Mathematics

AIMS Mathematics, 5(2): 811–827.
DOI: 10.3934/math.2020055
Received: 03 September 2019
Accepted: 19 December 2019
Published: 27 December 2019

http://www.aimspress.com/journal/Math

Research article

Closest reference point on the strong efficient frontier in data envelopment analysis

Akbar Moradi[1], Alireza Amirteimoori[1,*], Sohrab Kordrostami[2] and Mohsen Vaez-Ghasemi[1]

[1] Department of Applied Mathematics, Rasht Branch, Islamic Azad University, Rasht, Iran
[2] Department of Applied Mathematics, Lahijan Branch, Islamic Azad University, Lahijan, Iran

* **Correspondence:** Email: ateimoori@iaurasht.ac.ir; aamirteimoori@gmail.com.

Abstract: Data envelopment analysis (DEA) is a data-oriented procedure to evaluate the relative performances of a set of homogenous decision making units (DMUs) with multiple incommensurate inputs and outputs. Performance measurement using tools such as DEA needs to construct an empirical production technology set. In this analysis, DMUs are partitioned into two groups: efficient and inefficient. Inefficient DMUs are projected onto efficient frontier in such a way that their inputs are reduced and their outputs are increased. In this sense, finding a projection point with the shortest distance is important and it is a most frequently studied subject in the field of DEA. In this paper, a two-steps procedure is proposed to determine a projection point on the efficient frontier with closest distance. The reference point is constructed in such a way that it is located on the strong defining hyperplane of the DEA technology set. As we will show, the low computational efforts and the guarantee of determining an efficient projection point on the strong efficient frontier are the two important advantages of the proposed model.To show the applicability of the proposed approach, a real case on 28 international airlines is given.

Keywords: data envelopment analysis; efficiency; input-output; closest distance
Mathematics Subject Classification: 90B30

1. Introduction

Data envelopment analysis (DEA) is a powerful knowledge-based analytical method to evaluate the relative performance of a set of homogeneous decision making units (DMUs) that consumes multiple incommensurate inputs and outputs. DEA initiated in 1978 by Charnes, Cooper and Rhodes and extended by Banker, Charnes and Cooper [1]. In the last three decades, DEA has been applied in a wide range of applications in organizational units. See, for instances, Emrouznejad et al. [2,3].

The main result of DEA as a performance evaluation tool is to partition the DMUs into two groups: efficient and inefficient. Inefficient DMUs have to reduce their inputs and simultaneously increase their outputs to meet the efficient frontier of the production technology set. This means that inefficient DMUs have to lose part of their resources and simultaneously they should try to increase the level of their outputs production. In a rational sight, DMUs are interested to lose less resources and to increase less increment in outputs to meet the frontier. In this sense, determining an efficient projection point with minimal changes in inputs and outputs is an important and interesting subject in the field of DEA and it has attracted considerable attention among researchers in the last decade.

In what follows, we briefly review some of these studies on efficiency measure with closest reference point:

The first study on finding the distance to a reverse convex subset in a normed vector space is studied by Briec and Lemaire [4]. At the same time, Frei and Harker [5] have extended DEA methodology in two substantive ways. First, they developed a method to determine the least-norm projection from an inefficient DMU to the efficient frontier in both the input and output space simultaneously, and second, they introduced the notion of the “observable” frontier and its subsequent projection.

Another work that considered the shortest distance, have proposed by González and Álvarez [6]. They have studied the problem of efficiency improvement and how to identify appropriate benchmarks for inefficient firms to imitate. They argued that the most relevant benchmark is the closest reference firm on the efficient subset of the isoquant.

Lozano and Villa [7] had a different look at the shortest distance. They have advocated determining a sequence of targets, each one within an appropriate, short distance of the preceding. Their approach has two interesting features: (a) the sequence of targets ends in the efficient frontier and (b) the final efficient target is generally closer to the original unit than the one-step projection.

Amirteimoori and Kordrostami [8] proposed a Euclidean distance-based efficiency measure to evaluate the relative efficiency of a set of homogeneous DMU. An alternative Euclidean distance-based efficiency measure is defined in their work and it has been shown that the reference point on the efficient frontier has shortest distance to the original point. They applied their approach to a real case on gas companies. Aparicio et al. [9] have used the full dimensional efficient facets to propose an alternative Russell output measure of technical efficiency.

Aparicio and Pastor [10] have used two simple example to show a drawback of the approach proposed by Amirteimoori and Kordrostami [8]. They showed that in some cases, the reference point obtained from the work of Amirteimoori and Kordrostami [8] is not in the technology set. In order to overcome this drawback, they slightly modified the model introduced by Aparicio et al. [11].

In another study, Aparicio and Pastor [12] have shown that the least distance measures based on Hölder norms satisfy neither weak nor strong monotonicity on the strongly efficient frontier. Then, they provided a solution for output-oriented models that allows assuring strong monotonicity on the strongly efficient frontier.

An et al. [13] proposed a non-oriented DEA approach based on enhanced Russell [14] measure for measuring the environmental efficiency of DMUs and meanwhile, they provided the closest target for the DMU under evaluation to be efficient with less effort.

Aparicio et al. [15] have shown that the existing approaches for determining the least distance without identifying explicitly the frontier structure for graph measures do not work for oriented models. Then, they proposed a methodology for satisfactorily implementing these situations. Razipour et al. [16] have used the problem of closest reference target to find the closest targets in bank branches in Iran.

All the above studies show that the determination of closest efficient targets in production possibility set is an important subject in the field of DEA and it has attracted considerable interest among researchers in recent DEA literature. Despite of this, only a few studies exist that analyze the implications of using closest targets on the technical inefficiency measurement.

In this paper, a two-steps procedure is proposed to determine a projection point on the efficient frontier with closest distance. The reference point is constructed in such a way that it is located on the strong defining hyperplane of the DEA technology set. To do this, we first construct a strong defining hyperplane of the production set corresponding to each inefficient DMU, and then, the DMU is projected to this hyperplane in the direction of gradient vector.

The reminder of this paper is organized as follows: to start the study, the required preliminaries are given in next section. Our proposed approach appears in section 3. To illustrate the applicability of the approach, a real case on 28 international airlines in Asia-Australia is given in section 4. The paper ends with conclusions.

2. Preliminaries

Suppose there are n $DMU_j, j=1,...,n$ and each DMU_j uses m inputs to produce s outputs. Specially, DMU_j uses inputs $x_j=(x_{1j},...,x_{mj})\geq 0$ to produce the outputs $y_j=(y_{1j},...,y_{sj})\geq 0$. The technology set T is defined as the set of all feasible input–output combinations as

$$T=\{(x,y)\in R^m_{\geq 0}\times R^s_{\geq 0} \mid x\, can\ produce\ y\} \tag{1}$$

By accepting axiom such as constant return to scale (CRS), convexity, inclusion, free disposability of inputs and outputs and minimal extrapolation, T_c is constructed as follows: (Charnes et al. [17]).

$$T_c=\left\{(x,y)\mid \sum_{j=1}^{n}\lambda_j x_j\leq x,\ \sum_{j=1}^{n}\lambda_j y_j\geq y,\ \lambda_j\geq 0,\ j=1,\ldots,n\right\} \tag{2}$$

In a same way, if we ignore the CRS assumption, in variable returns to scale (VRS) framework, T_v is constructed as follows: (Banker et al. [1]).

$$T_v=\left\{(x,y)\mid \sum_{j=1}^{n}\lambda_j x_j\leq x,\ \sum_{j=1}^{n}\lambda_j y_j\geq y,\ \sum_{j=1}^{n}\lambda_j=1,\ \lambda_j\geq 0,\ j=1,\ldots,n\right\} \tag{3}$$

The input-oriented envelopment CCR model for evaluation efficiency of DMU_o as follows: (Charnes et al. [17]).

$$
\begin{aligned}
& Min\ \ \theta \\
& s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = \theta x_{io}, \quad i = 1,...,m, \\
& \qquad\ \ \sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro}, \quad r = 1,...,s, \\
& \qquad\ \ \lambda_j, s_i^-, s_r^+ \geq 0, \qquad for\ all\ i,\ j\ and\ r.
\end{aligned} \tag{4}
$$

The first m constraints in model (4) guarantee that the inputs of the new target unit do not exceed the inputs of DMU_o and the second s constraints guarantee that the outputs of the new target unit is not less than the outputs of DMU_o. In the above model, DMU_o is said to be efficient if and only if in all optimal solutions, we have $\theta^* = 1$ and all slack variables are equal to zero. In model (4), if we remove the slack variables and rewrite the constraints in inequality form, the dual formulation of model (4) (known as multiplier form) is expressed as follows:

$$
\begin{aligned}
& Max \sum_{r=1}^{s} u_r y_{ro} \\
& s.t. \quad \sum_{i=1}^{m} v_i x_{io} = 1, \\
& \qquad\ \ \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \quad j = 1,...,n \\
& \qquad\ \ u_r, v_i \geq 0, \qquad for\ all\ r\ and\ \ i,
\end{aligned} \tag{5}
$$

In model (5), DMU_o is said to be strong efficient if the optimal value of the objective function is equal to 1 and there exists at least one optimal solution $(u_1^*, u_2^*, \ldots, u_s^*, v_1^*, v_2^*, \ldots, v_m^*)$ with $u_r^* > 0$ and $v_i^* > 0$ for all *i* and *r*.

Definition 2.1: *Let* $H = \{(x, y) \mid u^T(y - \bar{y}) - v^T(x - \bar{x}) = 0\} \cap T_c$ *be a supporting hyperplane of* T_c *passing through a specific point* $(\bar{x}, \bar{y})$ *in* T_c. *Then, H is called strong defining hyperplane if and only if* $(u, v) > 0$.

Definition 2.2 (Pareto-Koopmans efficiency): *A DMU is said to be strong efficient if and only if it is not possible to improve any input or output without getting worse some other input or output.*

For an inefficient DMU_o, the reference set consists of all DMUs with $\lambda_j^* > 0$, in which λ_j^* is optimal solution to model (4). It is easy to show that all DMUs in the reference set of DMU_o are located on a unique supporting surface.

Definition 2.3 (Reference supporting surface)**:** *For a* DMU_o, *an efficient surface of* T_c *is called a reference Supporting surface, if it contains the reference units of* DMU_o.

Based on the structure of T_c or T_v, different strategies can be considered to project an inefficient point to the efficient frontier. The CCR model (4) evaluates the radial efficiency and it does not take the output shortfalls and input excesses in to consideration. However, the existence of nonzero slacks leads to incorrect estimation of efficiencies. Additive model deals directly with input excesses and output shortfalls to project an inefficient DMU to strong efficient frontier. The mathematical formulation of Additive model is as follows:

$$\begin{aligned}
&Max \quad z=\sum_{r=1}^{s} s_r^- + \sum_{i=1}^{m} s_i^+ \\
&s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io}, \quad i=1,\dots,m, \\
&\qquad\ \sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro}, \quad r=1,\dots,s, \\
&\qquad\ \lambda_j \geq 0,\ s_i^- \geq 0,\ s_r^+ \geq 0, \quad for\, all\, i, j, r.
\end{aligned} \tag{6}$$

In this model, DMU_o is said to be efficient if and only if the optimal objective value is equal to zero. The dual formulation of the additive model (6) is as follows:

$$\begin{aligned}
&Min \quad e_o=\sum_{i=1}^{m} v_i x_{io} - \sum_{r=1}^{s} u_r y_{ro} \\
&s.t. \qquad \sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \geq 0, \quad j=1,\dots,n, \\
&\qquad\qquad u_r, v_i \geq 1, \quad r=1,\dots,s \ \ and\ \ i=1,\dots,m.
\end{aligned} \tag{7}$$

The second constraint in (7) guarantees the positivity of the weights. So, if DMU_o prevails as inefficient, model (7) projects it to a strong defining hyperplane. Clearly, the projection point of an inefficient DMU_o is not necessarily the closest point on the frontier. It is important and interesting to find a point with closest distance to DMU_o under evaluation.

Tone [18] have augmented the additive model (6) by introducing an efficiency measure that is invariant to the units of the data. The slack-based measure (SBM) of efficiency introduced by Tone [18] as follows:

$$\begin{aligned}
&Min\ \rho_o=\frac{1-\frac{1}{m}\sum_{i=1}^{m}\frac{s_i^-}{x_{io}}}{1+\frac{1}{s}\sum_{i=1}^{m}\frac{s_r^+}{y_{ro}}} \\
&s.t. \quad x_{io}=\sum_{j=1}^{n}\lambda_j x_{ij}+s_i^-, \quad i=1,\dots,m, \\
&\qquad\ y_{ro}=\sum_{j=1}^{n}\lambda_j y_{rj}+s_r^+, \quad r=1,\dots,s, \\
&\qquad\ \lambda_j, s_i^-, s_r^+ \geq 0 \quad for\ all\ i,\ j\ and\ r.
\end{aligned} \tag{8}$$

In model (8), we assume that $x_{io}>0$ for all i. If $x_{io}=0$, we can remove the term $\frac{s_i^-}{x_{io}}$ from the summation. Furthermore, it can easily be seen that $0 \leq \rho_o \leq 1$.

Definition 2.4: *DMU_o is efficient if and only if $\rho_o^*=1$.*

Theorem 2.1: *If* DMU_A *dominates* DMU_B *so that* $x_A \le x_B$ *and* $y_A \ge y_B$ *then we have* $\rho_B^* \le \rho_A^*$.

Proof. See Tone [18].

Aparicio et al. [11] proposed a general two-step procedure to find minimum distance on the Pareto-efficient frontier. In the first step, efficient and inefficient DMUs are obtained by one of the classical radial models. Let *E* be the set of all extreme efficient units. In the second step, the following multiplier model (MADD) is proposed to the members of *E*:

$$
\begin{aligned}
Min\ & \sum_{i=1}^{m} s_{io}^{-} + \sum_{r=1}^{s} s_{ro}^{+} \\
s.t.\ & \sum_{j\in E} \lambda_j x_{ij} = x_{io} - s_{io}^{-},\ i=1,\dots,m \\
& \sum_{j\in E} \lambda_j y_{rj} = y_{ro} - s_{ro}^{+},\ r=1,\dots,s \\
& -\sum_{i=1}^{m} v_j x_{ij} + \sum_{r=1}^{s} \mu_r y_{rj} + d_j = 0,\ i \in E \\
& v_i \ge 1, \qquad i=1,\dots,m \\
& \mu_r \ge 1, \qquad r=1,\dots,s \\
& d_j \le M b_j, \qquad j \in E \\
& \lambda_j \le M(1-b_j), \qquad j \in E \\
& b_j \in \{0,1\}, \qquad j \in E \\
& d_j \ge 0, \qquad j \in E \\
& \lambda_j \ge 0, \qquad j \in E \\
& s_{io}^{-} \ge 0, \qquad i=1,\dots,m \\
& s_{ro}^{+} \ge 0, \qquad r=1,\dots,s
\end{aligned}
\tag{9}
$$

In which s_{ro}^{+} and s_{io}^{-} are slacks variables and M is a large positive number. In this model DMU_o is said to be efficient if and only if the optimal slacks in (MADD) are all zero. Model (9) is a mixed integer linear programming problem and if DMU_o is inefficient, its efficient projection on the frontier is closest point.

3. Closest reference point

As we stated before, in traditional DEA models, the reference point of an inefficient unit is calculated in input or output or jointly orientation. Clearly, the reference points in such orientations are not the closest reference points and we are interested to determine a reference point on the efficient frontier with minimal distance. In this section, an alternative shortest distance method has been developed by using gradient vectors.

Suppose the set of all *DMUs* is partitioned into four sets *E*, *NE*, *F* and *NF* in which *E* is the set of all strong efficient *DMUs*, *NE* is the set of all inefficient *DMUs*, but their reference point belongs to *E*, *F* is the set of weak efficient *DMUs* and *NF* is the set of inefficient *DMUs* but their reference point belongs to *F*. A procedure to partition *DMUs* into four sets *E*, *NE*, *F* and *NF* will be given later.

We are interested to find a projection point on the efficient frontier with minimal distance. We first employ the formulations (4) and (5) to determine efficient and inefficient *DMUs*. Efficient *DMUs* are belong to E and inefficient *DMUs* are belong to E'. The weak efficient *DMUs* are belong to *F*. also the inefficient *DMUs* that projection point are on the week supporting frontier, are belong to *NF*. Therefore $NE = E' - (F \bigcup NF)$. Also the formulation (7) is used to determine the closest reference supporting surface of $DMU_o \in E'$ as follows:

$$F_o = \{(U,V) \,|\, V^{*T}X_o - U^{*T}Y_o = 0\} \bigcap T_c$$

Which (U^*, V^*) is an optimal solution to model (7). Consider the gradient vector $(U^*, -V^*)$ and we now solve the following linear programming problem:

$$\begin{aligned} &Max \quad t \\ &s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij} \le x_{io} - v_i^* t \ , \ \ i = 1, \dots, m \\ &\qquad \sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{ro} + u_r^* t \ , \ \ r = 1, \dots, s \\ &\qquad \lambda_j \ge 0 \ , \qquad j = 1, \dots, n \end{aligned} \tag{10}$$

Suppose $DMU_o : (X_o, Y_o) \in NE$, we move from (X_o, Y_o) in the direction $(U^*, -V^*)$ and t is the step size. Clearly, in the optimality, t^* is the maximum step size and as we should expect, the obtained projected point is now calculated as

$$(X_o^*, Y_o^*) = (X_o - t^* V^*, Y_o + t^* U^*) \tag{11}$$

Now suppose $DMU_o : (X_o, Y_o) \in F \bigcup NF$. In this case, we solve the following linear quadratic formulation:

$$\begin{aligned} &Min \quad \sum_{i=1}^{m} (z_i - x_{io})^2 + \sum_{r=1}^{s} (w_r - y_{ro})^2 \\ &s.t. \quad \sum_{i=1}^{m} v_{io}^* z_i - \sum_{r=1}^{s} u_{ro}^* w_r = 0, \\ &\qquad \sum_{i=1}^{m} v_{ij}^* z_i - \sum_{r=1}^{s} u_{rj}^* w_r \ge 0, \quad j \ne o, \ \ j = 1, \dots, n, \\ &\qquad z_i, \ w_r \ge 0 \ \ for\, all\, i\, and\, r. \end{aligned} \tag{12}$$

In which z_i and w_r are respectively the *i-th* and *r-th* coordination of the *i-th* input and *r-th* output of the projected point. (U_j^*, V_j^*) for DMU_j, $j = 1, \dots, n$ is an optimal solution to model (7) when DMU_j is under evaluation. The first constraint guarantees that the new projection point is located on the efficient frontier and the second n-1 constraints are given to guarantee the feasibility of the new projection point.

Theorem 3.1: *The projection point obtained from the above mentioned procedure is the closest reference point on the strong efficient hyperplane.*

Proof. Case 1: let $DMU_o \in NE$ because the gradient vector is vertical on the reference hyperplane, therefore, move from (X_o, Y_o) in the gradient direction to vertical image on the reference hyperplane is the shortest distance.

Case 2: let $DMU_o \in F \bigcup NF$ because $\sum_{i=1}^{m} v_{io}^* z_i - \sum_{r=1}^{s} u_{ro}^* w_r = 0$ is the strong reference supporting surface of DMU_o, according to the objective function (square distance function) of formulation (12), the verdict is obvious.■

Let (X_o^*, Y_o^*) is the projection point of (X_o, Y_o) using the proposed approach and let (U_o^*, V_o^*) is the multiplier of closest reference supporting surface. Inspired by the efficiency index of Tone [18], we define the inefficiency index ρ_o as follows:

$$\rho_o = 1 - \frac{1 - \frac{1}{m+s}\left[\sum_{i=1}^{m} x_{io}^* / x_{io} + \sum_{i=1}^{m} y_{ro} / y_{ro}^*\right]}{1 + \frac{1}{m+s}\left[\sum_{i=1}^{m} x_{io}^* / x_{io} + \sum_{i=1}^{m} y_{ro} / y_{ro}^*\right]} \tag{13}$$

***Proposition 3.1*:** $DMU_o \in E$ *if and only if* $\rho_o = 1$.

Proof. Let $DMU_o \in E$ then $x_o^* = x_o$ and $y_o^* = y_o$ so it's obvious $\rho_o = 1$. If $\rho_o = 1$**,** then

$$1 - \frac{1}{m+s}\left(\sum_{i=1}^{m} \frac{x_{io}^*}{x_{io}} + \sum_{r=1}^{s} \frac{y_{ro}}{y_{ro}^*}\right) = 0$$

So,

$$\left(\sum_{i=1}^{m} \frac{x_{io}^*}{x_{io}} + \sum_{r=1}^{s} \frac{y_{ro}}{y_{ro}^*}\right) = m + s$$

On the other hand, because $x_o^* \le x_o$ and $y_o^* \ge y_o$ so we always have

$$\left(\sum_{i=1}^{m} \frac{x_{io}^*}{x_{io}} + \sum_{r=1}^{s} \frac{y_{ro}}{y_{ro}^*}\right) \le m + s$$

The above equality holds true if $x_o^* = x_o$ and $y_o^* = y_o$ this means that $DMU_o \in E$. ■

A point to be noted is that in all of the above discussion, the underlying technology set was constant returns to scale technology set. The procedure can easily be extended to variable returns to scale technology set.

At the end of this section, a simple example is used to illustrate the proposed approach. Suppose we have eight DMUs with two inputs and one output. We employ the formulations (4) and (5) to determine efficient and inefficient *DMUs*. The data set, the efficiency scores and the optimal weights obtained from models (4) and (5) are given in Table 1. As columns 5–11 of Table 1 show, four DMUs D1, D2, D3 and D4 are strong efficient and hence, $E = \{D1, D2, D3, D4\}$ and $E' = \{D5, D6, D7, D8\}$. As columns 5–8 of Table 1 shows, for D5 and D6 the efficiency scores are one, but the slack variables are not zero, so, $F = \{D5, D6\}$. Moreover, the projection point of *D8* is located on the weak supporting surface and hence $NF = \{D8\}$. Finally, $NE = \{D7\}$ $(NE = E' - F - NF)$. The Farrell cut of production possibility set is shown in Figure 1.

Table 1. The data set and results for simple example.

DMUs	x1	x2	y	Θ	s_1^-	s_2^-	s_1^+	v1	v2	u1
D1	1	5	1	1	0	0	0	1	0	1
D2	2	3	1	1	0	0	0	0.286	0.143	1
D3	3	2	1	1	0	0	0	0.2	0.2	1
D4	5	1	1	1	0	0	0	0	1	1
D5	7	1	1	1	2	0	0	0	1	1
D6	1	6	1	1	0	1	0	1	0	1
D7	5	3	1	0.64	0	0	0	0.091	0.182	0.636
D8	7	1.2	1	0.83	0.83	0	0	0	0.833	0.833

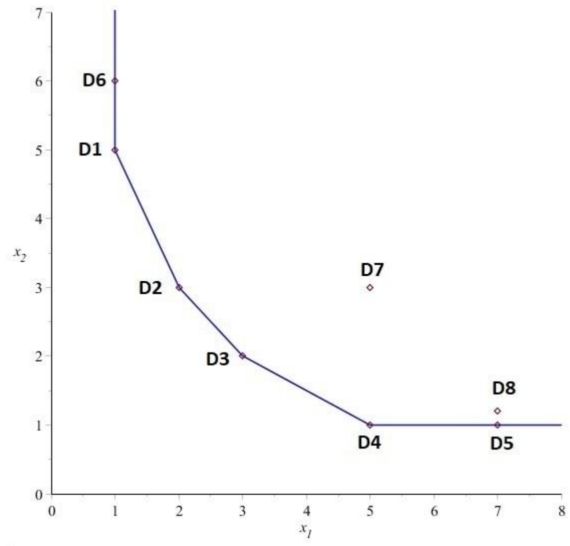

Figure 1. Production Possibility Set (PPS).

Now, we use model (7) to determine the closest reference supporting surface, the results are given in columns 3–5 of Table 2. Suppose $DMU_o \in NE = \{D7\}$. We have solved model (10) with $(U^*, V^*) = (5,1,1)$. The optimal value $t^*$is shows in column six of Table 2.

The projection point of $D7$ is obtained as $(4.895,\ 2.895,\ 1.525)$. Now, consider $DMU_8 \in F \bigcup NF$. We solve model (12) as follows:

$$\begin{aligned}
&\mathrm{Min}(z_1-7)^2+(z_2-1.2)^2+(w-1)^2\\
&s.t.\ \ z_1+2z_2-7w=0\\
&\qquad 2z_1+1z_2-7w\geq 0\\
&\qquad z_1+z_2-5w\geq 0\\
&\qquad z_1-w\geq 0\\
&\qquad z_2-w\geq 0\\
&\qquad z_1,z_2,w\geq 0
\end{aligned}$$

Table 2. The results for simple example.

DMUs	z	V1*	V2*	U1*	t*
D1	0	2	1	7	
D2	0	2	1	7	
D3	0	1	1	5	
D4	0	1	2	7	
D5	2	1	2	7	
D6	1	2	1	7	
D7	3	1	1	5	0.10526
D8	2.4	1	2	7	

The projection point of $D8$ is obtained as $(6.96,\ 1.11,\ 1.31)$.

The projection points of the inefficient *DMU*s are obtained in a similar manner and the results are shown in table 3. In each row, three different values are given, original data, the results of the CCR model and the results of our proposed model. The last column shows the Euclidean distance from original point to the new projection point (To this end, we have used the simple distance formulation $d=\sqrt{\sum_{i=1}^{m}(z_i-x_{io})^2+\sum_{r=1}^{s}(w_r-y_{ro})^2}$). As the results show, in all four DMUs, the distance obtained from our proposed approach is strictly less than the CCR model.

Table 3. The projection points to the inefficient DMUs.

DMUs		x1	x2	y	d
D5	Original	7	1	1	
	CCR	5	1	1	2
	GDM	6.852	1.37	1.37	0.54379
D6	Original	1	6	1	
	CCR	1	5	1	1
	GDM	1.185	5.926	1.185	0.27217
D7	Original	5	3	1	
	CCR	3.2	1.92	1	2.0991
	GDM	4.895	2.895	1.526	0.5470
D8	Original	7	1.2	1	
	CCR	4.98	0.996	1	2.03027
	GDM	6.89	1.378	1.378	0.43173

4. An empirical example

In this section, we apply the proposed procedure to a real data set consisting of 28 international airlines from Asia-Australia, Europe and North America. The data has been taken from Ray [19] and Aparicio et al. [11] have used this data set in their work. As Ray [19] and Aparicio et al. [11], we also used constant returns to scale technology. These 28 international airlines uses four inputs to generate two outputs. The inputs are as follows:

Number of employees (x1).

Millions of gallons fuel (x2).

Other kind of inputs (millions of U.S. dollar equivalent) excluding labor and fuel expenses (x3).

Capital, as the sum of maximum takeoff weights of all aircraft flown multiplied by the number of days flown (x4).

Outputs include:

Passenger-kilometers flown (y1).

Freight tonne-kilometers flown (y2).

The inputs/outputs data are given in Table 4. We first applied the dual formulation of the additive models, in constant returns to scale environment. The results are given in column 3 of Table 5. The input/output weights are also given in columns 4–9. As the results show, nine airlines are relatively efficient and all weights are strictly positive. The classification of *DMU*s are as follows:

$$E = \{\mathrm{JAL, QUANTAS, SAUDIA, SINGAPORE, FINNAIR, LUFTHANSA, SWISSAIR, PORTUGAL,} \\ \mathrm{AM. WEST}\}$$

$$NE = \{NIPPON, CATHAY, GARUDA, MALAYSIA, BRITISH, AUSTRIA, IBERIA, SAS, \\ AMERICAN, CANADIAN, DELTA, EASTERN, PANAM, TWA, USAIR\}$$

$$F \cup NF = \{AIR\ CANADA, CONTINENTAL,\ NORTHWEST, UNITED\}.$$

Table 4. The data related to 28 international airlines.

NUM	Name	Input				Output	
		x1	x2	x3	x4	y1	y2
1	NIPPON	12222	860	2008	6074	35261	614
2	CATHAY	12214	456	1492	4174	23388	1580
3	GARUDA	10428	304	3171	3305	14074	539
4	JAL	21430	1351	2536	17932	57290	3781
5	MALAYSIA	15156	279	1246	2258	12891	599
6	QUANTAS	17997	393	1474	4784	28991	1330
7	SAUDIA	24708	235	806	6819	18969	760
8	SINGAPORE	10864	523	1512	4479	32404	1902
9	AUSTRIA	4067	62	241	587	2943	65
10	BRITISH	51802	1294	4276	12161	67364	2618
11	FINNAIR	8630	185	303	1482	9925	157

Continued on next page

NUM	Name	Input				Output	
		x1	x2	x3	x4	y1	y2
12	IBERIA	30140	499	1238	3771	23312	845
13	LUFTHANSA	45514	1078	3314	9004	50989	5346
14	SAS	22180	377	1234	3119	20799	619
15	SWISSAIR	19985	392	964	2929	20092	1375
16	PORTUGAL	10520	121	831	1117	8961	234
17	AIR CANADA	22766	626	1197	4829	27676	998
18	AM. WEST	11914	309	611	2124	18378	169
19	AMERICAN	80627	2381	5149	18624	133796	1838
20	CANADIAN	16613	513	1051	3358	24372	625
21	CONTINENTAL	35661	1285	2835	9960	69050	1090
22	DELTA	61675	1997	3972	14063	96540	1300
23	EASTERN	21350	580	1498	4459	29050	245
24	NORTHWEST	42989	1762	3678	13698	85744	2513
25	PANAM	28638	991	2193	7131	54054	1382
26	TWA	35783	1118	2389	8704	62345	1119
27	UNITED	73902	2246	5678	18204	131905	2326
28	USAIR	53557	1252	3030	8952	59001	392

We then applied model (10) for DMUs in NE. The results are listed in tenth column of Table 5. Consider, for example, NIPPON AIR in NE. The optimal value of t^* is 138.3009. So, the projection point on the efficient frontier to NIPPON AIR is calculated as $(X^*,Y^*)=(11868.2, 721.7, 1869.7, 5935.7, 35399.3, 752.3)$. Now, consider CONTINENTAL AIR in $F \cup NF$, running the proposed approach to this unit and the projection point to this unit is calculated as (35651.37, 1257.374, 2590.639, 9950.374, 69065.55, 1099.626).

Table 5. The result of model (10) along with the input/output weights.

DMU	Name	Z	v1	v2	v3	v4	u1	u2	t
1	NIPPON	4333	2.558	1	1	1	1	1	138.3009
2	CATHAY	8127.4	1	33.367	1	1	1	1	1.823832
3	GARUDA	12434.6	1	33.367	1	1	1	1	3.478064
4	JAL	0	1	1	34.7	1	1	18.89	
5	MALAYSIA	13982.92	1	1	1	4.779	1	1	81.44536
6	QUANTAS	0	1	127.438	1	1	2.518	1	
7	SAUDIA	0	1	748.473	1	1	10.937	1	
8	SINGAPORE	0	1	33.367	1	1	1	1	
9	AUSTRIA	3955.76	1	33.367	1	1	1	1	0.6980644
10	BRITISH	41434.04	1	33.367	1	1	1	1	11.67467
11	FINNAIR	0	1	1	92.075	1	3.833	1	

Continued on next page

DMU	Name	Z	v1	v2	v3	v4	u1	u2	t
12	IBERIA	25351.36	1	1	12.196	1	1	1	29.74644
13	LUFTHANSA	0	1	1	1	14.989	1	25.043	
14	SAS	17280	1	1	1	4.779	1	1	64.18586
15	SWISSAIR	0	1	1	211.877	6.903	5.199	102.95	
16	PORTUGAL	0	1	232.907	1	22.338	7.17	1	
17	AIR CANADA	12900.43	1	1	25.385	1	1.615	1	
18	AM. WEST	0	1	1	1	15.442	2.474	1	
19	AMERICAN	14364.34	1	1	25.385	1	1.615	1	17.63091
20	CANADIAN	7167.57	1	1	25.385	1	1.615	1	7.136312
21	CONTINENTAL	6237.59	1	1	25.385	1	1.615	1	
22	DELTA	21311.39	1	1	25.385	1	1.615	1	27.33843
23	EASTERN	15363.26	1	748.473	1	1	10.937	1	0.2282259
24	NORTHWEST	10789.3	1	1	25.385	1	1.615	1	
25	PANAM	3727.28	1	1	25.385	1	1.615	1	4.824861
26	TWA	4417.27	1	1	25.385	1	1.615	1	4.796741
27	UNITED	23079.82	1	1	25.385	1	1.615	1	
28	USAIR	41231.32	1	1	1	4.779	1	1	318.3093

The projection points are calculated by three different approaches: CCR model, Aparicio et al. [11] and our proposed approach. The results are listed in Table 6. Now, let us compare the results of the proposed method with other approaches such as CCR model and Aparicio et al. [11]. In Table 6, in the first row of each airline the original data is given, the second row shows the projection points obtained from CCR model, the third row shows the results of Aparicio et al. [11], denoted by mERG, and the fourth row shows the results of our proposed method, denoted by GDM (Gradient Direction Method). As the results show, both approaches, our proposed and Aparicio et al. [11], provided closer projections than the CCR model.

To compare the results of these three different approaches, the distance of each projection point to original unit has been calculated and the results are given in the last column of Table 6. Consider the first airline, NIPPON. At the first and third inputs, the reduction level of our approach is better than one that proposed by Aparicio et al. [11] and for inputs two and four, the reduction level of Aparicio et al. [11] is better than ours. However, in whole sense, the distance from the original point to our projection point is 469.903, while this distance is 1552.975, in Aparicio et al. [11]. Comparing the results of the two approaches for other airlines, we have found that, except for AIR CANAD (DMU17) and AIR CONTINENTAL (DMU21), in all other airlines, the distances between projection points and observed airlines in our approach is less than the approach proposed by Aparicio et al. [11]. However, we checked the Air-Canada and AIR CONTINENTAL and it has been found that the projection points provided by Aparicio et al. [11] to these two DMUs are not efficient. It should be pointed out that we do not claim that our approach is better than the previous approach of Aparicio et al. [11], but, we provide another projection point with minimal distance. Moreover, we just use the results of Aparicio et al. [11] to confirm our results.

Table 6. The results obtained from different methods.

Name		Input				Output		
		x1	x2	x3	x4	y1	y2	d
NIPPON		12222	860	2008	6074	35261	614	
	CCR	11821.86	569.11	1645.31	4873.91	35261	2069.7	1983.822
	mERG	12222	620.9	1623.1	6074	35261	2099.4	1552.975
	GDM	11868.2	721.7	1869.7	5935.7	35399.3	752.3	469.903
CATHAY		12214	456	1492	4174	23388	1580	
	CCR	10687.44	399.01	1180.86	3434.26	23388	1580	1725.589
	mERG	12214	439.3	1139.6	4174	23388	1580	352.795
	GDM	12212.2	395.2	1490.2	4172.2	23389.8	1581.8	60.933
GARUDA		10428	304	3171	3305	14074	539	
	CCR	7063.86	205.93	691.06	2165.45	14074	720.78	4336.909
	mERG	10428	301.7	541.7	3305	14902.4	539	2756.714
	GDM	10424.5	187.9	3167.5	3301.5	14077.5	542.5	116.363
MALAYSIA		15156	279	1246	2258	12891	599	
	CCR	5318.84	210.21	558.78	1709.7	12891	599	9876.606
	mERG	14979.4	231.1	1246	2258	16559	599	3672.561
	GDM	15074.55	197.55	1164.55	1868.77	12972.45	680.45	429.733
AUSTRIA		4067	62	241	587	2943	65	
	CCR	2299.56	42.93	166.86	406.41	2943	84.45	1778.397
	mERG	3475.1	62	232.2	479	4014.4	65	1228.814
	GDM	4066.3	38.7	240.3	586.3	2943.7	65.7	23.353
BRITISH		51802	1294	4276	12161	67364	2618	
	CCR	40533.04	1014.53	3352.49	9534.51	67364	2618	11611.153
	mERG	51802	1294	2559.7	11407.2	67364	2618	1874.540
	GDM	51790.33	904.45	4264.32	12149.32	67375.67	2629.67	390.424
IBERIA		30140	499	1238	3771	23312	845	
	CCR	11124.94	383.38	945.94	2982.8	23312	845	19033.981
	mERG	23184.8	481.8	923.8	3771	25453.8	845	7284.307
	GDM	30110.3	469.3	875.2	3741.3	23341.7	874.7	368.828
SAS		22180	377	1234	3119	20799	619	
	CCR	13975.27	324.18	1061.11	2682	20799	619	8218.348
	mERG	17426	338.3	1234	3119	22862	619	5182.469
	GDM	22115.8	312.8	1169.8	2812.3	20863.2	683.2	338.634
AIR CANADA		22766	626	1197	4829	27676	998	
	CCR	19827.27	504.58	1042.49	4205.65	27676	998	3010.534
	mERG	22766	544	1019.4	4829	27726.8	998	202.105
	GDM	22646.14	506.1355	1138.032	2979.014	27970.74	1097.634	1884.531
AMERICAN		80627	2381	5149	18624	133796	1838	
	CCR	76484.91	2258.68	4884.48	16956.88	133796	2908.5	4600.772

Continued on next page

Name		Input				Output		
		x1	x2	x3	x4	y1	y2	d
	mERG	80627	2349.1	4698.8	18624	133796	2402.3	722.587
	GDM	80609.4	2363.4	4701.4	18606.4	133824.5	1855.6	449.886
CANADIAN		16613	513	1051	3358	24372	625	
	CCR	13264.87	404.33	918.92	3000.18	24372	625	3371.537
	mERG	16050.5	374.8	1051	3358	24496.6	625	592.478
	GDM	16605.9	505.9	869.8	3350.9	24383.5	632.1	182.119
CONTINENTAL		35661	1285	2835	9960	69050	1090	
	CCR	34411.55	1175.23	2735.67	9611.03	69050	2338.77	1806.723
	mERG	35661	1196.1	2680.3	9960	69050	2174.4	1098.981
	GDM	35573.8	1197.799	2747.799	8613.461	69265.72	1177.201	1374.816
DELTA		61675	1997	3972	14063	96540	1300	
	CCR	54899.41	1639.98	3535.64	12518.05	96540	2163.98	7025.655
	mERG	61675	1696.8	3338.7	12720.7	98100.1	1300	2174.137
	GDM	61647.7	1969.7	3278.0	14035.7	96584.2	1327.3	697.546
EASTERN		21350	580	1498	4459	29050	245	
	CCR	17364.35	471.72	1218.35	3626.59	29050	748.11	4113.558
	mERG	18832.4	488.4	965.8	3357.4	29050	267.1	2800.705
	GDM	21349.77	409.1798	1497.772	4458.772	29052.5	245.2282	170.839
NORTHWEST		42989	1762	3678	13698	85744	2513	
	CCR	40686.89	1491.5	3481.04	12964.46	85744	3294.21	2561.258
	mERG	42989	1533.8	3378.7	13698	85744	2998.2	614.064
	GDM	42972.42	1745.423	3257.203	13681.42	85770.78	2529.577	422.950
PANAM		28638	991	2193	7131	54054	1382	
	CCR	27617.88	896.7	2114.88	6876.99	54054	1678.14	1099.027
	mERG	28638	912.5	2096.8	7131	54054	1541.6	202.210
	GDM	28633.2	986.2	2070.5	7126.2	54061.8	1386.8	123.123
TWA		35783	1118	2389	8704	62345	1119	
	CCR	34678.89	1069.66	2315.29	8435.43	62345	1542.37	1215.813
	mERG	35783	1086.5	2266.5	8704	62345	1394.1	302.785
	GDM	35778.2	1113.2	2267.2	8699.2	62352.7	1123.8	122.420
UNITED		73902	2246	5678	18204	131905	2326	
	CCR	69949.77	2125.89	5374.34	17230.46	131905	4285.1	4529.085
	mERG	73902	2246	4877.6	17368.3	131905	3146	1418.251
	GDM	73866.54	2210.543	4777.909	18168.54	131962.3	2361.457	904.697
USAIR		53557	1252	3030	8952	59001	392	
	CCR	41428.43	968.47	2343.82	6924.72	59001	684.57	12322.699
	mERG	38248.9	992	1961.6	6818.9	59001	542.6	15495.799
	GDM	53238.7	933.7	2711.7	7430.8	59319.3	710.3	1679.484

As it is observed in the column 9 of Table 6, the distance of GDM method in all inefficient Airlines is evidently lower than the mERG method.

5. Conclusion

Benchmarking techniques, especially data envelopment analysis uses rational ideal evaluation to analyze the relative performances of decision making units. In this sense, a specific DMU is compared with a reference point on the efficient frontier of the production possibility set. In a rational sight, we may expect the reference point has the shortest distance to the DMU under consideration. So, finding a reference point to an inefficient DMU on the efficient frontier with closest distance is an important subject that recently has attracted considerable attention among researchers. This issue is important in the sense that inefficient DMUs could be efficient in an easiest manner. In this paper, we proposed an alternative procedure to determine a projection point with minimal changes and shortest distance on the strong efficient frontier. The gradient vectors of the reference supporting surfaces of the production technology set are used to determine closest reference points. The low computational efforts and the guarantee of determining an efficient projection point on the strong efficient frontier are the two important advantages of the proposed model. A real case on 28 internationals airlines from Asia-Australia, Europe and North America is given to show the real applicability of the proposed approach.

Conflict of interest

The authors declare no conflict of interest.

References

1. R. D. Banker, A. Charnes, W. W. Cooper, *Some models for estimating technical and scale inefficiency in data envelopment analysis*, Manage. Sci., **30** (1984), 1078–1092.
2. A. Emrouznejad, R. Banker, L. Neralić, *Advances in data envelopment analysis: Celebrating the 40th anniversary of DEA and the 100th anniversary of Professor Abraham Charnes' birthday*, Eur. J. Oper. Res., **278** (2019), 365–367.
3. 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.
4. W. Briec, B. Lemaire, *Technical efficiency and distance to a reverse convex set*, Eur. J. Oper. Res., **114** (1999), 178–187.
5. F. X. Frei, P. T. Harker, *Projections onto efficient frontiers: Theoretical and computational extensions to DEA*, J. Prod. Anal., **11** (1999), 275–300.
6. E. González, A. Álvarez, *From efficiency measurement to efficiency improvement: The choice of a relevant benchmark*, Eur. J. Oper. Res., **133** (2001), 512–520.
7. S. Lozano, G. Villa, *Determining a sequence of targets in DEA*, J. Oper. Res. Soc., **56** (2005), 1439–1447.
8. A. Amirteimoori, S. Kordrostami, *A Euclidean distance-based measure of efficiency in data envelopment analysis*, Optimization, **59** (2010), 985–996.
9. J. Aparicio, J. T. Pastor, *A well-defined efficiency measure for dealing with closest targets in DEA*, Appl. Math. Comput., **219** (2013), 9142–9154.

10. J. Aparicio, J. T. Pastor, *On how to properly calculate the Euclidean distance-based measure in DEA*, Optimization, **63** (2014), 421–432.
11. J. Aparicio, J. L. Ruiz, I. Sirvent, *Closest targets and minimum distance to the Pareto-efficient frontier in DEA*, J. Prod. Anal., **28** (2007), 209–218.
12. J. Aparicio, J. T. Pastor, *Closest targets and strong monotonicity on the strongly efficient frontier in DEA*, Omega, **44** (2014), 51–57.
13. Q. An, Z. Pang, H. Cen, et al. *Closest targets in environmental efficiency evaluation based on enhanced Russell measure*, Ecol. Indic., **51** (2015), 59–66.
14. R. R. Russell, *Measures of technical efficiencies*, J. Econ. Theor., **35** (1985), 109–126.
15. J. Aparicio, J. M. Cordero, J. T. Pastor, *The determination of the least distance to the strongly efficient frontier in data envelopment analysis oriented models: modelling and computational aspects*, Omega, **71** (2017), 1–10.
16. S. Razipour-GhalehJough, F. H. Lotfi, G. Jahanshahloo, et al. *Finding closest target for bank branches in the presence of weight restrictions using data envelopment analysis*, Ann. Oper. Res., **27** (2019), 1–33.
17. A. Charnes, W. W. Cooper, E. Rhodes, *Measuring the efficiency of decision making units*, Eur. J. Oper. Res., **2** (1978), 429–444.
18. K. Tone, *A slack-based measure of efficiency in DEA*, Eur. J. Oper. Res., **130** (2001), 498–509.
19. S. C. Ray, *Data Envelopment Analysis: Theory and Techniques for Economics and Operations Research*, Cambridge University Press, 2004.

AIMS Press