
A new approach on the lowest cost problem in data envelopment analysis

Xu Wang*

Department of Industrial and Management Systems Engineering,
Waseda University,
3-4-1 Okubo, Shinjuku-ku,
Tokyo 169-8555, Japan
Email: xu.wang@aoni.waseda.jp
*Corresponding author

Kuan Lu

NCE CTO Office,
Huawei Technologies Co., Ltd.,
No. 165 Beijing Rd, Z-park, Hai Dian Dist.,
Beijing 100094, China
Email: lvkuan1@huawei.com

Takashi Hasuike

Department of Industrial and Management Systems Engineering,
Waseda University,
3-4-1 Okubo, Shinjuku-ku,
Tokyo 169-8555, Japan
Email: thasuike@waseda.jp

Abstract: This paper aims at solving the lowest cost problem in data envelopment analysis (DEA), which is to provide an efficient target for an inefficient decision making unit (DMU) with the lowest adjustment costs. For this purpose, a new approach based on the least distance DEA model is proposed. Here, the marginal costs of adjusting the inputs and outputs are assumed to be known and symmetrical. For the practical merit, different with the existing studies, our approach is able to increase inputs and decrease outputs. Numerical experiments are conducted to compare the performance of the proposed approach with previous existing studies. The results show that the proposed approach can always provide an efficient target with no higher total adjustment costs than the costs of targets provided by previous approaches. Thus, the proposed approach is more practical and useful.

Keywords: data envelopment analysis; DEA; lowest cost problem; least distance model; symmetrical marginal costs.

Reference to this paper should be made as follows: Wang, X., Lu, K. and Hasuike, T. (2021) 'A new approach on the lowest cost problem in data envelopment analysis', *Asian J. Management Science and Applications*, Vol. 6, No. 1, pp.69–84.

Biographical notes: Xu Wang is a PhD student in Department of Industrial and Management Systems Engineering at Waseda University, Tokyo, Japan. He received his Master's degree from Tokyo University of Science, Tokyo, Japan. His research interests include data envelopment analysis, optimisation and operations research. He has published a paper in *Asia-Pacific Journal of Operational Research*.

Kuan Lu received his Master's degree from Tokyo University of Science and PhD from Tokyo Institute of Technology. His research interests include multiobjective optimisation, mathematical programming and operations research. He has published papers in journals such as *Optimization Letters*, *Asia-Pacific Journal of Operational Research* and *Journal of the Operations Research Society of Japan*.

Takashi Hasuike is a Professor in School of Creative Science and Engineering at Waseda University. His research interests include informatics, multiobjective and network optimisation and operations research. He has published papers in journal such as *Information Sciences*, *Journal of Global Optimization* and *Omega*.

This paper is a revised and expanded version of a paper entitled 'Minimizing target setting over the whole efficient frontier in data envelopment analysis' presented at 2019 Asian Conference of Management Science and Applications, Penglai, Shandong, China, 11–14 October 2019.

1 Introduction

Data envelopment analysis (DEA) has been widely used as a means of relative efficiency evaluation since the first DEA model was introduced in Charnes et al. (1978). It uses mathematical programming techniques and models to evaluate the relative efficiency of decision making units (DMUs) with multiple inputs and outputs, and generates 'what to do' information. Thus, it provides an efficient target for inefficient DMUs. There are a number of classical models that are widely applied for efficiency evaluation and target setting, such as Charnes, Cooper and Rhodes (CCR) model proposed in Charnes et al. (1978), Banker Charnes Cooper (BCC) model proposed in Banker et al. (1984) and slacks-based measure (SBM) model proposed in Tone (2001). However, because how to measure the difference between the DMU under evaluation and the efficient target is not defined in these classical models, the required adjustments of inputs and outputs to achieve efficiency are not considered. Brieç (1999) introduced ℓ_p -norm distance into DEA to measure this difference and proposed that the least distance based efficiency measures should be applied. This kind of DEA model minimises the distance between the DMU under evaluation and the efficient frontier, which can minimise the required adjustments of inputs and outputs. Least distance DEA models have been actively studied in recent years for their practicality, such as Ando et al. (2012), Fukuyama et al. (2014), Ando et al. (2017), Wang et al. (2020). Nonetheless, even if the required adjustments to achieve efficiency are the smallest, the total adjustment costs are not always the lowest, since the least distance models are cost-blind. To demonstrate this, we give a simple data example that has 4 DMUs and each DMU uses 2 inputs to generate 1 output in Table 1. In this data example, the marginal costs of adjusting one unit x_1 , x_2 and y are assumed as

7, 5 and 3. DMU_A , DMU_C and DMU_D can be determined to be efficient easily by using classical DEA models. When we evaluate inefficient DMU_B , the closest efficient target E and the efficient target with the lowest adjustment costs F for DMU_B are listed in Table 2¹. The approaches in Aparicio et al. (2007) and Wang et al. (2020) can provide the closest efficient target E for DMU_B in this example. Although the efficient target F 's required adjustments are larger, the total adjustment costs are lower than E 's. In practice, most organisations prefer to minimise adjustment costs when improving efficiency. Thus, a specific approach to solve the lowest cost problem in DEA is necessary. In fact, our proposed approach in Section 2 can provide the efficient target F with the lowest adjustment costs for DMU_B (see detailed computation process in Appendix).

Table 1 A simple data example

<i>DMU</i>	<i>Inputs</i>		<i>Outputs</i>	<i>Efficiency</i>
	x_1	x_2	y	
A	6	8	13	Yes
B	7	9	4	No
C	2	12	11	Yes
D	8	5	12	Yes

Table 2 The efficient target for DMU_B

<i>Efficient target</i>	x_1	x_2	y	<i>Required adjustments</i>	<i>Adjustment costs</i>
E	4.57	9	11.43	9.86	39.26
F	7	6.17	11.83	10.67	37.67

Recently, Pendharkar (2015) proposed two heuristic approaches based on the input-oriented BCC model for the lowest cost problem². The marginal costs of decreasing inputs and increasing outputs are assumed to be known and available. However, since the input-oriented BCC model can not always provide a closest efficient target, Pendharkar (2015)'s approach can not always provide a theoretically guaranteed efficient target with the lowest adjustment costs as well. The approach also has two other limitations:

- a the approach might be infeasible when output constraints are considered
- b the obtained targets may not always be efficient (on the efficient frontier).

In fact, it is not appropriate to develop an approach for the lowest cost problem based on the classical DEA models. To deal with above problems, Ang et al. (2019) proposed an approach based on the least distance model in Aparicio et al. (2007). This approach is no-oriented and can guarantee that the obtained targets are efficient. However, the feasible region of the model in Aparicio et al. (2007) is D_o , which is only a subset of the efficient frontier. Ang et al. (2019)'s approach also searched the target over D_o , so an efficient target with the lowest adjustment costs may not always be provided. This limitation is exemplified by the results of the numerical experiments in Section 3. Moreover, these studies do not consider that some inputs could be increased or that some outputs could be reduced. For example, it is very normal for a firm to hire more staff and increase investments in business practice. In terms of adjustments to outputs of production processes, in the real world, industrial production typically generates

undesirable outputs, such as smoke pollution or wastewater (see Seiford and Zhu, 2002; Cook and Zhu, 2013; Guo and Wu, 2013). These types of undesirable outputs should be reduced. Therefore, if the adjustments of inputs/outputs are allowed and adjustment costs are known, it may be better to increase some of them while decreasing others to ensure that overall goal of adjustment costs minimisation is achieved.

The aforementioned is the motivation for this paper. To propose a new approach, we extend the model in Wang et al. (2020). Wang et al. (2020) focused on the ℓ_p -norm distance from the assessed DMU to the efficient frontier and proposed a new least distance model that can provide the closest efficient target over the whole efficient frontier under constant return scale. However, the adjustment costs of inputs and outputs are not considered in the model. To fill in this gap, we extend the model in Wang et al. (2020) and propose a new approach in this paper. The new approach proposed in this paper can provide an efficient target with the lowest total adjustment costs over the whole efficient frontier under variable return scale. The marginal costs of adjusting the inputs and outputs are assumed to be known and available, as in the extant literature (Pendharkar, 2015; Ang et al., 2019). Since we consider the increase of inputs and decrease of outputs, the costs of adjustment are assumed to be symmetrical³, which is a common assumption in economics. Under these two assumptions, we can formulate the lowest cost problem into a mathematical programming with linear complementarity constraints. This problem can be solved easily using the mixed integer programming (MIP) approach. Our proposed approach is no-oriented and can guarantee that the obtained targets are efficient. To compare the performance with that of the previous approaches, numerical experiments with 48 simulated datasets are conducted. The 48 simulated datasets with different numbers of inputs/outputs and correlations between inputs are generated with the same procedure as Pendharkar (2015) and Ang et al. (2019). The results of the numerical experiments show that the approach proposed in this paper can always provide an efficient target with no higher adjustment costs compared with the previous approaches. Therefore, this paper's contributions can be summarised as follows:

- propose an approach to DEA that minimises the total adjustment costs incurred when transitioning an inefficient DMU to an efficient target
- enable the real world condition that some inputs could be increased or some outputs could be reduced to be reflected in the target setting process.

The reminder of the paper is organised as follows. We introduce several notations used in this paper and basic concepts in DEA, and develop a new approach to solve the lowest cost problem over the whole efficient frontier in Section 2. The results of numerical experiments are reported and compared with those of Ang et al. (2019) in Section 3. Conclusions and several remarks are presented in the last section.

2 A novel approach for lowest cost problem

In this section, several basic definitions in DEA are introduced as Ray (2004) and Cook and Zhu (2013). The efficient frontier under the variable returns to scale (VRS) assumption can be expressed as a system of linear complementarity constraints according to Wang et al. (2020). Using the explicit expression of the efficient frontier, we develop a

new approach for the lowest cost problem in DEA. Some of the notations used throughout this paper are introduced as follows.

Notation

- Let there be n DMUs, DMU_j ($j = 1, 2, \dots, n$) is the j^{th} DMU
- DMU_o is the DMU under evaluation
- $x_j = (x_{1j}, x_{2j}, \dots, x_{mj})^T$ is the input data vector of DMU_j
- $y_j = (y_{1j}, y_{2j}, \dots, y_{sj})^T$ is the output data vector of DMU_j
- $X = (x_1, x_2, \dots, x_n)$ is the input data matrix
- $Y = (y_1, y_2, \dots, y_n)$ is the output data matrix
- Let DMU_1 to DMU_k be the efficient DMUs

$$X_E = (x_1, x_2, \dots, x_k),$$

$$Y_E = (y_1, y_2, \dots, y_k),$$

- $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)^T$ represents the weight of each DMU
- 1 is a vector with all elements equal to one
- 0 is the zero vector
- Assume that the costs of the adjustment are symmetrical

$$c^I = (c_1^I, c_2^I, \dots, c_m^I)^T \text{ is the marginal cost vector of adjusting the inputs}$$

$$c^O = (c_1^O, c_2^O, \dots, c_s^O)^T \text{ is the marginal cost vector of adjusting the outputs}$$

- $s_o^- = (s_{1o}^-, s_{2o}^-, \dots, s_{mo}^-)^T$ is the input slacks vector of DMU_o
- $s_o^+ = (s_{1o}^+, s_{2o}^+, \dots, s_{so}^+)^T$ is the output slacks vector of DMU_o .

In DEA, each DMU_j uses m inputs x_j to generate s outputs y_j . A pair of inputs x and outputs y is called an activity which can be expressed by the notation (x, y) . DEA technology is represented by the following production possibility set T :

$$T = \{(x_j, y_j) \mid x_j \text{ can produce } y_j\}.$$

When we assume VRS, then T can be characterised as:

$$T = \{(x_j, y_j) \mid X\lambda \leq x_j, Y\lambda \geq y_j, 1^T \lambda = 1, \lambda \geq 0\},$$

which is used in this paper. On the other hand, the efficient frontier is a set of all efficient DMUs, which is defined as:

$$E = \{(x, y) \in T \mid (x', -y') \leq (x, -y), (x', -y') \neq (x, -y) \Rightarrow (x', -y') \notin T\}.$$

When a DMU_{*j*} (*x_j*, *y_j*) ∈ *T* is given, it can be easily determined whether it is efficient. However, it is difficult to describe the area of the efficient frontier because the definition is given in an implicit way. So, the original definition of the efficient frontier can not be easily exploited into an algorithm for computation. Based on the work of Wang et al. (2020), the following system can be used to define the efficient frontier under VRS assumption computation-friendly.

$$E = \left\{ (x_j, y_j) \left\{ \begin{array}{l} X_E \lambda - x_j = 0, Y_E \lambda - y_j = 0, 1^T \lambda = 1, \lambda \geq 0. \\ u_1 - u_3 + 1 = 0, \\ -u_2 + u_4 - 1 = 0, \\ X_E^T u_3 - Y_E^T u_4 - u_5 + 1u_6 = 0, \\ \lambda^T u_5 = 0, \\ u_1, u_2, u_3, u_4, u_5 \geq 0, \\ \text{has a solution } u_1, u_2, u_3, u_4, u_5, u_6, \lambda. \end{array} \right. \right\}.$$

The derivation process for this conclusion can be extended from Appendix in Wang et al. (2020). Hence, we omit it here. Notably, the efficient frontier used in this paper is different from that used in previous studies, such as those of Aparicio et al. (2007), Zhu et al. (2018), Ang et al. (2019) and Li et al. (2020). Those studies used a subset of the efficient frontier which dominates the DMU under evaluation. However, we used the whole efficient frontier in DEA, which enables the increase of inputs and decrease of outputs. Furthermore, using the whole efficient frontier can guarantee to obtain an efficient target with the lowest adjustment costs.

When DMU_{*o*} is under evaluation, the proposed approach for solving the lowest cost problem over the whole efficient frontier is as follows⁴.

$$\left| \begin{array}{ll} \text{minimise} & c^{I^T} |x_j - x_o| + c^{O^T} |y_j - y_o| \\ x_j, y_j, \lambda, u_1, u_2, u_3, u_4, u_5, u_6 & \\ \text{subject to} & X_E \lambda - x_j = 0, Y_E \lambda - y_j = 0, 1^T \lambda = 1, \\ & u_1 - u_3 + 1 = 0, \\ & -u_2 + u_4 - 1 = 0, \\ & X_E^T u_3 - Y_E^T u_4 - u_5 + 1u_6 = 0, \\ & \lambda^T u_5 = 0, \\ & \lambda, u_1, u_2, u_3, u_4, u_5 \geq 0, \end{array} \right. \tag{1}$$

The problem (1) with the linear complementarity constraints $\lambda^T u_5 = 0$ is equivalent to the following MIP problem. Additionally, the transformation is trivial. There are many approaches for solving this kind of problem, the MIP approach is used in this study for convenience.

$$\begin{array}{l}
 \text{minimise} \quad c^{I^T} |x_j - x_o| + c^{O^T} |y_j - y_o| \\
 x_j, y_j, \lambda, u_1, u_2, u_3, u_4, u_5, u_6 \\
 \text{subject to} \quad X_E \lambda - x_j = 0, Y_E \lambda - y_j = 0, 1^T \lambda = 1, \\
 u_1 - u_3 + 1 = 0, \\
 -u_2 + u_4 - 1 = 0, \\
 X_E^T u_3 - Y_E^T u_4 - u_5 + 1u_6 = 0, \\
 \lambda \leq Mb, u_5 \leq M(1-b), b \in \{0, 1\}^n, \\
 \lambda, u_1, u_2, u_3, u_4, u_5 \geq 0,
 \end{array} \tag{2}$$

where M is a sufficiently large positive value.

For any feasible solution $(x_j, y_j, \lambda, u_1, u_2, u_3, u_4, u_5, u_6, b)$ of (2), (x, y) is in E . Moreover, for any (x_j, y_j) in E , there exists $(x_j, y_j, \lambda, u_1, u_2, u_3, u_4, u_5, u_6, b)$ such that $(x_j, y_j, \lambda, u_1, u_2, u_3, u_4, u_5, u_6, b)$ is the feasible solution of problem (2). So, obviously, if we obtain the optimisation solution of (2), then we can obtain the efficient target with the lowest adjustment costs. Thus, the limitations of Pendharkar (2015)'s approach can be solved. Then, we compare our approach with Ang et al. (2019)'s approach, which can also be solve by a MIP approach and can be expressed as follows.

$$\begin{array}{l}
 \text{minimise} \quad c^{I^T} |x_j - x_o| + c^{O^T} |y_j - y_o| \\
 s_o^+, s_o^-, \lambda, v, u, d, b, u_o \\
 \text{subject to} \quad X_E \lambda = x_o - s_o^-, Y_E \lambda = y_o + s_o^+, \lambda^T 1 = 1, \\
 -X_E v + Y_E u + u_o + d = 0, \\
 d \leq Mb, \lambda \leq (1-b), b \in \{0, 1\}^n, \\
 v \geq 1, u \geq 1, \\
 \lambda, s_o^-, s_o^+, d \geq 0,
 \end{array} \tag{3}$$

where M is a sufficiently large positive value.

As noted, Ang et al. (2019)'s approach is based on the mADD model in Aparicio et al. (2007), whose feasible region is a subset of the efficient frontier D_o^5 , it can only decrease inputs and increase outputs. On the other hand, our approach is able to increase inputs and decrease outputs for adjustment costs minimisation. So we can search the target over the whole efficient frontier and obtain an efficient target with the lowest adjustment costs. This is exemplified by the results of the numerical experiments provided in the following section.

3 Numerical experiment

3.1 Experiment description

In this section, the numerical experiments with 48 simulated datasets are conducted to compare our approach with Ang et al. (2019)'s approach. The 48 simulated datasets are generated as Pendharkar (2015), which is based on a Cobb-Douglas production function

(Nataraja and Johnson, 2011). Each dataset has 500 DMUs. We assume that the adjustment costs of inputs and outputs are $c_i^I = 2i + 1, i \in \{1, \dots, m\}$ and $c_k^O = 2(k + m) + 1, k \in \{1, \dots, s\}$, respectively. All the numerical experiments are implemented with MATLAB.

3.2 Results

Table 3 illustrates the results of the numerical experiments with Ang et al. (2019)'s approach and our approach. The first column shows the number of datasets. The second column shows the correlation of the input variables: uncorrelated and correlated. The third and fifth column shows the number of inputs and outputs, respectively. The fourth column shows the two cases of technical inefficiency distribution: exponential and half-normal distribution. The sixth column shows the returns to scale relationship (RTSR) between the inputs and outputs: constant returns to scale (term CRS), increase returns to scale (term IRS) and decrease returns to scale (term DRS). The seventh and eighth column shows the adjustment costs provided by Ang et al. (2019)'s approach and ours. The last column is the reduction percentage of adjustment costs by using our approach.

In order to see the difference of the adjustment costs provided by our approach and previous approach, Figure 1, which is a 3D stacked column chart, is used to describe the results. In addition to that, Figure 2 describes the reduction (%) of the adjustment costs to make the inefficient DMUs efficient by our approach. Although Figure 1 shows that the difference is not very significant at some datasets, the adjustment costs can be significantly reduced at some datasets by using our approach according to Figure 2. Specifically, the results of the numerical experiments enable several observations to be made.

- There are four datasets: the 37th, 38th, 39th and 40th, for which the total adjustment costs provided by the two approaches are the same. This means that the efficient targets with the lowest adjustment costs lie within the subset of the efficient frontier (D_o).
- For all the other datasets, the approach proposed in this paper can provide an efficient target with lower total adjustment costs. The difference of the provided adjustment costs between Ang et al. (2019)'s approach and our approach is the largest as the 20th dataset (856,732.40 vs. 83,760.03). So, the reduction could be as high as 90.22% by using our approach at the 20th dataset. Hence, our new approach is more effective in helping inefficient DMUs to minimise the costs of adjustment.
- Though it is not shown in Table 3, our approach can provide an efficient target which has an adjustment cost that is either the same or less than that provided by Ang et al. (2019)'s approach for any single inefficient DMU.

Therefore, the results of the numerical experiments and the above methodological explanations show that our proposed approach for lowest cost problem outperforms the previous ones.

Figure 1 The adjustment costs by our approach and previous approach

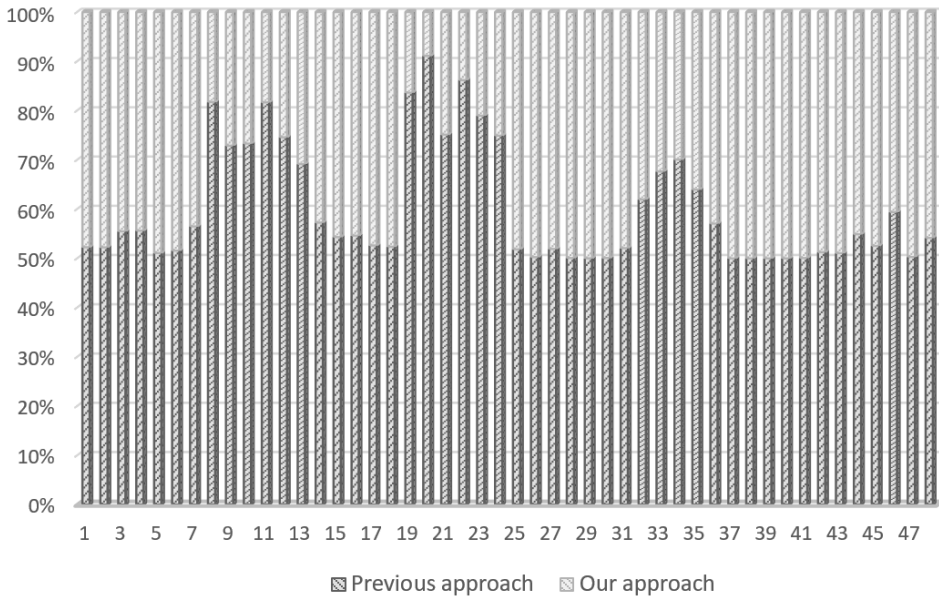
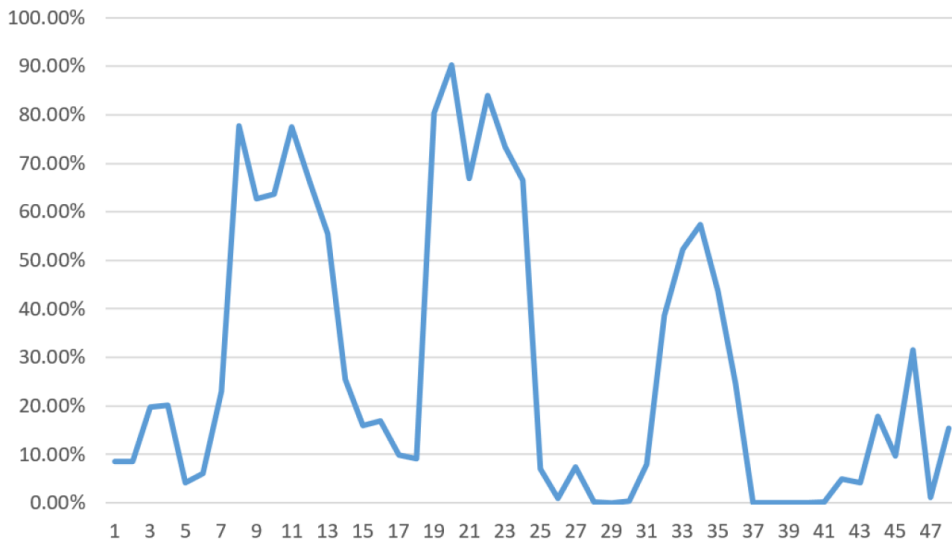


Figure 2 The reduction of the adjustment costs (see online version for colours)



3.3 Remark

Since Ang et al. (2019)'s approach can provide an efficient target with lower adjustment costs compared with that provided by the ADD (additive) model proposed in Charnes et al. (1985) and the approach in Pendharkar (2015), we can conclude that our proposed approach is also better than those approaches on the lowest cost problem.

Table 3 Results of the numerical experiments

Number	Dataset				Adjustment costs			
	Correlation	#Input	Distribution	#Output	RTSR	Ang et al. (2019)'s	Ours	Reduction (%)
1	Uncorrelated	3	Exponential	1	CRS	34,708.75	31,716.05	8.62
2	Uncorrelated	3	Exponential	1	IRS	23,009.84	21,040.60	8.56
3	Uncorrelated	3	Exponential	1	DRS	24,896.09	19,975.00	19.77
4	Uncorrelated	3	Half-normal	1	CRS	36,909.19	29,471.10	20.15
5	Uncorrelated	3	Half-normal	1	IRS	21,668.57	20,761.83	4.18
6	Uncorrelated	3	Half-normal	1	DRS	25,010.68	23,487.72	6.09
7	Uncorrelated	5	Exponential	1	CRS	44,519.78	34,313.31	22.93
8	Uncorrelated	5	Exponential	1	IRS	122,440.01	27,354.95	77.66
9	Uncorrelated	5	Exponential	1	DRS	108,547.09	40,418.42	62.76
10	Uncorrelated	5	Half-normal	1	CRS	79,492.01	28,956.84	63.57
11	Uncorrelated	5	Half-normal	1	IRS	122,093.74	27,418.26	77.54
12	Uncorrelated	5	Half-normal	1	DRS	102,710.03	35,007.70	65.92
13	Uncorrelated	3	Exponential	2	CRS	69,545.04	31,032.02	55.38
14	Uncorrelated	3	Exponential	2	IRS	67,460.25	50,358.34	25.35
15	Uncorrelated	3	Exponential	2	DRS	40,383.81	33,964.86	15.89

Table 3 Results of the numerical experiments (continued)

Number	Dataset			Adjustment costs				
	Correlation	#Input	Distribution	#Output	RTSR	Ang et al. (2019)'s	Ours	Reduction (%)
16	Uncorrelated	3	Half-normal	2	CRS	65,494.38	54,447.20	16.87
17	Uncorrelated	3	Half-normal	2	IRS	30,411.20	27,391.14	9.93
18	Uncorrelated	3	Half-normal	2	DRS	44,234.93	40,245.93	9.02
19	Uncorrelated	5	Exponential	2	CRS	219,385.09	42,897.28	80.45
20	Uncorrelated	5	Exponential	2	IRS	856,732.40	83,760.03	90.22
21	Uncorrelated	5	Exponential	2	DRS	335,638.27	111,177.14	66.88
22	Uncorrelated	5	Half-normal	2	CRS	356,569.14	57,162.97	83.97
23	Uncorrelated	5	Half-normal	2	IRS	239,905.70	63,878.79	73.37
24	Uncorrelated	5	Half-normal	2	DRS	166,240.53	55,684.97	66.50
25	Correlated	3	Exponential	1	CRS	317,623.16	295,134.08	7.08
26	Correlated	3	Exponential	1	IRS	898,442.64	889,628.29	0.98
27	Correlated	3	Exponential	1	DRS	518,602.99	480,508.77	7.35
28	Correlated	3	Half-normal	1	CRS	149,990.17	149,644.84	0.23

Table 4 Results of the numerical experiments

Number	Dataset			Adjustment costs				
	Correlation	#Input	Distribution	#Output	RTSR	Ang et al. (2019)'s	Ours	Reduction (%)
29	Correlated	3	Half-normal	1	IRS	392,462.91	392,445.29	0.01
30	Correlated	3	Half-normal	1	DRS	409,266.29	408,126.09	0.28
31	Correlated	5	Exponential	1	CRS	166,362.94	153,157.59	7.94
32	Correlated	5	Exponential	1	IRS	116,308.00	71,379.67	38.63
33	Correlated	5	Exponential	1	DRS	211,968.49	101,430.72	52.15
34	Correlated	5	Half-normal	1	CRS	345,202.41	147,485.00	57.28
35	Correlated	5	Half-normal	1	IRS	498,698.80	280,919.22	43.67
36	Correlated	5	Half-normal	1	DRS	279,072.34	209,974.74	24.76
37	Correlated	3	Exponential	2	CRS	110,569.45	110,569.45	0
38	Correlated	3	Exponential	2	IRS	759,480.82	759,480.82	0
39	Correlated	3	Exponential	2	DRS	283,059.85	283,059.85	0
40	Correlated	3	Half-normal	2	CRS	229,517.19	229,517.19	0
41	Correlated	3	Half-normal	2	IRS	225,340.40	224,763.60	0.26
42	Correlated	3	Half-normal	2	DRS	127,401.24	121,196.29	4.87
43	Correlated	5	Exponential	2	CRS	1,000,399.22	959,376.63	4.10
44	Correlated	5	Exponential	2	IRS	3,596,783.59	2,957,120.39	17.78
45	Correlated	5	Exponential	2	DRS	1,316,280.82	1,189,205.41	9.65
46	Correlated	5	Half-normal	2	CRS	864,198.75	590,977.64	31.62
47	Correlated	5	Half-normal	2	IRS	11,112,074.75	10,980,467.31	1.18
48	Correlated	5	Half-normal	2	DRS	238,182.17	201,654.12	15.34

4 Conclusions

This paper focuses on the lowest cost problem over the whole efficient frontier assuming that the marginal costs for adjusting inputs and outputs are known and symmetrical. In DEA, an inefficient DMU's efficiency can be improved by adjusting the inputs or outputs or both to reach the projection target on the efficient frontier. In this paper, to improve the approach in Ang et al. (2019), a new approach to provide an efficient target with the lowest adjustment costs is developed following Wang et al. (2020). In contrast to existing approaches, our approach considers that some inputs could be increased and some outputs could be reduced. The results of the numerical experiments show that the approach proposed in this paper can provide an efficient target with no higher adjustment costs than those of previous approaches. Thus, our proposed approach is more useful for decision makers.

Nonetheless, several limitations of this approach can be identified, and future research should aim to overcome these limitations. First, in the real world, the cost function may not comply with the assumptions in this paper and previous studies. There are two points to consider in this regard:

- a The marginal costs of adjusting inputs and outputs are always difficult to quantify in the real world. Though the marginal costs are available, they might be expressed as functions not constants.
- b The costs associated with increasing inputs and outputs might be different from those associated with decreasing them.

For example, retrenchment costs caused by reducing the labour input are usually different from training costs required to introduce new staff and increase the labour input. It then follows that the costs of adjusting inputs and outputs need to be considered on a case-by-case basis, and methods for achieving this should be researched more deeply in the future. Furthermore, the approach in this study does not set a budget constraint or an appropriate adjustment period. In the real world, firms are always subject to budget constraints and have a defined period over which adjustments need to be made. These should be considered in the models for minimising the total adjustment costs and determining the optimal paths to achieve efficiency, which is another interesting avenue for future research. To this end, more empirical studies applying the approach proposed in this paper should be conducted in the future.

Acknowledgements

The authors would like to thank the anonymous associate editor and reviewers for their valuable comments and suggestions. This research was supported in part by Grant-in-Aid for Scientific Research (C) 21K04536 of Japan Society for the Promotion of Science and Grant for Special Research Project (Project Numbers: 2021C-166, 2021E-018) of Waseda University.

References

- Ando, K., Kai, A., Maeda, Y. and Sekitani, K. (2012) 'Least distance based inefficiency measures on the Pareto-efficient frontier in DEA', *Journal of the Operations Research Society of Japan*, Vol. 55, No. 1, pp.73–91.
- Ando, K., Minamide, M., Sekitani, K. and Shi, J. (2017) 'Monotonicity of minimum distance inefficiency measures for data envelopment analysis', *European Journal of Operational Research*, Vol. 260, No. 1, pp.232–243.
- Ang, S., An, Q., Yang, F. and Ji, X. (2019) 'Target setting with minimum improving costs in data envelopment analysis: a mixed integer linear programming approach', *Expert Systems*, Vol. 36, No. 4, p.e12408.
- Aparicio, J., Ruiz, J.L. and Sirvent, I. (2007) 'Closest targets and minimum distance to the Pareto-efficient frontier in DEA', *Journal of Productivity Analysis*, Vol. 28, No. 3, pp.209–218.
- Banker, R.D., Charnes, A. and Cooper, W.W. (1984) 'Some models for estimating technical and scale inefficiencies in data envelopment analysis', *Management Science*, Vol. 30, No. 9, pp.1078–1092.
- Briec, W. (1999) 'Hölder distance function and measurement of technical efficiency', *Journal of Productivity Analysis*, Vol. 11, No. 2, pp.111–131.
- Charnes, A., Cooper, W. and Rhodes, E. (1978) 'Measuring the efficiency of decision making units', *European Journal of Operational Research*, Vol. 2, No. 6, pp.429–444.
- Charnes, A., Cooper, W.W., Golany, B., Seiford, L. and Stutz, J. (1985) 'Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions', *Journal of Econometrics*, Vol. 30, Nos. 1–2, pp.91–107.
- Cook, W.D. and Zhu, J. (2013) *Data Envelopment Analysis: Balanced Benchmarking*, CreateSpace Independent Publishing Platform, Charleston.
- De Mateo, F., Coelli, T. and O'Donnell, C. (2006) 'Optimal paths and costs of adjustment in dynamic DEA models: with application to Chilean department stores', *Annals of Operations Research*, Vol. 145, No. 1, pp.211–227.
- Fukuyama, H., Maeda, Y., Sekitani, K. and Shi, J. (2014) 'Input–output substitutability and strongly monotonic p-norm least distance DEA measures', *European Journal of Operational Research*, Vol. 237, No. 3, pp.997–1007.
- Guo, D. and Wu, J. (2013) 'A complete ranking of DMUS with undesirable outputs using restrictions in DEA models', *Mathematical and Computer Modelling*, Vol. 58, Nos. 5–6, pp.1102–1109.
- Li, X., Li, F., Zhao, N. and Zhu, Q. (2020) 'Measuring environmental sustainability performance of freight transportation seaports in China: a data envelopment analysis approach based on the closest targets', *Expert Systems*, Vol. 37, No. 4, p.e12334.
- Nataraja, N.R. and Johnson, A.L. (2011), 'Guidelines for using variable selection techniques in data envelopment analysis', *European Journal of Operational Research*, Vol. 215, No. 3, pp.662–669.
- Pendharkar, P.C. (2015) 'Cost minimizing target setting heuristics for making inefficient decision-making units efficient', *International Journal of Production Economics*, Vol. 162, pp.1–12.
- Ray, S.C. (2004) *Data Envelopment Analysis: Theory and Techniques for Economics and Operations Research*, Cambridge University Press, Cambridge.
- Seiford, L.M. and Zhu, J. (2002) 'Modeling undesirable factors in efficiency evaluation', *European Journal of Operational Research*, Vol. 142, No. 1, pp.16–20.
- Tone, K. (2001) 'A slacks-based measure of efficiency in data envelopment analysis', *European Journal of Operational Research*, Vol. 130, No. 3, pp.498–509.
- Treadway, A.B. (1970) 'Adjustment costs and variable inputs in the theory of the competitive firm', *Journal of Economic Theory*, Vol. 2, No. 4, pp.329–347.

Wang, X., Lu, K., Shi, J. and Hasuike, T. (2020) ‘A new MIP approach on the least distance problem in DEA’, *Asia-Pacific Journal of Operational Research*, Vol. 37, No. 06, p.2050027.
 Zhu, Q., Wu, J., Ji, X. and Li, F. (2018) ‘A simple MILP to determine closest targets in non-oriented DEA model satisfying strong monotonicity’, *Omega*, Vol. 79, pp.1–8.

Notes

- 1 The required adjustments are defined by ℓ_1 -norm distance and adjustments costs are defined by weighted ℓ_1 -norm distance.
- 2 This problem is also called the cost minimising target setting problem in Pendharkar (2015) and Ang et al. (2019).
- 3 The costs of adjustments are symmetrical: the marginal cost of increasing one unit input is the same as that of decreasing one unit input, and the cost of decreasing one unit output is the same as that of increasing one unit output (Treadway, 1970; De Mateo et al., 2006).
- 4 A simple data example to show the use of the proposed approach can be found in Appendix.
- 5 The remark of the difference between the E and D_o can be found in Wang et al. (2020).

Appendix

Here, the calculation process of our proposed approach based on the data in Table 1 is presented.

- Step 1. Determine the efficient DMUs by using classical DEA models.

Classical DEA models such as the additive model can be used to determine the efficient DMUs easily. When DMU_o is under evaluation, the simplest additive model, Charnes et al. (1985) can be expressed as follows.

$$\begin{array}{l}
 \text{maximum} \quad s_o^- + s_o^+ \\
 s_o^-, s_o^+, \lambda \\
 \text{subject to} \quad X\lambda + s_o^- = x_o, \\
 \quad \quad \quad Y\lambda - s_o^+ = y_o, \\
 \quad \quad \quad 1^T \lambda = 1, \\
 \quad \quad \quad s_o^- \geq 0, s_o^+ \geq 0, \lambda \geq 0,
 \end{array} \tag{4}$$

Suppose that $(s_o^{-*}, s_o^{+*}, \lambda^*)$ is the optimal solution to (4). DMU_o is efficient

if and only if $s_o^{-*} = s_o^{+*} = 0$. In this example, $X = \begin{pmatrix} 6 & 7 & 2 & 8 \\ 8 & 9 & 12 & 5 \end{pmatrix}$ and

$Y = (13 \quad 4 \quad 11 \quad 12)$. After calculating (4) four times, DMU_A, DMU_C, and DMU_D are determined to be efficient, and DMU_B is determined to be inefficient.

- Step 2. Determine the efficient target with the lowest adjustment costs for inefficient DMU_B using (2).

According to the results of Step 1, $X_E = \begin{pmatrix} 6 & 2 & 8 \\ 8 & 12 & 5 \end{pmatrix}$ and $Y_E = (13 \ 11 \ 12)$. By assumption, $c_I = (7 \ 5)$ and $c_O = 3$ in this example. Because (2) is an MIP problem, it can be solved easily by using MATLAB. After calculating (2), we obtain the target (optimal solution) F with the lowest adjustment costs for DMU_B . The total adjustment cost (objective function value) is 37.67.