Knowledge-Based Systems 89 (2015) 148–159

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

A fuzzy expected value approach under generalized data envelopment analysis

M.-R. Ghasemi [a,*], Joshua Ignatius [a,*], Sebastián Lozano [b], Ali Emrouznejad [c], Adel Hatami-Marbini [d]

[a] School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Minden, Penang, Malaysia
[b] Department of Industrial Management, University of Seville, Spain
[c] Aston Business School, Aston University, Brimingham, UK
[d] Louvain School of Management, Center of Operations Research and Econometrics (CORE), Université catholique de Louvain, 34 voie du roman pays, L1.03.01 B-1348 Louvain-la-Neuve, Belgium

ARTICLE INFO

Article history:
Received 20 January 2015
Received in revised form 2 June 2015
Accepted 18 June 2015
Available online 9 July 2015

Keywords:
Data envelopment analysis
Generalized data envelopment analysis
Fuzzy expected value
Super-efficiency
Symmetric and asymmetric fuzzy numbers

ABSTRACT

Fuzzy data envelopment analysis (DEA) models emerge as another class of DEA models to account for imprecise inputs and outputs for decision making units (*DMUs*). Although several approaches for solving fuzzy DEA models have been developed, there are some drawbacks, ranging from the inability to provide satisfactory discrimination power to simplistic numerical examples that handles only triangular fuzzy numbers or symmetrical fuzzy numbers. To address these drawbacks, this paper proposes using the concept of expected value in generalized DEA (GDEA) model. This allows the unification of three models – fuzzy expected CCR, fuzzy expected BCC, and fuzzy expected FDH models – and the ability of these models to handle both symmetrical and asymmetrical fuzzy numbers. We also explored the role of fuzzy GDEA model as a ranking method and compared it to existing super-efficiency evaluation models. Our proposed model is always feasible, while infeasibility problems remain in certain cases under existing super-efficiency models. In order to illustrate the performance of the proposed method, it is first tested using two established numerical examples and compared with the results obtained from alternative methods. A third example on energy dependency among 23 European Union (EU) member countries is further used to validate and describe the efficacy of our approach under asymmetric fuzzy numbers.

1. Introduction

Data envelopment analysis (DEA) was first proposed by Charnes et al. [7] and later become known as the CCR model. BCC model [6] extends the CCR model by accommodating for variable returns to scale. Concurrently, the Free Disposal Hull (FDH) model [11] was developed as an alternative DEA model which benefits from a mixed integer programming to calculate the relative efficiencies of decision making units (*DMUs*). In order to treat basic CCR, BCC and FDH models in a unified way, a generalized DEA model (GDEA) was proposed by Yun et al. [36]. Since traditional DEA models do not account for subjective input and output values, another class of DEA models emerged; that is, fuzzy DEA models [12,17].

Several solution approaches have been developed for fuzzy DEA models, which include: (1) the defuzzification approach [13,16,33], (2) the α-level based approach [3,4,27,28,37], (3) fuzzy ranking [5,15,18,19,32], (4) the possibility approach [22,25], (5) fuzzy arithmetic [34,35], and (6) the fuzzy random/type-2 fuzzy set [29–31]. Fuzzy ranking and α-cut approaches are the most popular as outlined in a survey on fuzzy DEA literature [17]. However, existing fuzzy DEA models exhibit some drawbacks.

The first major drawback of existing fuzzy DEA in the literature is the significant computational effort in solving the efficiency values. Guo and Tanaka's fuzzy ranking approach [15] needs two linear programming problems to obtain the efficiency value for any given *DMU*. The process involves feeding the optimal solution of the primary linear programming problem as coefficients of some fuzzy constraints into the second linear programming problem. The same computational complexity is also inherent in the fuzzy possibilitic approach proposed by Lertworasirikul et al. [25], where fuzzy constraints and objective function are defined across different possibility levels or α-cut. In the case of n *DMUs* and five levels

* Corresponding authors. Tel.: +60 4 6533384; fax: +60 4 6570910.
E-mail addresses: mreza.gsm@gmail.com (M.-R. Ghasemi), josh@usm.my (J. Ignatius), slozano@us.es (S. Lozano), a.emrouznejad@aston.ac.uk (A. Emrouznejad), adel.hatamimarbini@uclouvain.be (A. Hatami-Marbini).

http://dx.doi.org/10.1016/j.knosys.2015.06.025
0950-7051/

of possibility, there are 5^{n+2} linear programming problems to be solved, which remains computationally expensive. This problem also arises in α-level based approaches; it requires solving a sequence of linear programming models, thus leading to an increase in computational effort for obtaining fuzzy efficiencies of *DMU*s. Since there are different optimal solutions for each α-level, the decision maker (DM) is left to decide on which solution is the best for the scenario under his or her interpretation. In most cases, the decision analyst would decide based on the number of efficiencies that are generated across all α-cuts before deciding on the final ranking solution.

The second limitation in existing fuzzy DEA models is the focus on triangular fuzzy membership functions (see [24]) or symmetrical triangular fuzzy membership functions (see [15]). There is much left unexplored for inputs and outputs that are imprecise and do not conform to the said fuzzy membership functions.

The third drawback in existing fuzzy DEA models is its limited scope and much emphasis placed on the CCR model (see [33]). The unification of CCR, BCC and FDH comes under the category of GDEA model. Considering imprecision, Jahanshahloo et al. [21] are among the first authors to formulate the GDEA model with interval data (IGDEA); such that the upper bound efficiency value is obtained considering that the DM is optimistic for the DMU under evaluation (DMU_o), while pessimistic with the remaining *DMU*s in the evaluation set. Contrastingly, the lower bound efficiency values is obtained by considering that the DM is pessimistic for the *DMU* under evaluation (DMU_o), while optimistic with the remaining *DMU*s in the evaluation set. This is achieved by selecting only the extreme points in an interval for the input and output measures. It does not derive information using the form of a particular function, such as one expressed in fuzzy or possibilitic manner. In other words, the mid-values as appear in a fuzzy numbered dataset are effectively ignored and the results of efficiency covers a range comprising of an interval made up off overly optimistic and pessimistic in the proposed IGDEA model. Unlike previous models, our proposed fuzzy expected generalized DEA (FEGDEA) model solves both symmetrical and asymmetrical fuzzy numbers and requires less computational effort than competing models. We further propose a ranking method for efficient *DMU*s by adapting the FEGDEA and illustrate that our approach does not suffer from infeasibility issues as may be the case for existing methods.

In order to tackle the existing drawbacks in the fuzzy DEA literature, we propose the use of expected value approach for unifying all three models – fuzzy CCR, fuzzy BCC and fuzzy FDH models. In particular, our research process entails the following objectives. First, we investigate the performance of our method with existing method that handles symmetrical data. Second, we show that integrating the fuzzy expected value approach into the GDEA model outperforms integrating the fuzzy expected value in classical DEA models. Third, when efficient cases are to be ranked such as in super-efficiency analysis, the use of Andersen and Petersen [2] approach in FEGDEA model removes the issue of infeasibility, which occurs when it is applied to classical DEA models in certain cases. Fourth, we further show that having addressed all the above objectives, our proposed model is able to generate results under the CCR, BCC and FDH forms including ranking efficient units for both symmetrical and asymmetrical data.

The rest of the paper is structured as follows. Section 2 provides the preliminaries on the pertinent mathematical concepts on fuzzy DEA. Section 3 gives a brief description of the basic DEA models and GDEA model. Section 4 outlines the development of the proposed model. Section 5 illustrates a ranking method for the proposed model and suggests ways to discriminate those efficient *DMU*s. Section 6 describes the proposed method with two established numerical examples and a third example on an energy dependency case among 23 European Union (EU) member countries. The performance of our proposed model is compared to other existing methods for performance validation. Section 7 concludes the study.

2. Preliminary concepts

Definition 1. If X is a collection of objects denoted by x, called the universe, then a fuzzy set $\tilde{A}$ in X is a set of ordered pairs:

$$\tilde{A} = \{(x, \mu_{\tilde{A}})|x \in X\},$$

in which $\mu_{\tilde{A}}(x)$ is called the membership function of x in $\tilde{A}$ that $\mu_{\tilde{A}}(x) : X \to [0, 1]$.

Definition 2. The α-level (or α-cut) set of a fuzzy set $\tilde{A}$ is a crisp subset of X and is denoted by:

$$\tilde{A}_{(\alpha)} = \{x \in X|\mu_{\tilde{A}}(x) \geqslant \alpha\}.$$

Definition 3. A fuzzy set $\tilde{A}$ of set X is convex if

$$\mu_{\tilde{A}}(\lambda x_1 + (1-\lambda)x_2) \geqslant \min\{\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_1)\}, x_1, x_2 \in X, \ \lambda \in [0, 1].$$

Definition 4. A fuzzy number $\tilde{A}$ is a convex normalized fuzzy set $\tilde{A}$ of real line $\mathbb{R}$, in which there exists at least one $x_o \in \mathbb{R}$, with $\mu_{\tilde{A}}(x_o) = 1$ and $\mu_{\tilde{A}}(x)$ is piecewise continuous. A fuzzy number $\tilde{A} = (a^l, a^{m_1}, a^{m_2}, a^u)$ is a trapezoidal fuzzy number if

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{x-a^l}{a^{m_1}-a^l}, & a^l \leqslant x < a^{m_1}, \\ 1, & a^{m_1} \leqslant x \leqslant a^{m_2}, \\ \frac{a^u-x}{a^u-a^{m_2}}, & a^{m_2} < x \leqslant a^u, \\ 0, & otherwise. \end{cases}$$

The α-level set of the trapezoidal fuzzy number $\tilde{A}$ can be denoted as an interval, $[f^l(\alpha), f^u(\alpha)]$, in which $f^l(\alpha) = a^l + \alpha(a^{m_1} - a^l)$ and $f^u(\alpha) = a^u - \alpha(a^u - a^{m_2})$ where $\alpha \in [0, 1]$.

Remark 1. By assuming $a^m = a^{m_1} = a^{m_2}$ in a trapezoidal fuzzy number $\tilde{A} = (a^l, a^{m_1}, a^{m_2}, a^u)$ we obtain a triangular fuzzy number as $\tilde{A}' = (a^l, a^m, a^u)$. If we assume $a^{m_1} - a^l = a^u - a^{m_2}$ in the trapezoidal fuzzy number $\tilde{A}$ and $a^m - a^l = a^u - a^m$ in the triangular fuzzy number $\tilde{A}'$ we have symmetrical trapezoidal and triangular fuzzy numbers, respectively.

Definition 5 [20]. The expected interval (EI) and the expected value (EV) of a fuzzy number $\tilde{A}$ are defined as follows:

$$EI(\tilde{A}) = [E_1^A, E_2^A] = \left[\int_0^1 f^l(\alpha)d\alpha, \int_0^1 f^u(\alpha)d\alpha\right]; \quad EV(\tilde{A}) = \frac{E_1^A + E_2^A}{2}.$$

If we assume that $\tilde{A} = (a^l, a^{m_1}, a^{m_2}, a^u)$ is a trapezoidal fuzzy number then

$$EI(\tilde{A}) = \left[\frac{a^l + a^{m_1}}{2}, \frac{a^{m_2} + a^u}{2}\right]; \quad EV(\tilde{A}) = \frac{a^l + a^{m_1} + a^{m_2} + a^u}{4}.$$

If we further assume that $\tilde{A} = (a^l, a^m, a^u)$ is a triangular fuzzy number then