model (4) maximizes the efficiency score of DMU_p. With respect to constraints, the optimal value of the objective function in (4) will never exceed 1. Note that, the constraints (4a) guarantee the existence of DMUs in PPS, constraint (4b) is known as a normalization constraint, and constraints (4c) impose non-negativity on the weights.

To present the efficient and inefficient DMUs for the model (4), the following definition are needed.

Definition 2.1. DMU_p is efficient if and only if $\sum_{r=1}^{s} u_r^{*} y_{rp} = 1$ and there exists at least one optimal point (u^{*}, v^{*}) for (4) with $u^* > 0$ and $v^* > 0$.

The constraints (4c) may be converted to $u_r \ge \varepsilon, \ v_i \ge \varepsilon,$ for $r = 1,\dots,s$ and $i = 1,\dots,m$, where ε is the non-Archimedean infinitesimal value; see Cooper et al. [10] for a foundational development of this transformation and an interpretation of ε.

In order to evaluate the n DMUs' effiencies and also to compute the projection of them, the model (4) must be solved n times. We now propose the following integrated model which independently evaluates all DMUs and gives the projection of DMUs simultaneously by solving only one LP model:

$$
\begin{aligned}
\max \quad & \sum_{j=1}^{n}\theta_j = \sum_{j=1}^{n}\sum_{r=1}^{s} u_{rj} y_{rj} \\
\text{s.t.} \quad & \sum_{r=1}^{s} u_{rk} y_{rj} - \sum_{i=1}^{m} v_{ik} x_{ij} \le 0, && \forall j,k, \\
& \sum_{i=1}^{m} v_{ij} x_{ij} = 1, && \forall j, \\
& u_{rj} \ge 0,\ v_{ij} \ge 0, && \forall r,i,j.
\end{aligned}
\tag{5}
$$

Note that both models (4) and (5) produce the same results.

3. Integrated Model with Interval Data

Assume that there are n DMUs with interval inputs and outputs as $\tilde{x}_{ij} \in [x_{ij}^L, x_{ij}^U]$ and $\tilde{y}_{rj} \in [y_{rj}^L, y_{rj}^U]$, for $j = 1,\dots,n.$ In actual applications, we have some reasonable estimates for the mean of the inputs and outputs, say x_{ij} and y_{rj}, and their deviations, d_{ij}^x and d_{rj}^y, respectively. Indeed, the inputs $\tilde{x}_{ij}$ and the outputs $\tilde{y}_{rj}$ are independent, symmetric, and bounded random variables having unknown distributions with values in the intervals $[x_{ij} - d_{ij}^x, x_{ij} + d_{ij}^x]$ and $[y_{rj} - d_{rj}^y, y_{rj} + d_{rj}^y]$, respectively. Note that when d_{ij}^x and d_{rj}^y are allowed to be zero, then the x_{ij} and y_{rj} are called the nominal values of $\tilde{x}_{ij}$ and $\tilde{y}_{rj}$, respectively.

The following integrated model measures the efficiency of DMUs with the interval inputs and outputs:

$$
\begin{aligned}
\max \quad & \sum_{j=1}^{n}\rho_j = \sum_{j=1}^{n}\sum_{r=1}^{s} u_{rj} \tilde{y}_{rj} \\
\text{s.t.} \quad & \sum_{r=1}^{s} u_{rk} \tilde{y}_{rj} - \sum_{i=1}^{m} v_{ik} \tilde{x}_{ij} \le 0, && \forall j,k, \\
& \sum_{i=1}^{m} v_{ij} \tilde{x}_{ij} = 1, && \forall j, \\
& u_{rj} \ge 0,\ v_{ij} \ge 0, && \forall r,i,j,
\end{aligned}
\tag{6}
$$

where ρ_j is the efficiency of the jth DMU.

Since the inputs and outputs of (6) vary within intervals, the efficiency scores of the DMUs are not easily computed. Despotis and Smirlis [11] proposed two models to overcome this difficulty with interval data in DEA. Their models find the upper and lower bounds of the efficiency scores for DMUs as the optimistic and the pessimistic cases, respectively. The same idea may be applied to modify the model (6) into models (7) and (8) as follows:

$$
\begin{aligned}
\max \quad & \sum_{j=1}^{n} \rho_j^U = \sum_{j=1}^{n}\sum_{r=1}^{s} u_{rj} y_{rj}^U \\
\text{s.t.} \quad & \sum_{r=1}^{s} u_{rj} y_{rj}^U - \sum_{i=1}^{m} v_{ij} x_{ij}^L \le 0, && \forall j, \\
& \sum_{r=1}^{s} u_{rk} y_{rj}^L - \sum_{i=1}^{m} v_{ik} x_{ij}^U \le 0, && \forall j, k, k \ne j, \\
& \sum_{i=1}^{m} v_{ij} x_{ij}^L = 1, && \forall j, \\
& u_{rj} \ge 0,\ v_{ij} \ge 0, && \forall r, i, j,
\end{aligned}
\tag{7}
$$

$$
\begin{aligned}
\max \quad & \sum_{j=1}^{n} \rho_j^L = \sum_{j=1}^{n}\sum_{r=1}^{s} u_{rj} y_{rj}^L \\
\text{s.t.} \quad & \sum_{r=1}^{s} u_{rj} y_{rj}^L - \sum_{i=1}^{m} v_{ij} x_{ij}^U \le 0, && \forall j, \\
& \sum_{r=1}^{s} u_{rk} y_{rj}^U - \sum_{i=1}^{m} v_{ik} x_{ij}^L \le 0, && \forall j, k, k \ne j, \\
& \sum_{i=1}^{m} v_{ij} x_{ij}^U = 1, && \forall j, \\
& u_{rj} \ge 0,\ v_{ij} \ge 0, && \forall r, i, j.
\end{aligned}
\tag{8}
$$

One may realize that (7) and (8) are the optimistic and the pessimistic models and ρ_j^U and ρ_j^L are the maximum and minimum efficiencies of the jth DMU, respectively. Despotis and Smirlis [11] proved the following theorem to show that the efficiency scores lie within the upper and lower bounds.

Theorem 3.1. Let the optimal solutions of (6), (7), and (8) be (u_{rj}^*, v_{ij}^*), $(u_{rj}^{**}, v_{ij}^{**})$, and $(u_{rj}^{***}, v_{ij}^{***})$, respectively. Then, the solutions ρ_j of (6) lie between the solutions of ρ_j^U and ρ_j^L of the models (7) and (8). Thus, $\rho_j^L \le \rho_j \le \rho_j^U$, for $j = 1, \ldots, n$.

Proof: See Despotis and Smirlis [11].

4. Conservative Levels in Integrated Efficiency with Interval Inputs and Outputs

Here, a new model to measure the efficiencies of DMUs with interval inputs and outputs is formulated. Let us introduce the quantities γ_j^x and γ_j^y, for $j = 1, \ldots, n$, with values respectively in the intervals $[0, J_j^x]$ and $[0, J_j^y]$, where J_j^x and J_j^y are the number of uncertain inputs and outputs for DMU_j. Let us also name the quantities γ_j^x and γ_j^y as the conservative levels for inputs and outputs. So, a model being controlled by the levels of conservatism of the inputs and outputs is proposed to evaluate the efficiencies of the DMUs. Apparently, the proposed model is non-linear and assumes the same PPS for all DMUs. The advantage of the model is its capability in evaluating DMUs' efficiencies simultaneously. One must note that although the efficiencies of DMUs are not estimated independently, the projection is obtained concurrently. The proposed model is

$$
\begin{array}{lllr}
\max & \sum_{j=1}^{n} \eta_j = \sum_{j=1}^{n}\sum_{r=1}^{s} u_{rj} y_{rj} & & \\
\text{s.t.} & \sum_{i=1}^{m} v_{ij} x_{ij} = 1, & \forall j, & (9a) \\
& \sum_{r=1}^{s} u_{rj} y_{rk} - \sum_{i=1}^{m} v_{ij} x_{ik} \le 0, & \forall j,k, & (9b) \\
& y_{rj}^{U} - y_{rj} - z_{rj}^{y} \gamma_j^{y} - p_{rj} \ge 0, & \forall r,j, & (9c) \\
& y_{rj}^{L} - y_{rj} \le 0, & \forall r,j, & (9d) \\
& x_{ij}^{L} - x_{ij} + z_{ij}^{x} \gamma_j^{x} + q_{ij} \le 0, & \forall i,j, & (9e) \\
& x_{ij}^{U} - x_{ij} \ge 0, & \forall i,j, & (9f) \\
& z_j^{y} = \sum_{r=1}^{s} z_{rj}^{y}, & \forall j, & (9g) \\
& z_j^{x} = \sum_{i=1}^{m} z_{ij}^{x}, & \forall j, & (9h) \\
& z_j^{y} + p_{rj} \ge y_{rj}^{U} - y_{rj}^{L}, & \forall r,j, & (9i) \\
& z_j^{x} + q_{ij} \ge x_{ij}^{U} - x_{ij}^{L}, & \forall r,j, & (9l) \\
& x_{ij}, y_{rj}, z_{ij}^{x}, z_{rj}^{y}, z_j^{x}, z_j^{y}, q_{ij}, p_{rj} \ge 0, & \forall i,r,j, & (9t) \\
& v_{ij}, u_{rj} \ge \varepsilon, & \forall i,r,j, & (9w)
\end{array}
\qquad (9)
$$

where η_j is the effeciency and u_{rj} and v_{ij} are the weights assigned to the *r*th output and the *i*th input of the jth DMU, respectively.

The model (9), being called an integrated model, maximizes the sum of the efficiency scores of all DMUs. Since the efficiency score for each DMU lies in the intervals [0,1], then the optimal value of the objective function of (9) varies between 0 and *n*. In (9), the constraints (9*a*) are normalization constraints and the constraints (9*b*) guarantee that the DMUs are all in PPS. According to the constraints (9*c*) and (9*d*), it is obvious that $y_{rj}^L \le y_{rj} \le y_{rj}^U - (z_{rj}^y \gamma_j^y + p_{rj})$. Under the optimistic conditions, if $\gamma_j^y = 0$ then $y_{rj}^L \le y_{rj} \le y_{rj}^U - p_{rj}$. For $p_{rj} = 0$, we have $y_{rj}^L \le y_{rj} \le y_{rj}^U$. Investigation of constraints (9*i*) and (9*d*) reveals that for pessimistic conditions, when $\gamma_j^y = J_j^y$, the value of $z_{rj}^y \gamma_j^y + p_{rj}$ increases so that the $y_{rj}^U - (z_{rj}^y \gamma_j^y + p_{rj})$ equals y_{rj}^L. For $0 < \gamma_j^y < J_j^y$, (9) determines the y_{rj}^U as the maximum value of y_{rj}, considering the constraints (9*g*) and (9*i*). If there exists a t such that $z_{tj}^y = p_{tj} = 0$ and also the constraints (9*i*) are satisfied, then $z_j^y \ne 0$. Thus, in order to satisfy the constraints (9*g*), there exists the *l*th output of DMU_j such that $z_{lj}^y \ne 0$. Therefore, $z_{lj}^y \gamma_j^y + p_{lj} \ne 0$, and by constraints (9*c*), $y_{lj} < y_{lj}^U$. Note that by this the non-zero value of the parameter γ_j^y is imposed to the *l*th output of DMU_j.

Constraints (9*e*), (9*f*), (9*h*), and (9*l*) have interpretations respectively similar to constraints (9*c*), (9*d*), (9*g*), and (9*i*). Constraints (9*t*) impose non-negativity on the variables. Also, constraints (9*w*) show the lower bounds for the weights. Feasibility of model (9) is shown in Appendix A.

For more clarity, Figure 1 shows four DMUs with one certain input and one interval output. It also displays the PPSs of the proposed model and the Despotis and Smirlis's [11] models for DMU_1 in optimistic and pessimistic cases. Note that the model has the same PPS for all DMUs.

Theorem 4.1. Let us $(\bar{u}_{rj}, \bar{v}_{ij})$, $(\hat{u}_{rj}, \hat{v}_{ij})$ and $(u_{rj}^*, v_{ij}^*, z_j^{x*}, z_j^{y*}, z_{ij}^{x*}, z_{rj}^{x*}, q_{ij}^*, p_{rj}^*, x_{ij}^*, y_{rj}^*)$ be optimal solutions of the models (7), (8), and (9), respectively. Then, η_j, the efficiency score of the jth DMU, satisfies $\rho_j^L \le \eta_j \le \rho_j^U$.

Proof: Since $(u_{rj}^*, v_{ij}^*, z_j^{x*}, z_j^{y*}, z_{ij}^{x*}, z_{rj}^{x*}, q_{ij}^*, p_{rj}^*, x^*, y^*)$ is an optimal solution for (9), we can define the followings

$$\beta_j = \sum_{i=1}^{m} v_{ij}^* x_{ij}^L > 0,$$

$$\hat{u}_{rj} = \frac{u_{rj}^*}{\beta_j}, \ \ \forall r, j,$$

$$\hat{v}_{ij} = \frac{v_{ij}^*}{\beta_j}, \ \ \forall i, j.$$

By investigation of (9), we get

$$\beta_j = \sum_{i=1}^{m} v_{ij}^* x_{ij}^L \leq \sum_{i=1}^{m} v_{ij}^* x_{ij}^* = 1,$$

$$\sum_{i=1}^{m} \hat{v}_{ij} x_{ij}^L = \sum_{i=1}^{m} \frac{v_{ij}^*}{\beta_j} x_{ij}^L = \frac{1}{\beta_j} \sum_{i=1}^{m} v_{ij}^* x_{ij}^L = 1,$$

$$\sum_{r=1}^{s} \hat{u}_{rj} y_{rk}^L - \sum_{i=1}^{m} \hat{v}_{ij} x_{ik}^U = \frac{1}{\beta_j} (\sum_{r=1}^{s} u_{rj}^* y_{rk}^L - \sum_{i=1}^{m} v_{ij}^* x_{ik}^U)$$

$$\leq \frac{1}{\beta_j} (\sum_{r=1}^{s} u_{rj}^* y_{rk}^* - \sum_{i=1}^{m} v_{ij}^* x_{ik}^*) \leq 0, \ \forall j, k, k \neq j.$$

In addition, we have:

$$\sum_{r=1}^{s} \hat{u}_{rj} y_{rj}^U \leq 1 \quad \Rightarrow \quad \frac{1}{\beta_j} \sum_{r=1}^{s} u_{rj}^* y_{rj}^U \leq 1.$$

Referring to $\frac{1}{\beta_j} \sum_{i=1}^{m} v_{ij}^* x_{ij}^L = 1$, the following inequalities hold for all DMUs:

$$\frac{1}{\beta_j} (\sum_{r=1}^{s} u_{rj}^* y_{rj}^U - \sum_{i=1}^{m} v_{ij}^* x_{ij}^L) \leq 0,$$

$$\hat{u}_{rj} = \frac{u_{rj}^*}{\beta_j} \geq \frac{\varepsilon}{\beta_j} \geq 0, \ \forall r, j,$$

$$\hat{v}_{ij} = \frac{v_{ij}^*}{\beta_j} \geq \frac{\varepsilon}{\beta_j} \geq 0, \ \forall i, j.$$

It is obvious that $(u_{rj}^*, v_{ij}^*, z_j^{x*}, z_j^{y*}, z_{ij}^{x*}, z_{rj}^{x*}, q_{ij}^*, p_{rj}^*, x^*, y^*)$ is a feasible solution for (7). Thus, we have $\eta_j \leq \rho_j^U$ for all J_j^x and J_j^y.

Similarly, we can show that $\rho_j^L \leq \eta_j$ for J_j^x and J_j^y. Then, $\rho_j^L \leq \eta_j \leq \rho_j^U$.

Corollary 4.2. If DMU_j evaluated by (9) is efficient, then it is efficient for (7).

Next, the following lemma may be concluded from the foregoing theorems.

Lemma 4.3. If $(u_{rj}^*, v_{ij}^*, z_j^{x*}, z_j^{y*}, z_{ij}^{x*}, z_{rj}^{x*}, q_{ij}^*, p_{rj}^*, x_{ij}^*, y_{rj}^*)$ is the optimal solution of (9), then (u_{rj}^*, v_{ij}^*) is an optimal solution of (6) provided that $\tilde{x}_{ij} = x_{ij}^*$ and $\tilde{y}_{rj} = y_{rj}^*$.

Proof: We need to show that $\rho_j^* = \eta_j^*$.

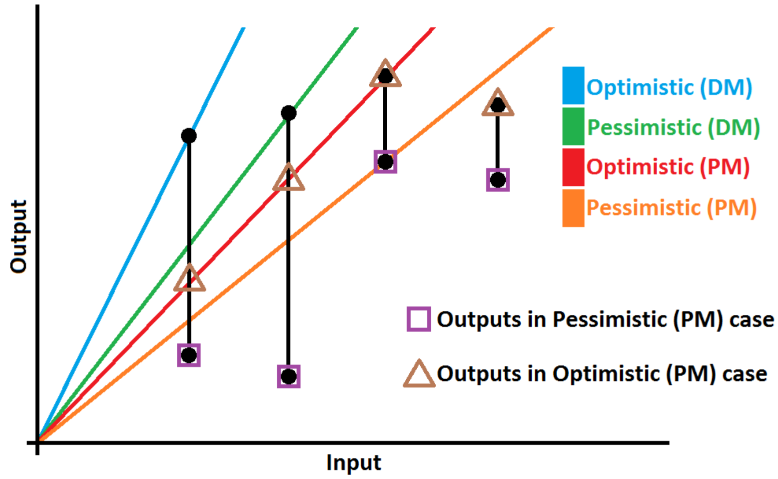


Figure 1. Comparison of the PPS of the model (9) and Despotis and Smirlis's approach for DMU_1 in the optimistic and pessimistic cases

(i) Since $\tilde{x}_{ij} = x^*_{ij}$ and $\tilde{y}_{rj} = y^*_{rj}$, it is easy to show that (u^*_{rj}, v^*_{ij}) is a feasible solution for (6). Thus, $\rho^*_j \leq \eta^*_j$.

(ii) Assume that $(\bar{u}_{rj}, \bar{v}_{ij})$ is an optimal solution for (6), $\tilde{x}_{ij} = x^*_{ij}$ and $\tilde{y}_{rj} = y^*_{rj}$, where $y^L_{rj} \leq y^*_{rj} \leq y^U_{rj} - (z^y_{rj}\gamma^y_j + p_{rj})$ and $x^L_{ij} + z^x_{ij}\gamma^x_j + q_{ij} \leq x^*_{ij} \leq x^U_{ij}$. So, there exist at least one y_{rj} and one x_{ij} such that the constraints (9*a*) and (9*b*) are satisfied. Hence, $\rho^*_j \geq \eta^*_j$.

From (i) and (ii), we conclude that $\rho^*_j = \eta^*_j$. Therefore, (u^*_{rj}, v^*_{ij}) is an optimal solution of (6).

Theorem 4.4. Let an optimal solution for Model (9) be $(u^*_{rj}, v^*_{ij}, z^{x*}_j, z^{y*}_j, z^{x*}_{ij}, z^{x*}_{rj}, q^*_{ij}, p^*_{rj}, x^*_{ij}, y^*_{rj})$. Then, we have: $\forall\, j \in \{1,...,n\},\ \exists l \in \{1,...,n\}$ such that $\sum_{r=1}^{s} u_{rj} y_{rl} - \sum_{i=1}^{m} v_{ij} x_{il} = 0$.

Proof: It is evident that, in Model (6), $\forall\, j \in \{1,...,n\},\ \exists l \in \{1,...,n\}$ such that $\sum_{r=1}^{s} u_{rj} y_{rl} - \sum_{i=1}^{m} v_{ij} x_{il} = 0$. So, by Lemma 4.3 the result follows.

Theorem 4.5. DMU_j is efficient if and only if $\eta_j = 1$.

Proof: Assume that $(u^*_{rj}, v^*_{ij}, z^{x*}_j, z^{y*}_j, z^{x*}_{ij}, z^{x*}_{rj}, q^*_{ij}, p^*_{rj}, x^*_{ij}, y^*_{rj})$ is the optimal solution for (9). (Only if part): By Theorem 4.5 and setting $l = j$, we have $\sum_{r=1}^{s} u^*_{rj} y_{rj} - \sum_{i=1}^{m} v^*_{ij} x_{ij} = 0$. Since $\sum_{i=1}^{m} v^*_{ij} x_{ij} = 1$, $\sum_{r=1}^{s} u^*_{rj} y_{rj} = 1$ and $\eta_j = 1$.

(If part): $\eta_j = 1 \Rightarrow \sum_{r=1}^{s} u_{rj}^{*} y_{rj} = 1 \Rightarrow \sum_{r=1}^{s} u_{rj}^{*} y_{rj} = \sum_{i=1}^{m} v_{ij}^{*} x_{ij} \Rightarrow \sum_{r=1}^{s} u_{rj}^{*} y_{rj} - \sum_{i=1}^{m} v_{ij}^{*} x_{ij} = 0$. In other words, DMU_j is efficient.

Remark 4.6. Let us define $\left(\underline{\hat{x}}_p^{*}, \underline{\hat{y}}_p^{*}\right)$ as the projection of DMU_p on the frontier. Then, $\left(\underline{\hat{x}}_p^{*}, \underline{\hat{y}}_p^{*}\right)$ which is an improved activity for any inefficient DMU_p, would be efficient when evaluated by Model (9). Applying $\left(\underline{\hat{x}}_p^{*}, \underline{\hat{y}}_p^{*}\right)$ in constraint (9*b*), we get

$$\begin{aligned} &\sum_{r=1}^{s} u_{rp} \hat{y}_{rp}^{*} - \sum_{i=1}^{m} v_{ip} \hat{x}_{ip}^{*} = \sum_{r=1}^{s} u_{rp} (y_{rp}^{*}) - \sum_{i=1}^{m} v_{ip} (\eta_p x_{ip}^{*}) \\ &= \sum_{r=1}^{s} u_{rp} y_{rp}^{*} - \eta_p \sum_{i=1}^{m} v_{ip} x_{ip}^{*}, \end{aligned} \tag{10}$$

$$\sum_{i=1}^{m} v_{ij} x_{ip}^{*} = 1.$$

The above relations imply $\sum_{r=1}^{s} u_{rj} y_{rp}^{*} - \eta_p = 0$.

Considering the efficiency score of any DMU which depends on the conservative level for the input and output parameters, the mean value of all efficiency scores for fixed $\gamma_j^x = \gamma^x$ and $\gamma_j^y = \gamma^y$ is given by $\theta_j(\Gamma)$, where $\Gamma = \gamma^x + \gamma^y$. Consequently, all DMUs may be divided into one of the following three classes.

Class 1: The DMUs which are efficient, for all γ_j^x and γ_j^y, that is, $E^{++} = \{ j \mid \forall \Gamma, \theta_j(\Gamma) = 1 \}$.

Class 2: The DMUs, which are efficient for some γ_j^x and γ_j^y, that is, $E^{+} = \{ j \mid \exists \bar{\Gamma}; \forall \Gamma \in [0, \bar{\Gamma}], \theta_j(\Gamma) = 1 \text{ and } \theta_j(\Gamma > \bar{\Gamma}) < 1 \}$.

Class 3: The DMUs, which are inefficient for $\gamma_j^x = \gamma_j^y = 0$, that is, $E^{-} = \{ j \mid \forall \Gamma = 0, \theta_j(\Gamma) < 1 \}$.

It is clear that the DMUs with the highest performances belong to the class E^{++} and those with least performances belong to the class E^{-}. Besides, the DMUs in class E^{+} have DMU performances in between the ones corresponding to classes E^{++} and E^{-}.

5. Numerical Examples

Here, two numerical examples are provided to show the appropriateness of the proposed model.

Example 5.1. Consider 5 DMUs each with only one interval input and one interval output. The inputs and outputs for all DMUs are the same as shown in Table 1 and Figure 2.

Table 1: Inputs and outputs of 5 DMUs

DMU_j	*X*	*Y*
1	[0.75,1.05]	[9.00,10.00]
2	[1.25,1.50]	[1.00,11.30]
3	[1.20,1.75]	[6.00,7.50]
4	[1.10,1.60]	[2.00,9.00]
5	[0.80, 1.40]	[3.00,8.00]

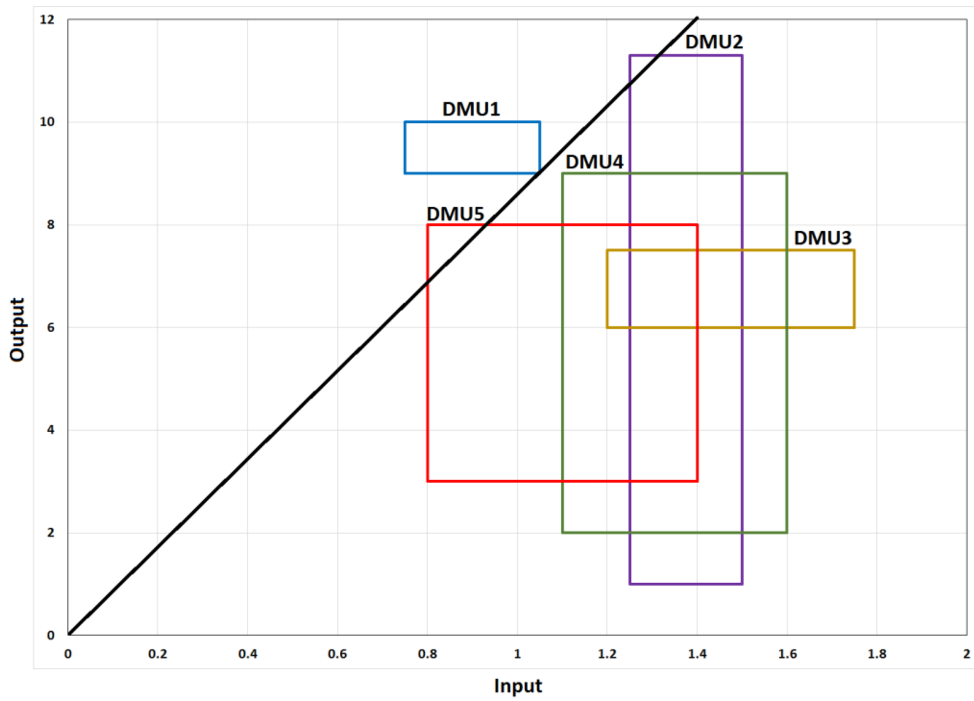
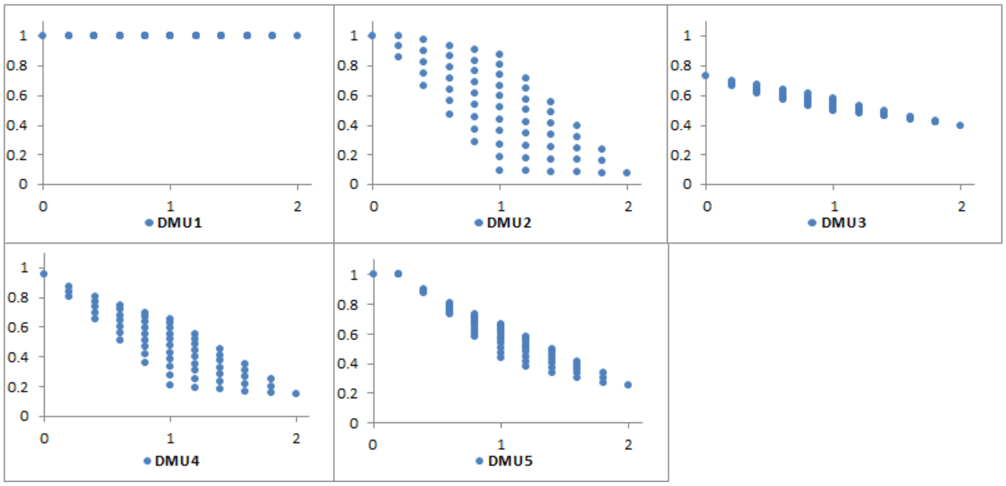


Figure 2. Inputs and outputs of 5 DMUs

Model (9) is run for different combinations of $\gamma_j^x = \gamma^x$ and $\gamma_j^y = \gamma^y$ for $j = 1, 2, ..., 5$, and a fixed $\Gamma = \gamma^x + \gamma^y$. The GAMS software package with $\varepsilon = 10^{-6}$ was used. In each case, the efficiency scores of the 5 DMUs were obtained as displayed in Figure 3 for all possible γ^x and γ^y such that $\Gamma = \gamma^x + \gamma^y$. For $\gamma^x = \gamma^y = 0$ and $C\ (\gamma^x = \gamma^y = 0)$, the global optimistic and pessimistic cases of the sum of efficiency scores for all DMUs are also shown in Figure 3.

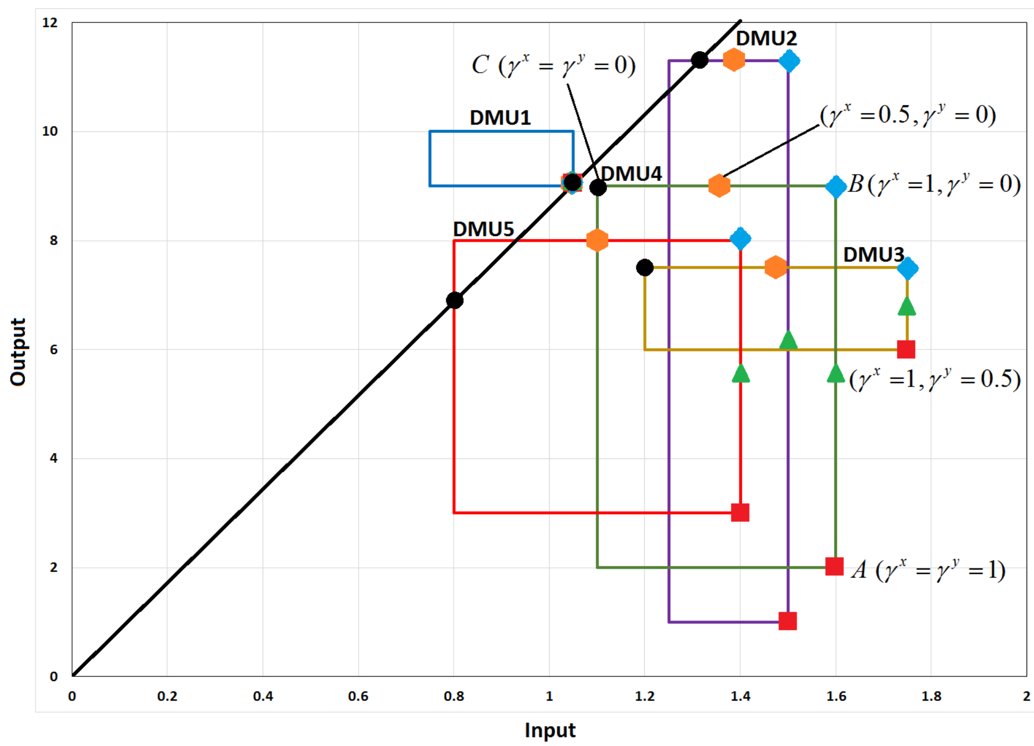
In Figure 4, the point A shows the pessimistic case for DMU$_4$, when $\gamma^x = \gamma^y = 1$. For $\gamma^x = 1$ and $\gamma^y = 0$, B is generated. For $\gamma^x = 1$ and $\gamma^y = 0.5$, the point falls on the line segment AB. For $\gamma^x = \gamma^y = 0$, the point C showing the optimistic case is obtained. Finally, for $\gamma^x = 0.5$ and $\gamma^y = 0$, the point falls on the line segment BC.

Note that $E^{++} = \{DMU_1\}$, $E^{+} = \{DMU_2, DMU_5\}$, and $E^{-} = \{DMU_3, DMU_4\}$. Hence, DMU_3 and DMU_4 are inefficient DMUs, for all Γs, DMU_2 is efficient, if $\Gamma \le 0.025$ and DMU_5 is efficient, if $\Gamma \le 0.1$.



(Axiom x and y show Γ and efficiency scores, respectively.)

Figure 3. The efficiency scores of 5 DMUs



(Axiom x and y show Γ and efficiency scores, respectively.)

Figure 4. Changing γ^x and γ^y of DMUs from pessimistic case to optimistic case

Example 5.2. The data set correspond to 24 branches of Bank Mellat in Iran. Each branch uses two interval inputs, the number of the staff and departments, to produce five interval outputs including long-term saving, short-term saving, saving account, Gharzol Hasaaneh savings account and Facility. The inputs and outputs are given in tables 2 and 3, respectively.

Table 2. The input data for 24 bank branches of Bank Mellat

DMU	INPUTS			
	Input 1		Input 2	
	L	U	L	U
1	5.31	17.71	15858	92482
2	11.45	15.48	7438	31538.52
3	12.05	18.41	10386.79	72224.6
4	13.57	22.02	8464.5	137725.5
5	10.13	15.76	18029.33	35331.83
6	9.82	14.92	5276.28	22054.14
7	9.79	14.98	4982.16	19385.33
8	21.26	34.9	82756	194775
9	11.37	17.03	2127	5143
10	15.18	20.3	1850	6852
11	15.74	20.02	7302	166435.5
12	13.43	17.47	21724	25584
13	11.53	15.6	8749.71	31709
14	9.32	13.7	16421	24383
15	14.53	21.38	12168	35371
16	13.23	18.07	3912	20405
17	14.84	20.22	9308.33	39126.33
18	15.99	22.62	6332	8105
19	9.86	16.07	167	14433
20	5.68	9.27	2705	19713
21	14.31	23.4	1366	7676
22	16.72	23.13	40907	64700.66
23	17.72	24.96	6837	42869
24	16.36	23.47	2	513

Model (9) is run for different combinations of γ_j^x and γ_j^y, and a fixed Γ with $\varepsilon = 10^{-8}$. In each case, the efficiency scores are obtained for the 24 DMUs. The efficiency scores of the 24 DMUs are displayed in Figure 5 for all possible γ_j^x and γ_j^y. In addition, all DMUs are classified into the following three classes:

$$\begin{cases} E^{++} = \{DMU_1, DMU_5, DMU_6, DMU_9, DMU_{10}, DMU_{12}, DMU_{15}, DMU_{24}\}. \\ \\ E^{+} = \{DMU_2, DMU_3, DMU_4, DMU_7, DMU_8, DMU_{11}, DMU_{13}, DMU_{14}, \\ \qquad DMU_{16}, DMU_{17}, DMU_{18}, DMU_{19}, DMU_{20}, DMU_{21}, DMU_{22}, DMU_{23}\}. \\ \\ E^{-} = \Phi. \end{cases}$$

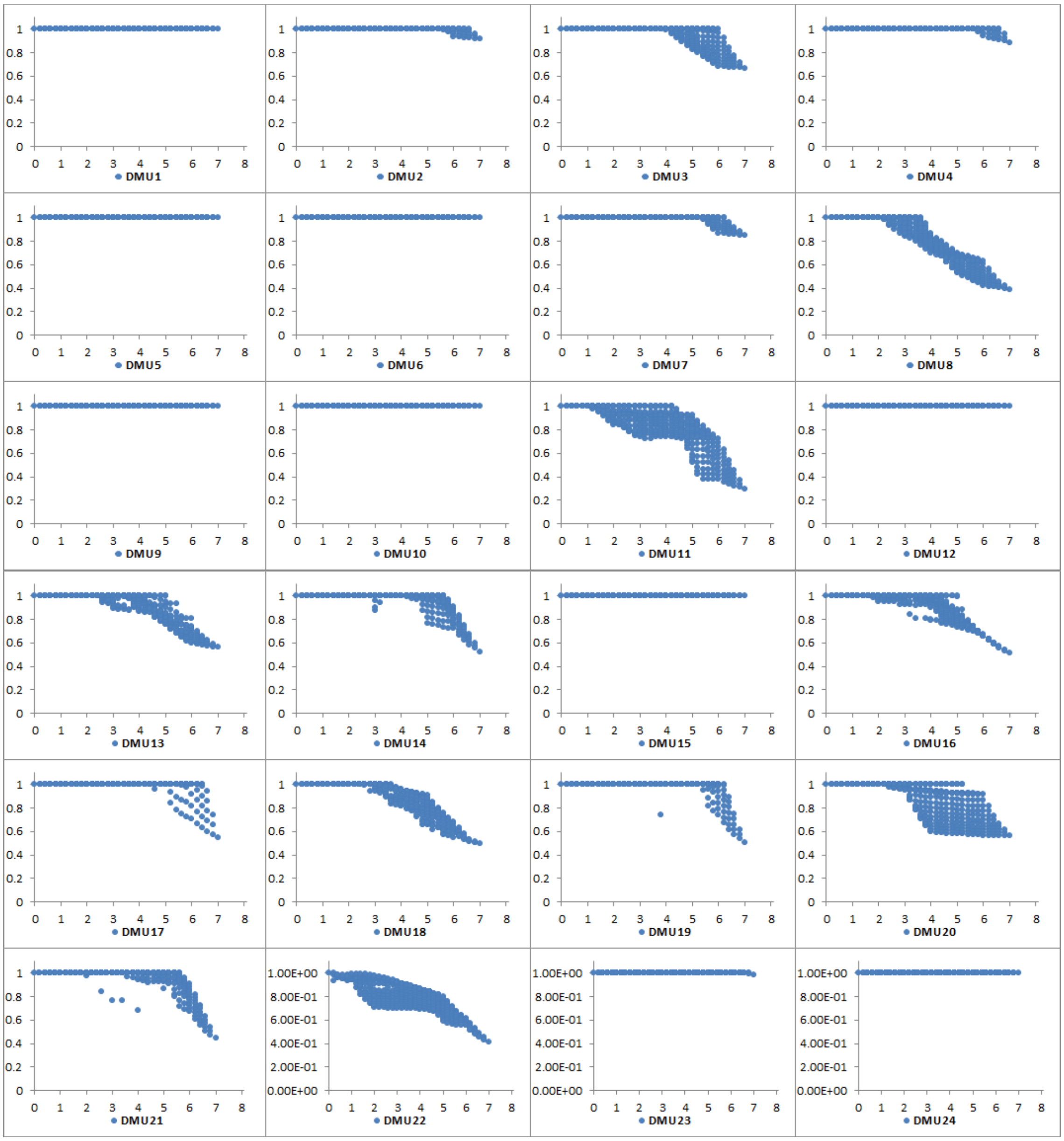
The DMUs belonging to the class E^{++} are efficient, for all Γs, and those in the class E^{+} are efficient, only for some Γs.

Table 3. The output data for 24 bank branches

DMU	OUTPUTS									
	Output 1		Output 2		Output 3		Output 4		Output 5	
	L	U	L	U	L	U	L	U	L	U
1	47439	72223	41786	186285	7056	67976	93428	325545	125042	463205
2	63955	80954	72961	152987	1817	3941	27982	48446	134204	252779
3	46796	73575	72033	138149	1084	2877	20262	41541	202161	308616
4	32408	69519	95560	384179	2862	13186	16132	96651	651237	1024448
5	129268	203921	71866	116872	3643	7485	24399	82857	570838	1196066
6	51881	87640	68286	89866	3356	7184	38329	59831	68824	107925
7	19195	51083	81111	180587	1326	3162	11324	55757	63335	114851
8	11899	86263	82933	206712	1783	13161	11528	42622	301321	713080
9	22206	38956	18192	29999	5208	8166	26076	39227	10475	17347
10	73291	135821	55705	103384	4598	42317	32215	74637	93208	229520
11	22687	39402	18887	39018	2076	5794	16817	24061	71279	277331
12	14436	29255	36250	56242	5916	7859	18054	36830	23953	29366
13	37300	59139	38230	71756	1540	3963	23124	38194	78984	121222
14	32074	44262	35889	55188	1689	3873	10497	21586	73194	123711
15	2`3867	164039	171853	323883	4244	16796	9146	69360	112884	226481
16	18885	37599	34335	69025	2304	3493	14114	35739	93197	119955
17	24038	33714	57322	86462	2841	10909	14544	26299	99800	228069
18	6151	13400	32055	67708	1966	4089	9380	23230	30231	47726
19	20310	48037	32549	49351	1803	5855	11425	39469	71825	155367
20	0	1290	42001	230890	415	677	447	1456	15088	42180
21	22418	29912	26226	108310	2364	5997	15621	22784	48989	96910
22	27844	49510	45112	74292	2724	4708	21905	39690	68866	159784
23	23207	55912	106992	157711	4487	27609	19638	64623	638265	1594075
24	17499	32972	35203	56413	3757	6870	12296	39317	99312	150623

6. Conclusion

In conventional DEA, the data are assumed to be specific numerical values. However, in reality the observed values of the inputs and outputs are mostly imprecise. The impreciseness of the data in the DEA modeling is dealt with in various ways in the literature. Here, a deterministic methodology was proposed to address the problem of measuring the DMU efficiencies when the inputs and outputs were supposed to lie in intervals. The concept of conservative level for inputs and outputs was used to propose a DEA model. The efficiency scores obtained by solving the suggested model were somewhere between the optimistic and the pessimistic cases introduced by Despotis and Smirlis [11].



(Axiom *x* and *y* show Γ and efficiency scores, respectively.)

Figure 5. The efficiency scores of 24 DMUs

When the conservative level is chosen at some certain values, the optimistic and the pessimistic results are also accessible. In our proposed approach, the inputs and outputs for each DMU were taken to have fixed values and the sum of efficiencies was maximized. Therefore, the DMUs were evaluated by the same production possibility set (PPS). While Despotis and Smirlis [11] considered the optimistic and pessimistic PPS for each DMU, the proposed model evaluates all DMUs by the same PPS; moreover, all DMUs are evaluated by solving only one model; these are the advantages of the proposed model. Ranking the units with respect to efficiency scores and computing the Malmquist index, when the data is uncertain, may be considered for future research.

References

[1] Amirteimoori, A., and Kordrostami, S. (2005), Multi-component efficiency measurement with imprecise data, *Applied Mathematics and Computation*, 162(3), 1265-1277.

[2] Banker, R.D., Charnes, A. and Cooper, W.W. (1984), Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science*, 30(9), 1078-1092.

[3] Ben-Tal, A. and Nemirovski, A. (1998), Robust convex optimization, *Mathematics of Operations Research*, 23(4), 769-805.

[4] Ben-Tal, A. and Nemirovski, A. (1999), Robust solutions of uncertain linear programs, *Operations Research Letters*, 25(1), 1-13.

[5] Ben-Tal, A. and Nemirovski, A. (2000), Robust solutions of linear programming problems contaminated with uncertain data, *Mathematical Programming*, 88(3), 411-424.

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

[7] Cooper, W.W., Park, K.S. and Yu, G. (1999), IDEA and AR-IDEA: models for dealing with imprecise data in DEA, *Management Science*, 45(4), 597-607.

[8] Cooper, W.W., Park, K.S. and Yu, G. (2001a), An illustrative application of IDEA (imprecise data envelopment analysis) to a Korean mobile telecommunication company. *Operations Research*, 49(6), 807-820.

[9] Cooper, W.W., Park, K.S. and Yu, G. (2001b), IDEA (imprecise data envelopment analysis) with CMDs (column maximum decision making units), *Journal of the Operational Research Society*, 52(2), 176-181.

[10] Cooper, W.W., Seiford, L.M. and Tone, K. (2007), Data Envelopment Analysis a Comprehensive Text with Models, Applications, References and DEA-Solver Software, Second Edition, Springer.

[11] Despotis, D.K. and Smirlis, Y.G. (2002), Data envelopment analysis with imprecise data, *European Journal of Operational Research*, 140(1), 24-36.

[12] El Ghaoui, L. and Lebret, H. (1997), Robust solutions to least-squares problems with uncertain data, *SIAM Journal on Matrix Analysis and Applications*, 18(4), 1035-1064.

[13] El Ghaoui, L., Oustry, F. and Lebret, H. (1998), Robust solutions to uncertain semidefinite programs, *SIAM Journal on Optimization*, 9(1), 33-52.

[14] Emrouznejad, A., Rostamy-Malkhalifeh, M., Hatami-Marbini, A., Tavana, M. and Aghayi, N. (2011), An overall profit Malmquist productivity index with fuzzy and interval data, *Mathematical and Computer Modelling*, 54(11), 2827-2838.

[15] Emrouznejad, A., Rostamy-Malkhalifeh, M., Hatami-Marbini, A. and Tavana, M. (2012), General and multiplicative non-parametric corporate performance models with interval ratio data, *Applied Mathematical Modelling*, 36(11), 5506-5514.

[16] Jahanshahloo, G.R., Hosseinzadeh Lofti, F. and Moradi, M. (2004a). Sensitivity and stability analysis in DEA with interval data, *Applied Mathematics and Computation*, 156(2), 463-477.

[17] Jahanshahloo, G.R., Matin, R.K. and Vencheh, A.H. (2004b), On FDH efficiency analysis with interval data, *Applied Mathematics and Computation*, 159(1), 47-55.

[18] Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., Rostamy Malkhalifeh, M. and Ahadzadeh Namin, M. (2009), A generalized model for data envelopment analysis with interval data, *Applied Mathematical Modelling*, 33(7), 3237-3244.

[19] Kao, C. (2006), Interval efficiency measures in data envelopment analysis with imprecise data, *European Journal of Operational Research*, 174(2), 1087-1099.

[20] Kim, S.H., Park, C.G. and Park, K.S. (1999), An application of data envelopment analysis in telephone offices evaluation with partial data, *Computers and Operations Research*, 26(1), 59-72.

[21] Park, K.S. (2010), Duality, efficiency computations and interpretations in imprecise DEA, *European Journal of Operational Research*, 200(1), 289-296.
[22] Sadjadi, S.J. and Omrani, H. (2008), Data envelopment analysis with uncertain data: an application for Iranian electricity distribution companies, *Energy Policy*, 36(11), 4247-4254.
[23] Sadjadi, S.J. and Omrani, H. (2010), A bootstrapped robust data envelopment analysis model for efficiency estimating of telecommunication companies in Iran, *Telecommunications Policy*, 34(4), 221-232.
[24] Sadjadi, S.J., Omrani, H., Abdollahzadeh, S., Alinaghian, M. and Mohammadi, H. (2011a), A robust super-efficiency data envelopment analysis model for ranking of provincial gas companies in Iran, *Expert Systems with Applications*, 38(9), 10875-10881.
[25] Sadjadi, S.J., Omrani, H., Makui, A. and Shahanaghi, K. (2011b), An interactive robust data envelopment analysis model for determining alternative targets in Iranian electricity distribution companies, *Expert Systems with Applications*, 38(8), 9830-9839.
[26] Sengupta, J.K. (1992), A fuzzy systems approach in data envelopment analysis, *Computers & Mathematics with Applications*, 24(8), 259-266.
[27] Shokouhi, A.H., Hatami-Marbini, A., Tavana, M. and Saati, S. (2010), A robust optimization approach for imprecise data envelopment analysis, *Computers & Industrial Engineering*, 59(3), 387-397.
[28] Shokouhi, A.H., Shahriari, H., Agrell, P.J. and Hatami-Marbini, A. (2014), Consistent and robust ranking in imprecise data envelopment analysis under perturbations of random subsets of data, *OR Spectrum*, 36(1), 133-160.
[29] Smirlis, Y.G., Maragos, E.K. and Despotis, D.K. (2006), Data envelopment analysis with missing values: an interval DEA approach, *Applied Mathematics and Computation*, 177(1), 1-10.
[30] Soyster, A.L. (1973), Technical note—convex programming with set-inclusive constraints and applications to inexact linear programming, *Operations Research*, 21(5), 1154-1157.
[31] Toloo, M., Aghayi, N. and Rostamy-Malkhalifeh, M. (2008), Measuring overall profit efficiency with interval data, *Applied Mathematics and Computation*, 201(1), 640-649.
[32] Wang, Y.M., Greatbanks, R. and Yang, J.B. (2005), Interval efficiency assessment using data envelopment analysis, *Fuzzy Sets and Systems*, 153(3), 347-370.
[33] Wang, K. and Wei, F. (2010), Robust data envelopment analysis based MCDM with the consideration of uncertain data, *Journal of Systems Engineering and Electronics*, 21(6), 981-989.
[34] Zhu, J. (2003), Imprecise data envelopment analysis (IDEA): a review and improvement with an application, *European Journal of Operational Research*, 144(3), 513-529.

Appendix A

Theorem 4.1. Model (9) is always feasible.

Proof: Let $z_{ij}^x = z_{rj}^y = z_j^x = z_j^y = 0, q_{ij} = x_{ij}^U - x_{ij}^L, p_{rj} = y_{rj}^U - y_{rj}^L$. Then, constraints (9e) and (9f) imply that $x_{ij} \geq x_{ij}^U$ and $x_{ij} \leq x_{ij}^U$, respectively, and also constraints (9*c*) and (9*d*) lead to $y_{rj} \leq y_{rj}^L$ and $y_{rj} \geq y_{rj}^L$, respectively. Hence, we have $x_{ij} = x_{ij}^U$ and $y_{rj} = y_{rj}^L$. Setting $v_{ij} = \frac{1}{x_{ij}^U}$ and $u_{rj} = \frac{1}{y_{rj}^L}$, completes the proof.