Journal of Nonlinear and Convex Analysis
Volume 25, Number 12, 2024, 3007–3017

Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

GENERALIZATION AND EXTENSION OF DATA ENVELOPMENT ANALYSIS

YEBOON YUN, JINGWEN SU, AND MIN YOON*

Abstract. Data Envelopment Analysis (DEA) is one of the most widely used methods to evaluate the relative efficiency of decision-making units (DMUs) with multiple-input and multiple-output in production systems. CCR (Charnes-Cooper-Rhodes), BCC (Banker-Charnes-Cooper), and FDH (Free Disposal Hull) models are well-known, and various models and efficiency measures have been proposed. In this study, we suggest a new efficiency measure and two models of the extended and inverted GDEA (generalized DEA) model. Through an illustrative example, the concept of the proposed assessment and models will be demonstrated, and compared with the basic DEA models.

1. Introduction

Data Envelopment Analysis (DEA) [4] is well-known as one of the methods for measuring the relative efficiency of decision-making units (DMUs) with multiple inputs and outputs in production systems. Efficiency evaluation is crucial for organizations seeking to optimize resource utilization, enhance productivity, and streamline processes. DEA can offer a comprehensive approach, considering multiple factors simultaneously, making it particularly valuable in complex decision-making environments. The conventional DEA models, for example, the CCR [3], the BCC [2], and the FDH [5] model estimate the efficiency of DMUs based on their distance from a so-called efficient frontier, which is determined under circumstance such as a decision maker's value judgment. The ratio of the weighted sum of the outputs to the weighted sum of the inputs, which is extended from Farell's concept [6], is utilized as the basic efficiency measure in these DEA models.

Conventional DEA models such as CCR, BCC, and FDH as mentioned above, estimate efficiency based on the distance from the frontier determined by decision makers' specific value judgments. However, the DEA model has inherent limitations of struggling to accommodate diverse value judgments of decision makers, which yields to make it difficult to select a suitable model according to the conditions. For that purpose, various evaluation models and efficiency measures have been developed [1, 7–10, 12, 13]. Among them, based on using domination cones, the models were suggested [7,12], which includes the CCR and BCC model. Especially, Yun *et al.* [13] suggested the generalized DEA (GDEA) model that comprehensively evaluates the efficiency of DMUs by adjusting a parameter in the model, and treats the conventional DEA models in a unified way. Since the GDEA-efficient values are

2020 *Mathematics Subject Classification.* 90B50, 90B99.
Key words and phrases. Data envelopment analysis, efficiency, extended GDEA, inverted GDEA.
*Corresponding Author.

associated with the given parameter, under such situation, their interpretation and analysis may not be uniform.

In this research, we first define a new efficiency measure to regularize whereby we try to regularize an efficient score of DMUs. DMUs with the efficient performance have that their score is one. To compare and rank these efficient DMUs, we present an extension of the GDEA model, determining a super-efficiency from Andersen and Petersen's idea [1]. This study also suggests an inverted GDEA model to evaluate DMUs from the aspect [11] that how inefficient they are. Through an illustrative example with 1-input and 1-output, the efficiencies by the proposed evaluation methods are demonstrated and finally, compablack with the basic DEA models. The rest of this paper is organized as follows: Chapter 2 presents some notations and a brief explanation of basic DEA models. In Chapter 3, a new evaluation method and two models based on the assessment will be introduced. Finally, the conclusion will be described in Chapter 4.

2. Basic DEA model

To begin with, for explanation, the notations used in this paper are summarized:

- ℓ, m, n : the number of DMUs, outputs and inputs, respectively
- DMU_i : the i-th DMU
  DMU_o : DMU of the object to be evaluated
- y_{ji} : the j-th output data of DMU_i
  x_{ki} : the k-th input data of DMU_i
- $Y \in \mathbb{R}^{m\times\ell}$: output matrix for all DMUs
  $X \in \mathbb{R}^{n\times\ell}$: input matrix for all DMUs
- $\boldsymbol{y}_i := (y_{1i}, \ldots, y_{mi})^T$: output vector of DMU_i
  $\boldsymbol{x}_i := (x_{1i}, \ldots, x_{ni})^T$: input vector of DMU_i
- ϵ : a sufficiently small positive number (for example, 10^{-7}).

Data envelopment analysis (DEA) [4] is a method for measuring the relative efficiency of DMUs performing similar tasks in a production system that consumes multiple inputs to produce multiple outputs. The basic efficiency in the DEA is measublack by the ratio of the weighted sum of outputs to the weighted sum of inputs without preassigned weights for the inputs and outputs, defined as follows:

$$\frac{\sum_{j=1}^{m} \mu_j y_j}{\sum_{k=1}^{n} \nu_k x_k}, \tag{2.1}$$

where μ_j and ν_k denotes a weight to the jth-output and kth-input, respectively.

Therefore, DEA is induced to the problem for finding the weights μ_j and ν_k to maximize the above ratio (2.1) for a DMU. By transforming the measure into a linear programming problem, the form of the CCR model [3] can be derived as the following:

$$\begin{aligned}
\underset{\mu_j,\nu_k}{\text{maximize}} \quad & \theta := \sum_{j=1}^{m} \mu_j y_{jo} && \text{(CCR)}\\
\text{subject to} \quad & \sum_{k=1}^{n} \nu_k x_{ko} = 1,\\
& \sum_{j=1}^{m} \mu_j y_{ji} - \sum_{k=1}^{n} \nu_k x_{ki} \leq 0,\ i = 1,\dots,\ell,\\
& \mu_j \geq \epsilon,\ \nu_k \geq \epsilon,\ j = 1,\dots,m; k = 1,\dots,n.
\end{aligned}$$

Solving the above problem (CCR) for a $\text{DMU}_o,\ o = 1,\dots,\ell$, the efficiency θ for the DMU_o can be assessed.

The dual form (CCR_D) to the above primal form (CCR) is derived as follows:

$$\begin{aligned}
\underset{\theta,\lambda,s_x,s_y}{\text{minimize}} \quad & \theta - \epsilon(\mathbf{1}^T \boldsymbol{s}_x + \mathbf{1}^T \boldsymbol{s}_y) && (\text{CCR}_D)\\
\text{subject to} \quad & X\boldsymbol{\lambda} - \theta \boldsymbol{x}_o + \boldsymbol{s}_x = \mathbf{0},\\
& Y\boldsymbol{\lambda} - \boldsymbol{y}_o - \boldsymbol{s}_y = \mathbf{0},\\
& \boldsymbol{\lambda} \geq \mathbf{0}, \boldsymbol{s}_x \geq \mathbf{0}, \boldsymbol{s}_y \geq \mathbf{0},\\
& \theta \in \mathbb{R}, \boldsymbol{\lambda} \in \mathbb{R}^\ell, \boldsymbol{s}_x \in \mathbb{R}^n, \boldsymbol{s}_y \in \mathbb{R}^m.
\end{aligned}$$

Adding the constraint $\mathbf{1}^T\boldsymbol{\lambda} = 1$ to the dual form (CCR_D) can lead to the linear form of the BCC model [2]

$$\begin{aligned}
\underset{\theta,\lambda,s_x,s_y}{\text{minimize}} \quad & \theta - \epsilon(\mathbf{1}^T \boldsymbol{s}_x + \mathbf{1}^T \boldsymbol{s}_y) && (\text{BCC}_D)\\
\text{subject to} \quad & X\boldsymbol{\lambda} - \theta \boldsymbol{x}_o + \boldsymbol{s}_x = \mathbf{0},\\
& Y\boldsymbol{\lambda} - \boldsymbol{y}_o - \boldsymbol{s}_y = \mathbf{0},\\
& \mathbf{1}^T\boldsymbol{\lambda} = 1,\\
& \boldsymbol{\lambda} \geq \mathbf{0}, \boldsymbol{s}_x \geq \mathbf{0}, \boldsymbol{s}_y \geq \mathbf{0},\\
& \theta \in \mathbb{R}, \boldsymbol{\lambda} \in \mathbb{R}^\ell, \boldsymbol{s}_x \in \mathbb{R}^n, \boldsymbol{s}_y \in \mathbb{R}^m,
\end{aligned}$$

and imposing the constraints $\lambda_i \in \{0, 1\}$ $(i = 1,\dots,\ell)$ as well, the form of the FDH model [5] is conducted below:

$$\begin{aligned}
\underset{\theta,\lambda,s_x,s_y}{\text{minimize}} \quad & \theta - \epsilon(\mathbf{1}^T \boldsymbol{s}_x + \mathbf{1}^T \boldsymbol{s}_y) && (\text{FDH}_D)\\
\text{subject to} \quad & X\boldsymbol{\lambda} - \theta \boldsymbol{x}_o + \boldsymbol{s}_x = \mathbf{0},\\
& Y\boldsymbol{\lambda} - \boldsymbol{y}_o - \boldsymbol{s}_y = \mathbf{0},\\
& \mathbf{1}^T\boldsymbol{\lambda} = 1; \lambda_i \in \{0,1\},\ i = 1,\dots,\ell,\\
& \boldsymbol{\lambda} \geq \mathbf{0}, \boldsymbol{s}_x \geq \mathbf{0}, \boldsymbol{s}_y \geq \mathbf{0},\\
& \theta \in \mathbb{R}, \boldsymbol{\lambda} \in \mathbb{R}^\ell, \boldsymbol{s}_x \in \mathbb{R}^n, \boldsymbol{s}_y \in \mathbb{R}^m.
\end{aligned}$$

3. An illustrative example

In order to illustrate DEA analysis by using the efficient frontier in the DEA models, we consider the numerical example shown in Table 1, where the aim is to evaluate six DMUs with 1-input and 1-output.

Table 1. Example with 1-input and 1-output

DMU	A	B	C	D	E	F
input	3	4	9	7	6	8
output	2	4	7	3	5	5.5

Basically, DEA models optimize the performance of all DMUs with respect to the efficient frontier (in a sense) derived from the observed data. From the constraint for $\boldsymbol{\lambda}$ in the DEA models, the feasible set (production possibility set) alters, which yields the diverse types of the efficient frontier (see Fig.1–3).

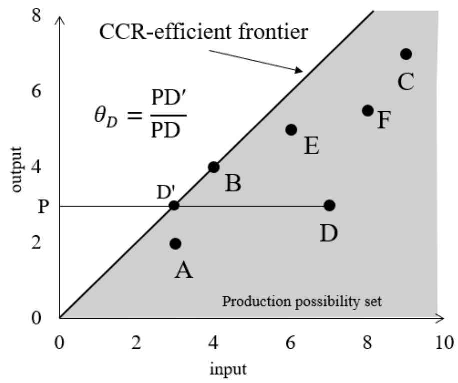


Figure 1. Efficient frontier in the CCR Model

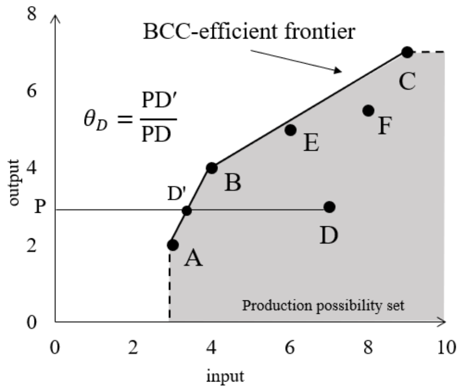


Figure 2. Efficient frontier in the BCC Model

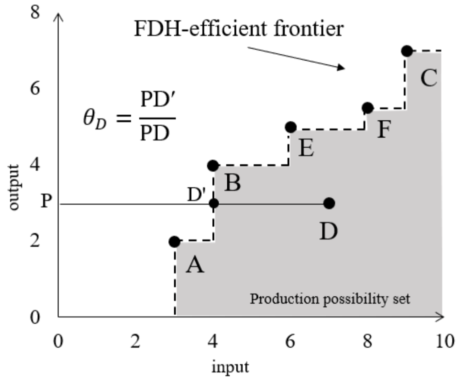


FIGURE 3. Efficient frontier in the FDH Model

Based on DMUs' location from the DEA-efficient frontier, DEA calculates their efficient values (scores). For example, the efficient value θ_D becomes $\theta_D = \mathrm{PD}'/\mathrm{PD}$. This value is obtained by solving the above DEA formulations, and for all DMUs, their efficient values are shown in Table 2.

TABLE 2. Efficient values by the basic DEA models for Table 1

DMU	A	B	C	D	E	F
CCR model	0.67	1	0.78	0.43	0.83	0.69
BCC model	1	1	1	0.50	0.94	0.81
FDH model	1	1	1	0.57	1	1

As seen from the figures of the efficient frontiers and the table of the values, all of the DMUs on the efficient frontier have the efficient values of one ($\theta = 1$), and are called DEA-*efficient*; conversely, the DMUs not located on the efficient frontier are defined as DEA-*inefficient*.

Later, the GDEA model [13] suggested, and treats the basic DEA models such as the CCR model, BCC model, and FDH model in a unified way by adjusting the parameter α. Table 3 shows various GDEA-efficient values, where the DMU with the efficient values of zero is defined GDEA-*efficient*; the GDEA-*efficient* DMUs have a positive value.

TABLE 3. Efficient values by GDEA model for Table 1

	DMU	A	B	C	D	E	F	DEA-efficiency
(i)	$\alpha = 15$ $(\boldsymbol{x}_o^T \boldsymbol{\nu} = \boldsymbol{y}_o^T \boldsymbol{\mu})$	11.00	0	4.50	31.00	4.46	9.61	CCR
(ii)	α=15	0	0	0	26.50	0.13	7.28	BCC
(iii)	α=5	0	0	0	9.33	0	1.75	
(iv)	α=1	0	0	0	2.67	0	0	FDH
(v)	α=0.1	0	0	0	1.17	0	0	

As shown in the above table, the GDEA model provides the same efficiencies by the basic DEA models, and particularly, the efficiency for the case of (iii) between the BCC model and the FDH model is evaluated by only the GDEA model (see Fig.4). However, even though the DMU_D for the case (iv) and (v) is GDEA-inefficient in terms of the FDH model, the efficient values of the DMU are different. Since the GDEA-efficient values are associated with the given value α, under such situation, their interpretation and analysis may not be uniform.

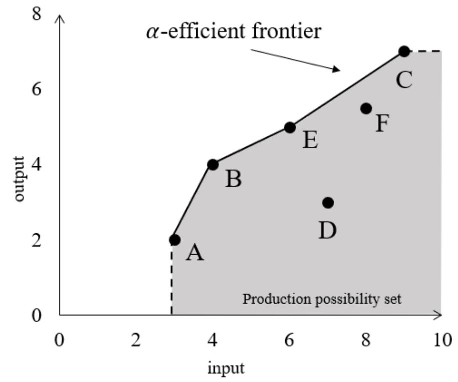


FIGURE 4. Efficient frontier in the GDEA Model ($\alpha = 5$)

4. GENERALIZED DEA MODEL

In this section, we define a new efficiency measure whereby the efficient value of DMUs is regularized not relying on the value of the parameter α. DMUs with the efficient performance have their score of one.

Extended GDEA model. To compare and rank these efficient DMUs, we present an extension of the GDEA model, using the concept a super-efficiency from Andersen and Petersen's idea [1]; for a given value $\alpha(\geq 0)$,

$$
\begin{aligned}
&\underset{\Delta,\mu_j,\nu_k}{\text{minimize}} \quad \Delta && \text{(ex-GDEA)} \\
&\text{subject to} \quad \Delta \geq \tilde{d}_i + \alpha\left(\sum_{j=1}^{m}\mu_j(y_{ji}-y_{jo}) + \sum_{k=1}^{n}\nu_k(-x_{ki}+x_{ko})\right), \\
&\qquad\qquad i = 1,\ldots,\ell,\ i \neq o, \\
&\qquad\qquad \sum_{j=1}^{m}\mu_j + \sum_{k=1}^{n}\nu_k = 1, \\
&\qquad\qquad \mu_j, \nu_k \geq \epsilon,\ j = 1,\ldots,m; k = 1,\ldots,n,
\end{aligned}
$$

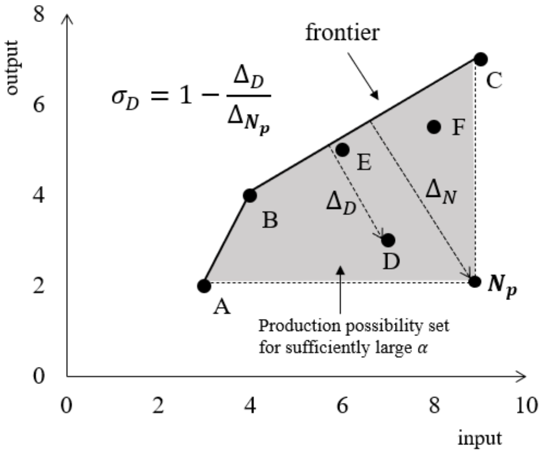
where $\tilde{d}_i = \min\ \{y_{ji} - y_{jo}, -x_{ki} + x_{ko}\}$, $i = 1,\ldots,\ell,\ i \neq o$.

In order to build a new assessment method of the efficiency in the above ex-GDEA model, we consider two virtual DMUs (see Fig.5):

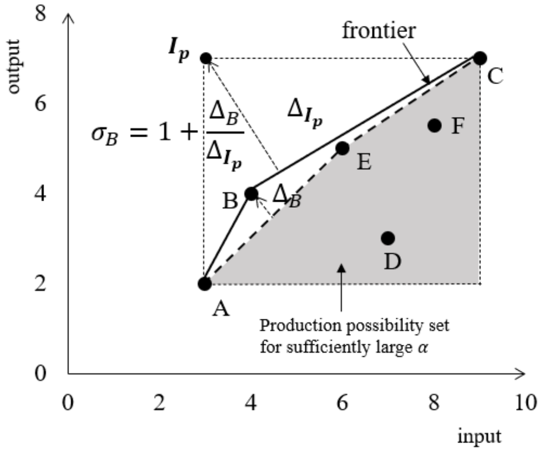
- DMU_{N_p} with the worst outputs and inputs
- DMU_{I_p} with the best outputs and inputs

Then, finding the optimal value $\Delta^*_{N_p}$, $\Delta^*_{I_p}$ and Δ^*_i $(i = 1, \dots, \ell)$, a DMU is evaluated based on the following measure:

$$\sigma^*_i := 1 - \text{sign}(\Delta^*_i) \max\left(\frac{\Delta^*_i}{\Delta^*_{N_p}}, \frac{\Delta^*_i}{\Delta^*_{I_p}}\right) \tag{4.1}$$



(a) the case for an inefficient DMU with $\Delta_i > 0$



(b) the case for an efficient DMU with $\Delta_i \leqq 0$

FIGURE 5. Efficiency evaluation in the ex-GDEA model

The formula (4.1) implies that a DMU_i with $\Delta_i^* > 0$ is evaluated by

$$\sigma_i^* = 1 - \frac{\Delta_i^*}{\Delta_{N_p}^*},$$

which indicates that it is *inefficient* in terms of the ex-GDEA-efficiency, as seen from Fig.5 (a).

For the case with $\Delta_i^* \leqq 0$ of Fig.5 (b), the formula (4.1) becomes

$$\sigma_i^* = 1 + \frac{\Delta_i^*}{\Delta_{I_p}^*},$$

which yields that a DMU is defined ex-GDEA-*efficient*.

By the first constraint of the form of the ex-GDEA model $i \neq o$, the efficient frontier is generated by the DMUs excluding the DMU_o, which allows an efficient DMU to take a value greater than one. For example, B is evaluated based on the frontier which is represented with the dotted line (A–E–C) in Fig.5(b).

Inverted ex-GDEA model. To also assess DMUs from the aspect that *how inefficient* a DMU is, this research considers an inverted ex-GDEA model by exchanging the positions of inputs and outputs. Fig.6 illustrates the meaning of the evaluation by the inverted ex-GDEA model (in-ex-GDEA), and for the example shown in Table 1, the efficient values by the ex-GDEA and in-ex-GDEA model are summarized in Table 4. Thus, if a value σ of a DMU is smaller than one, the DMU is efficient from the viewpoint of the assessment with the in-ex-GDEA model.

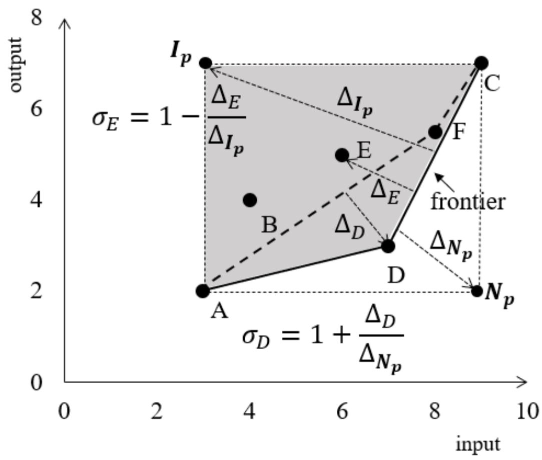


FIGURE 6. Efficiency evaluation in the in-ex-GDEA model

Table 4. Efficient values by the proposed GDEA models

Model	DMU	A	B	C	D	E	F
ex-GDEA	$\alpha = 15$ $(\boldsymbol{x}_o^T\boldsymbol{\nu} = \boldsymbol{y}_o^T\boldsymbol{\mu})$	0.82	1.35	0.93	0.49	0.93	0.84
	α=15	1.43	1.25	1.64	0.45	0.99	0.85
	α=5	1.41	1.31	1.61	0.48	1.08	0.90
	α=0.1	1.34	1.63	1.51	0.65	1.34	1.17
in-ex-GDEA	$\alpha = 15$ $(\boldsymbol{x}_o^T\boldsymbol{\nu} = \boldsymbol{y}_o^T\boldsymbol{\mu})$	0.78	0.52	0.71	1.58	0.62	0.76
	α=15	1.58	0.67	1.58	1.69	0.68	0.98
	α=5	1.55	0.70	1.55	1.78	0.70	1.04
	α=0.1	1.50	0.75	1.50	2.00	0.75	1.50

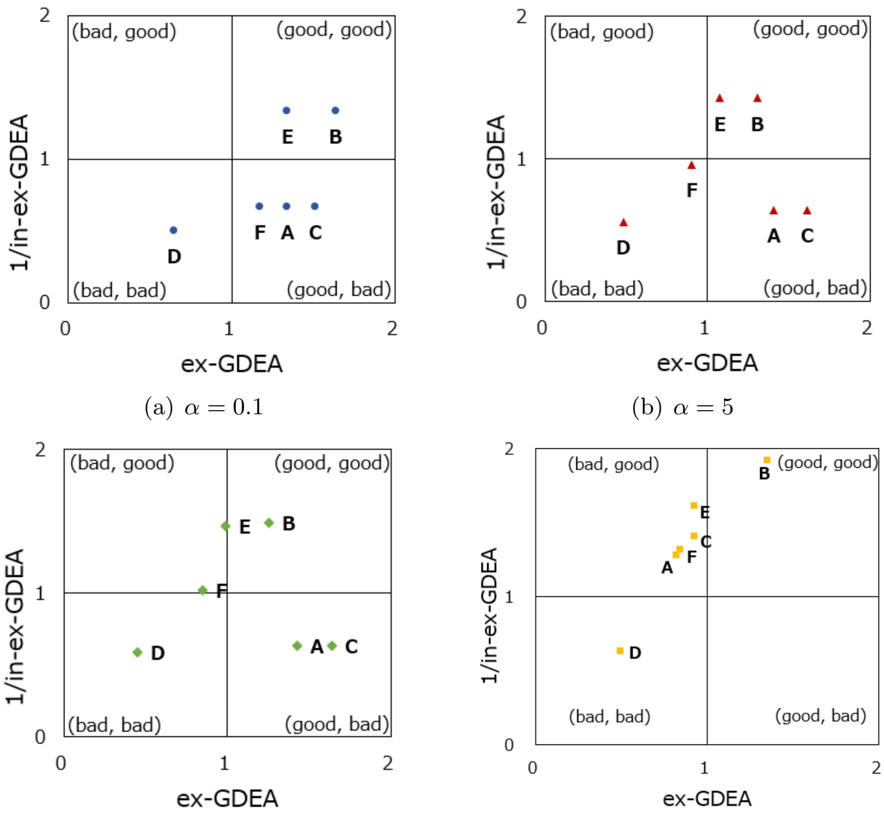
Comprehensive Assessment. To support a decision maker to comprehensively evaluate DMUs under her/his various value judgments, considering the evaluation by the proposed two models, we classify DMUs into four categories as follows. The scatter plot is shown in Fig.7, where the efficient values in the in-ex-GDEA model are taken as a reciprocal:

- Standard : good from the viewpoint of both ex-GDEA in-ex-GDEA
- Unique : good from the viewpoint of ex-GDEA and bad from in-ex-GDEA
- Ordinary : bad from the viewpoint of ex-GDEA and good from in-ex-GDEA
- Need to improve : bad from the viewpoint of both ex-GDEA and in-ex-GDEA

As also known from the figure, D is poorly appraised for all judgments, which indicates that it strongly needs to make an effort to blackuce inputs/increase outputs. B and E got a good evaluation because they have a good balance of output and input. A and C are evaluated as (good, bad), which means that they are evaluated with their characteristics.

5. Concluding remarks

In this study, we suggested two forms in the GDEA model, and defined the efficiency measures. The extensions of the GDEA model, the ex-GDEA model and in-ex-GDEA model, were proposed to evaluate the efficiency of DMUs corresponding to a decision maker's various value judgement. In addition, through an illustrative example, the efficiencies by the proposed evaluation method were demonstrated and compablack with ones by the basic DEA models. It can be expected that the proposed GDEA method will be helpful for decision makers to evaluate the efficiency more comprehensively in DEA. As a future work, we will apply for more and various cases, and investigate the effectiveness of our assessment measure.



(a) $\alpha = 0.1$ (b) $\alpha = 5$

(c) $\alpha = 15$ (d) $\alpha = 15$ ($\boldsymbol{x}_o^T \boldsymbol{\nu} = \boldsymbol{y}_o^T \boldsymbol{\mu}$)

FIGURE 7. Scatter plot with the efficient values by the proposed models

REFERENCES

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

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

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

[4] W. W. Cooper, L. M. Seiford and K. Tone, *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and Dea-Solver Software*, Kluwer Academic, 2000.

[5] D. Deprins, L. Simar and H. Tulkens, *Measuring labor-efficiency in post offices*, in: Performance of Public Enterprises: Concepts and Measurements, M. Marchand, P. Pestieau, and H. Tulkens (eds.), Amsterdam: North-Holland, Chapter 10, 1984, pp. 243–268.

[6] M. J. Farrell, *The measurement of productivity efficiency*, Journal of the Royal Statistical Society **120** (1957), 253–281.

[7] V. E. Krivonozhko, O. B. Utkin, M. M. Safin and A. V. Lychev, *On some generalization of the DEA models, Journal of the Operational Research Society* **60** (2009), 1518–1527.

[8] X. L. Meng and F. G. Shi, *An extended data envelopment analysis for the decision-making*, Journal of Inequalities and Applications **240** (2017): Article number 240.

[9] Q. Wei, H. Yan and L. Xiong, *A bi-objective generalized data envelopment analysis model and point-to-set mapping projection*, European Journal of Operational Research, **190** (2008), 855–876.
[10] Q. Wei, J. Zhang and X. Zhang, *An inverse DEA Model for inputs/outputs estimate*, European Journal of Operational Research **121** (2000), 151–163.
[11] Y. Yamada, T. Matsui and M. Sugiyama, *An inefficiency measurement method for management systems*, Journal of the Operations Research Society of Japan **37** (1994), 158–168 (in Japanese).
[12] G. Yu, Q. Wei and P. Brockett, *A generalized data envelopment analysis model: a unification and extension of existing methods for efficiency analysis of decision making units*, Annals of the Operations Research **66** (1996), 47–89.
[13] Y. B. Yun, H. Nakayama and T. Tanino, *A generalized model for data envelopment analysis*, European Journal of Operational Research **157** (2004), 87–105.

Manuscript received April 15, 2024
revised July 23, 2024

Yeboon Yun
Department of Civil, Environmental and Applied Systems Engineering, Kansai University, Japan
E-mail address: yeboon@kansai-u.ac.jp

Jingwen Su
Graduate School of Science and Engineering, Kansai University, Japan
E-mail address: sujingwen927@gmail.com

Min Yoon
Department of Applied Mathematics, Pukyong National University, Republic of Korea
E-mail address: myoon@pknu.ac.kr