Available online at www.sciencedirect.com
SCIENCE DIRECT®

Fuzzy Sets and Systems 153 (2005) 347–370

FUZZY sets and systems

www.elsevier.com/locate/fss

Interval efficiency assessment using data envelopment analysis☆

Ying-Ming Wang[a, b, c, *], Richard Greatbanks[a], Jian-Bo Yang[a]

[a] *Manchester Business School, The University of Manchester, P.O. Box 88, Manchester M60 1QD, UK*
[b] *School of Public Administration, Fuzhou University, Fuzhou, Fujian 350002, PR China*
[c] *Center for Accounting Studies of Xiamen University, Xiamen, Fujian 361005, PR China*

Received 15 March 2004; received in revised form 15 November 2004; accepted 20 December 2004
Available online 18 January 2005

Abstract

This paper studies how to conduct efficiency assessment using data envelopment analysis (DEA) in interval and/or fuzzy input–output environments. A new pair of interval DEA models is constructed on the basis of interval arithmetic, which differs from the existing DEA models handling interval data in that the former is a linear CCR model without the need of extra variable alternations and uses a fixed and unified production frontier (i.e. the same constraint set) to measure the efficiencies of decision-making units (DMUs) with interval input and output data, while the latter is usually a nonlinear optimization problem with the need of extra variable alternations or scale transformations and utilizes variable production frontiers (i.e. different constraint sets) to measure interval efficiencies. Ordinal preference information and fuzzy data are converted into interval data through the estimation of permissible intervals and α-*level* sets, respectively, and are incorporated into the interval DEA models. The proposed interval DEA models are developed for measuring the lower and upper bounds of the best relative efficiency of each DMU with interval input and output data, which are different from the interval formed by the worst and the best relative efficiencies of each DMU. A minimax regret-based approach (MRA) is introduced to compare and rank

☆ This research was supported by the project on Human Social Science of MOE, P.R.China under the Grant No. 01JA790082, the National Natural Science Foundation of China (NSFC) under the Grant No: 70271056, and also in part by the European Commission under the Grant No: IPS-2000-00030, the UK Engineering and Physical Science Research Council (EPSRC) under the Grant No: GR/R32413/01, and Fok Ying Tung Education Foundation under the Grant No: 71080.

* Corresponding author. Manchester Business School East, The University of Manchester, P.O. Box 88, Manchester M60 1QD, UK. Tel.: +44 161 2750788; fax: +44 161 2003505.

E-mail addresses: msymwang@hotmail.com, Yingming.Wang@Manchester.ac.uk (Y.-M. Wang).

0165-0114/$ - see front matter
doi:10.1016/j.fss.2004.12.011

the efficiency intervals of DMUs. Two numerical examples are provided to show the applications of the proposed interval DEA models and the preference ranking approach.

Keywords: Data envelopment analysis; Interval DEA model; Interval efficiency; Interval data; Minimax regret ranking

1. Introduction

Data envelopment analysis (DEA), as a very useful management and decision tool, has found surprising development in theory and methodology and extensive applications in the range of the whole world since it was first developed by Charnes et al. [1]. Traditional DEA models such as CCR and BBC models and so on do not deal with imprecise data and assume that all input and output data are exactly known. In real world situations, however, this assumption may not always be true. Due to the existence of uncertainty, DEA sometimes faces the situation of imprecise data, especially when a set of decision-making units (DMUs) contains missing data, judgment data, forecasting data or ordinal preference information. Generally speaking, uncertain information or imprecise data can be expressed in interval or fuzzy numbers. Therefore, how to evaluate the management or operation efficiency of a set of DMUs in interval and/or fuzzy environments is a worth-studying problem. This is the need of both the developments of DEA theory and methodology and its real applications.

Cooper et al. [2–4] were the first, to the best of our knowledge, to study how to deal with imprecise data such as bounded data, ordinal data and ratio bounded data in DEA. The resulting DEA model was called imprecise DEA (IDEA), which transformed a nonlinear programming problem into a linear programming (LP) problem equivalent through a series of scale transformations and variable alternations. The final efficiency score for each DMU was derived as a deterministic numerical value less than or equal to unity. Kim et al. [14] also use an analogous scale transformation and variable alternation method, but they did not take the interval data situation into account. Recently, Lee et al. [15] extended the idea of IDEA to the additive model. It is argued that Cooper et al.'s method makes the DEA model become very complicated because of great numbers of data transformations and variable alternations. On the one hand, their variable alternations make the numbers of decision variables dramatically increase from $(m + s)$ to $(m + s) \times n$, where m, s, n, respectively, represent the numbers of inputs, outputs and DMUs; on the other hand, their scale transformations also make both the exact data and imprecise information including preference data and interval data (bounded data) into constraints, which leads to a rapid increase in computation burden. For more discussions on their method, please refer to Zhu [23].

Despotis and Smirlis [5] also studied the problem of IDEA, but developed an alternative approach for dealing with imprecise data in DEA. Their approach was to transform a nonlinear DEA model to a LP equivalent by applying transformations only on the variables. The resulting efficiency scores were defined to be intervals. Based on their approach, Haghighat and Khorram [8] discussed the problem of maximum and minimum numbers of DEA efficient units and Jahanshahloo et al. [9–11] studied further the problems associated with return to scale, sensitivity and stability analysis, and FDH efficiency. It will be seen in next section that their DEA model in fact used variable production frontiers, i.e. different constraint sets, to measure the efficiencies of DMUs, which made them lack of comparability.

Entani et al. [6] proposed a DEA model with interval efficiencies measured from both the optimistic and the pessimistic viewpoints. Their model was first developed for crisp data and then extended to interval data and fuzzy data. In theory their model was able to deal with interval data and fuzzy data, but there is a drawback with their model. That is their model chooses only one input and one output data to capture the lower bound efficiency of each DMU no matter how many input and output data are involved in the model, which leads to their model suffering from the loss of information on the other input and output data of the DMU under evaluation. Moreover, their model also adopts variable production frontiers (i.e. different constraint sets) to measure the efficiency intervals of different DMUs. There is also a distinctive difference between their model and other interval DEA models. Their model uses the worst relative efficiency measured from the pessimistic viewpoint to constitute the lower bound of interval efficiency, while other DEA models dealing with interval data utilize the best relative efficiency in the most unfavorable situation to form the lower bound of interval efficiency. Their meanings and the premises for obtaining interval efficiency are totally different. Entani et al.'s DEA model can be used to measure the interval efficiency of a DMU with crisp or interval or fuzzy input–output data or their mixture, while the other interval DEA models mentioned previously cannot be used to measure the interval efficiency of a DMU with crisp input–output data and can only be used for interval data. It is certain that for interval data the efficiency intervals obtained by Entani et al.'s DEA model are wider than those obtained by other interval DEA models.

The main purpose of this paper is to develop a new pair of interval DEA models that can both overcome the shortcomings mentioned above and model imprecise data in a simple, rational and effective way. Different from Entani et al.'s interval efficiencies, the new pair of interval DEA models will be developed for interval input and output data rather than for crisp input and output data. The final efficiency score for each DMU will be characterized by an interval bounded by the best lower bound efficiency and the best upper bound efficiency of each DMU, which we refer to as interval efficiency or efficiency interval. A minimax regret-based approach is suggested to compare and rank the interval efficiencies of DMUs. Two numerical examples are provided to illustrate the applications of the proposed interval DEA models and the preference ranking approach.

The rest of the paper is organized as follows. In Section 2, we develop our new interval DEA models on the basis of interval arithmetic and compare their differences with other existing interval DEA models. Section 3 discusses how ordinal preference information and fuzzy data can be transformed into interval data through the estimation of permissible intervals and α-*level* sets, respectively. This is followed by the introduction of the minimax regret approach (MRA) for comparing and ranking interval efficiencies. In Section 5, we provide two numerical examples to illustrate the applications of the proposed interval DEA models and the preference ranking approach, MRA. The paper is concluded in Section 6.

2. Interval DEA models based on interval arithmetic

Assume that there are n DMUs to be evaluated. Each DMU consumes varying amounts of m different inputs to produce s different outputs. Especially, DMU_j consumes amounts $X_j = \{x_{ij}\}$ of inputs ($i = 1, 2, \ldots, m$) and produces amounts $Y_j = \{y_{rj}\}$ of outputs ($r = 1, 2, \ldots, s$). Without loss of generality, we assume that all the input and output data x_{ij} and x_{rj} ($i = 1, \ldots, m; r = 1, \ldots, s; j = 1, \ldots, n$) cannot be exactly obtained due to the existence of uncertainty. They are only known to lie within the upper and lower bounds represented by the intervals $[x_{ij}^{\mathrm{L}}, x_{ij}^{\mathrm{U}}]$ and $[y_{rj}^{\mathrm{L}}, y_{rj}^{\mathrm{U}}]$, where $x_{ij}^{\mathrm{L}} > 0$ and $y_{rj}^{\mathrm{L}} > 0$.

In order to deal with such an uncertain situation, the following pair of LP models has been developed to generate the upper and lower bounds of interval efficiency for each DMU (see Despotis and

Smirlis [5] for details):

$$\begin{aligned}
\text{Maximize } & H_{j_0}^{U} = \sum_{r=1}^{s} u_r y_{rj_0}^{U} \\
\text{subject to } & \sum_{i=1}^{m} v_i x_{ij_0}^{L} = 1, \\
& \sum_{r=1}^{s} u_r y_{rj_0}^{U} - \sum_{i=1}^{m} v_i x_{ij_0}^{L} \leqslant 0, \\
& \sum_{r=1}^{s} u_r y_{rj}^{L} - \sum_{i=1}^{m} v_i x_{ij}^{U} \leqslant 0, \quad j = 1, \ldots, n;\ j \neq j_0, \\
& u_r, v_i \geqslant \varepsilon \quad \forall r, i.
\end{aligned} \tag{1}$$

$$\begin{aligned}
\text{Maximize } & H_{j_0}^{L} = \sum_{r=1}^{s} u_r y_{rj_0}^{L} \\
\text{subject to } & \sum_{i=1}^{m} v_i x_{ij_0}^{U} = 1, \\
& \sum_{r=1}^{s} u_r y_{rj_0}^{L} - \sum_{i=1}^{m} v_i x_{ij_0}^{U} \leqslant 0, \\
& \sum_{r=1}^{s} u_r y_{rj}^{U} - \sum_{i=1}^{m} v_i x_{ij}^{L} \leqslant 0, \quad j = 1, \ldots, n;\ j \neq j_0, \\
& u_r, v_i \geqslant \varepsilon \quad \forall r, i,
\end{aligned} \tag{2}$$

where j_0 is the DMU under evaluation (usually denoted by DMU_0); u_r and v_i are the weights assigned to the outputs and inputs; $H_{j_0}^{U}$ and $H_{j_0}^{L}$ are the best possible relative efficiencies for DMU_0 under the most favorable and the most unfavorable situations, respectively, and ε is the non-Archimedean infinitesimal.

For convenience, we refer to the above pair of LP models as the upper and lower bounds DEA models, respectively. Zhu [23] has proved that Cooper et al.'s IDEA model [2] can be simplified as the above upper bound model in the case of interval data. Entani et al. [6] also utilized the above upper bound DEA model to measure the best possible relative efficiency of each DMU.

Carefully observing the above upper and lower bounds DEA models, we may find that the constraint sets used to measure the efficiencies of DMUs are different from one DMU to another and even the constraint sets utilized to measure the upper and lower bounds of efficiency of the same DMU are also different from each other. For example, the constraint set used to measure the upper bound efficiency of DMU_0 consists of the data set $\{(x_{ij_0}^{L}, y_{rj_0}^{U}), (x_{ij}^{U}, y_{rj}^{L})(j = 1, \ldots, n;\ j \neq j_0; i = 1, \ldots, m; r = 1, \ldots, s)\}$, while the constraint set utilized to measure its lower bound efficiency is composed of the data set $\{(x_{ij_0}^{U}, y_{rj_0}^{L}), (x_{ij}^{L}, y_{rj}^{U})(j = 1, \ldots, n;\ j \neq j_0; i = 1, \ldots, m; r = 1, \ldots, s)\}$. It is evident that these two data sets are different.

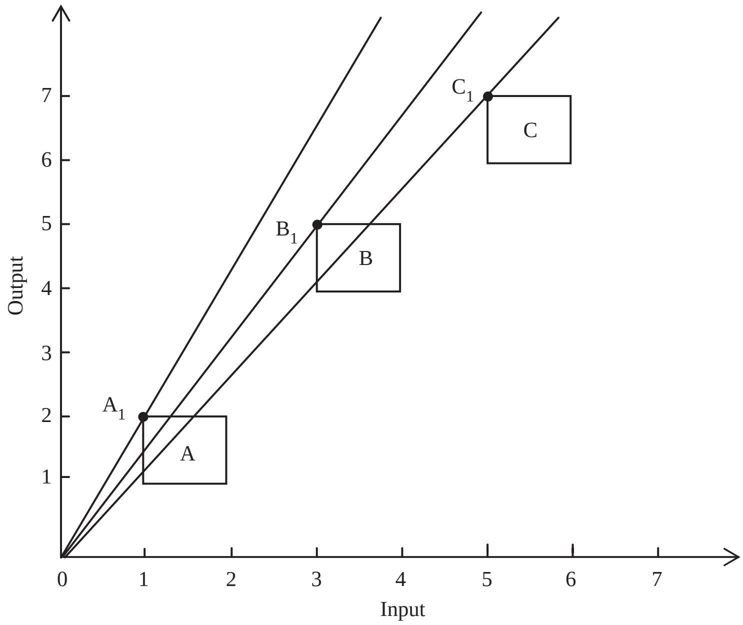

Fig. 1. The production frontiers used in models (1) and (2).

The main drawback of the use of different constraint sets to measure the efficiencies of DMUs is the lack of comparability among the efficiencies because different production frontiers were adopted in the process of efficiency measure. To better show this, let us consider the simplest case of one input and one output. Three DMUs, labeled as A, B and C in Fig. 1, use interval input [1,2], [3,4] and [5,6] to produce interval output [1,2], [4,5] and [6,7], respectively. When computing the upper bound efficiency of DMU_A, model (1) uses the data set $\{(1, 2), (4, 4), (6, 6)\}$, which forms the production frontier denoted by the radiate line OA_1 in Fig. 1; when calculating the upper bound efficiency of DMU_B, model (1) uses the data set $\{(2, 1), (3, 5), (6, 6)\}$, which yields the production frontier denoted by the radiate line OB_1 in Fig. 1; while computing the upper bound efficiency of DMU_C, model (1) utilizes the data set $\{(2, 1), (4, 4), (5, 7)\}$, which yields the production frontier denoted by the radiate line OC_1. The production frontiers used to compute the lower bound efficiencies of DMU_A, DMU_B and DMU_C are, respectively, the radiate lines OB_1, OA_1 and OA_1. Since the efficiency is calculated as the ratio of the actual output to the maximal output on production frontier, if the production frontier was not fixed and not unified, the comparisons among the efficiencies would become meaningless. In addition, we think *n* DMUs can only have one real production frontier. Since each DMU has the possibility of using the minimum inputs to produce the maximum outputs, the real production frontier should be yielded on the basis of the best production activity state of each DMU. The real production frontier in Fig. 1 is the radiate line OA_1, which is yielded on the basis of the data set $\{(1, 2), (3, 5), (5, 7)\}$.

In order to avoid the use of different production frontiers to measure the efficiencies of different DMUs, a new pair of interval DEA models will be developed. The models are based on the interval arithmetic and always use the same constraint set, which forms a fixed and unified production frontier, for all DMUs as well as for the measures of both the lower and upper bound efficiencies.

Let

$$\theta_j = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}}, \quad j = 1, \ldots, n$$

be the efficiency of DMU_j. According to the operation rules on interval data, we have

$$\theta_j = \frac{\sum_{r=1}^{s} u_r \left[y_{rj}^{\mathrm{L}}, y_{rj}^{\mathrm{U}}\right]}{\sum_{i=1}^{m} v_i \left[x_{ij}^{\mathrm{L}}, x_{ij}^{\mathrm{U}}\right]} = \frac{\left[\sum_{r=1}^{s} u_r y_{rj}^{\mathrm{L}}, \sum_{r=1}^{s} u_r y_{rj}^{\mathrm{U}}\right]}{\left[\sum_{i=1}^{m} v_i x_{ij}^{\mathrm{L}}, \sum_{i=1}^{m} v_i x_{ij}^{\mathrm{U}}\right]}$$
$$= \left[\frac{\sum_{r=1}^{s} u_r y_{rj}^{\mathrm{L}}}{\sum_{i=1}^{m} v_i x_{ij}^{\mathrm{U}}}, \frac{\sum_{r=1}^{s} u_r y_{rj}^{\mathrm{U}}}{\sum_{i=1}^{m} v_i x_{ij}^{L}}\right], \quad j = 1, \ldots, n.$$

It is obvious that θ_j should also be an interval number, which we denote by $[\theta_j^{\mathrm{L}}, \theta_j^{\mathrm{U}}]$ $(j = 1, \ldots, n)$.

Let

$$\theta_j = [\theta_j^{\mathrm{L}}, \theta_j^{\mathrm{U}}] = \left[\frac{\sum_{r=1}^{s} u_r y_{rj}^{\mathrm{L}}}{\sum_{i=1}^{m} v_i x_{ij}^{\mathrm{U}}}, \frac{\sum_{r=1}^{s} u_r y_{rj}^{\mathrm{U}}}{\sum_{i=1}^{m} v_i x_{ij}^{\mathrm{L}}}\right] \subseteq (0, 1], \quad j = 1, \ldots, n.$$

Then

$$\theta_j^{\mathrm{U}} = \frac{\sum_{r=1}^{s} u_r y_{rj}^{\mathrm{U}}}{\sum_{i=1}^{m} v_i x_{ij}^{\mathrm{L}}} \leqslant 1, \quad j = 1, \ldots, n,$$

$$\theta_j^{\mathrm{L}} = \frac{\sum_{r=1}^{s} u_r y_{rj}^{\mathrm{L}}}{\sum_{i=1}^{m} v_i x_{ij}^{\mathrm{U}}} > 0, \quad j = 1, \ldots, n.$$

In order to measure the upper and lower bounds of the efficiency of DMU_0, we construct the following pair of fractional programming models for DMU_0:

$$\begin{aligned} \text{Maximize } \ & \theta_{j_0}^{\mathrm{U}} = \frac{\sum_{r=1}^{s} u_r y_{rj_0}^{\mathrm{U}}}{\sum_{i=1}^{m} v_i x_{ij_0}^{\mathrm{L}}} \\ \text{subject to } \ & \theta_j^{\mathrm{U}} = \frac{\sum_{r=1}^{s} u_r y_{rj}^{\mathrm{U}}}{\sum_{i=1}^{m} v_i x_{ij}^{\mathrm{L}}} \leqslant 1, \quad j = 1, \ldots, n, \\ & u_r, v_i \geqslant \varepsilon \quad \forall r, i. \end{aligned} \tag{3}$$

$$\begin{aligned} \text{Maximize } \ & \theta_{j_0}^{\mathrm{L}} = \frac{\sum_{r=1}^{s} u_r y_{rj_0}^{\mathrm{L}}}{\sum_{i=1}^{m} v_i x_{ij_0}^{\mathrm{U}}} \\ \text{subject to } \ & \theta_j^{\mathrm{U}} = \frac{\sum_{r=1}^{s} u_r y_{rj}^{\mathrm{U}}}{\sum_{i=1}^{m} v_i x_{ij}^{\mathrm{L}}} \leqslant 1, \quad j = 1, \ldots, n, \\ & u_r, v_i \geqslant \varepsilon \quad \forall r, i. \end{aligned} \tag{4}$$

Using Charnes–Cooper transformation, the above pair of fractional programming models can be simplified as the following equivalent LP models:

$$\begin{aligned}
\text{Maximize } & \theta_{j0}^{\mathrm{U}} = \sum_{r=1}^{s} u_r y_{rj_0}^{\mathrm{U}} \\
\text{subject to } & \sum_{i=1}^{m} v_i x_{ij_0}^{\mathrm{L}} = 1, \\
& \sum_{r=1}^{s} u_r y_{rj}^{\mathrm{U}} - \sum_{i=1}^{m} v_i x_{ij}^{\mathrm{L}} \leqslant 0, \quad j = 1, \ldots, n, \\
& u_r, v_i \geqslant \varepsilon \quad \forall r, i.
\end{aligned} \tag{5}$$

$$\begin{aligned}
\text{Maximize } & \theta_{j0}^{\mathrm{L}} = \sum_{r=1}^{s} u_r y_{rj_0}^{\mathrm{L}} \\
\text{subject to } & \sum_{i=1}^{m} v_i x_{ij_0}^{\mathrm{U}} = 1, \\
& \sum_{r=1}^{s} u_r y_{rj}^{\mathrm{U}} - \sum_{i=1}^{m} v_i x_{ij}^{\mathrm{L}} \leqslant 0, \quad j = 1, \ldots, n, \\
& u_r, v_i \geqslant \varepsilon \quad \forall r, i,
\end{aligned} \tag{6}$$

where θ_{j0}^{U} stands for the best possible relative efficiency achieved by DMU_0 when all the DMUs are in the state of best production activity, while θ_{j0}^{L} stands for the lower bound of the best possible relative efficiency of DMU_0. They constitute a possible best relative efficiency interval $[\theta_{j0}^{\mathrm{L}}, \theta_{j0}^{\mathrm{U}}]$.

Note that model (5) determines the production frontier for all the DMUs and model (6) uses the production frontier as a benchmark to measure the lower bound efficiency of each DMU. Therefore, the meanings of θ_{j0}^{U} and θ_{j0}^{L} differ from the meanings of H_{j0}^{U} and H_{j0}^{L} in models (1) and (2). Moreover, models (1) and (2) come from the following pair of fraction programming problems:

$$\begin{aligned}
\text{Maximize } & H_{j0}^{\mathrm{U}} = \frac{\sum_{r=1}^{s} u_r y_{rj_0}^{\mathrm{U}}}{\sum_{i=1}^{m} v_i x_{ij_0}^{\mathrm{L}}} \\
\text{subject to } & \frac{\sum_{r=1}^{s} u_r y_{rj_0}^{\mathrm{U}}}{\sum_{i=1}^{m} v_i x_{ij_0}^{\mathrm{L}}} \leqslant 1, \\
& \frac{\sum_{r=1}^{s} u_r y_{rj}^{\mathrm{L}}}{\sum_{i=1}^{m} v_i x_{ij}^{\mathrm{U}}} \leqslant 1 \quad j = 1, \ldots, n;\ j \neq j_0, \\
& u_r, v_i \geqslant \varepsilon \quad \forall r, i.
\end{aligned} \tag{7}$$

$$\begin{aligned}
\text{Maximize } & H_{j_0}^{\mathrm{L}} = \frac{\sum_{r=1}^{s} u_r y_{rj_0}^{\mathrm{L}}}{\sum_{i=1}^{m} v_i x_{ij_0}^{\mathrm{U}}} \\
\text{subject to } & \frac{\sum_{r=1}^{s} u_r y_{rj_0}^{\mathrm{L}}}{\sum_{i=1}^{m} v_i x_{ij_0}^{\mathrm{U}}} \leqslant 1, \\
& \frac{\sum_{r=1}^{s} u_r y_{rj}^{\mathrm{U}}}{\sum_{i=1}^{m} v_i x_{ij}^{\mathrm{L}}} \leqslant 1, \quad j = 1, \ldots, n;\ j \neq j_0, \\
& u_r, v_i \geqslant \varepsilon \quad \forall r, i.
\end{aligned} \tag{8}$$

It can be seen from model (7) that the optimal weights most favorable for the upper bound efficiency of DMU_0 can only guarantee the lower bound efficiencies of the other DMUs to be less than or equal to unity, but cannot guarantee their upper bound efficiencies to be also less than or equal to unity. The similar phenomenon also exists in model (8). Such a phenomenon, however, does not appear in models (3)–(6) at all. No matter what values the inputs and outputs take for each DMU and no matter what weights the models use, the efficiencies of DMUs are all limited to less than or equal to one.

It is also very clear from models (3)–(6) that the constraint set used to measure the efficiencies of DMUs is completely the same, which is made up of the data set $\{(x_{ij}^{\mathrm{L}}, y_{rj}^{\mathrm{U}})\ (j = 1, \ldots, n;\ i = 1, \ldots, m;\ r = 1, \ldots, s)\}$.

However, it must be pointed out that both the best and the worst production activities of DMU_0 are considered in (4), one in the constraint set and the other in objective function. It seems quite confusing why two different data are used for DMU_0 in the same model. This can be understood from the following aspects:

First, $\theta_{j_0}^{\mathrm{L}}$ in (4) is measured not only relative to the other $(n-1)$ DMUs, but also relative to the best production activity of DMU_0 itself. The best production activity of DMU_0 plays a role of frame of reference in the model. This should not be understood as the coincidence of both the best and the worst production activities of DMU_0, just as an interval number $A = [a^{\mathrm{L}}, a^{\mathrm{U}}]$ can be expressed as $x = \lambda a^{\mathrm{L}} + (1-\lambda)a^{\mathrm{U}}$ $(0 \leqslant \lambda \leqslant 1)$, which does not mean that $x = a^{\mathrm{L}}$ and $x = a^{\mathrm{U}}$ occur at the same time.

Next, let DMU_0' be a virtual DMU that consumes the upper bound inputs of DMU_0 and produces only the lower bound outputs of DMU_0. Since DMU_0' represents the worst production activity of DMU_0, the best relative efficiency of DMU_0' can therefore be used to characterize the lower bound efficiency of DMU_0. Accordingly, model (4) can be considered as conventional DEA model evaluating DMU_0' using $(n+1)$ DMUs, where DMU_i $(i = 1, \ldots, n)$ consumes the least inputs to produce the most outputs, which leads to the efficiency of DMU_0' to be less than one automatically. It is no wonder that DMU_0 can be the reference DMU of DMU_0'.

Finally, production frontier is determined by the best production activities of the n DMUs regardless of their worst production activities. If the best production activity of DMU_0 were removed from the model, the production frontier would be changed and would be different from DMU to DMU, which would result in the efficiencies obtained incomparable.

About the relationship between $\theta_{j_0}^{\mathrm{U}}$ and $\theta_{j_0}^{\mathrm{L}}$ and the relationships between $\theta_{j_0}^{\mathrm{U}}$ and $H_{j_0}^{\mathrm{U}}$ as well as $\theta_{j_0}^{\mathrm{L}}$ and $H_{j_0}^{\mathrm{L}}$, we have the following two theorems.

Theorem 1. *If $\theta_{j0}^{U^*}$ and $\theta_{j0}^{L^*}$ are the optimum objective function values of models (5) and (6), respectively, then $\theta_{j0}^{L^*} \leqslant \theta_{j0}^{U^*}$ with the equality holding only when all the input and output data for* DMU_0 *degenerate from interval data to exact data.*

Proof. Suppose u_r^* and v_i^* $(r = 1, \ldots, s;\ i = 1, \ldots, m)$ are the optimal solution to model (6). Let

$$\beta_{j0} = \sum_{i=1}^{m} v_i^* x_{ij_0}^{L},$$

$$\tilde{u}_r = \frac{u_r^*}{\beta_{j0}}, \quad r = 1, \ldots, s,$$

$$\tilde{v}_i = \frac{v_i^*}{\beta_{j0}}, \quad i = 1, \ldots, m.$$

It follows that

$$\beta_{j0} = \sum_{i=1}^{m} v_i^* x_{ij_0}^{L} \leqslant \sum_{i=1}^{m} v_i^* x_{ij_0}^{U} = 1,$$

$$\sum_{i=1}^{m} \tilde{v}_i x_{ij_0}^{L} = \sum_{i=1}^{m} \left(\frac{v_i^*}{\beta_{j0}}\right) x_{ij_0}^{L} = \frac{1}{\beta_{j0}} \sum_{i=1}^{m} v_i^* x_{ij_0}^{L} = 1,$$

$$\sum_{r=1}^{s} \tilde{u}_r y_{rj}^{U} - \sum_{i=1}^{m} \tilde{v}_i x_{ij}^{L} = \frac{1}{\beta_{j0}} \left(\sum_{r=1}^{s} u_r^* y_{rj}^{U} - \sum_{i=1}^{m} v_i^* x_{ij}^{L}\right) \leqslant 0, \quad j = 1, \ldots, n,$$

$$\tilde{u}_r = \frac{u_r^*}{\beta_{j0}} \geqslant \frac{\varepsilon}{\beta_{j0}} \geqslant \varepsilon, \quad r = 1, \ldots, s,$$

$$\tilde{v}_i = \frac{v_i^*}{\beta_{j0}} \geqslant \frac{\varepsilon}{\beta_{j0}} \geqslant \varepsilon, \quad i = 1, \ldots, m.$$

It is obvious that $\tilde{u}_r$ and $\tilde{v}_i$ $(r = 1, \ldots, s;\ i = 1, \ldots, m)$ are a feasible solution to model (5). So, we have

$$\sum_{r=1}^{s} \tilde{u}_r y_{rj_0}^{U} \leqslant \theta_{j0}^{U^*},$$

$$\theta_{j0}^{L^*} = \sum_{r=1}^{s} u_r^* y_{rj_0}^{L} \leqslant \sum_{r=1}^{s} u_r^* y_{rj_0}^{U} = \sum_{r=1}^{s} (\beta_{j0} \tilde{u}_r) y_{rj_0}^{U} = \beta_{j0} \sum_{r=1}^{s} \tilde{u}_r y_{rj_0}^{U} \leqslant \beta_{j0} \theta_{j0}^{U^*} \leqslant \theta_{j0}^{U^*}.$$

The above equality holds only when $y_{rj_0}^{L} = y_{rj_0}^{U}$ and $x_{ij_0}^{U} = x_{rj_0}^{L}$ for all $r = 1, \ldots, s$ and $i = 1, \ldots, m$. □

Theorem 2. *If $\theta_{j_0}^{\mathrm{U}^*}$ and $\theta_{j_0}^{\mathrm{L}^*}$ are the optimum objective function values of models (5) and (6) and $H_{j_0}^{\mathrm{U}^*}$ and $H_{j_0}^{\mathrm{L}^*}$ are the optimum objective function values of models (1) and (2), respectively, then $\theta_{j_0}^{\mathrm{L}^*} \leqslant H_{j_0}^{\mathrm{L}^*}$ and $\theta_{j_0}^{\mathrm{U}^*} \leqslant H_{j_0}^{\mathrm{U}^*}$.*

Proof. Suppose u_r^* and v_i^* $(r = 1, \ldots, s; i = 1, \ldots, m)$ are the optimal solution to model (6). Then we have

$$\sum_{r=1}^{s} u_r^* y_{rj}^{\mathrm{U}} - \sum_{i=1}^{m} v_i^* x_{ij}^{\mathrm{L}} \leqslant 0, \quad j = 1, \ldots, n.$$

Especially for DMU$_{j_0}$, we have

$$\sum_{r=1}^{s} u_r^* y_{rj_0}^{\mathrm{L}} - \sum_{i=1}^{m} v_i^* x_{ij_0}^{\mathrm{U}} \leqslant \sum_{r=1}^{s} u_r^* y_{rj_0}^{\mathrm{U}} - \sum_{i=1}^{m} v_i^* x_{ij_0}^{\mathrm{L}} \leqslant 0,$$

which means u_r^* and v_i^* $(r = 1, \ldots, s; i = 1, \ldots, m)$ are also a feasible solution to model (2). So, we get

$$\theta_{j_0}^{\mathrm{L}^*} = \sum_{r=1}^{s} u_r^* y_{rj_0}^{\mathrm{L}} = H_{j_0}^{\mathrm{L}} \leqslant H_{j_0}^{\mathrm{L}^*}.$$

Similarly, if u_r^* and v_i^* $(r = 1, \ldots, s; i = 1, \ldots, m)$ are the optimal solution to model (5), then the following inequalities hold for all the DMUs:

$$\sum_{r=1}^{s} u_r^* y_{rj}^{\mathrm{U}} - \sum_{i=1}^{m} v_i^* x_{ij}^{\mathrm{L}} \leqslant 0, \quad j = 1, \ldots, n.$$

For DMU$_{j_0}$, we have

$$\sum_{r=1}^{s} u_r y_{rj_0}^{\mathrm{U}} - \sum_{i=1}^{m} v_i x_{ij_0}^{\mathrm{L}} \leqslant 0.$$

For DMU$_j$ $(j = 1, \ldots, n;\ j \neq j_0)$, we get

$$\sum_{r=1}^{s} u_r^* y_{rj_0}^{\mathrm{L}} - \sum_{i=1}^{m} v_i^* x_{ij_0}^{\mathrm{U}} \leqslant \sum_{r=1}^{s} u_r^* y_{rj_0}^{\mathrm{U}} - \sum_{i=1}^{m} v_i^* x_{ij_0}^{\mathrm{L}} \leqslant 0.$$

It is obvious that u_r^* and v_i^* $(r = 1, \ldots, s; i = 1, \ldots, m)$ are a feasible solution of model (1). Thus, we have the following inequality relation:

$$\theta_{j_0}^{\mathrm{U}^*} = \sum_{r=1}^{s} u_r^* y_{rj_0}^{\mathrm{U}} = H_{j_0}^{\mathrm{U}} \leqslant H_{j_0}^{\mathrm{U}^*}.$$

This completes the proof. □

In order to judge whether a DMU is DEA efficient or not, we give the following definition.

Definition 1. A DMU, DMU_0, is said to be DEA efficient if its best possible upper bound efficiency $\theta_{j_0}^{\mathrm{U}^*} = 1$; otherwise, it is said to be DEA inefficient if $\theta_{j_0}^{\mathrm{U}^*} < 1$.

Theorem 2 shows that DMU_0 is not necessarily DEA efficient in our interval DEA models (4)–(6) even if it was judged to be DEA efficient by models (1) and (2). So, the new interval DEA models can reduce the number of DMUs which are DEA efficient very significantly.

3. Incorporation of ordinal preference information and fuzzy data into the interval DEA models

In real decision-making and evaluation problems, ordinal preference information and/or fuzzy data are often encountered. DEA efficiency rating is no exception. How to deal with them becomes a key and important issue of decision and evaluation, which has aroused great interests of many experts and scholars. In this section, we discuss how to transform ordinal preference information and fuzzy data into interval data so that the interval DEA models developed in this paper can still work properly even in these situations.

3.1. The transformation of ordinal preference information

Suppose some input and/or output data for DMUs are given in the form of ordinal preference information. Usually, there may exist three types of ordinal preference information: (1) strong ordinal preference information such as $y_{rj} > y_{rk}$ or $x_{ij} > x_{ik}$, which can be further expressed as $y_{rj} \geqslant \chi_r y_{rk}$ and $x_{ij} \geqslant \eta_i x_{ik}$, where $\chi_r > 1$ and $\eta_i > 1$ are the parameters on the degree of preference intensity provided by decision maker (DM); (2) weak ordinal preference information such as $y_{rp} \geqslant y_{rq}$ or $x_{ip} \geqslant x_{iq}$; (3) indifference relationship such as $y_{rl} = y_{rt}$ or $x_{il} = x_{it}$. Since DEA model has the property of unit-invariance, the use of scale transformation to ordinal preference information does not change the original ordinal relationships and has no effect on the efficiencies of DMUs. Therefore, we may conduct a scale transformation to every ordinal input and output index so that its best ordinal datum is less than or equal to unity and then give an interval estimate for each ordinal datum.

Let us take the transformation of ordinal preference information about the output $y_{rj} (j = 1, \ldots, n)$ for example. The ordinal preference information about input and other output data can be converted in the same way.

For weak ordinal preference information $y_{r1} \geqslant y_{r2} \geqslant \cdots \geqslant y_{rn}$, we have the following ordinal relationships after scale transformation:

$$1 \geqslant \hat{y}_{r1} \geqslant \hat{y}_{r2} \geqslant \cdots \geqslant \hat{y}_{rn} \geqslant \sigma_r,$$

where σ_r is a small positive number reflecting the ratio of the possible minimum of $\{y_{rj} | j = 1, \ldots, n\}$ to its possible maximum. It can be approximately estimated by the DM. We refer to it as the ratio parameter for convenience. The resultant permissible interval for each $\hat{y}_{rj}$ is given by

$$\hat{y}_{rj} \in [\sigma_r, 1], \quad j = 1, \ldots, n.$$

Note that Zhu [23] ever transformed the weak ordinal preference information $y_{r1} \leqslant y_{r2} \leqslant \cdots \leqslant y_{rn}$ into the following intervals:

$$y_{rj} \in [0, 1] \quad \text{for DMU}_j \, (j = 1, \ldots, k-1),$$

$$y_{rk} = 1 \quad \text{for DMU}_k,$$

$$y_{rj} \in [1, M] \quad \text{for DMU}_j \, (j = k+1, \ldots, n),$$

where DMU_k is the DMU under evaluation and M is a sufficiently large number. Zhu utilized the lower bound value for each DMU_j ($j = 1, \ldots, n$; $j \neq k$) to evaluate the best possible relative efficiency of DMU_k. For example, when DMU_n was under evaluation, Zhu set $y_{r1} = \cdots = y_{r,n-1} = 0$ and $y_{rn} = 1$ and when DMU_1 was under evaluation, Zhu set $y_{r1} = 1$ and $y_{r2} = \cdots = y_{rn} = 1$. Therefore, Zhu in fact used [0, 1] for all the DMUs except DMU_n, for which he always utilized $y_{rn} = 1$. Here two problems arise from Zhu's approach. One is the use of too many zero inputs and outputs in the process of efficiency evaluation. The other is why y_{rn} is an exact number while all the others are imprecise data. It seems more reasonable that we use interval $[\sigma_r, 1]$ for all the DMUs.

For strong ordinal preference information $y_{r1} > y_{r2} > \cdots > y_{rn}$, we have the following ordinal relationships after scale transformation:

$$1 \geqslant \hat{y}_{r1}, \quad \hat{y}_{rj} \geqslant \chi_r \hat{y}_{r,j+1} \; (j = 1, \ldots, n-1) \quad \text{and} \quad \hat{y}_{rn} \geqslant \sigma_r,$$

where χ_r is a preference intensity parameter satisfying $\chi_r > 1$ provided by the DM and σ_r is the ratio parameter also provided by the DM. The resultant permissible interval for each $\hat{y}_{rj}$ can be derived as follows:

$$\hat{y}_{rj} \in [\sigma_r \chi_r^{n-j}, \chi_r^{1-j}], \quad j = 1, \ldots, n \quad \text{with } \sigma_r \leqslant \chi_r^{1-n}.$$

Note that Zhu [23] utilized the following data in his paper for strong ordinal preference information $y_{r1} < y_{r2} < \cdots < y_{rn}$:

$$y_{rj} = \varepsilon_j \approx 0 \quad \text{for DMU}_j \, (j = 1, \ldots, k),$$

$$y_{rk} = 1 \quad \text{for DMU}_k,$$

$$y_{rj} = \chi_r^{j-k} \quad \text{for DMU}_j \, (j = k+1, \ldots, n).$$

When evaluating DMU_n, Zhu set $y_{rj} = \varepsilon_j \approx 0$ ($j = 1, \ldots, n-1$) and $y_{rn} = 1$ and when DMU_1 was under evaluation, Zhu set $y_{r1} = 1$ and $y_{rj} = \chi_r^{j-1}$ ($j = 2, \ldots, n$). Although Zhu did not transform the strong ordinal preference information into interval data, he did use different numerical values for the evaluation of different DMUs. What he utilized is in fact the interval data $y_{rj} \in [\varepsilon_j, \chi_r^{j-1}]$ for DMU_j ($j = 1, \ldots, n-1$) and $y_{rn} \in [1, \chi_r^{n-1}]$ for DMU_n. Here arise a question again. Why was the lower bound for $\hat{y}_{rn}$ set to one while the lower bounds for all the others set close to zero?

Finally, for indifference relationship, the permissible intervals are the same as those obtained for weak ordinal preference information.

Through the scale transformation above and the estimation of permissible intervals, all the ordinal preference information is converted into interval data and can thus be incorporated into interval DEA models.

According to the simplest order relation between two interval numbers (see property 1 in Section 4), i.e. $A \leqslant B$ if and only if $a_{\mathrm{L}} \leqslant b_{\mathrm{L}}$ and $a_{\mathrm{U}} \leqslant b_{\mathrm{U}}$, where $A = [a_{\mathrm{L}}, a_{\mathrm{U}}]$ and $B = [b_{\mathrm{L}}, b_{\mathrm{U}}]$ are two interval numbers, the transformed interval data still reserve the original ordinal preference relationships.

3.2. The transformation of fuzzy data

When all or part of input and output data are fuzzy data expressed in triangular and trapezoidal fuzzy numbers, several approaches have been proposed to deal with the fuzzy data in the framework of DEA. Sengupta [20] proposed a fuzzy mathematical programming approach in which fuzziness was incorporated into DEA model by defining tolerance levels on both objective function and constraint violations. Triantis and Girod [21] suggested a mathematical programming approach through transforming fuzzy input and output data into crisp data using membership function values. Efficiency scores were computed for different values of membership functions and then averaged. Guo and Tanaka [7] proposed a fuzzy CCR model in which fuzzy constraints including fuzzy equalities and fuzzy inequalities were all converted into crisp constraints by predefining a possibility level and using the comparison rule for fuzzy numbers. Based on the same idea, León et al. [16] suggested a fuzzy BCC model. Lertworasirikul et al. [17] proposed a possibility approach in which fuzzy constraints were treated as fuzzy events and fuzzy DEA model was transformed into possibility DEA model by using possibility measures on fuzzy events. In the special case that fuzzy data are trapezoidal fuzzy numbers, possibility DEA model becomes LP model. Kao and Liu [12,13] suggested transforming fuzzy data into interval data by applying the α-*level* sets (also called α-*cuts*) so that a family of conventional crisp DEA models could be utilized. Note that the crisp DEA models they used are models (1) and (2), which have the shortcomings we have analyzed before. The α-*cut* approach was also adopted by Saati et al. [19], who defined the fuzzy CCR model as a possibilistic-programming problem and transformed it into an interval programming using the concept of α-*cuts*, which could be solved as a crisp LP model with the help of some variable substitutions and produced a crisp efficiency score for each DMU and for each given α-*cut*. Such an approach was further extended recently in [18] so that all DMUs could be evaluated using a common set of weights under a given α-*cut*.

In order to enhance the capability of interval DEA models (3)–(6) to deal with imprecise data, fuzzy data will be transformed into interval data in this paper by using the α-*level* sets [24]. Let the inputs $\tilde{x}_{ij}$ and outputs $\tilde{y}_{rj}$ be fuzzy data with membership functions $\mu_{\tilde{x}_{ij}}$ and $\mu_{\tilde{y}_{rj}}$, respectively, and $S(\tilde{x}_{ij})$ and $S(\tilde{y}_{rj})$ be the support of $\tilde{x}_{ij}$ and $\tilde{y}_{rj}$, respectively. Then the α-*level* sets of $\tilde{x}_{ij}$ and $\tilde{y}_{rj}$ can be defined as

$$
\begin{aligned}
(x_{ij})_\alpha &= \{x_{ij} \in S(\tilde{x}_{ij}) | \mu_{\tilde{x}_{ij}}(x_{ij}) \geqslant \alpha\} \\
&= \left[\min_{x_{ij}} \{x_{ij} \in S(\tilde{x}_{ij}) | \mu_{\tilde{x}_{ij}}(x_{ij}) \geqslant \alpha\}, \max_{x_{ij}} \{x_{ij} \in S(\tilde{x}_{ij}) | \mu_{\tilde{x}_{ij}}(x_{ij}) \geqslant \alpha\} \right] \quad \forall i, j,
\end{aligned}
$$

$$(y_{rj})_\alpha = \{y_{rj} \in S(\tilde{y}_{rj}) | \mu_{\tilde{y}_{rj}}(y_{rj}) \geqslant \alpha\}$$
$$= \left[\min_{y_{rj}} \{y_{rj} \in S(\tilde{y}_{rj}) | \mu_{\tilde{y}_{rj}}(y_{rj}) \geqslant \alpha\}, \max_{y_{rj}} \{y_{rj} \in S(\tilde{y}_{rj}) | \mu_{\tilde{y}_{rj}}(y_{rj}) \geqslant \alpha\}\right] \quad \forall r, j,$$

where $0 < \alpha \leqslant 1$. By setting different levels of confidence, namely $1 - \alpha$, fuzzy data are accordingly transformed into different *α-level* sets $\{(x_{ij})_\alpha | 0 < \alpha \leqslant 1\}$ and $\{(y_{rj})_\alpha | 0 < \alpha \leqslant 1\}$, which are all intervals. The widest input and output intervals will be $(x_{ij})_0 = \{x_{ij} \in S(\tilde{x}_{ij}) | \mu_{\tilde{x}_{ij}}(x_{ij}) > 0\} = [x_{ij}^{\mathrm{L}}, x_{ij}^{\mathrm{U}}]$ and $(y_{rj})_0 = \{y_{rj} \in S(\tilde{y}_{rj}) | \mu_{\tilde{y}_{rj}}(y_{rj}) > 0\} = [y_{rj}^{\mathrm{L}}, y_{rj}^{\mathrm{U}}]$, where x_{ij}^{L}, x_{ij}^{U}, y_{rj}^{L} and y_{rj}^{U} are the lower and upper bounds of fuzzy data $\tilde{x}_{ij}$ and $\tilde{y}_{rj}$, respectively. The production frontier will obviously be determined by interval data $[x_{ij}^{\mathrm{L}}, x_{ij}^{\mathrm{U}}]$ and $[y_{rj}^{\mathrm{L}}, y_{rj}^{\mathrm{U}}]$ $(i = 1, \ldots, m; j = 1, \ldots, n; r = 1, \ldots, s)$. Any *α-level* sets input and output data $(x_{ij})_\alpha = [(x_{ij})_\alpha^{\mathrm{L}}, (x_{ij})_\alpha^{\mathrm{U}}]$ and $(y_{rj})_\alpha = [(y_{rj})_\alpha^{\mathrm{L}}, (y_{rj})_\alpha^{\mathrm{U}}]$ should be measured using the identical production frontier. So, the interval DEA models for fuzzy input and output data will be as follows:

$$\begin{aligned}
\text{Maximize} \quad & (\theta_{j_0})_\alpha^{\mathrm{U}} = \sum_{r=1}^{s} u_r (y_{rj_0})_\alpha^{\mathrm{U}} \\
\text{subject to} \quad & \sum_{i=1}^{m} v_i (x_{ij_0})_\alpha^{\mathrm{L}} = 1 \\
& \sum_{r=1}^{s} u_r y_{rj}^{\mathrm{U}} - \sum_{i=1}^{m} v_i x_{ij}^{\mathrm{L}} \leqslant 0, \quad j = 1, \ldots, n, \\
& u_r, v_i \geqslant \varepsilon \quad \forall r, i.
\end{aligned} \tag{9}$$

$$\begin{aligned}
\text{Maximize} \quad & (\theta_{j_0})_\alpha^{\mathrm{L}} = \sum_{r=1}^{s} u_r (y_{rj_0})_\alpha^{\mathrm{L}} \\
\text{subject to} \quad & \sum_{i=1}^{m} v_i (x_{ij_0})_\alpha^{\mathrm{U}} = 1 \\
& \sum_{r=1}^{s} u_r y_{rj}^{\mathrm{U}} - \sum_{i=1}^{m} v_i x_{ij}^{\mathrm{L}} \leqslant 0, \quad j = 1, \ldots, n, \\
& u_r, v_i \geqslant \varepsilon \quad \forall r, i,
\end{aligned} \tag{10}$$

where $(\theta_{j_0})_\alpha^{\mathrm{U}}$ and $(\theta_{j_0})_\alpha^{\mathrm{L}}$ are, respectively, the upper and lower bounds of the best possible relative efficiency for DMU_0 under given *α-level* sets, which form an efficiency interval denoted by $(\theta_{j_0})_\alpha = [(\theta_{j_0})_\alpha^{\mathrm{L}}, (\theta_{j_0})_\alpha^{\mathrm{U}}]$.

Note that we use one production frontier for every *α-level* rather than different production frontiers for different *α-levels*. The main reason is that the real production frontier, we think, should not vary with *α-level* sets. If variable production frontiers were used for different *α-levels*, then the efficiencies under different *α-levels* would become incomparable.

Restricting ourselves to the cores and the supports of fuzzy input and output data, we may construct the following fuzzy DEA models to capture the fuzzy efficiency of DMU_0:

$$\begin{aligned} \text{Maximize } & \theta_{j_0}^{U} = \frac{\sum_{r=1}^{s} u_r y_{rj_0}^{U}}{\sum_{i=1}^{m} v_i x_{ij_0}^{L}} \\ \text{subject to } & \theta_{j}^{U} = \frac{\sum_{r=1}^{s} u_r y_{rj}^{U}}{\sum_{i=1}^{m} v_i x_{ij}^{L}} \leqslant 1, \quad j = 1, \ldots, n, \\ & u_r, v_i \geqslant \varepsilon \quad \forall r, i. \end{aligned} \tag{11}$$

$$\begin{aligned} \text{Maximize } & \theta_{j_0}^{M} = \frac{\sum_{r=1}^{s} u_r y_{rj_0}^{M}}{\sum_{i=1}^{m} v_i x_{ij_0}^{M}} \\ \text{subject to } & \theta_{j}^{U} = \frac{\sum_{r=1}^{s} u_r y_{rj}^{U}}{\sum_{i=1}^{m} v_i x_{ij}^{L}} \leqslant 1, \quad j = 1, \ldots, n, \\ & u_r, v_i \geqslant \varepsilon \quad \forall r, i. \end{aligned} \tag{12}$$

$$\begin{aligned} \text{Maximize } & \theta_{j_0}^{L} = \frac{\sum_{r=1}^{s} u_r y_{rj_0}^{L}}{\sum_{i=1}^{m} v_i x_{ij_0}^{U}} \\ \text{subject to } & \theta_{j}^{U} = \frac{\sum_{r=1}^{s} u_r y_{rj}^{U}}{\sum_{i=1}^{m} v_i x_{ij}^{L}} \leqslant 1, \quad j = 1, \ldots, n, \\ & u_r, v_i \geqslant \varepsilon \quad \forall r, i. \end{aligned} \tag{13}$$

The optimal objective function values of the above three fractional programming models, $\theta_{j_0}^{U*}$, $\theta_{j_0}^{M*}$ and $\theta_{j_0}^{L*}$, form the fuzzy efficiency of DMU_0, denoted by $[\theta_{j_0}^{L*}, \theta_{j_0}^{M*}, \theta_{j_0}^{U*}]$, which can be seen approximately as a triangular fuzzy number. These three models will be detailed and analyzed in another paper because a full comparison with the other existing fuzzy DEA models needs to be conducted and an effective fuzzy ranking approach needs to be developed to compare and rank the fuzzy efficiencies of DMUs.

It is clear from models (11)–(13) that the upper and lower bounds as well as modal value of a fuzzy efficiency are all determined by the constraints $(\sum_{r=1}^{s} u_r y_{rj}^{U} / \sum_{i=1}^{m} v_i x_{ij}^{L}) \leqslant 1$ $(j = 1, \ldots, n)$, which determine a fixed production frontier for all the DMUs with fuzzy inputs and outputs. So, another reason for us to use a fixed production frontier for all the α-*levels* is to keep the results obtained by two different modeling methodologies consistent with each other.

4. A minimax regret-based approach for comparing and ranking interval efficiencies

In interval efficiency assessment, since the final efficiency score for each DMU is characterized by an interval, a simple yet practical ranking approach is thus needed for comparing and ranking the efficiencies of different DMUs. A few approaches have already been developed to rank interval numbers, but they all have some shortcomings. Especially, when the interval numbers have the same center but different widths, they all fail to distinguish one from another. Interested readers may refer to Wang et al. [22] for more discussions on the existing approaches.

Here we introduce the minimax regret approach (MRA) developed by Wang et al. [22]. The approach has some attractive features and can be used to compare and rank the efficiency intervals of DMUs even if they are equi-centered but different in widths. The approach is summarized as follows.

Let $A_i = [a_i^{\mathrm{L}}, a_i^{\mathrm{U}}] = \langle m(A_i), w(A_i)\rangle$ $(i = 1, \ldots, n)$ be the efficiency intervals of n DMUs, where $m(A_i) = \frac{1}{2}(a_i^{\mathrm{R}} + a_i^{\mathrm{L}})$ and $w(A_i) = \frac{1}{2}(a_i^{\mathrm{R}} - a_i^{\mathrm{L}})$ are their midpoints (centers) and widths. Without loss of generality, suppose $A_i = [a_i^{\mathrm{L}}, a_i^{\mathrm{U}}]$ is chosen as the best efficiency interval. Let $b = \max_{j\neq i}\{a_j^{\mathrm{U}}\}$. Obviously, if $a_i^{\mathrm{L}} < b$, the DM might suffer the loss of efficiency (also called the loss of opportunity or regret) and feel regret. The maximum loss of efficiency he/she might suffer is given by

$$\max(r_i) = b - a_i^{\mathrm{L}} = \max_{j\neq i}\{a_j^{\mathrm{U}}\} - a_i^{\mathrm{L}}.$$

If $a_i^{\mathrm{L}} \geqslant b$, the DM will definitely suffer no loss of efficiency and feel no regret. In this situation, his/her regret is defined to be zero, i.e. $r_i = 0$. Combining the above two situations, we have

$$\max(r_i) = \max\left[\max_{j\neq i}(a_j^{\mathrm{U}}) - a_i^{\mathrm{L}}, 0\right].$$

Thus, the minimax regret criterion will choose the efficiency interval satisfying the following condition as the best (most desirable) efficiency interval:

$$\min_i\{\max(r_i)\} = \min_i\left\{\max\left[\max_{j\neq i}(a_j^{\mathrm{U}}) - a_i^{\mathrm{L}}, 0\right]\right\}.$$

Based on the analysis above, we give the following definition for comparing and ranking efficiency intervals.

Definition 2. Let $A_i = [a_i^{\mathrm{L}}, a_i^{\mathrm{U}}] = \langle m(A_i), w(A_i)\rangle (i = 1, \ldots, n)$ be a set of efficiency intervals. The maximum loss of efficiency (also called maximum regret) of each efficiency interval A_i is defined as

$$R(A_i) = \max\left[\max_{j\neq i}(a_j^{\mathrm{U}}) - a_i^{\mathrm{L}}, 0\right] = \max\left[\max_{j\neq i}\{m(A_j) + w(A_j)\} - (m(A_i) - w(A_i)), 0\right],$$
$$i = 1, \ldots, n.$$

It is evident that the efficiency interval with the smallest maximum loss of efficiency is the most desirable efficiency interval.

Since the maximum losses of efficiency are relative numbers. They are calculated according to the maximum efficiency among all the other efficiency intervals. So, they can only be used to choose the most desirable efficiency interval from among a set of efficiency intervals. But they cannot be used to rank them directly. To be able to generate a ranking for a set of efficiency intervals using the maximum losses of efficiency, the following eliminating steps are suggested:

Step 1: Calculate the maximum loss of efficiency of each efficiency interval and choose a most desirable efficiency interval that has the smallest maximum loss of efficiency (regret). Suppose A_{i_1} is selected, where $1 \leqslant i_1 \leqslant n$.

Step 2: Eliminate A_{i_1} from the consideration, recalculate the maximum loss of efficiency of every efficiency interval and determine a most desirable efficiency interval from the remaining $(n-1)$ efficiency intervals. Suppose A_{i_2} is chosen, where $1 \leqslant i_2 \leqslant n$ but $i_2 \neq i_1$.

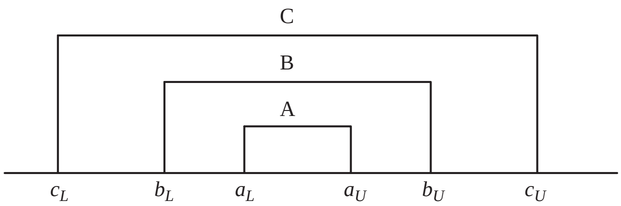

Fig. 2. Three equi-centered efficiency intervals.

Step 3: Eliminate A_{i_2} from the further consideration, re-compute the maximum loss of efficiency of each efficiency interval and determine a most desirable efficiency interval A_{i_3} from the remaining $(n-2)$ efficiency intervals.

Step 4: Repeat the above eliminating process until only one efficiency interval A_{i_n} is left. The final ranking is $A_{i_1} \succ A_{i_2} \succ \cdots \succ A_{i_n}$, where the symbol '$\succ$' means 'is superior to'.

The above ranking approach is referred to as the MRA. About the MRA, there exist the following properties (see Wang et al. [22] for proof).

Property 1. Let $A = [a_{\mathrm{L}}, a_{\mathrm{U}}]$ and $B = [b_{\mathrm{L}}, b_{\mathrm{U}}]$ be two efficiency intervals. If $a_{\mathrm{L}} \leqslant b_{\mathrm{L}}$ and $a_{\mathrm{U}} \leqslant b_{\mathrm{U}}$, then $R(A) \geqslant R(B)$.

Property 2. Let $A = [a_{\mathrm{L}}, a_{\mathrm{U}}] = \langle m(A), w(A) \rangle$ and $B = [b_{\mathrm{L}}, b_{\mathrm{U}}] = \langle m(B), w(B) \rangle$ be two efficiency intervals. If A is included in B, i.e. $a_{\mathrm{L}} \geqslant b_{\mathrm{L}}$ but $a_{\mathrm{U}} \leqslant b_{\mathrm{U}}$, then

(1) $R(A) > R(B)$ if $m(A) < m(B)$;
(2) $R(A) = R(B)$ if $m(A) = m(B)$;
(3) $R(A) < R(B)$ if $m(A) > m(B)$.

Property 3. Let $A = [a_{\mathrm{L}}, a_{\mathrm{U}}] = \langle m(A), w(A) \rangle$, $B = [b_{\mathrm{L}}, b_{\mathrm{U}}] = \langle m(A), w(B) \rangle$ and $C = [c_{\mathrm{L}}, c_{\mathrm{U}}] = \langle m(A), w(C) \rangle$ be three equi-centered efficiency intervals. If $w(A) < w(B) < w(C)$ (see Fig. 2), then $R(A) < R(B)$ and $R(A) < R(C)$.

Property 1 shows that for two non-nested efficiency intervals, the one with bigger, lower and upper bounds is preferred so the other. Property 2 shows how the MRA compares and ranks two efficiency intervals if one efficiency interval is included in another. In this situation, the order relationship generated using the MRA depends only on their centers if they are not the same. But if they are equi-centered, the MRA needs further to use Property 3 to compare and rank them. Property 3 shows that the efficiency interval with the same center but the smallest width is most desirable.

5. Illustrative examples

In this section, we examine two performance rating problems using the interval DEA models developed in this paper. One uses interval data; the other uses a mixture of exact data, interval data, fuzzy

Table 1
Data for seven DMUs with two inputs and one output

DMU	Inputs		Output
	Capital	Labor	Gross output value
1	[564 403, 621 755]	[674 111, 743 281]	[806 549, 866 063]
2	[614 371, 669 665]	[685 943, 742 345]	[917 507, 985 424]
3	[762 203, 798 427]	[762 207, 805 677]	[1 117 142, 1 195 562]
4	[862 016, 937 044]	[779 894, 846 496]	[1 206 179, 1 261 031]
5	[1 016 898, 1 082 662]	[799 714, 877 137]	[1 381 315, 1 462 543]
6	[1 164 350, 1 267 970]	[807 172, 889 416]	[1 497 679, 1 652 787]
7	[1 731 916, 1 816 008]	[818 090, 895 746]	[1 702 249, 1 812 655]

data and ordinal preference information. They can both be evaluated by using the proposed interval DEA models.

Example 1. Consider a performance measurement problem of manufacturing industry, in which there are seven manufacturing industries from different cities (DMUs) participating in the evaluation, each consuming two inputs (Capital and Labor) and producing one output (Gross output value). The data are all estimated and are thus imprecise and only known within the prescribed bounds, which are listed in Table 1.

Using the interval DEA models (5) and (6), we obtain the rating results listed in the second column of Table 2. Models were implemented in an MS-Excel worksheet and were solved by using the Excel Solver. The non-Archimedean infinitesimal was set to be $\varepsilon = 10^{-10}$. For comparison, in Table 2 we also give the interval efficiencies obtained by using the interval DEA models (1) and (2) developed by Despotis and Smirlis [5]. As can be seen from Table 2 that due to the use of variable production frontiers to measure the efficiencies of different DMUs, Despotis and Smirlis' interval DEA models assess all the seven DMUs to be DEA efficient. But in fact, when we utilize the fixed and unified production frontier to measure the efficiencies of all the seven DMUs, only DMU_1 and DMU_2 can probably be rated to be DEA efficient (efficient in scale). If they are in the state of the best production activity, they are DEA efficient; otherwise, they are also DEA inefficient.

Contrasting the lower bound efficiencies of each DMU obtained by two different interval DEA models, we find that the two different models (6) and (2) both generate the same lower bound efficiency for each DMU except for DMU_2. This is true for most of DMUs because the difference between (6) and (2) is only for DMU_0.

In order to compare and rank the efficiencies of the seven DMUs, we adopt the MRA to compute the maximum loss of efficiency for each DMU as follows:

$$R(\text{DMU}_1) = \max[\max(1, 0.9829, 0.9205, 0.9097, 0.9016, 0.6713) - 0.8088, 0] = 0.1912,$$

$$R(\text{DMU}_2) = \max[\max(1, 0.9829, 0.9205, 0.9097, 0.9016, 0.6713) - 0.8545, 0] = 0.1455,$$

$$R(\text{DMU}_3) = \max[\max(1, 1, 0.9205, 0.9097, 0.9016, 0.6713) - 0.8764, 0] = 0.1236,$$

Table 2
The interval efficiencies for the seven DMUs

DMU	Interval DEA models (5) and (6)		Interval DEA models (1) and (2)	
	$[\theta_{j0}^{L}, \theta_{j0}^{L}]$	Rank	$[H_{j0}^{L}, H_{j0}^{U}]$	Rank
1	[0.8088, 1.0000]	3	[0.8088, 1.0000]	4
2	[0.8545, 1.0000]	2	[0.8735, 1.0000]	2
3	[0.8764, 0.9829]	1	[0.8764, 1.0000]	1
4	[0.8100, 0.9205]	4	[0.8100, 1.0000]	3
5	[0.8062, 0.9097]	5	[0.8062, 1.0000]	5
6	[0.7499, 0.9016]	6	[0.7499, 1.0000]	6
7	[0.6007, 0.6713]	7	[0.6007, 1.0000]	7

$$R(\mathrm{DMU}_4) = \max[\max(1, 1, 0.9829, 0.9097, 0.9016, 0.6713) - 0.8100, 0] = 0.1900,$$

$$R(\mathrm{DMU}_5) = \max[\max(1, 1, 0.9829, 0.9205, 0.9016, 0.6713) - 0.8062, 0] = 0.1938,$$

$$R(\mathrm{DMU}_6) = \max[\max(1, 1, 0.9829, 0.9205, 0.9097, 0.6713) - 0.7499, 0] = 0.2501,$$

$$R(\mathrm{DMU}_7) = \max[\max(1, 1, 0.9829, 0.9205, 0.9097, 0.9016) - 0.6007, 0] = 0.3993.$$

Obviously, DMU_3 has the smallest maximum loss of efficiency. So, DMU_3 is rated as the best DMU and eliminated from the further consideration. Therefore, the remaining DMUs are DMU_1, DMU_2, DMU_4, DMU_5, DMU_6 and DMU_7, whose maximum losses of efficiency are recalculated as follows:

$$R(\mathrm{DMU}_1) = \max[\max(1, 0.9205, 0.9097, 0.9016, 0.6713) - 0.8088, 0] = 0.1912,$$

$$R(\mathrm{DMU}_2) = \max[\max(1, 0.9205, 0.9097, 0.9016, 0.6713) - 0.8545, 0] = 0.1455,$$

$$R(\mathrm{DMU}_4) = \max[\max(1, 1, 0.9097, 0.9016, 0.6713) - 0.8100, 0] = 0.1900,$$

$$R(\mathrm{DMU}_5) = \max[\max(1, 1, 0.9205, 0.9016, 0.6713) - 0.8062, 0] = 0.1938,$$

$$R(\mathrm{DMU}_6) = \max[\max(1, 1, 0.9205, 0.9097, 0.6713) - 0.7499, 0] = 0.2501,$$

$$R(\mathrm{DMU}_7) = \max[\max(1, 1, 0.9205, 0.9097, 0.9016) - 0.6007, 0] = 0.3993.$$

Among the above regrets, the maximum loss of efficiency of DMU_2 is the smallest, so DMU_2 is rated as the second best DMU and eliminated from the further consideration. The remaining DMUs are DMU_1, DMU_4, DMU_5, DMU_6 and DMU_7, whose maximum losses of efficiency are recomputed and shown below:

$$R(\mathrm{DMU}_1) = \max[\max(0.9205, 0.9097, 0.9016, 0.6713) - 0.8088, 0] = 0.1117,$$

$$R(\mathrm{DMU}_4) = \max[max(1, 0.9097, 0.9016, 0.6713) - 0.8100, 0] = 0.1900,$$

$$R(\mathrm{DMU}_5) = \max[\max(1, 0.9205, 0.9016, 0.6713) - 0.8062, 0] = 0.1938,$$

$$R(\mathrm{DMU}_6) = \max[\max(1, 0.9205, 0.9097, 0.6713) - 0.7499, 0] = 0.2501,$$

$$R(\mathrm{DMU}_7) = \max[\max(1, 0.9205, 0.9097, 0.9016) - 0.6007, 0] = 0.3993.$$

Table 3
Data for eight DMUs with two inputs and two outputs

DMU	Inputs		Outputs	
	Purchase cost	Number of employees	Gross output value	Product quality[a]
1	2166	1875	[14 548, 14 950]	2
2	1455	1342	[12 468, 13 045, 13 584]	7
3	2562	2359	[17 896, 18 452]	1
4	2346	2018	[14 968, 15 673, 15 900]	3
5	1517	1548	[13 980, 14 638]	8
6	2034	1760	[14 026, 14 324, 14 582]	6
7	2256	1982	[16 542, 17 169]	5
8	2465	2254	[17 600, 18 256]	4

[a]Ordinal scale from 1 = best to 8 = worst with the preference intensity parameter $\chi_2 = 1.12$ and the ratio parameter $\sigma_2 = 0.1$.

Among the above regrets, DMU_1 has the smallest maximum loss of efficiency. So, DMU_1 is rated as the third best DMU and eliminated from the further consideration. The remaining DMUs are now DMU_4, DMU_5, DMU_6 and DMU_7, whose maximum losses of efficiency need recalculating again:

$$R(\text{DMU}_4) = \max[\max(0.9097, 0.9016, 0.6713) - 0.8100, 0] = 0.0997,$$

$$R(\text{DMU}_5) = \max[\max(0.9205, 0.9016, 0.6713) - 0.8062, 0] = 0.1143,$$

$$R(\text{DMU}_6) = \max[\max(0.9205, 0.9097, 0.6713) - 0.7499, 0] = 0.1706,$$

$$R(\text{DMU}_7) = \max[\max(0.9205, 0.9097, 0.9016) - 0.6007, 0] = 0.3198.$$

Since DMU_4 has the smallest maximum loss of efficiency, it is rated as the best among the four DMUs. Repeating the above process, we finally get the ranking order of the seven DMUs as $\text{DMU}_3 \succ \text{DMU}_2 \succ \text{DMU}_1 \succ \text{DMU}_4 \succ \text{DMU}_5 \succ \text{DMU}_6 \succ \text{DMU}_7$. In the same way we get the ranking order of the seven DMU_s as $\text{DMU}_3 \succ \text{DMU}_2 \succ \text{DMU}_4 \succ \text{DMU}_1 \succ \text{DMU}_5 \succ \text{DMU}_6 \succ \text{DMU}_7$ for the efficiencies generated by Despotis and Smirlis' interval DEA models, which is slightly different from ours in the relationship between DMU_1 and DMU_4.

Example 2. Consider another performance-rating problem of manufacturing industry, in which there are eight manufacturing enterprises (DMUs) joining in the performance rating. Each manufacturing enterprise manufactures the same type of product, but the qualities are different. Therefore, both the gross output value (GOV) and the product quality (PQ) are considered as outputs. The inputs include purchase cost (PC) and the number of employees (NOE), whose data are known exactly. The data about the gross output values, however, are imprecise due to the unavailability at the moment and are thus estimated. Some of them are given as interval numbers and some as triangular fuzzy numbers. The product quality is a qualitative index and is given as strong ordinal preference information that is obtained from the evaluation of customers to their products. The data are presented in Table 3.

Table 4
The input–output data for the eight DMUs after the transformation of ordinal preference information

DMU	Inputs		Outputs	
	PC	NOE	GOV	PQ
1	2166	1875	[14 548, 14 950]	[0.1973823, 0.892857]
2	1455	1342	[12 468, 13 045, 13 584]	[0.1120000, 0.506631]
3	2562	2359	[17 896, 18 452]	[0.2210681, 1.000000]
4	2346	2018	[14 968, 15 673, 15 900]	[0.1762342, 0.797194]
5	1517	1548	[13 980, 14 638]	[0.1000000, 0.452349]
6	2034	1760	[14 026, 14 324, 14 582]	[0.1254400, 0.567427]
7	2256	1982	[16 542, 17 169]	[0.1404928, 0.635518]
8	2465	2254	[17 600, 18 256]	[0.1573519, 0.711780]

Suppose the preference intensity parameter and the ratio parameter about the strong ordinal preference information are given (or estimated) as $\chi_2 = 1.12$ and $\sigma_2 = 0.1$, respectively. Using the transformation technique described in Section 3.1, we can derive an interval estimate for the product quality of each DMU, which is shown in the last column of Table 4.

Since the GOV index for DMU_2, DMU_4 and DMU_6 is given in the form of triangular fuzzy number, i.e. $\text{GOV}_j = (\text{GOV}_j^{\text{L}}, \text{GOV}_j^{\text{M}}, \text{GOV}_j^{\text{U}})(j = 2, 4, 6)$, their membership functions can be expressed as

$$\mu_{\text{GOV}_j}(x_j) = \begin{cases} \dfrac{x_j - \text{GOV}_j^{\text{L}}}{\text{GOV}_j^{\text{M}} - \text{GOV}_j^{\text{L}}}, & \text{GOV}_j^{\text{L}} \leqslant x_j \leqslant \text{GOV}_j^{\text{M}} \\ \dfrac{\text{GOV}_j^{\text{U}} - x_j}{\text{GOV}_j^{\text{U}} - \text{GOV}_j^{\text{M}}}, & \text{GOV}_j^{\text{M}} \leqslant x_j \leqslant \text{GOV}_j^{\text{U}} \\ 0, & x_j \notin [\text{GOV}_j^{\text{L}}, \text{GOV}_j^{\text{U}}] \end{cases}, \quad j = 2, 4, 6$$

where GOV_j^{L}, GOV_j^{M} and GOV_j^{U} are the lower bound, most likely and upper bound values of GOV_j, respectively. For a given α-*level*, the corresponding α-*level* sets are given by

$$\begin{aligned}(\text{GOV}_j)_\alpha &= \{x_j \in S(\text{GOV}_j) | \mu_{\text{GOV}_j}(x_j) \geqslant \alpha\} \\ &= [(\text{GOV}_j)_\alpha^{\text{L}}, (\text{GOV}_j)_\alpha^{\text{U}}] \\ &= [\text{GOV}_j^{\text{L}} + \alpha(\text{GOV}_j^{\text{M}} - \text{GOV}_j^{\text{L}}), \text{GOV}_j^{\text{U}} - \alpha(\text{GOV}_j^{\text{U}} - \text{GOV}_j^{\text{M}})], \quad j = 2, 4, 6.\end{aligned}$$

As for exact data, they can be viewed as a special case of interval data with the lower and upper bounds being equal. Therefore, all the input and output data are now transformed into interval numbers and can be evaluated using interval DEA models. Table 5 reports the results of efficiency assessments for the eight DMUs under five different α-*levels* sets obtained by using interval DEA models (5) and (6).

It can be seen from Table 5 that the efficiency intervals for DMU_1, DMU_3, DMU_5, DMU_7 and DMU_8 do not vary with the α-*level* sets. This is because α-*level* has no impact on the exact data, interval data and ordinal data and the efficiencies of DMUs are measured on the basis of a fixed and unified production frontier, which consists of the best production activity of each DMU. From the upper bound efficiencies under $\alpha = 0$ level sets, we know that DMU_1, DMU_2 and DMU_5 all have the possibility to be DEA

Table 5
The efficiency intervals for the eight DMUs under different α-*level* sets

DMU	$\alpha = 0$	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 0.75$	$\alpha = 1$
1	[0.7665, 1.0000]	[0.7665,1.0000]	[0.7665,1.0000]	[0.7665, 1.0000]	[0.7665, 1.0000]
2	[0.9178, 1.0000]	[0.9285, 0.9959]	[0.9391, 0.9918]	[0.9497, 0.9877]	[0.9603, 0.9836]
3	[0.7495, 0.9764]	[0.7495, 0.9764]	[0.7495, 0.9764]	[0.7495, 0.9764]	[0.7495, 0.9764]
4	[0.7328, 0.8985]	[0.7414, 0.8970]	[0.7500, 0.8955]	[0.7587, 0.8939]	[0.7673, 0.8924]
5	[0.9550, 1.0000]	[0.9550, 1.0000]	[0.9550, 1.0000]	[0.9550, 1.0000]	[0.9550, 1.0000]
6	[0.7873, 0.8344]	[0.7915, 0.8324]	[0.7957, 0.8304]	[0.7999, 0.8284]	[0.8040, 0.8264]
7	[0.8245, 0.8558]	[0.8245, 0.8558]	[0.8245, 0.8558]	[0.8245, 0.8558]	[0.8245, 0.8558]
8	[0.7714, 0.8164]	[0.7714, 0.8164]	[0.7714, 0.8164]	[0.7714, 0.8164]	[0.7714, 0.8164]

Table 6
The ranking order for the eight DMUs under different α-*level* sets

DMU	$\alpha = 0$	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 0.75$	$\alpha = 1$
1	4	5	5	5	5
2	2	2	2	2	1
3	5	6	6	6	6
4	6	7	7	7	7
5	1	1	1	1	2
6	7	4	4	4	4
7	3	3	3	3	3
8	8	8	8	8	8

efficient. If they are able to use the minimum inputs to produce the maximum outputs, they are DEA efficient (efficient in scale); otherwise, they are not DEA efficient. Although DMU_1, DMU_2 and DMU_5 all have the possibility to be DEA efficient, due to the differences in the lower bound efficiencies, their performances are in fact different. Table 6 shows the ranking orders of the eight DMUs under different α-*level* sets obtained by using the minimax regret approach.

From Table 6 we see that DMU_8 is always ranked at the last place under any α-*level* sets and DMU_5 and DMU_2 are always ranked at the first two places. Therefore, we are certain that DMU_8 performs the worst among the eight DMUs while DMU_5 and DMU_2 perform better than any others. The rankings for the other DMUs almost have no or little changes except for DMU_6. At $\alpha = 0$ level, DMU_6 is ranked at the seventh place, but at other α-*levels*, it is ranked at the fourth place. Since the membership function $\mu_{GOV_6}(x)$ takes zero at $\alpha = 0$, we think DMU_6 has very little possibility to be ranked at the seventh place. It ranks at the fourth place seems more believable.

But generally speaking, the rankings under different α-*levels* might be quite different. In this situation, the overall ranking cannot be easily observed intuitively. In order to generate an overall ranking, the following alternative ways could be suggested: (1) choosing a trade-off between the precision and the confidence. High α means precision of the interval chosen and low α means high confidence in the result. A risk-averse assessor or DM might choose a high alpha because of strong dislike of uncertainty (fuzziness), while a risk-taking assessor or DM might prefer a low alpha because of seeking of risk; (2) averaging

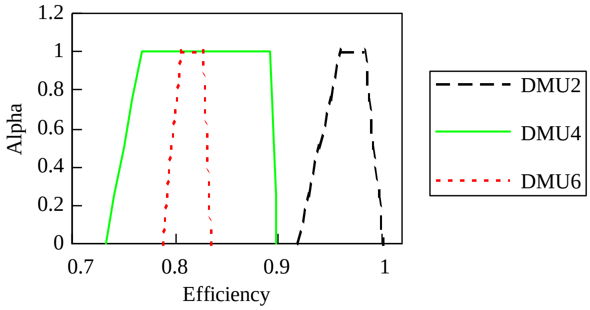

Fig. 3. Membership functions of efficiency for three DMUs.

the efficiencies using alpha as weights. This gives precise efficiency intervals more weights, but takes no account of confidence; (3) constructing the membership functions through the efficiencies under different α-*levels* and comparing them using fuzzy ranking approaches. Fig. 3 shows the membership functions of efficiency for DMU_2, DMU_4 and DMU_6, which are generated from Table 5. The in-depth discussions about fuzzy ranking approaches will be discussed in another paper of ours.

6. Conclusions

In this paper we have developed a new pair of interval DEA models for dealing with imprecise data such as interval data, ordinal preference information, fuzzy data and their mixture. Compared with the IDEA model developed by Cooper et al. [2–4], our interval DEA models are much easier to understand and more convenient to use. First of all, there exist no scale transformations and variable alternations, which makes our models more concise. Next, our interval DEA models keep the advantage of original DEA model and do not impose any extra constraints on it, which makes our models easier and more convenient to use. Finally, except calculating the best possible relative efficiency, our DEA models also compute the best possible lower bound efficiency, which makes our models more powerful and more practical. Compared with the interval DEA models developed by Despotis and Smirlis [5], our interval DEA models utilize a fixed and unified production frontier as a benchmark to measure the efficiencies of all DMUs, which makes our models more rational and more reliable. The use of a fixed and unified production frontier also simplifies to a great extent the computation of the efficiencies of those DMUs without any fuzziness because α-*level* has no impact on their efficiencies, there is no need to recalculate them for different α-*levels*. Moreover, the means we treat ordinal preference information also seems more reasonable than the way Zhu [23] did. Two numerical examples have illustrated the advantages, potential and applications of our interval DEA models.

Acknowledgements

The authors would like to thank the Editors-in-Chief, D. Dubois and H. Prade, and two anonymous referees for their constructive comments and suggestions that have helped to improve the quality of the paper to its current standard.

References

[1] A. Charnes, W.W. Cooper, E. Rhodes, Measuring the efficiency of decision making units, European J. Oper. Res. 2 (1978) 429–444.

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

[3] W.W. Cooper, K.S. Park, G. Yu, An illustrative application of IDEA (imprecise data envelopment analysis) to a Korean mobile telecommunication company, Oper. Res. 49 (2001) 807–820.

[4] W.W. Cooper, K.S. Park, G. Yu, IDEA (imprecise data envelopment analysis) with CMDs (column maximum decision making units), J. Oper. Res. Soc. 52 (2001) 176–181.

[5] D.K. Despotis, Y.G. Smirlis, Data envelopment analysis with imprecise data, European J. Oper. Res. 140 (2002) 24–36.

[6] T. Entani, Y. Maeda, H. Tanaka, Dual models of interval DEA and its extension to interval data, European J. Oper. Res. 136 (2002) 32–45.

[7] P. Guo, H. Tanaka, Fuzzy DEA: a perceptual evaluation method, Fuzzy Sets and Systems 119 (2001) 149–160.

[8] M.S. Haghighat, E. Khorram, The maximum and minimum number of efficient units in DEA with interval data, Appl. Math. Comput. (2004), in press.

[9] G.R. Jahanshahloo, F. Hosseinzadeh Lofti, M. Moradi, Sensitivity and stability analysis in DEA with interval data, Appl. Math. Comput. 156 (2004) 463–477.

[10] G.R. Jahanshahloo, R.K. Matin, A.H. Vencheh, On return to scale of fully efficient DMUs in data envelopment analysis under interval data, Appl. Math. Comput. 154 (2004) 31–40.

[11] G.R. Jahanshahloo, R.K. Matin, A.H. Vencheh, On FDH efficiency analysis with interval data, Appl. Math. Comput. 159 (2004) 47–55.

[12] C. Kao, S.T. Liu, Fuzzy efficiency measures in data envelopment analysis, Fuzzy Sets and Systems 113 (2000) 427–437.

[13] C. Kao, S.T. Liu, A mathematical programming approach to fuzzy efficiency ranking, Internat. J. Production Econom. 86 (2003) 45–154.

[14] S.H. Kim, C.G. Park, K.S. Park, An application of data envelopment analysis in telephone offices evaluation with partial data, Comput. Oper. Res. 26 (1999) 59–72.

[15] Y.K. Lee, K.S. Park, S.H. Kim, Identification of inefficiencies in an additive model based IDEA (imprecise data envelopment analysis), Comput. Operat. Res. 29 (2002) 1661–1676.

[16] T. León, V. Liern, J.L. Ruiz, I. Sirvent, A fuzzy mathematical programming approach to the assessment of efficiency with DEA models, Fuzzy Sets and Systems 139 (2003) 407–419.

[17] S. Lertworasirikul, S.C. Fang, J.A. Joines, H.L.W. Nuttle, Fuzzy data envelopment analysis (DEA): a possibility approach, Fuzzy Sets and Systems 139 (2003) 379–394.

[18] S. Saati, A. Memariani, Reducing weight flexibility in fuzzy DEA, Appl. Math. Comput. 161 (2005) 611–622.

[19] S. Saati, A. Memariani, G.R. Jahanshahloo, Efficiency analysis and ranking of DMUs with fuzzy data, Fuzzy Optim. Decis. Mak. 1 (2002) 255–267.

[20] J.K. Sengupta, A fuzzy systems approach in data envelopment analysis, Comput. Math. Appl. 24 (1992) 259–266.

[21] K. Triantis, O. Girod, A mathematical programming approach for measuring technical efficiency in a fuzzy environment, J. Productivity Anal. 10 (1998) 85–102.

[22] Y.M. Wang, J.B. Yang, D.L. Xu, Two approaches for ranking interval numbers based on decision making under uncertainty, Decis. Support System, submitted for publication.

[23] J. Zhu, Imprecise data envelopment analysis (IDEA): a review and improvement with an application, European J. Oper. Res. 144 (2003) 513–529.

[24] H.J. Zimmermann, Fuzzy Set Theory and its Applications, second ed., Kluwer-Nijhoff, Boston, 1991.