Applied Soft Computing 46 (2016) 204–219

Contents lists available at ScienceDirect

Applied Soft Computing

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

An integrated data envelopment analysis and free disposal hull framework for cost-efficiency measurement using rough sets

Rashed Khanjani Shiraz[a], Hirofumi Fukuyama[b], Madjid Tavana[c,d,*], Debora Di Caprio[e,f]

[a] School of Mathematics Science, University of Tabriz, Tabriz, Iran
[b] Faculty of Commerce, Fukuoka University, Fukuoka, Japan
[c] Business Systems and Analytics Department, Distinguished Chair of Business Analytics, La Salle University, Philadelphia, PA 19141, USA
[d] Business Information Systems Department, Faculty of Business Administration and Economics, University of Paderborn, D-33098 Paderborn, Germany
[e] Department of Mathematics and Statistics, York University, Toronto M3J 1P3, Canada
[f] Polo Tecnologico IISS G. Galilei, Via Cadorna 14, 39100 Bolzano, Italy

ARTICLE INFO

Article history:
Received 9 November 2014
Received in revised form 8 April 2016
Accepted 30 April 2016
Available online 6 May 2016

Keywords:
Data envelopment analysis (DEA)
Free disposal hull (FDH)
Rough sets
Cost-efficiency

ABSTRACT

Traditional cost-efficiency analysis methods require exact and precise values for inputs, outputs and input prices. However, this is not the case in many real-life applications. This study proposes a rough cost-efficiency approach to the problem of ranking efficient decision making units (DMUs). Based on rough set theory, a nonparametric methodology for cost-efficiency analysis is developed. The merits of this methodology include computational ease and the capacity to incorporate data uncertainty. Furthermore, it applies to both convex data envelopment analysis (DEA) and non-convex free disposal hull (FDH) technologies under different returns-to-scale assumptions. A numerical example and a real-life case study in the Japanese banking industry demonstrate the applicability of the proposed framework. In particular, the rankings of the DMUs resulting from the proposed models are compared with those obtained using the maximum technical efficiency loss index.

1. Introduction

In this study, we develop two cost-efficiency models based on data envelopment analysis (DEA) and free disposal hull (FDH), respectively. The purpose of the proposed models is to provide a method to evaluate the cost-efficiency of decision-making units (DMUs) when there exists a significant level of impreciseness in the data while allowing for different returns-to-scale assumptions.

1.1. Traditional DEA and FDH models versus imprecise data

The traditional DEA model, initially proposed by Charnes et al. [7] and extended by Banker et al. [3], is a nonparametric mathematical programming method for evaluating the relative efficiency of decision-making units characterized by crisp multiple inputs and outputs. It consists of solving a fractional linear programming problem through an equivalent linear programming formulation assuming convexity and constant returns-to-scale (CRS).

The traditional DEA model has received considerable attention in both theory and applications since the very beginning (see Ref. [18] or Ref. [12] for a comprehensive bibliography) quickly becoming an important research tool in management science, operations research, and decision theory. Regarding, in particular, the study of cost-efficiency and DMUs' performances, Färe et al. [15] operationalized Farrell's [16] cost-efficiency notion. This cost-efficiency measure requires operating with crisp input and output data.

* Corresponding author at: Business Systems and Analytics Department, Distinguished Chair of Business Analytics, La Salle University, Philadelphia, PA 19141, USA. Fax: +1 2672952854.

E-mail addresses: Rashed.shiraz@gmail.com (R. Khanjani Shiraz), fukuyama@fukuoka-u.ac.jp (H. Fukuyama), tavana@lasalle.edu (M. Tavana), dicaper@mathstat.yorku.ca, debora.dicaprio@istruzione.it (D. Di Caprio).

http://dx.doi.org/10.1016/j.asoc.2016.04.043
1568-4946/

The traditional FDH model, proposed by Deprins et al. [8], is an other nonparametric deterministic method useful to evaluate the technical efficiency of DMUs. It exploits the input–output disposability without imposing the convexity assumption. Among others, Tulkens [52] presented a FDH-based mixed integer linear programming while Kerstens and Vanden Eeckaut [27] introduced various returns-to-scale specifications, namely, non-increasing, non-decreasing, and constant returns-to-scale. Agrell and Tind [1] suggested a linear programming FDH model that was later extended by Leleu [30] to various returns-to-scale technologies using cost functions.

Both the traditional DEA and FDH models require the values of inputs and outputs to be known precisely. However, quite often, the observed data are imprecise, vague or taken at one point in time, which may not reflect the overall distribution of the data (such as, for example, the total working hours of employees). In this regard, Hougaard [21] stated three reasons due to which the traditional DEA model is inadequate to model some real life situations and evaluate the relative efficiency scores. First, the efficiency scores are very sensitive to changes in the data and to errors in the estimation of the efficient frontier of the technology. Second, the relative quality differences among decision-making units (DMUs) in terms of inputs and outputs, may distort the true efficiency performance landscape. Third, efficiency scores are simply approximations of the DMU's unknown preferences. The traditional FDH model shows the same limitations.

1.2. Main shortcomings of interval models with imprecise data

Many performance measurement techniques based on interval values have been proposed to account for the impreciseness and vagueness of data relative to production technologies.

Entani et al. [14] suggested evaluating the efficiency of a DMU by an interval delimited by a pessimistic and an optimistic measure. Despotis and Smirlis [9] introduced an interval DEA method to deal with imprecise data. Extending Despotis and Smirlis' [9] method, Kao [26] constructed a two-level mathematical programming model to facilitate the calculation of efficiency intervals for ordinal data. Kao [26] also supported the interval approach from a psychological viewpoint arguing that DMUs are willing to accept interval measures better than crisp ones since the former ones do not directly imply that the performance of a DMU is worse than that of others. Inuiguchi and Mizoshita [22] examined some DEA models with interval input-output data discussing how to obtain lower and upper bounds for the efficiency scores.

Although frequently used to model real-life problems, the interval models and the corresponding solution methods may not be versatile enough to provide a satisfactory evaluation of DMUs such as banks or manufacturing firms.

Three different approaches have been suggested through the years for handling vagueness and impreciseness in DEA models: (1) stochastic, (2) fuzzy, and (3) rough.

The stochastic DEA approach makes it possible to replace crisp data with statistical or probabilistic values. In this approach, chance-constrained formulations are usually introduced and the uncertainty relative to the available data incorporated by interpreting inputs and outputs as random variables whose cumulative distribution functions and probability density functions are known. However, solving stochastic DEA with various cumulative and probability distributions is often very complex from the computational and, hence, practical viewpoint.

The fuzzy DEA approach uses membership functions to model uncertainty and imprecise data sets while the rough DEA approach centers on lower and upper approximations of crisp sets.

All the DEA approaches described above focus on constructing interval values. Note though that, properly speaking, only stochastic and fuzzy DEA are interval methods since they require the existence of lower and upper bounds of inputs and outputs. On the other hand, the rough DEA approach allows for a more general interpretation of data uncertainty since rough variables are used to define lower and the upper approximations of the interval values corresponding to the available inputs and outputs.

Hence, the main shortcoming of interval models, that is, the fact that they can be applied only to situations where inputs and outputs are already endowed with both a lower and an upper bound, is directly inherited from stochastic and fuzzy DEA while it remains a tangential issue in rough DEA.

Besides the aforementioned shortcomings, rough DEA also has the same limitations of a classical DEA model, that is, it applies only to convex technologies and usually assumes constant returns-to-scale. Thus, in order to complement our analysis and extend the efficiency evaluation problem to the non-convex case, a rough FDH approach is necessary to integrate the DEA framework. Moreover, a rough FDH model also allows for considering different returns-to-scale assumptions.

1.3. Contribution

In summary, there already exists a very ample literature on DEA and FDH, most of which accounting for the uncertainty characterizing the inputs and outputs of many real-life situations where a measure of the relative efficiencies of a set of DMUs is required. In particular, the flexibility of the DEA approach has been exploited to solve a myriad of input–output efficiency evaluation models, including chance-constrained ones, whose inputs and outputs are fuzzy, random, rough, random-rough or fuzzy-rough variables.

All the models proposed in the previous studies are at some extend based on the interval approach and each of them presents difficulties and advantages when dealing with specific situations.

The current study focuses on:

- Developing a cost-efficiency model for convex technologies able to incorporate the impreciseness of inputs, outputs and input prices within the DEA framework while reducing the computational complexity of the traditional approach;
- Developing a rough FDH approach to the cost-efficiency evaluation problem for non-convex technologies so as to complement rough DEA cost-efficiency measures when non-constant returns-to-scale are assumed.

An attempt to propose rough DEA as the most appropriate framework to deal with efficiency evaluation problems was made by Xu et al. [54] who developed a DEA model with rough parameters to evaluate the performance of real supply chains. However, to the best of our knowledge, none of the existing studies has proposed a cost-efficiency model where imprecise inputs, outputs and input prices are all dealt with simultaneously.

The proposed rough DEA and rough FDH models use the α-pessimistic and α-optimistic operators to produce a cost-efficiency interval for each DMU such that only one of the extremes suffices to establish whether the DMU is efficient.

We show that the cost-efficiency evaluation intervals resulting from the proposed models contain the corresponding Farrell and standard FDH cost-efficiency values. Thus, a DMU is efficient according to our models if it is efficient according to both the Farrell and FDH approaches. This makes the proposed rough DEA and rough FDH models reliable extensions of the traditional ones. Finally, a simple variation of the maximum technological loss index of Wang et al. [53] and Xu et al. [54] allows validating the rankings of the efficient DMUs obtained by implementing the proposed rough models.

A numerical example and a real-life case study in the Japanese banking industry are discussed to show the efficacy and applicability of the proposed models.

The remainder of the paper is organized as follows. Section 2 provides a selective survey on cost-efficiency evaluation and imprecise data. Section 3 presents the basics of the traditional DEA and FDH models, and some background notions of rough set theory. Sections 4 and 5 develop the proposed DEA and FDH cost-efficiency models with rough variables, respectively. Sections 6 and 7 illustrate the proposed methods by means of a numerical example and a real-life case study, respectively. Section 8 presents the conclusions.

2. Literature review: cost-efficiency evaluation and imprecise data

The first study dealing with imprecise data was proposed by Thompson et al. [51], who obtained an upper cost-efficiency bound by restricting the multiplier weights. Kuosmanen and Post [28,29] showed how to compute upper and lower cost-efficiency bounds by assuming that prices are random variables. Camanho and Dyson [6] developed some models for estimating cost-efficiency bounds in complex uncertainty scenarios where input prices are assumed to be expressed as ranges. Mostafaee and Saljooghi [35] carried out a deeper analysis of Camanho and Dyson's [6] model proposing a two-level mathematical programming problem for calculating the lower and upper bounds of Farrell cost-efficiency. Finally, Bagherzadeh Valami [4] extended the traditional Farrell cost-efficiency model to a framework where input prices are treated as triangular fuzzy numbers.

Fuzzy DEA is usually computationally complicated due to the use of nonlinear programming. Lertworasirikul et al. [31] classified fuzzy DEA papers into four groups according the approach followed: (1) tolerance approach, (2) α-level based approach, (3) fuzzy ranking approach, and (4) possibility approach. Sengupta [48] developed the tolerance approach where equalities and inequalities are fuzzified without directly dealing with fuzzy coefficients. The α-level based approach aims at converting a fuzzy DEA model into a pair of parametric programs to find lower and upper bounds for the membership functions of efficiency scores at the predetermined α-level. See, among others, Saati et al. [44], who also propose a ranking method for fuzzy DMUs using the fuzzy DEA approach. In the fuzzy ranking approach, fuzzy efficiency scores are obtained by means of fuzzy linear programs. Finally, the possibility approach is based on Zadeh's [58] fuzzy set theory and the possibility theory extensions proposed by Dubois and Prade [10]. Relative to the last approach, see also Guo et al. [19].

Hatami-Marbini et al. [20] added a fifth group to those identified by Lertworasirikul et al. [31] and referred to it as "other developments". Among the thirteen papers cited as examples of other developments, it deserves to mention Hougaard [21] and Sheth and Triantis [49]. Hougaard [21] modeled DEA efficiency measures via fuzzy intervals while treating the observations as exact values. Sheth and Triantis [49] introduced a fuzzy goal DEA model for assessing efficiency and effectiveness. For additional references on fuzzy DEA classifications, see also Emrouznejad and Tavana [13].

Finally, Ma and Cui [34] proposed a novel approach that tries to overcome the shortages concerning various aspects of fuzzy DEA, such as α-cut approaches, types of fuzzy numbers, and ranking techniques. They proposed a fuzzy DEA model where fuzzy sample DMUs are tested through five evaluation approaches and ranking methods.

An alternative valid theoretical framework that can be used to account for imprecise input-output data is provided by rough set theory (Pawlak [41,42]). For a detailed review on rough set theory, the readers can refer to Liu [33], among others.

Rough sets and rough variables have been extensively used and often combined with random or fuzzy approaches. Among others, Pawlak and Skowron [38–40] extensively studied rough sets and their applications. Dubois and Prade [11] extended rough set theory into the fuzzy framework. Tao and Xu [50] presented a rough multiple objective programming model for a solid transportation problem. Nguyen [36] successfully applied rough set theory to feature selection, attribute reduction, and rule learning problems.

Xu et al. [54] proposed a rough DEA approach comparing it with the standard DEA approach when solving problems where impreciseness and vagueness are inherent in the decisions making process. Xu and Yao [55,56] examined a class of multiobjective programming problems with random rough variables discussing how to transform a chance-constrained model with random rough variables into a crisp equivalent model. They used random rough simulations to deal with general random rough objective functions and random rough constraints, which are usually hard to transform into a deterministic form.

Xu and Zhao [57] used fuzzy rough variables to deal with the imprecise data of an inventory problem. More recently, Khanjani et al. [25] proposed a DEA model with fuzzy rough inputs and outputs.

3. Preliminaries

Let DMU$_j$ $(j = 1, \ldots, J)$ be the DMUs to evaluate. Assume each DMU$_j$ to have N inputs and M outputs denoted by x_{nj} $(n = 1, \ldots, N)$ and y_{mj} $(m = 1, \ldots, M)$, respectively. All observed inputs and outputs are assumed positive. Let p_{nj} $(n = 1, \ldots, N; \quad j = 1, \ldots, J)$ be the price of the input x_{nj} of DMU$_j$. Let u_m $(m = 1, \ldots, M)$ and v_n $(n = 1, \ldots, N)$ be the multiplier weights of the m-th output and n-th input, respectively. These weights are common to all DMUs. Finally, let * indicate the optimal solution of an optimization problem.

3.1. Traditional DEA and Farrell cost-efficiency

This section provides the basics of traditional DEA and cost-efficiency models. The traditional input-oriented DEA model is a fractional linear programming problem that is solved by transforming it into the following linear programming problem (known as the CCR model):

$$\begin{aligned}&\text{Max}\sum_{m=1}^{M}u_m y_{mk}\\&\text{subject to :}\\&\sum_{n=1}^{N}v_n x_{nk}=1,\ \sum_{m=1}^{M}u_m y_{mj}-\sum_{n=1}^{N}v_n x_{nj}\le 0,\ j=1,\dots,J\\&u_m,v_n\ge 0,\ m=1,\dots,M;\ n=1,\dots,N\end{aligned}\tag{1}$$

The dual of Eq. (1) is the following input-oriented model:

$$\begin{aligned}&\text{Min}\theta\\&\text{subject to :}\\&\sum_{j=1}^{J}\lambda_j x_{nj}\le\theta x_{nk},\ n=1,\dots,N\\&\sum_{j=1}^{J}\lambda_j y_{mj}\ge y_{mk},\ m=1,\dots,M\\&\lambda_j\ge 0,\ j=1,\dots,J\end{aligned}\tag{2}$$

If $\theta^*=1$, DMU_k is technically input-efficient; if $\theta^*\ne 1$, DMU_k is technically input-inefficient.

Farrell [16] proposed a cost-efficiency model with crisp data. This model assesses the ability of a DMU to produce the current outputs at minimal cost given its input prices. The Farrell cost-efficiency score is obtained in two steps: (a) find the optimal value for the standard cost function by solving the following problem:

$$\begin{aligned}&\text{Min}\sum_{n=1}^{N}p_{nk}x_n\\&\text{subject to :}\\&\sum_{j=1}^{J}\lambda_j y_{mj}\ge y_{mk},\ m=1,\dots,M\\&\sum_{j=1}^{J}\lambda_j x_{nj}\le x_n,\ n=1,\dots,N\\&\lambda_j\ge 0,\ j=1,\dots,J,\ x_n\ge 0,\ n=1,\dots,N\end{aligned}\tag{3}$$

where $x_n\ \ (n=1,\dots,N)$ and $\lambda_j\ \ \ (j=1,\dots,J)$ are the decision variables; (b) compute the ratio of the minimum cost (i.e., the optimal solution to Model (3)) at the current prices of DMU_k, as follows:

$$\text{Farrell cost-efficiency}=\frac{\sum_{n=1}^{N}p_{nk}x_n^*}{\sum_{n=1}^{N}p_{nk}x_{nk}}\tag{4}$$

A DMU with a Farrell cost-efficiency score of one is Farrell cost-efficient; otherwise, it is Farrell cost-inefficient.

As an alternative to Model (3), Camanho and Dyson [6] presented a cost-efficiency model by directly imposing weight restrictions in a multiplier linear programming problem, that is:

$$\begin{aligned}&\text{Max}\sum_{m=1}^{M}u_m y_{mk}\\&\text{subject to :}\\&\sum_{m=1}^{M}u_m y_{mj}-\sum_{n=1}^{N}v_n x_{nj}\le 0,\ j=1,\dots,J\\&\sum_{n=1}^{N}v_n x_{nk}=1\\&\frac{v_{n'}}{v_n}=\frac{p_{n'k}}{p_{nk}},\ n'<n,\ n=l+1,\dots,N,\ n'=l+1,\dots,N\\&u_m\ge\varepsilon,\ m=1,\dots,M,\ v_n\ge\varepsilon,\ n=1,\dots,N\end{aligned}\tag{5}$$

where ε is a non-Archimedean positive number whose use guarantees that all input-output weights are positive.

Jahanshahloo et al. [23] and Jahanshahloo et al. [24] modified Model (3) presenting the following cost-efficiency model:

$$\begin{aligned}&\text{Min}\,\theta\\&\text{subject to :}\\&\sum_{j=1}^{J}\lambda_j\left(\sum_{n=1}^{N}p_{nk}x_{nj}\right)\le\theta\left(\sum_{n=1}^{N}p_{nk}x_{nk}\right)\\&\sum_{j=1}^{J}\lambda_j y_{mj}\ge y_{mk},\quad m=1,\dots,M\\&\lambda_j\ge 0,\ j=1,\dots,J\end{aligned}\tag{6}$$

3.2. FDH technologies and returns-to-scale assumptions

The conventional FDH technology is defined under the VRS assumption and it is commonly referred to as VRS-FDH technology:

$$T_{\text{FDH}}^{\text{VRS}} = \left\{ (x, y) : \sum_{j=1}^{J} \lambda_j x_{nj} \le x_n \forall n, \sum_{j=1}^{J} \lambda_j y_{mj} \ge y_m \forall m, \sum_{j=1}^{J} \lambda_j = 1, \lambda_j \in \{0, 1\} \right\} \tag{7}$$

Leleu [30] introduced the NIRS, NDRS and CRS specifications by modifying the VRS technology as follows:

$$T_{\text{FDH}}^{\Gamma} = \left\{ (x', y') : (x', y') = (1+\delta)(x, y), (x, y) \in T_{\text{FDH}}^{\text{VRS}}, \delta \in \Gamma \right\} \tag{8}$$

where $(1+\delta)$ is a scaling parameter introducing different returns-to-scale assumptions. Indeed, in Eq. (8),

$$\Gamma \in \{\text{VRS}, \text{NIRS}, \text{NDRS}, \text{CRS}\}$$

with VRS = $\{\delta : \delta = 0\}$, NIRS = $\{\delta : -1 \le \delta \le 0\}$, NDRS = $\{\delta : \delta \ge 0\}$ and CRS = $\{\delta : \delta \ge -1\}$.

Note that the FDH-CRS technology is the union of the FDH-NIRS technology with the FDH-NDRS technology, while the FDH-VRS technology is a subset of the intersection of the FDH-NIRS technology with the FDH-NDRS technology, i.e., $T_{\text{FDH}}^{\text{CRS}} = T_{\text{FDH}}^{\text{NIRS}} \cup T_{\text{FDH}}^{\text{NDRS}}$ and $T_{\text{FDH}}^{\text{VRS}} \subseteq T_{\text{FDH}}^{\text{NIRS}} \cap T_{\text{FDH}}^{\text{NDRS}}$. For more on the relationships among the four technologies described by Eq. (8), see Kerstens and Vanden Eechaut [27] and Briec et al. [5].

Given the above returns-to-scale assumptions, the following linear programming FDH technical efficiency model is formulated:

$$\begin{aligned}
&\min \sum_{j=1}^{J} \theta_j \\
&\text{subject to :} \\
&x_{nj}(\lambda_j + \omega_j) \le x_{nk}\theta_j \;\; n = 1, \ldots, N; j = 1, \ldots, J \\
&y_{mj}(\lambda_j + \omega_j) \ge \lambda_j y_{mk} \; m = 1, \ldots, M; j = 1, \ldots, J \\
&\sum_{j=1}^{J} \lambda_j = 1, ,\lambda_j \ge 0, j = 1, \ldots, J
\end{aligned} \tag{9}$$

where for every $j = 1, \ldots, J$, the variable ω_j is the scaling factor of DMU$_j$ such that:

$$\begin{aligned}
&\omega_j \in \Gamma_j, \Gamma_j \in \left\{ \text{VRS}, \text{NIRS}, \text{NDRS}, \text{CRS} \right\} \\
&\text{VRS} = \{\omega_j : \omega_j = 0\}, \text{NIRS} = \left\{ \omega_j : \omega_j \le 0 \right\} \\
&\text{NDRS} = \left\{ \omega_j : \omega_j \ge 0 \right\}, \text{CRS} = \left\{ \omega_j : \omega_j \text{unconstrained} \right\}
\end{aligned} \tag{10}$$

As noted by Leleu [30], Model (9) is appealing because it allows for the dual formulation and the shadow profit interpretation. By manipulating Model (9), Leleu [30] introduced the following linear program for evaluating FDH cost-efficiency:

$$\begin{aligned}
&\text{Min} \sum_{n=1}^{N} \sum_{j=1}^{J} p_{nk} \hat{x}_{nj} \\
&\text{subject to :} \\
&\left(\lambda_j + \omega_j\right) y_{mj} \ge \lambda_j y_{mk}, \quad m = 1, \ldots, M; \;\; j = 1, \ldots, J, \\
&\left(\lambda_j + \omega_j\right) x_{nj} \le \hat{x}_{nj}, \; n = 1, \ldots, N; j = 1, \ldots, J, \\
&\sum_{j=1}^{J} \lambda_j = 1, \quad \lambda_j \ge 0, \quad \omega_j \in \Gamma_j, \; \hat{x}_{nj} \ge 0, \quad n = 1, \ldots, N; \; j = 1, \ldots, J
\end{aligned} \tag{11}$$

where $\hat{x}_n \;\; (n = 1, \ldots, N)$ and $\lambda_j \quad (j = 1, \ldots, J)$ are the decision variables. The FDH cost-efficiency score is obtained as follows:

$$\text{FDH cost-efficiency} = \frac{\sum_{n=1}^{N} \sum_{j=1}^{J} p_{nk} \hat{x}_{nj}^{*}}{\sum_{n=1}^{N} p_{nk} x_{nk}} \tag{12}$$

Recently, Paryab et al. [43] showed that the FDH cost-efficiency score can be alternatively obtained by solving the following linear programming problem:

$$\begin{aligned}
&\text{Min} \sum_{j=1}^{J} \bar{C}_j \left(\lambda_j + \omega_j\right) \\
&\text{subject to :} \\
&\left(\lambda_j + \omega_j\right) y_{mj} \ge \lambda_j y_{mk}, m = 1, \ldots, M; j = 1, \ldots, J \\
&\sum_{j=1}^{J} \lambda_j = 1, \lambda_j \ge 0; \omega_j \in \Gamma_j; j = 1, \ldots, J
\end{aligned} \tag{13}$$

where $\bar{C}_j = \sum_{n=1}^{N} p_{nk}x_{nj} / \sum_{n=1}^{N} p_{nk}x_{nk}$ is the fixed value computed as the ratio of DMU$_k$'s total cost to DMU$_k$'s total cost based on the input prices faced by DMU$_k$. The ratios $\bar{C}_j \quad (j = 1, .., J)$ are used as weights in the objective function of Model (13). Thus, letting $\theta_j = \bar{C}_j(\lambda_j + \omega_j)$, Model (13) is transformed into the following equivalent linear programming model:

$$
\begin{aligned}
&\text{Min} \textstyle\sum_{j=1}^{J} \theta_j \\
&\text{subject to :} \\
&(\lambda_j + \omega_j)c_j = \theta_j c_k, j = 1, \ldots, J \\
&(\lambda_j + \omega_j)y_{mj} \geq \lambda_j y_{mk}, m = 1, \ldots, M; j = 1, \ldots, J \\
&\textstyle\sum_{j=1}^{J} \lambda_j = 1, \lambda_j \geq 0, \omega_j \in \Gamma_j, j = 1, \ldots, J
\end{aligned}
\tag{14}
$$

where $c_j = \sum_{n=1}^{N} p_{nk}x_{nj}$ is DMU$_k$'s total cost in terms of DMU$_k$'s input prices. The optimal objective function value of Model (14) is obtained from the following model:

$$
\begin{aligned}
&\text{Min} \textstyle\sum_{j=1}^{J} \theta_j \\
&\text{subject to :} \\
&(\lambda_j + \omega_j)c_j \leq \theta_j c_k, j = 1, \ldots, J \\
&(\lambda_j + \omega_j)y_{mj} \geq \lambda_j y_{mk}, m = 1, \ldots, M; j = 1, \ldots, J, \\
&\textstyle\sum_{j=1}^{J} \lambda_j = 1, \lambda_j \geq 0, \omega_j \in \Gamma_j, j = 1, \ldots, J.
\end{aligned}
\tag{15}
$$

Model (15) provides another expression for the evaluation of FDH technical cost-efficiency.

3.3. Background on rough set theory

Pawlak [41] and Liu [32] respectively introduced rough sets and rough variables.

Definition 1 *(Liu [33])*. Let Λ be a nonempty set, $\mathcal{A}$ a σ-algebra of subsets of Λ, and Δ an element in $\mathcal{A}$. Denote by π the set function on $\mathcal{A}$ satisfying the following axioms:

Axiom 1. $\pi\left\{\Lambda\right\} < \infty$.

Axiom 2. $\pi\left\{\Delta\right\} > 0$.

Axiom 3. $\pi\left\{A\right\} \geq 0$, for any $A \in \mathcal{A}$.

Axiom 4. For every countable sequence of mutually disjoint events $\left\{A_i\right\}_{i=1}^{\infty}$ in $\mathcal{A}$, we have:

$$\pi\left\{\bigcup_{i=1}^{\infty} A_i\right\} = \sum_{i=1}^{\infty} \pi\left\{A_i\right\}$$

The tuple $(\Lambda, \Delta, \mathcal{A}, \pi)$ is a *rough space*. If $\pi\{\emptyset\} = 0$, then π is a measure on $(\Lambda, \mathcal{A})$ and the tuple $(\Lambda, \mathcal{A}, \pi)$ is called a *measure space*.
The definitions of rough set and rough variable are stated below.

Definition 2 *(Liu [33])*. Let the pair (U, A) be an information system, where U is a non-empty finite set of objects called the universe and A is a finite set of attributes. Let $X \subseteq U$. Then, based on the information contained in A, X can be approximated by its lower and upper approximations, denoted by $\underline{X}$ and $\bar{X}$, respectively. The pair $(\underline{X}, \bar{X})$ is called a rough set.

Definition 3 *(Liu [33])*. A rough variable ξ is a measurable function from a rough space $(\Lambda, \Delta, \mathcal{A}, \pi)$ to the set of real numbers.

Definition 4 *(Liu [33])*. Suppose that the rough space $(\Lambda, \Delta, \mathcal{A}, \pi)$ satisfies Axioms 1–3. The trust of $K \in \mathcal{A}$ is defined by $\text{Tr}\{K\} = (1/2)\left[\overline{\text{Tr}}\{K\} + \underline{\text{Tr}}\{K\}\right]$, where $\overline{\text{Tr}}\{K\} = \pi\{K\}/\pi\{\Lambda\}$ and $\underline{\text{Tr}}\{K\} = \pi\{K \cap \Delta\}/\pi\{\Delta\}$ are the upper trust and the lower trust, respectively.

Assume that $\Lambda = \left\{z : c \leq z \leq d\right\}$ and $\Delta = \left\{z : a \leq z \leq b\right\}$, where $c \leq a < b \leq d$. Then, $\mathcal{A}$ coincides with the Borel algebra on Λ, π is the Lebesgue measure, and a rough variable can be represented as $\xi = ([a, b], [c, d])$ with $c \leq a < b \leq d$. The trust $\text{Tr}\left\{\xi \geq r\right\}$ is given by:

$$
K = [r, d] \Rightarrow \text{Tr}\left\{\xi \geq r\right\} = \begin{cases} 0 & \text{if } d \leq r \\ \dfrac{d-r}{2(d-c)} & \text{if } b \leq r \leq d \\ \dfrac{1}{2}\left(\dfrac{d-r}{d-c} + \dfrac{b-r}{b-a}\right) & \text{if } a \leq r \leq b \\ \dfrac{1}{2}\left(\dfrac{d-r}{d-c} + 1\right) & \text{if } c \leq r \leq a \\ 1 & \text{if } r \leq c \end{cases}
\tag{16}
$$

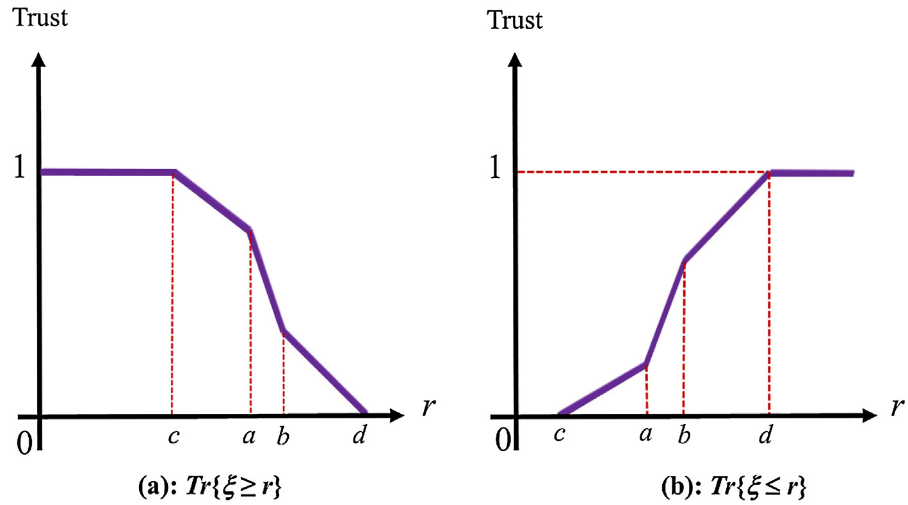

Fig. 1. Step function of rough variables.

Similarly, the trust $\mathrm{Tr}\left\{\xi \leq r\right\}$ is given by:

$$K=[a,r] \Rightarrow \mathrm{Tr}\left\{\xi \leq r\right\} = \begin{cases} 0 & \text{if } r \leq c \\ \dfrac{r-c}{2(d-c)} & \text{if } c \leq r \leq a \\ \dfrac{1}{2}\left(\dfrac{r-c}{d-c}+\dfrac{r-a}{b-a}\right) & \text{if } a \leq r \leq b \\ \dfrac{1}{2}\left(\dfrac{r-c}{d-c}+1\right) & \text{if } b \leq r \leq d \\ 1 & \text{if } d \leq r \end{cases} \tag{17}$$

Fig. 1 provides a graphical representation of $\mathrm{Tr}\left\{\xi \geq r\right\}$ and $\mathrm{Tr}\left\{\xi \leq r\right\}$.

Definition 5 *(Liu [33])*. Let ξ be a rough variable. For $\alpha \in (0.5, 1]$, the α-optimistic and the α-pessimistic values of ξ are defined, respectively, by:

$$\xi_{\inf}(\alpha) = \inf\left\{r : \mathrm{Tr}\left\{\xi \leq r\right\} \geq \alpha\right\} \tag{18}$$

and

$$\xi_{\sup}(\alpha) = \sup\left\{r : \mathrm{Tr}\left\{\xi \geq r\right\} \geq \alpha\right\} \tag{19}$$

For $\alpha \in [0, 0.5)$, $\xi_{\inf}(\alpha)$ and $\xi_{\sup}(\alpha)$ define the α-pessimistic and the α-optimistic values, respectively.

Henceforth, only the case where $\alpha \in [0.5, 1]$ will be considered. The analysis for $\alpha \in [0, 0.5)$ is similar.

By Eqs. (16) and (17) and the definition of $\xi_{\inf}(\alpha)$ and $\xi_{\sup}(\alpha)$, the α-optimistic value of $\xi = ([a, b], [c, d])$ is given by:

$$\xi_{\inf}(\alpha) = \begin{cases} (1-2\alpha)c + 2\alpha d & \text{if } 0 \leq \alpha \leq \dfrac{a-c}{2(d-c)} \\ 2(1-\alpha)c + (2\alpha-1)d & \text{if } \dfrac{b+d-2c}{2(d-c)} \leq \alpha \leq 1 \\ \dfrac{c(b-a)+a(d-c)+2\alpha(b-a)(d-c)}{(b-a)+(d-c)} & \text{if } \dfrac{a-c}{2(d-c)} < \alpha < \dfrac{b+d-2c}{2(d-c)} \end{cases}$$

while its α-pessimistic value is given by:

$$\xi_{\sup}(\alpha) = \begin{cases} (1-2\alpha)d + 2\alpha c \text{if} 0 \leq \alpha \leq \dfrac{d-b}{2(d-c)} \\ 2(1-\alpha)d + (2\alpha-1)c \text{ if} \dfrac{2d-a-c}{2(d-c)} \leq \alpha \leq 1 \\ \dfrac{d(b-a)+b(d-c)-2\alpha(b-a)(d-c)}{(b-a)+(d-c)} \text{if } \dfrac{d-b}{2(d-c)} < \alpha < \dfrac{2d-a-c}{2(d-c)} \end{cases}$$

Theorem 1 in Xu et al. [54] outlines the main properties of $\xi_{\inf}(\alpha)$ and $\xi_{\sup}(\alpha)$ as functions of α.

Remark 1. Given a rough variable ξ, for every $\alpha \in [0.5, 1]$, $\xi_{\inf}(\alpha) \geq \xi_{\sup}(\alpha)$. Thus, using the α-pessimistic and α-optimistic operators, an interval value is associated with ξ, namely, $\left[\xi_{\sup}(\alpha), \xi_{\inf}(\alpha)\right]$. See, among others, Liu [33] and Xu et al. [54].□

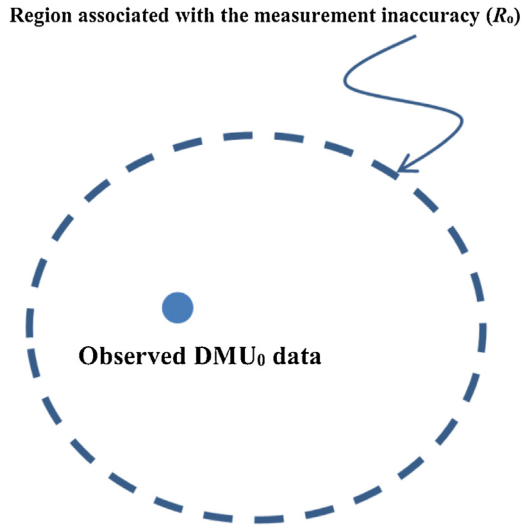

Fig. 2. Visualization of rough data.

Definition 6 *(Liu [33]).* Let ξ be a rough variable defined on $(\Lambda, \Delta, \mathcal{A}, \pi)$. The expected value of ξ is defined by $E(\xi) = \int_0^{+\infty} \text{Tr}\left\{\xi \geq r\right\} dr - \int_{-\infty}^{0} \text{Tr}\left\{\xi \leq r\right\} dr$, provided that at least one of the two integrals is finite.

If, in particular, $\xi = ([a, b], [c, d])$, with $c \leq a < b \leq d$, then $E\left[\xi\right] = (1/4)(a + b + c + d)$.

Definition 7 *(Liu [32]).* Let $f : \Re^n \to \Re$ be a measurable function and let $\xi_1, \xi_2, \ldots, \xi_n$ be rough variables on $(\Lambda, \Delta, \mathcal{A}, \pi)$. Then, $\xi = f\left(\xi_1, \xi_2, \ldots, \xi_n\right)$ is a rough variable defined on $(\Lambda, \Delta, \mathcal{A}, \pi)$:

$$\forall(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \Lambda, \xi(\lambda_1, \lambda_2, \ldots, \lambda_n) = f\left(\xi_1(\lambda_1), \xi_2(\lambda_2), \ldots, \xi_n(\lambda_n)\right)$$

In particular, given two rough variables ξ_1 and ξ_2 defined on $(\Lambda, \Delta, \mathcal{A}, \pi)$, both $\xi_1 + \xi_2$ and $\xi_1 \times \xi_2$ are rough variables defined on $(\Lambda, \Delta, \mathcal{A}, \pi)$:

$$\forall(\lambda_1, \lambda_2) \in \Lambda, (\xi_1 + \xi_2)(\lambda_1, \lambda_2) = \xi_1(\lambda_1) + \xi_2(\lambda_2) \text{ and} (\xi_1 \times \xi_2)(\lambda_1, \lambda_2) = \xi_1(\lambda_1) \times \xi_2(\lambda_2)$$

Example 1 *(Alefeld and Herzberger [2], Liu [33]).* Let $\xi = ([a_1, a_2], [a_3, a_4])$ and $\eta = ([b_1, b_2], [b_3, b_4])$ be two rough variables where $a_3 \leq a_1 < a_2 \leq a_4$ and $b_3 \leq b_1 < b_2 \leq b_4$. Then,

$$\xi + \eta = ([a_1 + b_1, a_2 + b_2], [a_3 + b_3, a_4 + b_4])$$
$$\xi - \eta = ([a_1 - b_2, a_2 - b_1], [a_3 - b_4, a_4 - b_3])$$
$$\xi \times \eta = ([a_1 \times b_1, a_2 \times b_2], [a_3 \times b_3, a_4 \times b_4]), \text{ if } a_3 \geq 0, b_3 \geq 0$$

$$\beta\xi = \begin{cases} \left(\left[\beta a_1, \beta a_2\right], \left[\beta a_3, \beta a_4\right]\right), \text{ if } \beta \geq 0 \\ \left(\left[\beta a_2, \beta a_1\right], \left[\beta a_4, \beta a_3\right]\right), \text{ if } \beta < 0 \end{cases}$$

4. Introducing rough DEA cost-efficiency

A rough interpretation of the inputs and outputs of traditional DEA models allows to account for the ambiguity and imprecision that characterize the data of many real-life problems.

For every $n = 1, \ldots, N$, $m = 1, \ldots, M$ and $j = 1, \ldots, J$, let:

$$\tilde{x}_{nj} = \left(\left[x_{nj}^a, x_{nj}^b\right], \left[x_{nj}^c, x_{nj}^d\right]\right), \quad 0 < x_{nj}^c \leq x_{nj}^a < x_{nj}^b \leq x_{nj}^d$$

$$\tilde{y}_{mj} = \left(\left[y_{mj}^a, y_{mj}^b\right], \left[y_{mj}^c, y_{mj}^d\right]\right), \quad 0 < y_{mj}^c \leq y_{mj}^a < y_{mj}^b \leq y_{mj}^d$$

$$\tilde{p}_{nj} = \left(\left[p_{nj}^a, p_{nj}^b\right], \left[p_{nj}^c, p_{nj}^d\right]\right), \quad 0 < p_{nj}^c \leq p_{nj}^a < p_{nj}^b \leq p_{nj}^d,$$

represent the n-th rough input, the m-th rough output and the rough price of the n-th input of DMU_j, respectively.

Fig. 2 provides a visual representation of how rough input-output data can be interpreted. Consider the observed data of a fixed DMU, DMU_o. Let T be the conceptual production possibility set, and let the observed point be in the region R_o. There are three possible cases:

(a) $R_o \cap T = \emptyset$, (b) $R_o \subsetneq T$, (c) $R_o \cap T \neq \emptyset$ and $R_o \not\subset T$. If $R_o \cap T = \emptyset$, then DMU_o is efficient (this evaluation is conducted based on the super-efficiency concept). If $R_o \subsetneq T$, then DMU_o is inefficient. Finally, if $R_o \cap T \neq \emptyset$ and $R_o \not\subset T$, then the estimated efficiency score should be related to a rough variable because of the impreciseness of the data. A similar interpretation is given in Xu et al. [54].

We propose the following cost-efficiency model with rough data:

$$
\begin{aligned}
&\text{Min } \theta \\
&\text{subject to :} \\
&\textstyle\sum_{j=1}^{J} \lambda_j \left(\sum_{n=1}^{N} \tilde{p}_{nj}\tilde{x}_{nj} \right) \le \theta \left(\sum_{n=1}^{N} \tilde{p}_{nk}\tilde{x}_{nk} \right) \\
&\textstyle\sum_{j=1}^{J} \lambda_j \tilde{y}_{mj} \ge \tilde{y}_{mk}, \quad m = 1, \dots, M \\
&\lambda_j \ge 0, \quad j = 1, \dots, J; \ \theta : \text{free}
\end{aligned}
\tag{21}
$$

Model (21) has been obtained from Model (6) by replacing all the crisp values of inputs, outputs and prices with the corresponding rough data. The solution $(\theta_k^*)_{\mathrm{RDEA}}$ of Model (21) will be referred to as the Rough DEA (RDEA) *cost-efficiency* of DMU_k .

Since $\tilde{p}_{nj}$ and $\tilde{x}_{nj}$ are rough variables, then $\tilde{p}_{nj}\tilde{x}_{nj}$ is a rough variable (see Definition 7). Moreover, by Example 1:

$$
\tilde{p}_{nj}\tilde{x}_{nj} = \left(\left[p_{nj}^a, p_{nj}^b\right], \left[p_{nj}^c, p_{nj}^d\right] \right) \times \left(\left[x_{nj}^a, x_{nj}^b\right], \left[x_{nj}^c, x_{nj}^d\right] \right) \qquad = \left(\left[p_{nj}^a x_{nj}^a, p_{nj}^b x_{nj}^b\right], \left[p_{nj}^c x_{nj}^c, p_{nj}^d x_{nj}^d\right] \right) \tag{22}
$$

Thus, letting $\bar{x}_{nj} = \tilde{p}_{nj}\tilde{x}_{nj} = \left(\left[p_{nj}^a x_{nj}^a, p_{nj}^b x_{nj}^b\right], \left[p_{nj}^c x_{nj}^c, p_{nj}^d x_{nj}^d\right] \right)$ and implementing the α-optimistic and α-pessimistic operators relative to the rough variables $\bar{x}_{nj}$ $(n = 1, \dots, N, \quad j = 1, \dots, J)$ and $\tilde{y}_{mj}$ $(m = 1, \dots, M, \quad j = 1, \dots, J)$ where $\alpha \in [0.5, 1]$, Model (21) is converted into the following interval linear programming problem (see also Remark 1):

$$
\begin{aligned}
&\text{Min } \theta \\
&\text{subject to :} \\
&\textstyle\sum_{j=1}^{J} \sum_{n=1}^{N} \left(\sum_{n=1}^{N} \lambda_j \left[\bar{x}_{nj}^{\sup(\alpha)}, \bar{x}_{nj}^{\inf(\alpha)} \right] \right) \le \theta \sum_{n=1}^{N} \left[\bar{x}_{nk}^{\sup(\alpha)}, \bar{x}_{nk}^{\inf(\alpha)} \right] \\
&\textstyle\sum_{j=1}^{J} \lambda_j \left[y_{mj}^{\sup(\alpha)}, y_{mj}^{\inf(\alpha)} \right] \ge \left[y_{mk}^{\sup(\alpha)}, y_{mk}^{\inf(\alpha)} \right], \quad m = 1, \dots, M \\
&\lambda_j \ge 0, \quad j = 1, \dots, J; \ \theta : \ \text{free}
\end{aligned}
\tag{23}
$$

Model (23) determines a pair of linear programming problems whose solutions respectively provide the α-optimistic value (the upper bound) and the α-pessimistic value (the lower bound) for the RDEA cost-efficiency score of DMU_k.

The α-optimistic value is obtained by solving the following:

$$
\begin{aligned}
&\left(\theta_k^*\right)_{\mathrm{RDEA}}^{\inf(\alpha)} = \text{Min } \theta^{\inf(\alpha)} \\
&\text{subject to:} \\
&\textstyle\sum_{j=1, j \neq k}^{J} \lambda_j \left(\sum_{n=1}^{N} \bar{x}_{nj}^{\inf(\alpha)} \right) + \lambda_k \left(\sum_{n=1}^{N} \bar{x}_{nk}^{\sup(\alpha)} \right) \le \theta^{\inf(\alpha)} \left(\sum_{n=1}^{N} \bar{x}_{nk}^{\sup(\alpha)} \right) \\
&\textstyle\sum_{j=1, j \neq k}^{J} \lambda_j y_{mj}^{\sup(\alpha)} + \lambda_k y_{mk}^{\inf(\alpha)} \ge y_{mk}^{\inf(\alpha)}, \ m = 1, \dots, M \\
&\lambda_j \ge 0, j = 1, \dots, J, \theta : \text{free}
\end{aligned}
\tag{24}
$$

while the α-pessimistic value is obtained by solving the following:

$$
\begin{aligned}
&\left(\theta_k^*\right)_{\mathrm{RDEA}}^{\inf(\alpha)} = \text{Min } \theta^{\inf(\alpha)} \\
&\text{subject to:} \\
&\textstyle\sum_{j=1, j \neq k} \lambda_j \left(\sum_{n=1}^{N} \bar{x}_{nj}^{\sup(\alpha)} \right) + \lambda_k \left(\sum_{n=1}^{N} \bar{x}_{nj}^{\inf(\alpha)} \right) \le \theta^{\sup(\alpha)} \left(\sum_{n=1}^{N} \bar{x}_{nk}^{\inf(\alpha)} \right) \\
&\textstyle\sum_{j=1, j \neq k} \lambda_j \left(\sum_{n=1}^{N} \bar{x}_{nj}^{\inf(\alpha)} \right) + \lambda_k y_{mk}^{\sup(\alpha)} \ge y_{mk}^{\sup(\alpha)}, \ m = 1, \dots, M \\
&\lambda_j \ge 0, j = 1, \dots, J, \theta : \text{free}
\end{aligned}
\tag{25}
$$

Definition 8. DMU_k is said to be RDEA *cost-efficient* if one of the following holds:

(a) $\alpha \in (0.5, \quad 1]$ and $(\theta_k^*)_{\mathrm{RDEA}}^{\inf(\alpha)} = 1$;
(b) $\alpha \in [0, \quad 0.5)$ and $(\theta_k^*)_{\mathrm{RDEA}}^{\sup(\alpha)} = 1$;
(c) $\alpha = 0.5$ and $(\theta_k^*)_{\mathrm{RDEA}}^{\inf(\alpha)} = (\theta_k^*)_{\mathrm{RDEA}}^{\sup(\alpha)} = 1$.

DMU_k is said to be RDEA *cost-inefficient* if it is not RDEA cost-efficient.

The following proposition shows the monotonicity of the RDEA cost-efficiency measure with respect to α.

Proposition 1. Let $\alpha_1, \quad \alpha_2 \in [0, \quad 1]$ with $\alpha_1 \ge \alpha_2$.Then, for DMU_k,

$$(a)\left(\theta_k^*\right)_{RDEA}^{\inf(\alpha_1)} \geq \left(\theta_k^*\right)_{RDEA}^{\inf(\alpha_2)} (b)\left(\theta_k^*\right)_{RDEA}^{\sup(\alpha_2)} \geq \left(\theta_k^*\right)_{RDEA}^{\sup(\alpha_1)}$$

Proof. Assume that $\alpha_1 \geq \alpha_2$. Theorem 1 in Xu et al. [54] implies that:

$$\begin{cases} y_{mj}^{\inf(\alpha_1)} \geq y_{mj}^{\inf(\alpha_2)} \\ y_{mj}^{\sup(\alpha_1)} \leq y_{mj}^{\sup(\alpha_2)} \end{cases} \quad \text{and} \quad \begin{cases} \sum_{n=1}^{N} \bar{x}_{nj}^{\inf(\alpha_2)} \leq \sum_{n=1}^{N} \bar{x}_{nj}^{\inf(\alpha_1)} \\ \sum_{n=1}^{N} \bar{x}_{nj}^{\sup(\alpha_2)} \geq \sum_{n=1}^{N} \bar{x}_{nj}^{\sup(\alpha_1)} \end{cases} \tag{26}$$

At the same time, the constraints of Model (24) can be written as follows:

$$\sum_{j=1,\ \ j\neq k} \lambda_j \left(\sum_{n=1}^{N} \bar{x}_{nj}^{\inf(\alpha)}\right) \leq \left(\theta^{\inf(\alpha)} - \lambda_k\right) \left(\sum_{n=1}^{N} \bar{x}_{nk}^{\sup(\alpha)}\right) \tag{27}$$

$$\sum_{j=1,\ \ j\neq k} \lambda_j y_{mj}^{\sup(\alpha)} \geq (1-\lambda_k) y_{mk}^{\inf(\alpha)}, \quad m = 1, \ldots, M \tag{28}$$

Since the left hand side of Eq. (27) is non-negative for positive data, $\theta^{\inf(\alpha)} \geq \lambda_k$. Furthermore, $\theta^{\inf(\alpha)} = \lambda_k = 1$ is a feasible solution of Model (24) and, hence, $1 \geq \theta^{\inf(\alpha)}$. It follows that $1 \geq \lambda_k$. The relationships in Eqs. (26)–(28) indicate that the feasible region determined by the constraints of Model (24) for α_1 is not bigger than the one for α_2. Hence, (a) holds. The proof of (b) is similar.□

Remark 2. This study explicitly considers rough cost-efficiency measures for the constant returns-to-scale DEA technology. Nonetheless, assuming $\sum_{j=1}^{J} \lambda_j = 1$, $\sum_{j=1}^{J} \lambda_j \leq 1$, or $\sum_{j=1}^{J} \lambda_j \geq 1$, the VRS (variable returns-to-scale), the NIRS (non-increasing returns-to-scale), or the NDRS (non-decreasing returns-to-scale) versions can be obtained, respectively. The corresponding α-optimistic and α-pessimistic cost-efficiency values can be evaluated reasoning in a similar manner as for the constant returns-to scale case.□

5. Introducing rough FDH cost-efficiency

The previous section discussed cost-efficiency for convex technologies (polyhedral or DEA technology). This section develops a rough cost-efficiency approach for non-convex FDH technologies with all the data being rough. Again, the trust level α will be assumed to vary in $[0.5,\quad 1]$.

Replacing x_{nj}, y_{mj} and p_{nj} with $\tilde{x}_{nj}, \tilde{y}_{mj}$ and $\tilde{p}_{nj}$, respectively, Model (15) gives raise to an interval cost-efficiency model with rough inputs, outputs and input prices. Considering the different returns-to-scale assumptions discussed in Subsection 3.2, the cost constraints for the α-optimistic model can be written as follows:

$$\begin{array}{l} \left(\lambda_j + \omega_j\right) \sum_{n=1}^{N} \bar{x}_{nj}^{\inf(\alpha)} \leq \theta_j^{\inf(\alpha)} \sum_{n=1}^{N} \bar{x}_{nk}^{\sup(\alpha)}, j = 1, \ldots, J, j \neq k \\ (\lambda_k + \omega_k) \sum_{n=1}^{N} \bar{x}_{nk}^{\sup(\alpha)} \leq \theta_k^{\inf(\alpha)} \sum_{n=1}^{N} \bar{x}_{nk}^{\sup(\alpha)} \end{array} \tag{29}$$

while the output constraints become:

$$\begin{array}{l} \left(\lambda_j + \omega_j\right) y_{mj}^{\sup(\alpha)} \geq \lambda_j y_{mk}^{\inf(\alpha)}, m = 1, \ldots, M, j = 1, \ldots, J; j \neq k \\ (\lambda_k + \omega_k) y_{mk}^{\inf(\alpha)} \geq \lambda_k y_{mk}^{\inf(\alpha)}, m = 1, \ldots, M \end{array} \tag{30}$$

Constraints (29) and (30) simplify as follows:

$$\begin{array}{l} \left(\lambda_j + \omega_j\right) \sum_{n=1}^{N} \bar{x}_{nj}^{\inf(\alpha)} \leq \theta_j^{\inf(\alpha)} \sum_{n=1}^{N} \bar{x}_{nk}^{\sup(\alpha)}, j = 1, \ldots, J, j \neq k \\ (\lambda_k + \omega_k) \leq \theta_k^{\inf(\alpha)} \\ \left(\lambda_j + \omega_j\right) y_{mj}^{\sup(\alpha)} \geq \lambda_j y_{mk}^{\inf(\alpha)}, m = 1, \ldots, M, j = 1, \ldots, J; j \neq k \\ \omega_k \geq 0 \end{array} \tag{31}$$

Hence, the α-optimistic model associated with the rough version of Model (15) is the following:

$$\begin{array}{l} \left(\theta^*\right)_{RFDH}^{\inf(\alpha)} = \text{Min} \sum_{j=1}^{J} \theta_j^{\inf(\alpha)} \\ \text{subject to :} \\ \left(\lambda_j + \omega_j\right) \sum_{n=1}^{N} \bar{x}_{nj}^{\inf(\alpha)} \leq \theta_j^{\inf(\alpha)} \sum_{n=1}^{N} \bar{x}_{nk}^{\sup(\alpha)}, j = 1, \ldots, J, j \neq k \\ \lambda_k + \omega_k \leq \theta_k^{\inf(\alpha)}, \\ \left(\lambda_j + \omega_j\right) y_{mj}^{\sup(\alpha)} \geq \lambda_j y_{mk}^{\inf(\alpha)}, m = 1, \ldots, M, j = 1, \ldots, J; j \neq k \\ \omega_k \geq 0, \sum_{j=1}^{J} \lambda_j = 1, \lambda_j \geq 0, \omega_j \in \Gamma_j, j = 1, \ldots, J \end{array} \tag{32}$$

Table 1
Rough inputs and rough outputs.

DMU	Input 1	Input 2	Input Price 1	Input Price 2	Output 1
1	([98,116], [82,134])	([234,275], [205,294])	([19,29], [15,43])	([20,29], [13,37])	([478,496], [456,532])
2	([100,130], [85,145])	([297,324], [264,365])	([19,29], [6,42])	([18,28], [11,40])	([464,495], [443,543])
3	([55,75], [35,76])	([205,238], [175,288])	([24,37], [17,25])	([17,25], [8,39])	([87,116], [86,135])
4	([112,136], [96,176])	([329,369], [301,395])	([22,31], [14,42])	([21,28], [13,40])	([724,754], [704,786])
5	([94,106], [84,126])	([286,303], [255,342])	([15,17], [13,19])	([23,31], [16,38])	([368,395], [344,423])

The α-pessimistic model associated with the rough version of Model (15) is obtained in a similar way:

$$\begin{aligned}
&\left(\theta_k^*\right)_{\mathrm{RFDH}}^{\sup(\alpha)} = \mathrm{Min}\ \textstyle\sum_{j=1}^{J}\theta_j^{\sup(\alpha)}\\
&\text{subject to :}\\
&\left(\lambda_j+\omega_j\right)\textstyle\sum_{n=1}^{N}\bar{x}_{nj}^{\sup(\alpha)} \le \theta_j^{\sup(\alpha)}\left(\textstyle\sum_{n=1}^{N}\bar{x}_{nk}^{\inf(\alpha)}\right), j=1,\ldots,J; j\ne k\\
&\lambda_k+\omega_k \le \theta_k^{\sup(\alpha)}\\
&\left(\lambda_j+\omega_j\right)y_{mj}^{\inf(\alpha)} \ge \lambda_j y_{mk}^{\sup(\alpha)}, m=1,\ldots,M; j=1,\ldots,J; j\ne k\\
&\omega_k \ge 0, \textstyle\sum_{j=1}^{J}\lambda_j = 1, \lambda_j \ge 0; \omega_j \in \Gamma_j; j=1,\ldots,J
\end{aligned} \tag{33}$$

The solutions $(\theta_k^*)_{\mathrm{RFDH}}^{\inf(\alpha)}$ and $(\theta_k^*)_{\mathrm{RFDH}}^{\sup(\alpha)}$ of Models (32) and (33), respectively, represent the upper and lower bounds for the Rough FDH (RFDH) *cost-efficiency* of DMU_k.

Definition 9. DMU_k is said to be RFDH *cost-efficient* if one of the following holds:

(d) $\alpha \in (0.5,\ 1]$ and $(\theta_k^*)_{\mathrm{RFDH}}^{\inf(\alpha)} = 1$;
(e) $\alpha \in [0,\ 0.5)$ and $(\theta_k^*)_{\mathrm{RFDH}}^{\sup(\alpha)} = 1$;
(f) $\alpha = 0.5$ and $(\theta_k^*)_{\mathrm{RFDH}}^{\inf(\alpha)} = (\theta_k^*)_{\mathrm{RFDH}}^{\sup(\alpha)} = 1$.

DMU_k is said to be RFDH *cost-inefficient* if it is not RFDH cost-efficient.
The following proposition states the monotonicity of the RFDH cost-efficiency measure with respect to the trust level α.

Proposition 2. Let $\alpha_1,\ \alpha_2 \in [0,\ 1]$ with $\alpha_1 \ge \alpha_2$. Then, for DMU_k,

(a) $\left(\theta_k^*\right)_{RFDH}^{\inf(\alpha_1)} \ge \left(\theta_k^*\right)_{RFDH}^{\inf(\alpha_2)}$ (b) $\left(\theta_k^*\right)_{RFDH}^{\sup(\alpha_2)} \ge \left(\theta_k^*\right)_{RFDH}^{\sup(\alpha_1)}$

Proof. Theorem 1 in Xu et al. [54] implies that:

$$\begin{cases} y_{mj}^{\inf(\alpha_1)} \ge y_{mj}^{\inf(\alpha_2)} \\ y_{mj}^{\sup(\alpha_1)} \le y_{mj}^{\sup(\alpha_2)} \end{cases} \quad \text{and} \quad \begin{cases} \sum_{n=1}^{N}\bar{x}_{nj}^{\inf(\alpha_2)} \le \sum_{n=1}^{N}\bar{x}_{nj}^{\inf(\alpha_1)} \\ \sum_{n=1}^{N}\bar{x}_{nj}^{\sup(\alpha_2)} \ge \sum_{n=1}^{N}\bar{x}_{nj}^{\sup(\alpha_1)} \end{cases} \tag{34}$$

The validity of (a) and (b) is now shown reasoning as in the proof of Proposition 1.□

6. Numerical example

To illustrate the proposed framework, consider a hypothetical bank with five branches/DMUs each of which uses two inputs to produce one output. The two inputs used by each one of the branches are:

$x_1 \overset{def}{=}$ number of branch managers,

$x_2 \overset{def}{=}$ number of administrative and commercial staff members,

while the output produced is:

$y_1 \overset{def}{=}$ number of general service transactions.

Suppose that the input and output data are available for the last four consecutive years.
As stated by Paradi et al. [37], it is very difficult to accurately measure a bank branch efficiency since: (1) bank branches also provide services that are not directly paid for, and (2) bank branches are subject to complex government regulations, which are likely to affect costs, input prices and service transactions. Although bank management is mainly concerned with the performance of all its branches (Schaffnit et al. [46]), the bank may grant to each branch management some discretionary power when setting acceptable input-output levels. That is, within the common bounds decided by the bank, the single branch is allowed to establish its own lower and upper bounds for inputs and outputs.
Table 1 shows the rough data assumed for the five hypothetical DMUs in this example. The input prices p_{nj} $(n = 1, 2;\ \ j = 1, \ldots, 5)$ are computed as the ratio of the annual salaries and fringe benefits of the staff members to the n-th input x_{nj}. Since the costs deriving from the salaries and fringe benefits can fluctuate substantially due, for example, to the regional economic conditions and the availability of

Table 2
RDEA versus standard DEA cost-efficiency results under CRS for the trust level $\alpha = 0.9$.

DMU	RDEA cost-efficiency	Farrell cost-efficiency
1	[0.6584,1.0000]	0.8757
2	[0.5460,0.9893]	0.7326
3	[0.1513,0.3484]	0.2349
4	[0.8192,1.0000]	1.0000
5	[0.4575,0.8029]	0.6150

Table 3
RFDH versus standard FDH cost-efficiency results under CRS.

DMU	RFDH cost-efficiency	Standard FDH cost-efficiency
1	[0.6824,1.0000]	0.8757
2	[0.5511,1.0000]	0.7326
3	[0.1441,0.3976]	0.2349
4	[0.8223,1.0000]	1.0000
5	[0.4588,0.8185]	0.6150

Table 4
Bank information.

DMU	Name	Type	Assets (millions of yen)
1	Mizuho Bank	City	73,823,446
2	Mizuho Corporate Bank	City	76,364,244
3	Bank of Tokyo-Mitsubishi UFJ	City	155,399,430
4	Resona Bank	City	27,620,554
5	Sumitomo Mitsui Banking	City	111,235,993
6	Mitsubishi UFJ Trust and Banking	Trust	26,533,972
7	Mizuho Trust & Banking	Trust	11,753,015
8	Chuo Mitsui Trust & Banking	Trust	28,101,014
9	Sumitomo Trust and Banking	Trust	20,226,273

experienced branch account managers, the bank sets a price interval for each DMU using the minimum and maximum data available over the last four years. Afterwards, the branch manager can provide his estimation for the same interval, which makes p_{nj} a rough variable.

Consider, for instance DMU_1. Suppose that the bank has set the interval for p_{11} to be $[15, \quad 43]$. Based on his experience and, hence, on the third and second largest data available for input 1, the branch manager of DMU_1 may estimate the interval to be $[19, \quad 29]$. By Definition 6, the expected value of p_{11} for DMU_1 is $\left(1/4\right)(15+19+29+43) = 26.5$. The intervals for p_{21} are constructed in a similar manner.

Tables 2 and 3 show the lower and upper bounds obtained for the RDEA and RFDH cost-efficiency measures of the DMUs, respectively. These scores have been computed by implementing Models (15), (16), (32) and (33), respectively. Moreover, Table 2 compares the RDEA results with the Farrell cost-efficiency measures, while Table 3 compares the RFDH results with the standard FDH cost-efficiency scores. The Farrell and the standard FDH cost-efficiency values have been computed using the expected values of the rough data in Table 1. All cost-efficiency measures have been performed under CRS and at the threshold level $\alpha = 0.9$. The bank has been assumed to choose the α-level subjectively, from a managerial perspective.

The cost-efficiency scores obtained in the RDEA case are very close to those obtained in the RFDH case. Moreover, the cost-efficiency values obtained using the standard DEA (Farrell) and the standard FDH approaches fall in the corresponding intervals produced by the RDEA and the RFDH cost-efficiency models, respectively.

Note, in particular, that the only Farrell cost-efficient DMU, DMU_4, is also the most efficient among all the RDEA cost-efficient DMUs. See Table 2. Table 3 shows that the same result is true when considering the FDH cost-efficiency values.

7. Case study: Japanese city and trust banks

To further illustrate the proposed RDEA–RFDH cost-efficiency approach, consider the following case study where the data on five City banks and four Trust banks operating in Japan are used. The names of the banks and total assets are given in Table 4. The case study follows the intermediation approach of Sealey and Lindley [47] and assumes that each bank produces loans (y_1) and securities investments (y_2) using labor (x_1), physical capital (x_2), and deposits (x_3). See also Fukuyama and Weber [17].

Four different values are needed to obtain a rough variable expressed by two approximation intervals such as $([a, \quad b], [c, \quad d])$, where $c \leq a < b \leq d$. Since the sample data cover four consecutive years, there are four values associated with each variable. This allows transforming the available data in rough ones whose lower and upper approximation intervals are shown in Table 5.

Remark 3. Note that the real data relative to the nine Japanese banks considered in the case study are not precise values. As said above, following Sealey and Lindley's [47] intermediation approach a bank is regarded as a financial intermediary converting labor, physical capital and raised funds into outputs associated with lending and investment activities. Thus, the number of workers can be used as a proxy for labor, the balance sheet values of premises and real estate as proxies for physical capital, and the deposits as proxies for the raised funds. The proxies of the two output activities are represented by the balance sheet values of the total of the loans and securities investments. That is, the inputs and outputs used here are all proxies. Moreover, deflating all financial data by the GDP deflator, they can actually be thought of as estimated values.□

Table 5
Japanese bank data.[a]

DMU	Rough inputs data		
	Input 1	Input 2	Input 3
1	([18145,18943], [17271,18969])	([822887,948435], [752821,1028855])	([57180669,58327503], [59980118,55819338])
2	([7900,8147], [7619,8307])	([190987,209322], [190435,214478])	([20080606,20262691], [22866455,19677151])
3	([33827,34902], [33280,34902])	([1262719,1268998], [1247756,1348718])	([104366346,108761738], [112851470,103521670])
4	([8152,8966], [8053,9246])	([286607,287932], [281902,314070])	([20103542,20853084], [22187525,19758953])
5	([21816,22460], [17886,22524])	([849948,876945], [801784,916726])	([71797518,73700069], [78930138,68050471])
6	([7069,7090], [6989,7144])	([247154,247607], [246456,250365])	([13395242,12519996], [13395242,13087922])
7	([3138,3327], [2964,3332])	([51946,51951], [50875,52955])	([3016634,2466767], [2763194,2624138]
8	([6371,6373], [6173,6373])	([124684,126082], [122959,136213])	([9953271,8535668], [9249971,9228211])
9	([6026,6084], [5869,6085])	([143880,146009], [141024,153173])	([13116649,12100633], [12778714,12299614])

DMU	Rough outputs data	
	Output 1	Output 2
1	([34575616,35582385], [33961974,38353938])	([20576426,21202088], [13818237,21202088])
2	([28110635,29138936], [27568671,30900193])	([23391626,24888149], [15916168,24888149])
3	([72128898,72287264], [69276882,76225726])	([54464833,62157046], [34007269,62157046])
4	([17997403,18008724], [17597528,18329680])	([5033178,5700837], [4047936,5700837])
5	([58888713,59224956], [58358415,62232713])	([29849582,42487667], [23317870,42487667])
6	([10729829,10818471], [10019341,11289036])	([7245742,9934501], [7245742,11394224])
7	([3553296,3566966], [3464442,3617072])	([1716068,2010307], [1613765,2198584])
8	([8865505,9350182], [8040281,9447311])	([4690076,4701419], [3925799,5035947])
9	([11868478,12470163], [11304553,12957029])	([5011409,5259314], [4680299,5277188])

DMU	Rough input prices data		
	Input price 1	Input price 2	Input price 3
1	([0.01036,0.01084], [0.00946,0.01107])	([0.37895,0.46789], [0.35195,0.53879])	([0.00273,0.00460], [0.00193,0.00524])
2	([0.01100,0.01143], [0.01063,0.01219])	([0.77228,0.78642], [0.72605,0.80142])	([0.01752,0.03860], [0.01259,0.05765])
3	([0.01129,0.01116], [0.01018,0.01136])	([0.49274,0.53171], [0.48430,0.53239])	([0.00465,0.01013], [0.00349,0.01420])
4	([0.00942,0.00947], [0.00820,0.00969])	([0.48727,0.50383], [0.46515,0.51905])	([0.00331,0.00519], [0.00242,0.00652])
5	([0.01144,0.01183], [0.01122,0.01325])	([0.48102,0.51806], [0.47852,0.57856])	([0.00474,0.01065], [0.00394,0.01348])
6	([0.0888,0.00987], [0.00551,0.01001])	([0.50045,0.51061], [0.48145,0.61200])	([0.00752,0.01217], [0.00595,0.01390])
7	([0.01074,0.01100], [0.00933,0.01100])	([1.05713,1.11044], [1.05623,1.23610])	([0.01135,0.01351], [0.00879,0.01760])
8	([0.00726,0.00790], [0.00620,0.00790])	([0.42249,0.43704], [0.40892,0.43704])	([0.00781,0.01079], [0.00639,0.01185])
9	([0.00852,0.00863], [0.00801,0.00863])	([0.53506,0.55246], [0.51364,0.60779])	([0.00899,0.01697], [0.00728,0.02050])

[a] Input–output values (except for labor) are shown in billions of yen.

Table 6
RDEA and RFHD cost-efficiency measures under CRS.

DMUS	RDEA cost-efficiency				
	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$	$\alpha = 1$
1	[0.8506,1.0000]	[0.7295,1.0000]	[0.6429,1.0000]	[0.5283,1.0000]	[0.3890,1.0000]
2	[0.6740,0.9633]	[0.5392,1.0000]	[0.3936,1.0000]	[0.2781,1.0000]	[0.1825,1.0000]
3	[0.8963,1.0000]	[0.7177,1.0000]	[0.5584,1.0000]	[0.4216,1.0000]	[0.2903,1.0000]
4	[1.0000,1.0000]	[1.0000,1.0000]	[0.9074,1.0000]	[0.7567,1.0000]	[0.5560,1.0000]
5	[0.8246,1.0000]	[0.7184,1.0000]	[0.6228,1.0000]	[0.4950,1.0000]	[0.3559,1.0000]
6	[0.8003,1.0000]	[0.6762,1.0000]	[0.5627,1.0000]	[0.4632,1.0000]	[0.3275,1.0000]
7	[0.5605,0.6779]	[0.5124,0.7645]	[0.4524,0.8841]	[0.3794,1.0000]	[0.2841,1.0000]
8	[**0.9594**,1.0000]	[0.8483,1.0000]	[0.7319,1.0000]	[0.6138,1.0000]	[0.4404,1.0000]
9	[0.6331,0.7984]	[0.5663,0.9065]	[0.5043,1.0000]	[0.4206,1.0000]	[0.3037,1.0000]

DMUS	RFDH cost-efficiency				
	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$	$\alpha = 1$
1	[0.8740,1.0000]	[0.7367,1.0000]	[0.6520,1.0000]	[0.5412,1.0000]	[0.4040,1.0000]
2	[0.6740,0.9633]	[0.5601,1.0000]	[0.3972,1.0000]	[0.2819,1.0000]	[0.1879,1.0000]
3	[0.9188,1.0000]	[0.7232,1.0000]	[0.5638,1.0000]	[0.4274,1.0000]	[0.2903,1.0000]
4	[1.0000,1.0000]	[1.0000,1.0000]	[0.9074,1.0000]	[0.7567,1.0000]	[0.5560,1.0000]
5	[0.8485,1.0000]	[0.7333,1.0000]	[0.6263,1.0000]	[0.5040,1.0000]	[0.3559,1.0000]
6	[0.8003,1.0000]	[0.7059,1.0000]	[0.5748,1.0000]	[0.4657,1.0000]	[0.3275,1.0000]
7	[0.5606,0.6802]	[0.5142,0.7645]	[0.4550,0.8841]	[0.3824,1.0000]	[0.2841,1.0000]
8	[**1.0000**,1.0000]	[0.8974,1.0000]	[0.7687,1.0000]	[0.6329,1.0000]	[0.4404,1.0000]
9	[0.6723, 0.8608]	[0.5965,0.9794]	[0.5287,1.0000]	[0.4350,1.0000]	[0.3037,1.0000]

The RDEA and RFDH cost-efficiency measures have been computed by implementing Models (15), (16), (32) and (33), respectively. The computations have been performed under the CRS assumption and for $\alpha = 0.6, \quad 0.7, \quad 0.8, \quad 0.9, \quad 1$ using the GAMS software. The results obtained are shown in Table 6. Recall that for $\alpha \in (0.5, \quad 1]$, the RDEA and the RFDH cost-efficiency intervals are represented by $\left[\left(\theta_k^*\right)_{\text{RDEA}}^{\sup(\alpha)}, \left(\theta_k^*\right)_{\text{RDEA}}^{\inf(\alpha)}\right]$ and $\left[\left(\theta_k^*\right)_{\text{RFDH}}^{\sup(\alpha)}, \left(\theta_k^*\right)_{\text{RFDH}}^{\inf(\alpha)}\right]$, respectively.

Consider, in particular, DMU_7. The α-optimistic values obtained by the RDEA models for $\alpha = 0.6, \alpha = 0.7, \alpha = 0.8, \alpha = 0.9$ and $\alpha = 1$ are 0.5605, 0.5124, 0.4524, 0.3794 and 0.2841, respectively. The corresponding α-pessimistic values are 0.6779, 0.7645, 0.8841 and 1, respectively. This fact shows that the lower (upper) bound defines a non-increasing (non-decreasing) function of the trust level α, that is, Proposition 1 is verified. The validity of Proposition 2 can be checked similarly considering the α-optimistic and α-pessimistic values obtained by the RFDH models.

Table 6 shows that the results obtained in the RDEA case are very similar to those obtained in the RFDH case. In particular, DMUs 1, 3, 4, 5, 6, and 8 are all RDEA and RFDH cost-efficient independently from the trust level (see Definitions 8 and 9). Furthermore, for $\alpha = 0.6$ and $\alpha = 0.7$, DMU_4 is RDEA cost-efficient according to both the α-optimistic and α-pessimistic model, whereas the same it is true for both DMU_4 and DMU_8 in the RFDH case. Finally, except for DMU_4, all DMUs are RDEA cost-inefficient with respect to the α-optimistic model for all the trust levels considered, the same being true, with the exception of DMU_4 and DMU_8, in the RFDH case.

Clearly, the main difference between the two sets of results is in the behavior of DMU_8 which is RDEA cost-inefficient but RFDH cost-efficient for $\alpha = 0.6$ and $\alpha = 0.7$. This difference, signaled in bold in Table 6, arises from the RFDH and the RDEA cost-efficiency approaches being designed for non-convex and convex technologies, respectively.

The numerical results in Table 6 allow to rank the efficient DMUs by RDEA and by RFDH with respect to each of the α levels considered. The efficient DMUs are ranked on the basis of the their α-optimistic values.

In particular, in the $\alpha = 0.6$ case, DMU_4 is more efficient than DMU_8 according to the RDEA approach, while they are equally efficient in the RFDH approach. In the $\alpha = 0.9$ case, the positions of DMU_3 and DMU_9 switch when passing from the RDEA ranking to the RFDH ranking. Note that for $\alpha = 0.9$ and $\alpha = 1$, all the DMUs are ranked, while as α decreases, the number of DMUs to rank decreases: when $\alpha = 0.6$, six out of the nine DMUs are ranked in both approaches.

Finally, consider ranking the DMUs by their overall efficiency, that is, taking into account the entire efficiency intervals obtained by the RDEA and FDH models. In order to do so, it is necessary to use a *maximum cost-efficiency loss index*. The proposed index is a variation of the maximum technical efficiency loss index of Wang et al. [53] and Xu et al. [54].

To fix the ideas, consider the trust level $\alpha = 0.9$. All the RDEA cost-efficient DMUs have the score of one, that is, $\left(\theta_k^*\right)_{\text{RDEA}}^{\inf(0.9)} = 1$. The maximum cost-efficiency loss index of each DMU is computed below.

$$R(DMU_1) = \max[\max\{1, 1, \ldots, 1, 1\} - 0.5283, 0] = 0.4588$$

$$R(DMU_2) = \max[\max\{1, 1, \ldots, 1, 1\} - 0.2781, 0] = 0.7181$$

$$R(DMU_3) = \max[\max\{1, 1, \ldots, 1, 1\} - 0.4274, 0] = 0.5726$$

$$R(DMU_4) = \max[\max\{1, 1, \ldots, 1, 1\} - 0.7567, 0] = 0.2433$$

$$R(DMU_5) = \max[\max\{1, 1, \ldots, 1, 1\} - 0.5040, 0] = 0.4960$$

$$R(DMU_6) = \max[\max\{1, 1, \ldots, 1, 1\} - 0.4657, 0] = 0.5343$$

$$R(DMU_7) = \max[\max\{1, 1, \ldots, 1, 1\} - 0.3824, 0] = 0.6176$$

$$R(DMU_8) = \max[\max\{1, 1, \ldots, 1, 1\} - 0.6329, 0] = 0.3671$$

$$R(DMU_9) = \max[\max\{1, 1, \ldots, 1, 1\} - 0.4350, 0] = 0.5650$$

Since DMU_4 has the smallest maximum loss index, DMU_4 is to be rated as the best performer. The complete ranking of the DMUs is as follows:

Resona (DMU_4) $\succ$ Chuo Mitsui Trust (DMU_8) $\succ$ Mizuho (DMU_1) $\succ$

Sumitomo Mitsui (DMU_5) $\succ$ Mitsubishi UFJ (DMU_6) $\succ$ Tokyo Mitsubishi UFJ (DMU_3) $\succ$

Sumitomo Trust (DMU_9) $\succ$ Mizuho Trust (DMU_7) $\succ$ Mizhuho Corporate (DMU_2).

The ranking obtained for the RFDH cost-efficient DMUs when $\alpha = 0.9$ is the same except for the fact that DMU_9 performs better than DMU_3. Table 7 shows the rankings obtained for the RDEA and the RFDH cost-efficient DMUs for the different values of α. The differences between two rankings corresponding to the same α value are indicated in bold characters.

Comparing Tables 6 and 7, it is clear that the rankings based on the maximum loss index and those based on the α-pessimistic values are the same. This further proves the efficacy of the proposed methodology.

Remark 4. One may ask if it is realistic to use data from a four years range in order to construct a rough set. Indeed, in today's economy the situation of banks can change dramatically over a span of four years. For instance, the four consecutive data values collected for an input could have been (year, value): (2012, 19), (2013, 20), (2014, 21), (2015, 10), leading to the rough set ([19, 20], [10, 21]), where the lower approximation is far from the observed current proxy (see Remark 3) of 10. The Japanese economy was rather stable during the sample years considered in the case study (fiscal years 2007–2010) due to a prolonged economic stagnation. Thus, even if some variations due to the use of balance sheet items on the inputs and outputs was recorded, this remark does not apply to our study case. However, the use of Xu et al. [54] maximum loss index actually allows for an evaluation of the performance of the banks according to ranking and sensitivity analysis. In our real world example, the data over the four years range are simultaneously used to evaluate the overall performance of

Table 7
Rankings of RDEA and RFDH cost-efficient DMUs under CRS.[a]

α	RDEA Ranking
0.6	**DMU 4** ≻ **DMU 8** ≻ DMU 3 ≻ DMU 1 ≻ DMU 5 ≻ DMU 6
0.7	DMU 4 ≻ DMU 8 ≻ DMU 1 ≻ DMU 5 ≻ DMU 3 ≻ DMU 6 ≻ DMU 2
0.8	DMU 4 ≻ DMU 8 ≻ DMU 1 ≻ DMU 5 ≻ DMU 6 ≻ DMU 3 ≻ DMU 9 ≻ DMU 2
0.9	DMU 4 ≻ DMU 8 ≻ DMU 1 ≻ DMU 5 ≻ DMU 6 ≻ **DMU 3** ≻ **DMU 9** ≻ DMU 7 ≻ DMU 2
1.0	DMU 4 ≻ DMU 8 ≻ DMU 1 ≻ DMU 5 ≻ DMU 6 ≻ DMU 9 ≻ DMU 3 ≻ DMU 7 ≻ DMU 2
α	RFDH Ranking
0.6	**DMU 4** ∼ **DMU 8** ≻ DMU 3 ≻ DMU 1 ≻ DMU 5
0.7	DMU 4 ≻ DMU 8 ≻ DMU 1 ≻ DMU 5 ≻ DMU 3 ≻ DMU 6 ≻ DMU 2
0.8	DMU 4 ≻ DMU 8 ≻ DMU 1 ≻ DMU 5 ≻ DMU 6 ≻ DMU 3 ≻ DMU 9 ≻ DMU 2
0.9	DMU 4 ≻ DMU 8 ≻ DMU 1 ≻ DMU 5 ≻ DMU 6 ≻ **DMU 9** ≻ **DMU 3** ≻ DMU 7 ≻ DMU 2
1.0	DMU 4 ≻ DMU 8 ≻ DMU 1 ≻ ≻DMU 5 ≻ DMU 6 ≻ DMU 9 ≻ DMU 3 ≻ DMU 7 ≻ DMU 2

[a] The differences between the RDEA and RFDH rankings corresponding to the same α are shown in bold.

DMUs. Rather than estimating technical crisp-data based efficiency scores for each year, the rough approach, together with Xu et al. [54] maximum loss index, provides an overall performance evaluation based on the degree of regret faced by the DMUs when losing efficiency. Finally, note that, half-year data may be used in which case seasonal adjustments may be needed. This is why we used the values at the end of fiscal year (the Japanese fiscal year ends at the end of March).□

Remark 5. Although this does not happen in the case study, it is possible to have less than four years of data or less than four distinct values available for each input, output or input price relative to one or more banks. For instance, one of the banks could register the values 1, 18, 18, 100 for one of the inputs over a four years range. Technically speaking, these kind of data do not really allow to define a rough variable anymore. The definition of rough variable requires the existence of a non-trivial lower approximation (that is, the intermediate interval $[a, \quad b]$ must not reduce to a singleton). Moreover, the trust measure as well as the α-optimistic and α-pessimistic values, that are indeed computed using the trust measure, cannot be defined if $a = b$. Nevertheless, an intermediate approach is possible that allows solving the cost-efficiency evaluation problem when both properly defined rough variables and the above-discussed case of "pseudo-rough" variables must be implemented altogether. Relative to the inputs, outputs or input prices for which less than three distinct observations are available, the following equations can be used in place of Eqs. (16) and (17), respectively.

$$\mathrm{Tr}\left\{\xi \geq r\right\} = \begin{cases} 0 \text{ if } d \leq r \\ \dfrac{1}{2}\left(\dfrac{a-r}{d-a}\right) + \dfrac{1}{2} \text{if } a \leq r \leq d \\ \dfrac{1}{2}\left(\dfrac{c-r}{a-c}\right) + 1 \ \text{ if } c \leq r \leq a \\ 1 \ \text{ if } r \leq c \end{cases} \tag{35}$$

$$\mathrm{Tr}\left\{\xi \leq r\right\} = \begin{cases} 0 \text{ if } r \leq c \\ \dfrac{1}{2}\left(\dfrac{r-c}{a-c}\right) \text{if } c \leq r \leq a \\ \dfrac{1}{2}\left(\dfrac{r-a}{d-a}\right) + \dfrac{1}{2} \text{ if } a \leq r \leq d \\ 1 \text{ if } d \leq r \end{cases} \tag{36}$$

The α-optimistic and α-pessimistic values can be modified accordingly as well as all the other notions recalled in the preliminaries. This leads to a parallel and integrated version of the interval-based methodology proposed in the previous sections, which can be used to provide performance evaluations for the DMUs and rank them in a similar manner as it has been done for the case study analyzed in this section.□

8. Conclusions

The traditional Farrell cost-efficiency model requires the data relative to inputs and outputs as well as to input prices to be known exactly. However, in real-world applications, the exact values of these data may not be available, in which case the Farrell cost-efficiency cannot be estimated. Furthermore, in the Farrell cost-efficiency model, production possibility sets are assumed to be convex. Convexity requires all inputs and outputs to be divisible, but many commodities are indivisible in practical situations, so that convexity is often violated.

To deal with these kind of problems, the present paper has developed two rough cost-efficiency approaches, a rough DEA and a rough FDH cost-efficiency model, based on the convexity and non-convexity assumptions, respectively. Acknowledging the importance of implementing different returns-to-scale assumptions, the proposed cost-efficiency models have been designed so as to apply to constant, non-increasing, non-decreasing, and variable returns-to-scale technologies. Moreover, the proposed models can be transformed into linear programming problems reducing the computational difficulty of the traditional cost-efficiency frameworks.

The results obtained have shown that managerial preferences can be implemented using the α-pessimistic and the α-optimistic operators and that cost-efficient DMUs can be stably ranked using interval values deriving from interpreting inputs, outputs and input prices as rough sets.

A numerical example and a real-life case study in the Japanese banking industry have been discussed to show the efficacy and applicability of the proposed models. In particular, the rankings of cost-efficient DMUs achieved with the implementation of the proposed rough models have been compared with those obtained using the maximum technical efficiency loss index of Wang et al. [53] and Xu et al. [54].

If, in addition to input–output and input price data, information on output prices is available in the form of imprecise data, a rough revenue efficiency and profit efficiency approach can be developed similarly. Future extensions of the current study include developing environmental efficiency models, directional distance models, and noncompetitive cost-efficiency models (e.g., Sahoo and Tone [45]) with rough data.

Acknowledgement

The authors would like to thank the anonymous reviewers and the editor-in-chief for their constructive comments and suggestions.

References

[1] P.J. Agrell, J. Tind, A dual approach to nonconvex frontier models, J. Prod. Anal. 16 (2001) 129–147.
[2] G. Alefeld, J. Herzberger, Introduction to Interval Computations, Academic Press, New York, 1983.
[3] R.D. Banker, A. Charnes, W.W. Cooper, Some method for estimating technical and scale inefficiencies in data envelopment analysis, Manag. Sci. 30 (1984) 1078–1092.
[4] H. Bagherzadeh Valami, Cost efficiency with triangular fuzzy number input prices: an application of DEA, Chaos Solut. Fractals 42 (2009) 1631–1637.
[5] W. Briec, K. Kerstens, P. Vanden Eeckaut, Non-convex technologies and cost functions: definitions, duality and nonparametric tests of convexity, J. Econ. 81 (2004) 155–192.
[6] A.S. Camanho, R.G. Dyson, Cost efficiency measurement with price uncertainty: a DEA application to bank branch assessments, Eur. J. Oper. Res. 161 (2005) 432–446.
[7] A. Charnes, W.W. Cooper, E. Rhodes, Measuring the efficiency of decision-making units, Eur. J. Oper. Res. 2 (1978) 429–444.
[8] D. Deprins, L. Simar, H. Tulkens, Measuring labor efficiency in post offices, in: M. Marchand, P. Pestieau, H. Tulkens (Eds.), The Performance of Public Enterprises Concepts and Measurements, Elsevier, Amsterdam, 1984, pp. 247–263.
[9] D.K. Despotis, Y.G. Smirlis, Data envelopment analysis with imprecise data, Eur. J. Oper. Res. 140 (2002) 24–36.
[10] D. Dubois, H. Prade, Possibility Theory: An Approach to Computerized Processing of Uncertainty, Plenum Press, New York, 1988.
[11] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, Int. J. Gen. Syst. 17 (1990) 191–209.
[12] A. Emrouznejad, B.R. Parker, G. Tavares, Evaluation of research in efficiency and productivity: a survey and analysis of the first 30 years of scholarly literature in DEA, Socio Econ. Plan. Sci. 42 (2008) 151–157.
[13] A. Emrouznejad, M. Tavana (Eds.), Performance Measurement with Fuzzy Data Envelopment Analysis, Springer, New York, 2014.
[14] T. Entani, Y. Maeda, H. Tanaka, Dual models of interval DEA and its extension to interval data, Eur. J. Oper. Res. 136 (2002) 32–45.
[15] R. Färe, S. Grosskopf, C.A.K. Lovell, The Measurement of Efficiency of Production, Klumer-Nijhoff Publishing, Boston, 1985.
[16] M.J. Farrell, The measurement of productive efficiency, J. R. Stat. Soc. Ser. A 120 (1957) 253–281.
[17] H. Fukuyama, W.L. Weber, Estimating output gains by means of Luenberger efficiency measures, Eur. J. Oper. Res. 165 (2005) 535–547.
[18] G. Gattoufi, M. Oral, A. Reisman, Data envelopment analysis literature: a bibliography update (1951–2001), Socio Econ. Plan. Sci. 38 (2004) 115–232.
[19] P. Guo, H. Tanaka, M. Inuiguchi, Self-organizing fuzzy aggregation models to rank the objects with multiple attributes, IEEE Trans. Syst. Man Cybern. Part A: Syst. Hum. 30 (2000) 573–580.
[20] A. Hatami-Marbini, A. Emrouznejad, M. Tavana, A taxonomy and review of the fuzzy data envelopment analysis literature: two decades in the making, Eur. J. Oper. Res. 214 (2011) 457–472.
[21] J.L. Hougaard, Fuzzy scores of technical efficiency, Eur. J. Oper. Res. 115 (1999) 529–541.
[22] M. Inuiguchi, F. Mizoshita, Qualitative and quantitative data envelopment analysis with interval data, Ann. Oper. Res. 195 (2012) 189–220.
[23] G. Jahanshahloo, S.M. Mirdehghan, J. Vakili, An interpretation of the cost model in data envelopment analysis, J. Appl. Sci. 11 (2011) 389–392.
[24] G.R. Jahanshahloo, M. Soleimani-damaneh, A. Mostafaee, A smplified version of the DEA cost efficiency model, Eur. J. Oper. Res. 184 (2008) 814–815.
[25] R. Khanjani Shiraz, V. Charles, L. Jalalzadeh, Fuzzy rough DEA model: a possibility and expected value approaches, Expert Syst. Appl. 1 (2) (2014) 434–444.
[26] C. Kao, Interval efficiency measures in data envelopment analysis with imprecise data, Eur. J. Oper. Res. 174 (2006) 1087–1099.
[27] K. Kerstens, P. Vanden Eechaut, Estimating return to scale using nonparametric deterministic technologies: a new method based on goodness-of-fit, Eur. J. Oper. Res. 134 (1999) 43–58.
[28] T. Kuosmanen, T. Post, Measuring economic efficiency with incomplete price information: with an application to European commercial banks, Eur. J. Oper. Res. 134 (2001) 43–58.
[29] T. Kuosmanen, T. Post, Measuring economic efficiency with incompleteprice information, Eur. J. Oper. Res. 144 (2003) 454–457.
[30] H. Leleu, A linear programming framework for free disposal hull technologies and cost functions: primal and dull models, Eur. J. Oper. Res. 168 (2006) 340–344.
[31] S. Lertworasirikul, F. Shu-Cherng, J.A. Joines, H.L.W. Nuttle, Fuzzy data envelopmentanalysis (DEA): a possibility approach, Fuzzy Sets Syst. 139 (2003) 379–394.
[32] B. Liu, Theory and Practice of Uncertain Programming, Physica Verlag, New York, 2002.
[33] B. Liu, Uncertainty Theory: An Introduction to its Axiomatic Foundations, Springer, Berlin, 2004.
[34] Z. Ma, W. Cui, Fuzzy data envelopment analysis approach based on sample decision making units, J. Syst. Eng. Electron. 23 (3) (2012) 399–407.
[35] A. Mostafaee, F.H. Saljooghi, Cost efficiency measures in data envelopment analysis with data uncertainty, Eur. J. Oper. Res. 202 (2010) 595–603.
[36] H.S. Nguyen, Approximate boolean reasoning: foundations and applications in data mining, Trans. Rough Sets V Lect. Notes Comput. Sci. 4100 (2006) 344–523.
[37] J. Paradi, S. Rouatt, J. Zhu, Two-stage evaluation of bank branch efficiency using data envelopment analysis, Omega 39 (2011) 99–109.
[38] Z. Pawlak, A. Skowron, Rudiments of rough sets, Inf. Sci. 177 (2007) 3–27.
[39] Z. Pawlak, A. Skowron, Rough sets: some extensions, Inf. Sci. 177 (2007) 28–40.
[40] Z. Pawlak, A. Skowron, Rough sets and boolean reasoning, Inf. Sci. 177 (2007) 41–73.
[41] Z. Pawlak, Rough sets, Int. J. Inf. Comput. Sci. 11 (1982) 341–356.
[42] Z. Pawlak, Rough set theory and its applications, J. Telcommun. Inf. Technol. 3 (2002) 7–10.
[43] K. Paryab, R. Khanjani Shiraz, L. Jalalzadeh, An improvement model for assessing FDH-cost efficiency, Asia Pac. J. Oper. Res. 29 (4) (2012) 1250022.
[44] S.M. Saati, A. Memariani, G.R. Jahanshahloo, Efficiency analysis and ranking of DMUs with fuzzy data, Fuzzy Optim. Decis. Making 1 (3) (2002) 255–267.
[45] B.K. Sahoo, K. Tone, Non-parametric measurement of economies of scale and scope in non-competitive environment with price uncertainty, Omega 41 (2013) 97–111.
[46] C. Schaffnit, D. Rosen, J.C. Paradi, Best practice analysis of bank branches: an application of DEA in a large Canadian bank, Eur. J. Oper. Res. 98 (1997) 269–289.
[47] C.W. Sealey, J.T. Lindley, Inputs, outputs and a theory of production and cost at depository financial institutions, J. Finance 32 (1977) 1251–1266.
[48] J.K. Sengupta, A fuzzy systems approach in data envelopment analysis, Comput. Math. Appl. 24 (1992) 259–266.
[49] N.K. Sheth, K. Triantis, Measuring and evaluating efficiency and effectiveness using goal programming and data envelopment analysis in a fuzzy environment, Yugoslav J. Oper. Res. 13 (2003) 35–60.
[50] Z. Tao, J. Xu, A class of rough multiple objective programming and its application to solid transportation problem, Inf. Sci. 188 (2012) 215–235.
[51] R.G. Thompson, P.S. Dharmapala, D.B. Humphrey, W.M. Taylor, R.M. Thrall, Computing DEA/AR efficiency and profit ratio measures with an illustrative bank application, Ann. Oper. Res. 68 (1996) 303–327.
[52] H. Tulkens, On FDH analysis: some methodological issues and applications to retail banking, courts and urban transit, J. Prod. Anal. 4 (1993) 183–210.
[53] Y.M. Wang, R. Greatbanks, J.B. Yang, Interval efficiency assessment using data envelopment analysis, Fuzzy Sets Syst. 153 (2005) 347–370.
[54] J. Xu, B. Li, D. Wu, Rough data envelopment analysis and its application to supply chain performance evaluation, Int. J. Prod. Econ. 122 (2009) 628–638.
[55] J. Xu, L. Yao, A class of expected value multi-objective programming problems with random rough coefficients, Math. Comput. Model. 50 (1–2) (2009) 141–158.
[56] J. Xu, L. Yao, A class of multiobjective linear programming models with random rough coefficients, Math. Comput. Model. 49 (1–2) (2009) 189–206.
[57] J.P. Xu, L.H. Zhao, A class of fuzzy rough expected value multi-objective decision making model and its application to inventory problems, Comput. Math. Appl. 56 (2008) 2107–2119.
[58] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets Syst. 1 (1978) 3–28.