RAIRO-Oper. Res. 56 (2022) 2159–2179
https://doi.org/10.1051/ro/2022100

RAIRO Operations Research
www.rairo-ro.org

IMPROVING WEAK EFFICIENCY FRONTIER IN A VARIABLE RETURNS TO SCALE STOCHASTIC DATA ENVELOPMENT ANALYSIS MODEL

Davood Forghani[1], Mustapha D. Ibrahim[2,3,*] and Sahand Daneshvar[1]

Abstract. The conventional stochastic data envelopment analysis (SDEA) model suffers from biased efficiency scores for units located at the weak efficient frontier or compared to the weak frontier. This study modifies the weak efficient hyperplane(s) while maintaining the general production function by restricting the gradients of weak efficient hyperplanes in the original model using facet analysis. Empirical analysis on environmental efficiency of sustainable development goals validates the results of the modification. Results of the modified model compared to the conventional model show change in efficiency scores of weak efficient units and those compared to the weak part of the frontier while the efficiency scores of the strong efficient frontier remain the same. Furthermore, the proposed model shows greater discriminatory power compared to the conventional model, hence, providing a reliable benchmark and improvement strategy post efficiency analysis.

Mathematics Subject Classification. 90B50, 90C05, 90C90.

Received January 31, 2022. Accepted June 11, 2022.

1. Introduction

Performance management is an important part of operations research and management science. There are many techniques for performance management in operations research, however, data envelopment analysis (DEA) is an effective technique [28] that has been applied in key sectors such as finance [20], energy management and sustainability [16,21] among others. DEA evaluates relative efficiency of entities known as decision-making units (DMUs) with multiple inputs and outputs. DEA was introduced by Charnes *et al.* [8] under constant returns to scale (CRS), and was later enhanced by Banker *et al.* [4] under variable returns to scale (VRS). DEA creates a reference technology set called production possibility set (PPS), in which a frontier distinguishes comparatively the most efficient DMUs. A DMU is categorized as efficient or inefficient depending on its location relative to the affirmed frontier [17].

Keywords. Efficiency, stochastic data envelopment analysis (SDEA), weak efficient frontier, facet analysis, sustainable development goals.

[1] Department of Industrial Engineering, Faculty of Engineering, Eastern Mediterranean University, Famagusta, TRNC, Via Mersin 10, Famagusta, Turkey.

[2] Industrial Engineering Technology, Higher Colleges of Technology, Sharjah, UAE.

[3] CERIS – Civil Engineering Research and Innovation for Sustainability, Instituto Superior, Técnico, University of Lisbon, Lisbon, Portugal.

*Corresponding author: mibrahim1@hct.ac.ae; dalahs2011@gmail.com

Since the emergence of DEA, notable amounts of research have been carried out on identifying efficient frontiers and improvement targets [33]. However, there is inadequate attention paid to the weak efficient frontier and the corresponding DMUs despite their role in estimating efficiency and identifying the level of inefficiency, which has both technical and managerial implication in performance management, hence, the motivation of this study. The weak efficient frontier exists due to the mandatory postulate of satisfying the convexity constraint, which is considered a drawback in basic DEA models [10]. A biased efficiency score can be proposed for units on the weak part of the frontier and DMUs compared to the weak part of the frontier. This will impact the interpretation and implementation of the results. Few studies have developed approaches toward handling weak efficient frontier. Takeda and Nishino [30] adopted inner product norms to assess sensitivity in efficiency classification based on minimizing the distance to the weak efficient frontier. Daneshvar *et al.* [10] introduced a modified VRS model using facet analysis, which results in stability region and a new benchmark for scoring formerly weak efficient DMUs and inefficient DMUs compared to the weak efficient frontier. Similarly, the limitation of DEA as a result of weak efficient frontier is observed in stochastic DEA (SDEA). To the best of our knowledge, this issue is yet to be addressed in SDEA, hence, the gap in literature and novelty of this study.

SDEA is an important extension of DEA in which stochastic models based on the possibility of random variations in input-output data are considered [6]. In traditional DEA, the data is considered to be deterministic. Consequently, efficiency measurements do not assume to have stochastic properties. Therefore, the cause of inefficiency in DEA is only due to technical inefficiency [27]. Huang and Li [15] stated that there is no place for stochastic variations in the data when using DEA. The sample size in traditional DEA is also a possible drawback. The frontier is sensitive to the sample size [29]. If the number of observations in the sample increases, the technical efficiency of the DMUs often decreases because there is a possibility that the DMUs which are placed close to the frontier increase. Banker [3] stated that for a monotone increasing and concave production function, under certain conditions, the estimated frontier by DEA can reach the true frontier asymptotically. By inserting observations into the sample, efficiency scores either decrease or stay constant. Therefore, while many other statistical methods are affected by sample size, frontier estimation methods that work only based on extremal points are affected even more severely.

Measurement error has a major impact on efficiency estimation [13, 15]. Frontier estimation methods are in general are based only on extremal points. Outliers as a result of measurement error can influence the frontier estimation. Sengupta [29] proposed identifying the outliers and removing them from the data set. This method is based on central location measures and emphasizes that almost all of the deviation is a result of measurement errors. The notion of a severely deterministic frontier in which efficiency scores are fixed is logical under certain conditions [26]. Effects of sampling, the observations quantity, and other error forms makes interpretation of the deterministic efficiency scores challenging. Consequently, SDEA models are proposed to mitigate such limitations. Unlike traditional DEA, SDEA accommodates stochastic variations in data, such as data entry and measurement errors. As a result, SDEA reflects these variations in efficiency scores. A DMU that is rated as efficient relative to others in DEA, could be inefficient in SDEA, and *vice versa* [23, 32].

In DEA, there are many cases where DMUs are compared to weak efficient frontier or the whole frontier is made of only weak efficient units [11]. The weak frontier has significant effect on the final efficiency scores [31]. When the weight of an input or output becomes zero, this infers the input or outputs has no contribution to the efficiency of the unit under evaluation, which is not practical. Studies have tried mitigating this by adding a non-Archimedean value ε as the lower bound (strict positivity condition) for the weights of the model [7, 9]. However, this comes with certain drawbacks such as biased efficiency score for units outside the strong efficient frontier and output improvement restrictions which result to conflicting improvement for units with the same inputs. Furthermore, the managerial implication of estimating a biased efficiency score include developing improvement strategies based on inaccurate efficiency scores.

In this study, the weak hyperplanes of a basic SDEA model with the above mentioned drawbacks are modified to obtain an unbiased efficiency score. In order to modify the model, gradients of weak efficient hyperplanes are modified by adding constraints. These modifications are made without violating concavity and monotonicity, hence the production function properties remain the same. Moreover, the modified model can easily be implemented using mathematical optimization modeling systems such as GAMS software.

The remainder of the article is organized as follows: Section 2 presents the background of Banker's SDEA and technical validation of the proposed model. Section 3 discusses the proposed model with illustrative examples in Section 3.3. A case study on environmental efficiency analysis of sustainable development goals is used to validate the proposed model in Section 4. The paper is concluded in Section 5.

2. Banker's stochastic data envelopment analysis model

There are multiple notions of stochastic DEA that drives the technique in several directions [26]. These directions are grouped into three categories. The first direction employs DEA to handle estimated deviations from frontier as random deviations. The second develops DEA to handle random noise in the form of either measurement errors or specification errors. The third direction utilize DEA to formulate the PPS as a random PPS, based on the random variation in data. Two important approaches within the field of SDEA contain all the three directions mentioned above: Stochastic frontier analysis (SFA) and chance-constrained DEA (CCDEA) [26]. In this research, a proposed model in a work of Banker [2] is studied, which can be considered as the foundation of SFA models. Banker *et al.* published some of the results of [2] in [6] and performed sensitivity and stability analysis. Application of the Banker's model of SDEA can be found in [5].

Banker [2] introduced a basic model of SDEA. In the model, a symmetric two-sided deviation term peculiar to random factors (such as model specification and measurement errors) in company with the one-sided deviation term related to DMU's inefficiency is developed. As a result, only the single output case is considered because the multiple output case results in nonlinear programming.

To express the relationship between this model and conventional DEA, consider the assumptions of variable return to scale model of Banker [4] for estimating the PPS from observed data on output vectors y_j and input vectors $x_j, j = 1, \ldots, n$. For the single output case the postulates for estimating the production frontier correspondence $y = f(x)$ relating the single output y to the input vector $x, f : X \to R$ where X is the convex hull of x_j, can be specified as with the following postulates:

– Monotonicity of production frontier

$$\text{if } y = f(x), y' = f(x') \text{ and } x \geq x', \text{ then } y \geq y'. \tag{2.1}$$

– Concavity of production frontier

$$\text{if } y = f(x), y' = f(x') \text{ and } 0 \leq \lambda \leq 1, \tag{2.2}$$

$$\text{then } (1-\lambda)y + \lambda y' \leq f((1-\lambda)x + \lambda x'). \tag{2.3}$$

– Envelopment of observed data

$$\text{for each observation } j = 1, \ldots, n, \quad y_j \leq f(x_j). \tag{2.4}$$

– Minimum extrapolation

$$\text{if } g : X \to R \text{ satisfies postulates 1, 2 and 3 then } g(x) \geq f(x) \text{ for all } x \in X.$$

By considering these four postulates, the state of monotone increasing and concave production frontier is satisfied. To estimate this production frontier by stochastic DEA, the possible effect of uncontrolled random factors must be incorporated. Hence, such deviations caused by random factors and their stochasticity impact on the specification of the model is represented by the term u_j which can be expressed as:

$$u_j = u_j^+ - u_j^- \text{ with } u_j^+, u_j^- \geq 0. \tag{2.5}$$

The random deviations u_j are supposed to be symmetric. Therefore, it is captured in the constraint below:

$$\sum_{j=1}^{n} u_j^+ = \sum_{j=1}^{n} u_j^- . \tag{2.6}$$

Along with deviations due to random factors, as in traditional DEA, the inefficiency of the DMU may cause a shortfall in output compared to the predicted output level. Such deviations due to DMU inefficiency are shown by a nonnegative term v_j. To sum up, the actual output level could be represented as follows:

$$y_j = f(x_j) - u_j^+ + u_j^- - v_j \tag{2.7}$$

where $f(x_j)$ is the function of the estimated frontier. Equation (2.7) indicates that an efficient DMU is the one with $v_j = 0$ regardless of being placed under the frontier ($u_j^+ = 0$), or above the frontier ($u_j^- = 0$), or on the frontier ($u_j^+ = u_j^- = 0$). Then, the production frontier values are approximated by minimizing a weighted sum of the two deviations subject to the following constraints (for details, see Banker [2], pp. 4–9):

$$\begin{aligned}
&\text{Minimize } \sum_{j=1}^{n}\left(u_j^+ + u_j^- + cv_j\right)\\
&\text{subject to}\\
&\qquad \text{for each } j = 1,\ldots,n \text{ and for all } k = 1,\ldots,n \text{ and } k \neq j\\
&\qquad (x_k - x_j)w_j + (v_j - v_k) + \left(u_j^+ - u_j^- - u_k^+ + u_k^-\right) \geq y_k - y_j\\
&\qquad \sum_{j=1}^{n}\left(u_j^+ - u_j^-\right) = 0\\
&\qquad w_j \geq 0, v_j, u_j^+, u_j^- \geq 0
\end{aligned} \tag{2.8}$$

u_j^+ : the deviation of DMU_j from the frontier due to error (negative residual);
u_j^-: the deviation of DMU_j from the frontier due to error (positive residual);
v_j: the deviation of DMU_j from the frontier due to technical inefficiency (negative residual);
w_j: the slope of the estimated monotone increasing concave frontier at the point of the efficient output-oriented DMU_j.

The SDEA model with m inputs and n observations has $(m + 3)n$ variables and $n^2 - n + 1$ constraints and obviously is not infeasible since a feasible solution can be obtained from basic DEA when $u_j^+ = u_j^- = 0$.

The weight $c > 0$ in the objective function is a pre-specified constant which by giving different values, different estimates of the production function may be obtained. The model represents a combination of the minimum absolute deviation (MAD) model (due to random factors) and the basic DEA model (due to inefficiencies).

The correlation between the constant c and the contributions of the MAD and the DEA models are formalized.

Theorem. *For any given data set $\{(x_j, y_j)|j = 1,\ldots,n\}$, there exist c^M and c^D with $1/n \leq c^M \leq c^D \leq 2$ such that the model reduces to a minimum absolute deviation model (i.e. $v_j^* = 0$ for all j) for all $c > c^D$, and to the basic DEA formulation (i.e. $u^{+*} = u^{-*} = 0$ for all j) for all $c < c^M$ (for proof, see Banker [2], pp. 9–12).*

Lemma 2.1. *If $c > 2 \rightarrow v_j^* = 0$ for all $j = 1,\ldots,n$.*

Lemma 2.2. *If $c < 1/n \rightarrow u_{j^*}^+ = u_{j^*}^- = 0$ for all $j = 1,\ldots,n$.*

3. Proposed stochastic data envelopment analysis model

Banker's SDEA model presents two tools for estimating production function (MAD & DEA). Each has its individual drawbacks when considered separately. However, when combined, they support the drawback associated with each individual technique. For example, MAD is a deterministic method that estimates only the average performance while DEA evaluates efficiency relative to a production frontier that measures the best obtainable performance. Furthermore, the problem with the regression-based parametric methods is that by specifying a particular parametric form, a considerable arbitrary and restrictive structure is imposed on the input-output correspondence. In contrast, DEA imposes a minimal structure of monotonicity and convexity on the PPS. On the other hand, DEA method only allows for one-sided inefficiency deviations whereas regression gives the possibility of having a two-sided deviations component due to random errors. Given the above mentioned advantages, the model does not account for the effect of weak efficient frontier which tends to affect the efficiency scores of all related DMUs.

3.1. Weak efficient frontier

In this section, the problem definition which is associated to the weak efficient frontier is discussed. In Banker's stochastic model, since all the deviations are measured vertically (the direction of the deviations' vector due to MAD and DEA has to be the same so they can be summed for total deviation), starting from each DMU to the frontier line, the combined DEA model should be considered as output-oriented BCC.

In output-oriented BCC, the weak efficient frontier is the horizontal line projected from the efficient DMU with the highest amount of input as shown in Figure 1. Multidimensional frontier visualization of production possibility sets was developed by Afanasiev *et al.* [1] to visualize the efficiency frontier. Weak efficient frontier has a significant effect on all DMUs. The relative efficiency will be evaluated only by its output rather than the ratio of output over input. A biased efficiency score is proposed as a result of that relative comparison. This has an effect on recommendations and improvement strategies based on the estimated efficiency scores. In this study, a modified SDEA model is proposed to mitigate the effect of the weak efficient frontier thus presenting an unbiased efficiency score for all related DMUs. A single-output multiple-input model is considered, although the presented stochastic model by Banker can be extended to multiple outputs in a similar way. However, the first set of constraints would include nonlinear terms. Hence, the calculations will be less controllable than for the single output.

Previous studies [6, 18, 19] pay less attention to the variable w_j, which shows the slope of the frontier in an one-input and one-output stochastic model (2D space), or the normal vector of the frontier hyperplane in a multiple dimensional space. A non-Archimedean value ε as the lower bound (strict positivity condition) for the weights of the model has been used to mitigate this effect [7, 9]. However, non-Archimedean value ε results to a strict vertical frontier for input orientation, and horizontal plane for output orientation. This penalizes DMUs compared to the weak part of the frontier and estimate biased efficiency score for units outside the strong part of the efficiency frontier. Furthermore, when improving inefficient DMUs such DMUs A, B, and C in Figure 1, different target outputs will be proposed for inefficient DMUs with the same input. Similarly for large scale (high input) inefficient DMUs seeking to decompose, different input targets will be proposed for DMUs with the same outputs. This has major implication for performance improvement in addition to the biased efficiency score.

In this study, the focus is made on the variable w_j and its potential to improve the model. The steps described in Section 3.3 illustrate the proposed method to modify the weak efficient frontier toward the best practice frontier as a piecewise linear monotonically increasing and concave production function. But first, Section 3.2 describes facet analysis.

3.2. Facet analysis

Based on the constraints (refer to the classic model), the supporting hyperplanes for PPS of Banker's stochastic DEA model can be rewritten as follows:

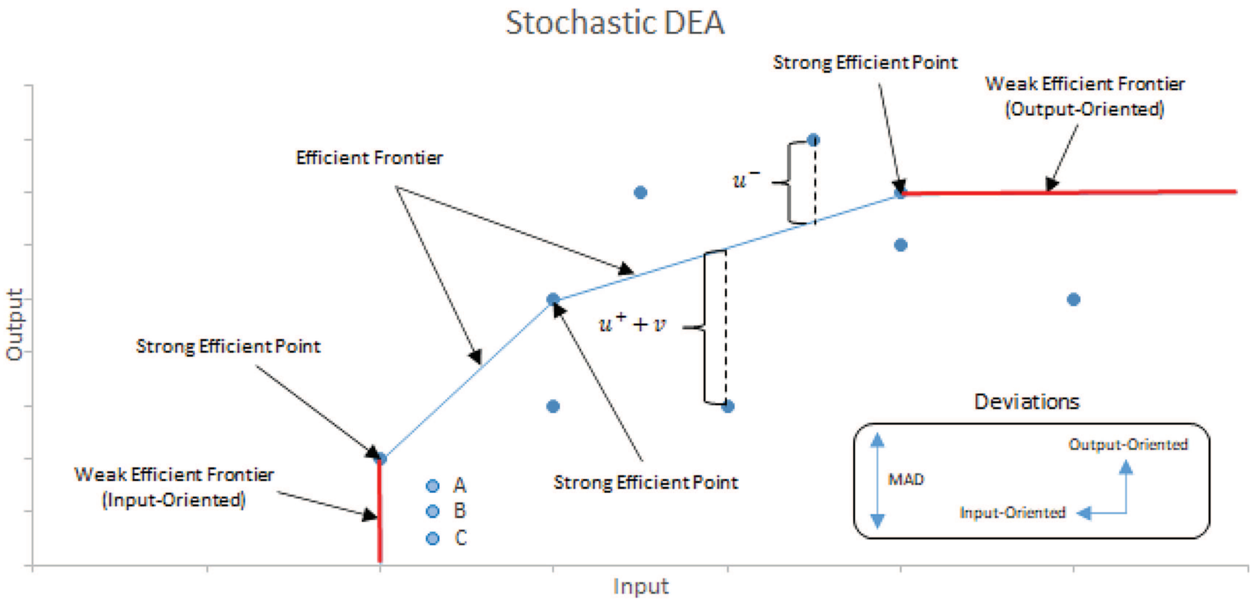

Figure 1. Stochastic DEA frontier.

$$(w_{1j}, w_{2j}, \ldots, w_{nj}, v_j, v_k, u_j^+, u_j^-, u_k^+, u_k^-) \times \begin{pmatrix} (x_{1k} - x_{1j}) \\ (x_{2k} - x_{2j}) \\ \vdots \\ (x_{nk} - x_{nj}) \\ 1 \\ -1 \\ 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \geq y_k - y_j. \tag{3.1}$$

By considering all the deviation terms as one variable u_0 and modifying (3.1), then we have:

$$(-w_{1j}, -w_{2j}, \ldots, -w_{nj}, 1) \times \begin{pmatrix} (x_{1k} - x_{1j}) \\ (x_{2k} - x_{2j}) \\ \vdots \\ (x_{nk} - x_{nj}) \\ y_k - y_j \end{pmatrix} \leq u_0. \tag{3.2}$$

Let us call the vector $(-w_{1j}, -w_{2j}, \ldots, -w_{nj}, 1), P_j$. P_j is the normal vector to efficient facet in multiplier side of output-oriented BCC model which can be interpreted as the scaled price vector associated with a DMU in the relative interior of an efficient facet [25]. If at least one of the elements of the vector P_j is equal to zero, then the corresponding facet is weak facet. Otherwise, it is an efficient facet. For example, if $w_{11} \geq 0$, then, the price related to input 1 of DMU 1 could be equal to zero, and if $w_{11} = 0$, it means that no matter how much you increase that input, the efficiency score will not change. This is the biased result of DEA for DMUs compared

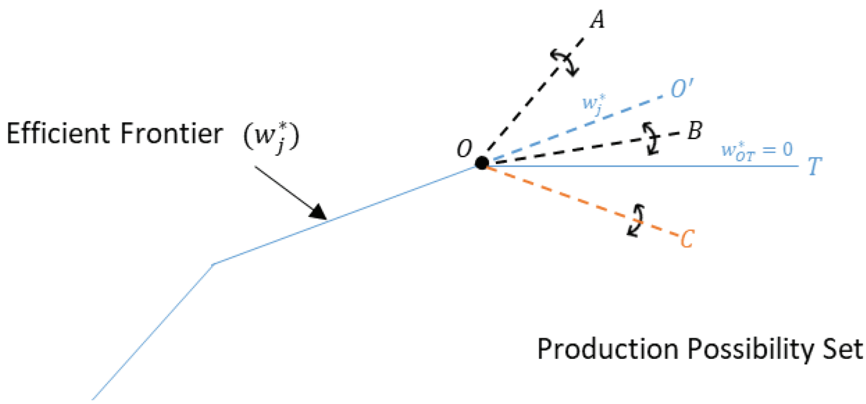

FIGURE 2. Possible scenario to modify weak efficient frontier.

with weak efficient frontier. However, assigning a minimum attainable price to w_{ij} seems logical to change full dimensional weak facets (FDWFs) to full dimensional efficient facets (FDEFs). In this way, contribution of all inputs of all DMUs are guaranteed while assumptions behind DEA method and production function like concavity are satisfied. Furthermore, the possibility set in this case is expected to expand while frontier hyperplanes adjust themselves to keep the concavity and monotonically increasing property.

3.2.1. One-input one-output model

The following steps illustrate the modification of the weak efficient frontier in SDEA.

Step 1. Solve the observed set of inputs and output using model 8.
Note: the constant c could be defined by the examiner considering the importance of the type of deviations as well as its limits given before (stochastic possibility: $\frac{1}{n} < c < 2$, Definitely MAD: $c \geq 2$, Definitely Classic DEA $\leq 1/n$). Different values of "c" can be used to investigate the effect.
Step 2. Find the smallest non-zero value for w_j and naming it β.
Note: if there are any w_j with value zero, it shows that the corresponding DMU$_j$ is placed on or compared with the weak efficient frontier.
Step 3. Adding a constraint stating that w_j have to be greater than or equal to β.
Note: To modify the weak efficient frontier in an output-oriented approach, corresponding w_j which is zero should not violate the assumptions and constraints made beforehand. Figure 2 shows different scenarios. However, only one is feasible to be applied. As it is illustrated in Figure 2, line OT is the weak efficient frontier with $w_j^* = w_{OT}^* = 0$. To modify the slope, different options are available:
(1) $\overline{OC}$: it violates the assumption of monotonically increasing function.
(2) $\overline{OA}$: it violates the assumption of concavity.
(3) $\overline{OB}$: it doesn't violate any assumption but it's subjective to decide which value between zero and w_j^*.
(4) $\overline{OO'}$: It doesn't violate any assumption. Also, it's vital to the consistency of the method.
By applying this steps, it will guarantee that the frontier will not be horizontal. Also, since the monotonically increasing concave frontier is already satisfied by other constraints, it is definite that the so-called weak efficient frontier will be modified to continue with the same slope as the closest efficient frontier with the minimum slope.
Step 4. Solve the set of inputs and outputs with the modified SDEA model (3.3) by defining beta as the minimum strictly positive slope of the frontier.

Note: the non-negativity constraint for w_j is redundant. This means that the number of constraints stays the same.

$$\begin{aligned}
&\text{Minimize } \sum_{j=1}^{n}\left(u_j^+ + u_j^- + cv_j\right)\\
&\text{subject to}\\
&\qquad \text{for each } j=1,\dots,n \text{ and for all } k=1,\dots,n \text{ and } k\neq j\\
&\qquad (x_k - x_j)w_j + (v_j - v_k) + \left(u_j^+ - u_j^- - u_k^+ + u_k^-\right) \geq y_k - y_j\\
&\qquad \sum_{j=1}^{n}\left(u_j^+ - u_j^-\right) = 0\\
&\qquad w_j \geq \beta \text{ for each } j=1,\dots,n\\
&\qquad v_j, u_j^+, u_j^- \geq 0.
\end{aligned} \tag{3.3}$$

3.2.2. *Two-inputs one-output model*

The classic SDEA model is also applicable in a one-output multiple-input system. However, the case of multiple outputs becomes nonlinear [6]. Hence, this study demonstrates the modified SDEA model for multiple-inputs one-output system.

To begin with, the stochastic model presented by Banker could be developed for two inputs case as bellow:

$$\begin{aligned}
&\text{Minimize } \sum_{j=1}^{n}\left(u_j^+ + u_j^- + cv_j\right)\\
&\text{subject to}\\
&\qquad \text{for each } j=1,\dots,n \text{ and for all } k=1,\dots,n \text{ and } k\neq j\\
&\qquad \sum_{i=1}^{2}[(x_{ik} - x_{ij})w_{ij}] + (v_j - v_k) + \left(u_j^+ - u_j^- - u_k^+ + u_k^-\right) \geq y_k - y_j\\
&\qquad \sum_{j=1}^{n}\left(u_j^+ - u_j^-\right) = 0\\
&\qquad w_{ij} \geq 0, \quad v_j, u_j^+, u_j^- \geq 0.
\end{aligned} \tag{3.4}$$

u_j^+: the deviation of DMU_j from the frontier due to error (negative residual);
u_j^-: the deviation of DMU_j from the frontier due to error (positive residual);
v_j: the deviation of DMU_j from the frontier due to technical inefficiency (negative residual);
w_{ij}: the element of the normal vector of the estimated monotone increasing concave frontier plane at the point of the efficient output-oriented DMU_j;
i: index of inputs ($i = 1, 2$);
j: index of the DMUs ($j = 1, 2, \dots, n$).

The implementation steps are similar to one inputs-output system. However, since we have two w's for each $\text{DMU}_j(w_{1j}, w_{2j})$, two constraints will be added to the base model.

Step 1. Solving the model as it is given above with observed inputs and outputs.
Step 2. Find the smallest non-zero values for w_{1j} and w_{2j} and naming them respectively β_1 and β_2.
Note: if there are any DMU with $w_{1j} = w_{2j} = 0$, it shows that the corresponding DMU_j is placed on or compared with the weak efficient frontier (parallel to the Inputs surface in the coordinate system).

Step 3. Add the following constraints to the basic model:

$$w_{1j} \geq \beta_1$$
$$w_{2j} \geq \beta_2.$$

Step 4. Solve the proposed model below:

$$\begin{aligned}
&\text{Minimize } \sum_{j=1}^{n}\left(u_j^+ + u_j^- + cv_j\right)\\
&\text{subject to}\\
&\qquad \text{for each } j = 1,\ldots,n \text{ and for all } k = 1,\ldots,n \text{ and } k \neq j\\
&\qquad \sum_{i=1}^{2}[(x_{ik} - x_{ij})w_{ij}] + (v_j - v_k) + \left(u_j^+ - u_j^- - u_k^+ + u_k^-\right) \geq y_k - y_j\\
&\qquad \sum_{j=1}^{n}\left(u_j^+ - u_j^-\right) = 0\\
&\qquad w_{1j} \geq \beta_1\\
&\qquad w_{2j} \geq \beta_2\\
&\qquad v_j, u_j^+, u_j^- \geq 0.
\end{aligned} \tag{3.5}$$

3.3. Illustrative examples

To illustrate the practicality of the proposed model, the data set used by Banker [2] is utilized. Banker's data set does not have any DMU located at the weak efficient frontier or compared to the weak efficient frontier. Therefore, new DMUs are added to the PPS.

Example 3.1. Table 1 presents a single input single output data set from a production company. Outputs are production units and inputs are labor in hours. The first twelve months were given by Banker and the last three are added to examine the modified model.

Model (2.8) is applied to the data set. Different ranges of "c" are selected. Figure 3 illustrates the sensitivity of the estimates on deviations caused by inefficiency or random factors. By increasing the weight c, the estimated production frontier tends to move further down and as a result, symmetric deviation due to random factors appears. Concavity and monotonically increasing postulates are reasonably considered since by increasing the labor hour, the labor productivity tends to decrease due to the traffic and complications at upper capacity operation, but it doesn't always mean that the productivity level stops at a point (in this example point number 4) and from that point by increasing the labor, production doesn't increase (points 13, 14 and 15 are evaluated by weak frontier).

By applying the modified SDEA model (3.3), and the steps described. The estimated production function will continue to increase with the least possible slope giving a more realistic approximation by turning the weak efficient frontier into an efficient frontier. It is obvious that by increasing the number of DMUs and finding an appropriate corresponding c value, the possibility of having a smooth frontier similar to logarithmic production functions rises. Figure 4 shows the modified weak frontier.

Points 13, 14, and 15 are not compared to a weak frontier anymore. Moreover, the evaluation criteria have changed in many cases. For example, in the case $c = 0.55$, after solving the banker's traditional SDEA model by GAMS software, the minimum nonzero w_j equals 0.7. Hence, $\beta = 0.7$, and by adding the new constraint to the model, new values for the slope and the deviations will be obtained. As it can be seen in Figures 3 and 4, and also Appendix A, DMUs 12, 7, and 9 that were strong efficient points, stay the same in the modified model.

Table 1. One input and one output data.

Month	Input	Output
1	416.092	2089.51
2	349.785	1919.35
3	399.403	1974.43
4	455.73	2117.16
5	360.803	1792.88
6	396.241	1818.82
7	272.435	1537.66
8	312.949	1598.41
9	314.229	1701.87
10	416.09	1868.95
11	290.686	1554.14
12	260.762	1436.05
13	500	2117.16
14	520	2000
15	475	2050

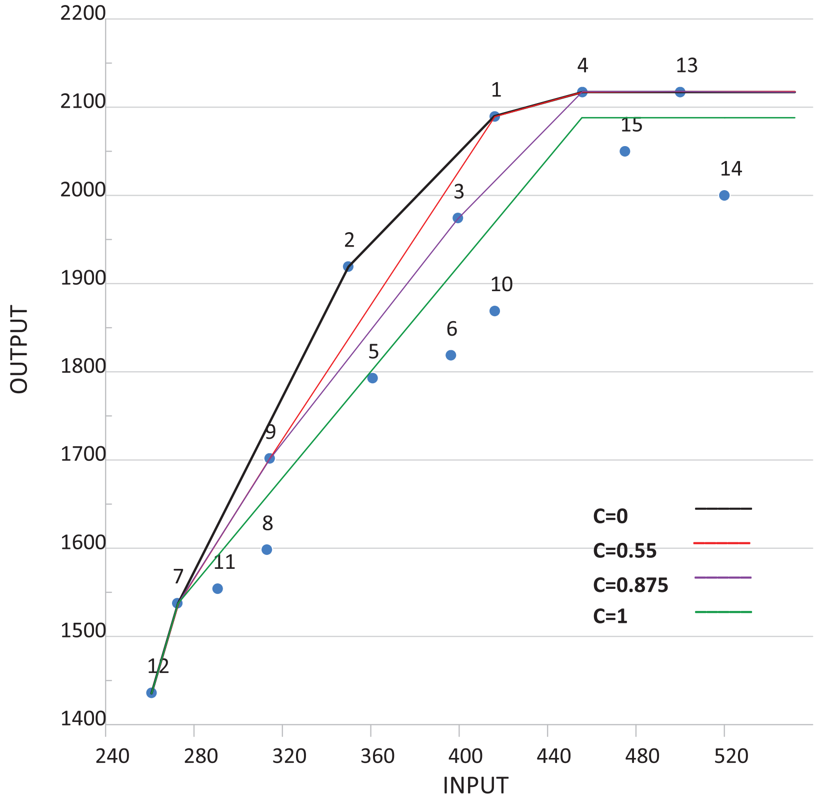

Figure 3. SDEA estimated production frontier.

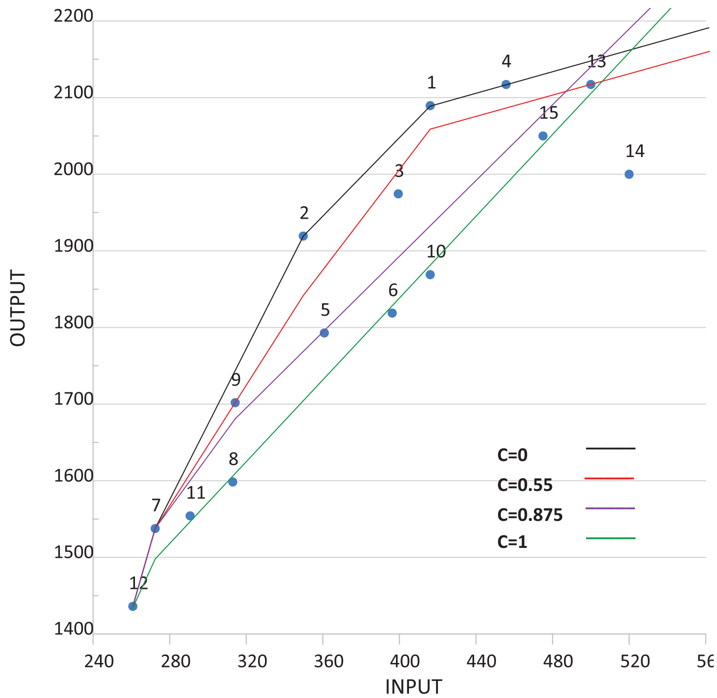

FIGURE 4. Modified SDEA estimated production frontier.

However, DMUs 1 and 4 which were strong efficient points before, are efficient but with error deviation in the modified model. Hence, they were located on the frontier because they had some error in their measurement and the actual frontier passes under points 1 and 4. Although it implies that PPS shrunk, yet for DMUs 14 and 15, which were compared with weak efficient frontier, the PPS has become larger. This assigns them smaller efficiency scores, but better opportunity to be improved. Weak efficient DMU 13 turns to a strong efficient point and the other DMUs can be compared with the new frontier. This means that another criterion has been obtained for comparing or improving the DMUs. The related results for the classic and modified model are given in Appendix A. The deviation ratio $\frac{y_j}{\hat{y}_j} = \frac{y_j}{y_j+u^++v-u^-}$ and the efficiency estimates $1-\left(\frac{v_j}{y_j}\right)$ are also reported.

Example 3.2. The following data were generated illustrated the two inputs one output modified model. Table 2 presents the data set. The steps discussed in model development are applied considering four different values for the weight c. Table 3 shows the results obtained for $c = 0.750$ and $c = 0.875$. The constant c in the model has changed between Tables 3a and 3b, however, independent change did not affect the outcomes of efficiency whether it changes between the classic and modified model; leading us to conclude that the modified model is able to give more realistic result regardless of the c value. It could be observed from Table 3a that some of the efficiency estimates have remained the same. Although, the balance between u^+ and u^- has changed which is projected in deviation ratio. Nevertheless, some changes similar to DMU 12 show that it is efficient in the classical model but not in the modified one. This indicates that DMU 12 is actually inefficient and could be improved because its deviation from the frontier is due to a combination of inefficiency and error not only error. In DMU 11 the efficiency estimate increases to its maximum value while the distance to the frontier stays almost the same. The reason is that the deviation due to inefficiency in the classical model is in fact due to errors and random factors. Considering a higher value for c (Tab. 3b), taking DMU 11, compared to Table 3a, an increase in efficiency of the classical model is detected since by increasing the weight c, the model moves toward strict

Table 2. Two input and one output data.

	Input 1	Input 2	Output
DMU 1	1	7	2
DMU 2	3	4	2.5
DMU 3	3.5	4.3	3.32
DMU 4	8	1	1
DMU 5	1.5	10	3
DMU 6	4.5	6	5.75
DMU 7	7.5	9	6
DMU 8	12	1.5	2
DMU 9	8	10	6
DMU 10	1.6	6	1.8
DMU 11	8	7	2
DMU 12	6.4	6.7	4
DMU 13	10	11.6	5
DMU 14	4	4.5	3.21
DMU 15	1.5	8	2.12

Table 3. Two input and one output results.

	(a)																
	Classical model										Modified model						
DMU	w_{1j}	w_{2j}	u^+	u^-	v	c	Deviation Ratio	Efficiency Estimates	β_1	β_2	w_{1j}	w_{2j}	u^+	u^-	v	Deviation Ratio	Efficiency Estimates
1	2.00	0.00	0.00	0.00	0.00		1.000	1.000			0.90	0.18	0.00	0.00	0.00	1.000	1.000
2	0.67	1.62	0.00	0.00	0.00		1.000	1.000			0.67	1.62	0.00	0.00	0.00	1.000	1.000
3	0.14	0.89	0.00	0.00	0.00		1.000	1.000			0.14	0.78	0.00	0.00	0.00	1.000	1.000
4	0.00	2.00	0.00	0.00	0.00		1.000	1.000			0.14	0.89	0.00	0.00	0.00	1.000	1.000
5	2.00	0.00	0.00	0.00	0.00		1.000	1.000			0.90	0.18	0.00	0.00	0.00	1.000	1.000
6	0.00	0.34	0.00	0.78	0.00		1.156	1.000			0.14	0.18	0.00	0.97	0.00	1.202	1.000
7	0.00	0.00	0.00	0.00	0.00		1.000	1.000			0.14	0.18	0.00	0.25	0.00	1.044	1.000
8	0.00	2.00	0.00	0.00	0.00	0.750	1.000	1.000	0.139	0.183	0.14	0.89	0.00	0.00	0.00	0.999	0.999
9	0.00	0.00	0.00	0.00	0.00		1.000	1.000			0.14	0.18	0.00	0.00	0.00	1.000	1.000
10	5.50	3.50	0.00	0.00	0.00		1.000	1.000			5.50	3.50	0.00	0.00	0.00	1.000	1.000
11	0.00	0.34	0.13	0.00	0.59		0.865	0.873			0.14	0.18	0.85	0.00	0.00	0.844	1.000
12	0.14	0.89	0.64	0.00	0.00		0.861	1.000			0.14	0.78	0.37	0.00	0.17	0.880	0.957
13	0.00	0.00	0.00	0.00	1.00		0.833	0.800			0.14	0.18	0.00	0.00	1.57	0.761	0.686
14	0.14	0.89	0.00	0.00	0.36		0.900	0.888			0.14	0.78	0.00	0.00	0.34	0.906	0.896
15	0.90	0.18	0.00	0.00	0.51		0.805	0.758			0.85	0.19	0.00	0.00	0.50	0.810	0.766
	(b)																
1	2.00	0.00	0.00	0.00	0.00		1.000	1.000			0.75	0.21	0.00	0.00	0.00	1.000	1.000
2	0.67	1.62	0.00	0.00	0.00		1.000	1.000			0.49	1.32	0.00	0.00	0.00	1.000	1.000
3	0.05	0.61	0.00	0.00	0.00		1.000	1.000			0.05	0.54	0.00	0.18	0.00	1.058	1.000
4	0.00	2.00	0.00	0.00	0.00		1.000	1.000			0.05	1.64	0.00	0.00	0.00	1.000	1.000
5	2.00	0.00	0.00	0.00	0.00		1.000	1.000			0.75	0.21	0.00	0.00	0.00	1.000	1.000
6	0.00	0.53	0.00	1.35	0.00		1.307	1.000			0.05	0.51	0.00	1.64	0.00	1.400	1.000
7	0.00	0.00	0.00	0.00	0.00		1.000	1.000			0.05	0.21	0.00	0.23	0.00	1.040	1.000
8	0.00	2.00	0.00	0.00	0.00	0.875	1.000	1.000	0.045	0.208	0.05	1.64	0.00	0.00	0.00	1.000	1.000
9	0.00	0.00	0.00	0.00	0.00		1.000	1.000			0.05	0.21	0.00	0.00	0.00	1.000	1.000
10	5.50	3.50	0.00	0.00	0.00		1.000	1.000			5.50	3.50	0.00	0.00	0.00	1.000	1.000
11	0.00	0.53	0.00	0.00	0.33		0.932	0.928			0.05	0.51	0.17	0.00	0.00	0.964	1.000
12	0.05	0.61	0.24	0.00	0.00		0.943	1.000			0.05	0.54	0.00	0.00	0.00	1.000	1.000
13	0.00	0.00	0.86	0.00	0.15		0.833	0.971			0.05	0.21	1.42	0.00	0.00	0.778	1.000
14	0.05	0.61	0.25	0.00	0.00		0.927	1.000			0.05	0.54	0.01	0.00	0.05	0.982	0.985
15	0.75	0.21	0.00	0.00	0.46		0.820	0.781			0.69	0.22	0.44	0.00	0.00	0.827	1.000

regression and the inefficiency share decreases. For the same reason (less effect of DEA), in DMU 12, only the balance between error terms changes. Comparing the two models, on the other hand, shows that DMU 14 went through a decrease in its efficiency score. That leads to the conclusion that it is truly inefficient. If a DMU is

TABLE 4. Descriptive statistics of Latin America and Caribbean (LAC) countries.

		Labor force (total)	Capital ($billion)	Energy Consumption (TJ)	EPI
2012	Mean	12 703 765.52	52.61	403.74	55.06
	Std. Dev.	22 638 542.54	115.15	810.25	5.76
	Min	140 788	0.21	2.59	41.15
	Max	97 597 798	488.35	3246.74	69.03
2014	Mean	13 023 934.1	53.91	418.40	50.84
	Std. Dev.	23 161 740.18	116.19	845.37	9.67
	Min	151 507	0.28	2.85	19.01
	Max	99 932 834	494.98	3444.85	69.93
2016	Mean	13 460 976.86	48.70	422.82	72.32
	Std. Dev.	23 796 036	95.71	823.02	8.05
	Min	164 047	0.37	3.09	43.28
	Max	102 508 951	374.28	3310.70	80.03
2018	Mean	13 983 141.43	49.81	432.46	57.99
	Std. Dev.	24 514 148.16	96.39	833.25	7.17
	Min	173 515	0.33	3.05	33.74
	Max	105 542 232	378.96	3370.65	67.85
2020	Mean	14 380 421.57	49.08	445.73	45.83
	Std. Dev.	25 010 345.34	93.66	846.14	7.21
	Min	183 771	0.35	3.25	27.00
	Max	107 371 779	362.23	3440.01	55.30

efficient in the standard model and inefficient in the modified model, then it is truly inefficient. If it is inefficient in the standard model, and efficient in the modified model, then the inefficiency in the standard model is as a result of random factors such as measurement errors.

4. Environmental efficiency of sustainable development goals case study

In this section, we consider a case study with a comparative example to demonstrate the applicability and robustness of the proposed model. Environmental efficiency of sustainable development goals (SDGs) of 21 Latin America and Caribbean (LAC) countries are considered from 2012 to 2020. LAC is a region in transition economically [23], and environmentally [24]. The energy and environmental issues facial LAC is crucial to its ability to attain the SDGs. Furthermore, environmental efficiency is imperative for attainment of SDGs. Studies have shown direct or indirect relationship between environmental efficiency and other SDGs [14, 22]. Defining environmental efficiency is a complex task due to the undesirable output such as environmental degradation which is a byproduct of human activities. However, environmental performance index (EPI), a composite indicator that carefully compounds multiple environmental dimensions into an index that articulates the environmental implication was developed by Yale Center for Environmental Law & Policy and Center for International Earth Science Information Network Earth Institute, Columbia University [12]. Literature reveals that assessing environmental efficiency is generating growing attention among researchers. Numerous models have been used including the conventional DEA models. A drawback of such models includes minimal discriminatory power of the models, and identifying the units that are weak efficient. An advantage of identifying weak efficient units is implementing precautionary actions to prevent future possible inefficient performance. Therefore, utilizing the proposed modified SDEA that presents an opportunity to discriminate between efficient and inefficient countries, and identify those that appear efficient but are not really efficient.

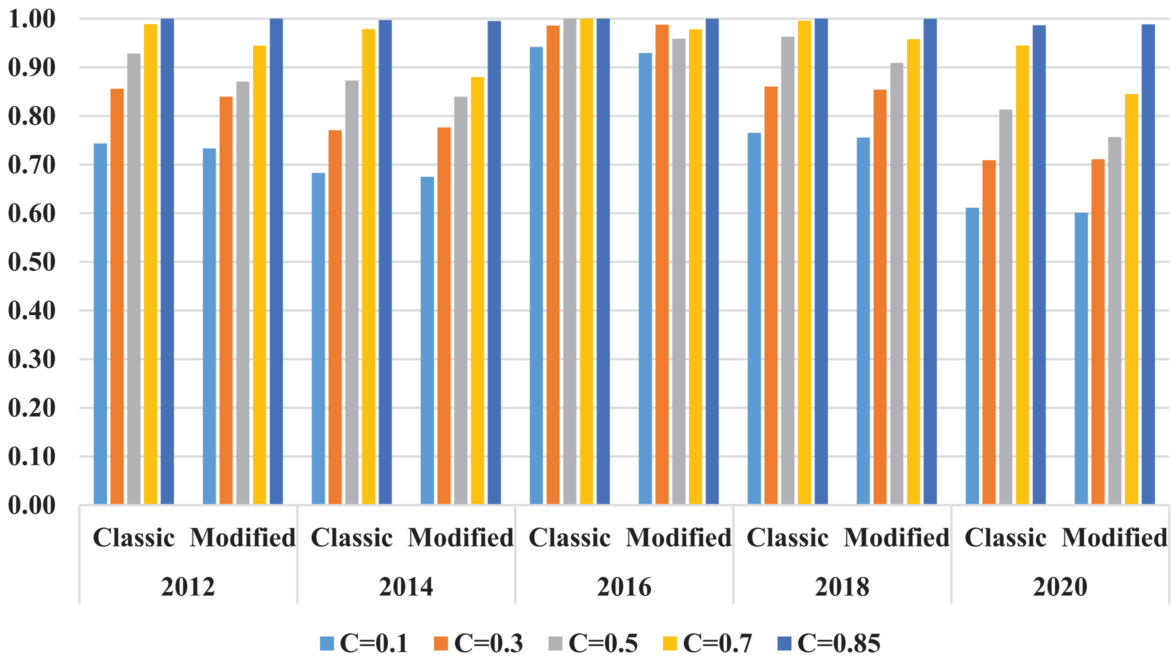

FIGURE 5. Average environmental efficiency of classic and modified SDEA models.

TABLE 5. Comparison between numbers of efficient units.

Year	2012		2014		2016		2018		2020	
SDEA	Classic	Modified	Classic	Modified	Classic	Modified	Classic	Modified	Classic	Modified
$C = 0.1$	1	1	0	0	7	7	0	0	0	0
$C = 0.3$	6	2	0	2	19	18	3	2	0	1
$C = 0.5$	7	8	5	7	21	18	15	13	4	3
$C = 0.7$	17	15	16	9	21	18	19	15	12	8
$C = 0.85$	21	21	19	20	21	21	21	20	18	18

Public available data similar to previous studies of [16] on environmental efficiency are considered. These countries utilize three inputs Labor ($x1$), Capital ($x2$), and Energy consumption ($x3$) for its general sustainability. Since environmental efficiency is the primary focus, Environmental Performance Index (EPI) ($y1$) is considered as output. The composite nature of EPI makes it suitable indicator for environmental sustainability output. The descriptive statistics of the inputs and outputs are presented in Table 4. The average labor force shows a steady increase over the evaluated period. Similarly, average capital and energy consumption surge continuously. Average environmental performance showed a stable increase from 2012 to 2016, with a dip in performance in 2018 and 2020. An extended three input and output version of model (3.4) for the classic SDEA, and model (3.5) for the modified SDEA are employed. $C = 0.1$, $C = 0.3$, $C = 0.5$, $C = 0.7$, and $C = 0.85$ are used in the evaluation.

Using GAMS and the data in the data in brief package attached to this paper. The environmental efficiency of LAC countries were estimated. Figure 5 presents the average efficiency scores as a summary comparison between the classic SDEA and the modified SDEA model. Results show 2016 to be the most efficient period. In

Table 6. Environmental efficiency of Latin America and the Caribbean countries.

	2012 ($C = 0.7$)		2014 ($C = 0.85$)		2016 ($C = 0.7$)		2018 (0.85)	
Country Name	Classic	Modified	Classic	Modified	Classic	Modified	Classic	Modified
Argentina	1.00	0.71	1.00	1.00	1.00	1.00	1.00	1.00
Belize	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00
Bolivia	0.94	0.94	1.00	0.93	1.00	1.00	1.00	0.99
Brazil	1.00	0.78	1.00	1.00	1.00	0.09	1.00	0.63
Chile	1.00	0.83	1.00	1.00	1.00	1.00	1.00	0.94
Colombia	0.97	0.86	1.00	1.00	1.00	1.00	1.00	1.00
Costa Rica	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Cuba	0.94	0.96	1.00	0.97	1.00	1.00	1.00	1.00
Dominican Republic	0.88	0.92	1.00	0.93	1.00	1.00	1.00	1.00
Ecuador	0.95	0.98	1.00	0.99	1.00	1.00	1.00	1.00
El Salvador	0.92	0.93	1.00	1.00	1.00	1.00	1.00	1.00
Guatemala	1.00	0.91	1.00	1.00	1.00	1.00	1.00	1.00
Haiti	0.81	0.79	1.00	0.57	0.83	0.81	1.00	0.74
Honduras	1.00	0.93	1.00	1.00	1.00	1.00	1.00	0.94
Jamaica	1.00	0.96	1.00	1.00	1.00	1.00	1.00	1.00
Mexico	1.00	1.00	1.00	1.00	1.00	0.44	1.00	1.00
Nicaragua	1.00	1.00	1.00	0.94	1.00	1.00	1.00	1.00
Panama	1.00	0.98	1.00	0.98	1.00	1.00	1.00	1.00
Paraguay	1.00	0.90	1.00	1.00	1.00	1.00	1.00	0.96
Peru	1.00	0.79	0.78	1.00	1.00	1.00	1.00	1.00
Uruguay	1.00	0.98	1.00	0.95	1.00	1.00	1.00	1.00

all evaluated period and different version of "C". The modified SDEA model present a lesser average efficiency compared to the conventional SDEA model.

The discriminatory power of the proposed modified SDEA model is illustrated in Table 5, showing the number of units identified as efficient across the evaluated periods out of the 21 countries evaluated. Number of DMUs identified as inefficient consistently decreased for most periods in the modified SDEA model compared to the classic model, with the exception of $C = 0.5$ (2012 and 2014), and $C = 0.85$ (2016). To numerically illustrate the robustness and discriminatory power of the model, Table 6 presents the estimated efficiency for 2012 ($C = 0.7$), 2014 ($C = 0.85$), 2016 ($C = 0.7$), 2018 ($C = 0.85$), and 2020 ($C = 0.7$). For detail efficiency scores of all values of C across all periods for classic and model, see Appendix B.

5. Conclusion

Efficiency evaluation in DEA infers inefficiency to be purely technical. The conventional DEA models requires inputs/outputs to be relatively perfect. In reality, errors in data collection or entry occur which has an effect on efficiency evaluation and improvement strategies. An important part is SDEA in which various stochastic models are studied according to the possibility of random deviations in observations. Among different approaches to SDEA, Banker's SDEA model is the foundation. Effects of the weak efficient frontier that persist in conventional DEA still exists in SDEA. Biased efficiency scores are allocated to DMUs located at the weak frontier or DMUs compared to the weak part of the frontier. This has significant effect on interpretation of results.

The main goal of this study is to address the weak efficient frontier considering the production function postulates in order to imitate the underlying true frontier. Facet analysis was employed to generate constraints for the weak efficient hyperplanes. Results of the numerical example show a new benchmark for estimating relative efficiency that not only follows general production function properties such as nonnegativity, monotonicity (nondecreasing), and concavity but also seems to be able to show a smoother frontier, especially in the weak efficient area. Compared to the conventional SDEA model, if a DMU is efficient in the standard model and inefficient in modified model, then it is truly inefficient. If it is inefficient in the standard model, and efficient in the modified model, then the inefficiency in the standard model is as a result of random factors such as measurement errors. Thus, results should be interpreted carefully. An empirical study on environmental efficiency of SDGs for LAC countries was performed which further validates the proposed model. Environmental efficiency is a typical example of stochastic data due to the nature of data collection. Results of the study shows an improved discriminatory power of the modified model. The empirical application of the model is limited to single output for the definition of environmental efficiency inspite of the robustness of EPI. Future study could consider multiple outputs for environmental efficiency and undesirable outputs such as CO_2 emissions. As a preface to the future research direction, Banker [2] introduced alternative models and extensions to deal with multiple outputs case which can be examined by applying the proposed method. However, the model becomes a nonlinear programming problem when multiple outputs are considered. Thus, the linearization methods could be helpful to address this issue. Another potential area of investigation is sensitivity and stability analysis. The situation of perturbations of all inputs, output, and simultaneously all inputs and output could be studied on the modified models. Moreover, finding a method for specifying the weight "c" based on the importance of the deviations, and a specific desired estimate of the frontier could be studied as well.

Data availability statement

The datasets analyzed during the current study are available in the [NAME] repository, [M.D. Ibrahim, Environmental SDG. Mendeley Data, V1 (2021). DOI: 10.17632/969b2ybwrj.1].

Appendix A. 1 input 1 output illustrated application of Banker's SDEA and suggested modified model for different values of the weight "c"

Observations			Classic Model								Modified Model							
DMU	X	Y	w	u^+	u^-	v	c	z	Deviation Ratio	Efficiency Estimates	β	w	u^+	u^-	v	z	Deviation Ratio	Efficiency Estimates
1	416.09	2089.51	0.70	0.00	0.00	0.00			1.000	1.000		2.57	0.00	0.00	0.00		1.000	1.000
2	349.79	1919.35	2.57	0.00	0.00	0.00			1.000	1.000		2.57	0.00	0.00	0.00		1.000	1.000
3	399.40	1974.43	2.57	0.00	0.00	72.25			0.965	0.963		2.57	0.00	0.00	72.25		0.965	0.963
4	455.73	2117.16	0.00	0.00	0.00	0.00			1.000	1.000		0.70	0.00	0.00	0.00		1.000	1.000
5	360.80	1792.88	2.57	0.00	0.00	154.75			0.921	0.914		2.57	0.00	0.00	154.75		0.921	0.914
6	396.24	1818.82	2.57	0.00	0.00	219.75			0.892	0.879		2.57	0.00	0.00	219.75		0.892	0.879
7	272.44	1537.66	4.94	0.00	0.00	0.00			1.000	1.000		4.94	0.00	0.00	0.00		1.000	1.000
8	312.95	1598.41	4.94	0.00	0.00	139.17	0.00	0.00	0.920	0.913	0.70	4.94	0.00	0.00	139.17	0.00	0.920	0.913
9	314.23	1701.87	4.94	0.00	0.00	42.03			0.976	0.975		4.94	0.00	0.00	42.03		0.976	0.975
10	416.09	1868.95	2.57	0.00	0.00	220.56			0.894	0.882		2.57	0.00	0.00	220.56		0.894	0.882
11	290.69	1554.14	4.94	0.00	0.00	73.58			0.955	0.953		4.94	0.00	0.00	73.58		0.955	0.953
12	260.76	1436.05	8.71	0.00	0.00	0.00			1.000	1.000		8.71	0.00	0.00	0.00		1.000	1.000
13	500.00	2117.16	0.00	0.00	0.00	0.00			1.000	1.000		0.70	0.00	0.00	30.92		0.986	0.985
14	520.00	2000.00	0.00	0.00	0.00	117.16			0.945	0.941		0.70	0.00	0.00	162.04		0.925	0.919
15	475.00	2050.00	0.00	0.00	0.00	67.16			0.968	0.967		0.70	0.00	0.00	80.63		0.962	0.961

Observations			Classic Model								Modified Model							
DMU	X	Y	w	u^+	u^-	v	c	z	Deviation Ratio	Efficiency Estimates	β	w	u^+	u^-	v	z	Deviation Ratio	Efficiency Estimates
1	416.09	2089.51	0.70	0.00	0.00	0.00			1.000	1.000		0.70	0.00	30.92	0.00		1.015	1.000
2	349.79	1919.35	3.74	0.00	77.78	0.00			1.042	1.000		3.27	0.00	77.78	0.00		1.042	1.000
3	399.40	1974.43	3.74	0.00	0.00	52.68			0.974	0.973		3.27	0.00	0.00	29.54		0.985	0.985
4	455.73	2117.16	0.00	0.00	0.00	0.00			1.000	1.000		0.70	0.00	30.90	0.00		1.015	1.000
5	360.80	1792.88	3.74	0.00	0.00	89.89			0.952	0.950		3.27	0.00	0.00	84.75		0.955	0.953
6	396.24	1818.82	3.74	77.78	0.00	118.68			0.903	0.935		3.27	0.00	0.00	174.80		0.912	0.904
7	272.44	1537.66	3.93	0.00	0.00	0.00			1.000	1.000		3.93	0.00	0.00	0.00		1.000	1.000
8	312.95	1598.41	3.93	0.00	0.00	98.43	0.55	606.44	0.942	0.938	0.70	3.93	0.00	0.00	98.43	649.69	0.942	0.938
9	314.23	1701.87	3.93	0.00	0.00	0.00			1.000	1.000		3.93	0.00	0.00	0.00		1.000	1.000
10	416.09	1868.95	3.74	0.00	0.00	220.55			0.894	0.882		3.27	139.60	0.00	50.04		0.908	0.973
11	290.69	1554.14	3.93	0.00	0.00	55.23			0.966	0.964		3.93	0.00	0.00	55.23		0.966	0.964
12	260.76	1436.05	8.71	0.00	0.00	0.00			1.000	1.000		8.71	0.00	0.00	0.00		1.000	1.000
13	500.00	2117.16	0.00	0.00	0.00	0.00			1.000	1.000		0.70	0.00	0.00	0.00		1.000	1.000
14	520.00	2000.00	0.00	0.00	0.00	117.16			0.945	0.941		0.70	0.00	0.00	131.12		0.938	0.934
15	475.00	2050.00	0.00	0.00	0.00	67.16			0.968	0.967		0.70	0.00	0.00	49.71		0.976	0.976

Observations			Classic Model								Modified Model							
DMU	X	Y	w	u^{+}	u^{-}	v	c	z	Deviation Ratio	Efficiency Estimates	β	w	u^{+}	u^{-}	v	z	Deviation Ratio	Efficiency Estimates
1	416.09	2089.51	2.47	0.00	70.39	0.00			1.035	1.000		2.47	0.00	156.71	0.00		1.081	1.000
2	349.79	1919.35	2.68	0.00	77.78	0.00			1.042	1.000		2.47	0.00	150.53	0.00		1.085	1.000
3	399.40	1974.43	2.68	0.00	0.00	0.00			1.000	1.000		2.47	0.00	82.90	0.00		1.044	1.000
4	455.73	2117.16	0.00	0.00	0.00	0.00			1.000	1.000		2.47	0.00	86.34	0.00		1.043	1.000
5	360.80	1792.88	2.68	0.00	0.00	78.19			0.958	0.956		2.47	3.19	0.00	0.00		0.998	1.000
6	396.24	1818.82	2.68	147.14	0.00	0.00			0.925	1.000		2.47	64.89	0.00	0.00		0.966	1.000
7	272.44	1537.66	3.93	0.00	0.00	0.00			1.000	1.000		3.43	0.00	0.00	0.00		1.000	1.000
8	312.95	1598.41	3.93	0.00	0.00	98.43	0.875	790.99	0.942	0.938	2.47	3.43	78.09	0.00	0.00	994.92	0.953	1.000
9	314.23	1701.87	3.93	0.00	0.00	0.00			1.000	1.000		3.43	0.00	20.98	0.00		1.012	1.000
10	416.09	1868.95	2.68	1.03	0.00	149.13			0.926	0.920		2.47	63.84	0.00	0.00		0.967	1.000
11	290.69	1554.14	3.93	0.00	0.00	55.23			0.966	0.964		3.43	46.07	0.00	0.00		0.971	1.000
12	260.76	1436.05	8.71	0.00	0.00	0.00			1.000	1.000		8.71	0.00	0.00	0.00		1.000	1.000
13	500.00	2117.16	0.00	0.00	0.00	0.00			1.000	1.000		2.47	23.14	0.00	0.00		0.989	1.000
14	520.00	2000.00	0.00	0.00	0.00	117.16			0.945	0.941		2.47	189.76	0.00	0.00		0.913	1.000
15	475.00	2050.00	0.00	0.00	0.00	67.16			0.968	0.967		2.47	28.48	0.00	0.00		0.986	1.000

Observations			Classic Model								Modified Model							
DMU	X	Y	w	u^{+}	u^{-}	v	c	z	Deviation Ratio	Efficiency Estimates	β	w	u^{+}	u^{-}	v	z	Deviation Ratio	Efficiency Estimates
1	416.09	2089.51	3.00	0.00	120.84	0.00			1.061	1.000		3.00	0.00	148.23	0.00		1.076	1.000
2	349.79	1919.35	3.00	0.00	149.62	0.00			1.085	1.000		3.00	0.00	176.99	0.00		1.102	1.000
3	399.40	1974.43	3.00	0.00	55.83	0.00			1.029	1.000		3.00	0.00	83.22	0.00		1.044	1.000
4	455.73	2117.16	0.00	0.00	29.56	0.00			1.014	1.000		3.00	0.00	56.70	0.00		1.028	1.000
5	360.80	1792.88	3.00	9.91	0.00	0.00			0.995	1.000		3.00	0.00	17.47	0.00		1.010	1.000
6	396.24	1818.82	3.00	90.30	0.00	0.00			0.953	1.000		3.00	62.90	0.00	0.00		0.967	1.000
7	272.44	1537.66	3.00	0.00	0.00	0.00			1.000	1.000		3.00	0.00	27.35	0.00		1.018	1.000
8	312.95	1598.41	3.00	60.80	0.00	0.00	1.00	848.42	0.963	1.000	3.00	3.00	33.44	0.00	0.00	1152.85	0.980	1.000
9	314.23	1701.87	3.00	0.00	38.82	0.00			1.023	1.000		3.00	0.00	66.18	0.00		1.040	1.000
10	416.09	1868.95	3.00	99.72	0.00	0.00			0.949	1.000		3.00	72.32	0.00	0.00		0.963	1.000
11	290.69	1554.14	3.00	38.28	0.00	0.00			0.976	1.000		3.00	10.92	0.00	0.00		0.993	1.000
12	260.76	1436.05	8.71	0.00	0.00	0.00			1.000	1.000		6.36	0.00	0.00	0.00		1.000	1.000
13	500.00	2117.16	0.00	0.00	29.56	0.00			1.014	1.000		3.00	75.84	0.00	0.00		0.965	1.000
14	520.00	2000.00	0.00	87.60	0.00	0.00			0.958	1.000		3.00	253.00	0.00	0.00		0.888	1.000
15	475.00	2050.00	0.00	37.60	0.00	0.00			0.982	1.000		3.00	68.00	0.00	0.00		0.968	1.000

Appendix B. Detail environmental efficiency of Latin America and the Caribbean countries

B.1. $C = 0.1$

	2012		2014		2016		2018		2020	
Country Name	Classic	Modified	Classic	Modified	Classic	Modified	Classic	Modified	Classic	Modified
Argentina	0.721	0.709	0.634	0.622	1	1	0.755	0.743	0.667	0.655
Belize	1	1	0.883	0.935	1	1	0.882	0.882	0.65	0.65
Bolivia	0.735	0.734	0.677	0.676	0.931	0.931	0.736	0.736	0.587	0.587
Brazil	0.775	0.667	0.676	0.564	1	0.901	0.773	0.671	0.654	0.553
Chile	0.713	0.707	0.896	0.889	0.992	1	0.739	0.733	0.712	0.705
Colombia	0.8	0.787	0.655	0.641	0.969	0.955	0.835	0.82	0.681	0.665
Costa Rica	0.895	0.896	0.762	0.763	1	1	0.874	0.875	0.682	0.683
Cuba	0.744	0.743	0.724	0.724	1	1	0.818	0.818	0.63	0.63
Dominican Republic	0.709	0.708	0.706	0.705	0.99	0.988	0.857	0.855	0.623	0.62
Ecuador	0.781	0.779	0.756	0.754	0.857	0.854	0.742	0.739	0.661	0.659
El Salvador	0.712	0.712	0.609	0.609	0.91	0.909	0.731	0.731	0.453	0.453
Guatemala	0.687	0.686	0.636	0.636	0.905	0.905	0.685	0.684	0.499	0.498
Haiti	0.588	0.587	0.31	0.309	0.611	0.61	0.489	0.488	0.403	0.402
Honduras	0.712	0.712	0.666	0.666	0.924	0.923	0.693	0.693	0.651	0.651
Jamaica	0.743	0.744	0.791	0.791	1	1	0.793	0.793	0.719	0.719
Mexico	0.628	0.573	0.702	0.646	0.934	0.876	0.76	0.701	0.504	0.445
Nicaragua	0.806	0.805	0.693	0.692	0.863	0.862	0.752	0.751	0.659	0.659
Panama	0.759	0.759	0.741	0.741	1	1	0.809	0.809	0.575	0.574
Paraguay	0.699	0.697	0.53	0.529	0.994	0.916	0.708	0.706	0.613	0.611
Peru	0.652	0.645	0.586	0.579	0.935	0.927	0.797	0.789	0.561	0.553
Uruguay	0.752	0.752	0.707	0.707	0.962	0.962	0.847	0.847	0.653	0.654

B.2. $C = 0.3$

Country Name	2012 Classic	2012 Modified	2014 Classic	2014 Modified	2016 Classic	2016 Modified	2018 Classic	2018 Modified	2020 Classic	2020 Modified
Argentina	0.786	0.79	0.7	0.703	1	1	0.821	0.825	0.733	1
Belize	1	1	0.963	1	1	1	1	1	0.801	0.801
Bolivia	1	0.831	0.756	0.755	1	1	0.809	0.807	0.66	0.658
Brazil	0.841	0.802	0.742	0.702	1	1	0.839	0.807	0.72	0.689
Chile	0.776	0.779	0.958	0.961	1	1	0.802	0.806	0.775	0.778
Colombia	0.861	0.864	0.716	0.718	1	1	0.895	0.898	0.741	0.744
Costa Rica	1	0.99	0.856	1	1	1	0.959	0.958	0.767	0.766
Cuba	0.815	0.813	0.796	0.793	1	1	0.889	0.885	0.701	0.697
Dominican Republic	0.793	0.791	0.772	0.77	1	1	0.902	0.9	0.672	0.67
Ecuador	0.85	0.848	0.824	0.822	0.925	1	0.81	0.808	0.73	0.728
El Salvador	0.835	0.835	0.732	0.732	1	1	0.842	0.842	0.561	0.562
Guatemala	0.773	1	0.719	0.718	1	0.98	0.762	0.76	0.575	0.572
Haiti	0.765	0.764	0.483	0.483	0.779	0.779	0.651	0.65	0.558	0.558
Honduras	1	0.821	0.77	0.771	1	1	0.784	0.784	0.958	0.74
Jamaica	1	0.882	0.925	0.925	1	1	1	0.923	0.847	0.847
Mexico	0.694	0.679	0.768	0.753	1	0.983	0.826	0.809	0.57	0.553
Nicaragua	1	0.957	0.837	0.837	1	1	0.887	0.887	0.809	0.857
Panama	0.837	0.836	0.814	0.812	1	1	1	1	0.644	0.641
Paraguay	0.79	0.79	0.615	0.614	1	1	0.787	0.786	0.693	0.691
Peru	0.717	0.717	0.651	0.651	1	1	0.862	0.862	0.627	0.627
Uruguay	0.843	0.841	0.787	0.786	1	1	0.939	0.938	0.75	0.749

B.3. $C = 0.5$

Country Name	2012 Classic	2012 Modified	2014 Classic	2014 Modified	2016 Classic	2016 Modified	2018 Classic	2018 Modified	2020 Classic	2020 Modified
Argentina	0.88	0.808	0.794	0.724	1	1	0.916	1	0.827	0.758
Belize	1	1	0.984	0.974	1	1	1	1	1	0.826
Bolivia	0.925	0.928	0.854	1	1	1	1	1	0.747	1
Brazil	0.936	0.256	0.837	0.143	1	0.583	1	0.343	0.814	0.241
Chile	1	0.837	1	1	1	1	1	1	1	0.835
Colombia	0.953	1	0.809	0.717	1	1	1	0.891	0.836	0.729
Costa Rica	1	1	0.913	1	1	1	1	1	0.831	0.87
Cuba	0.907	0.929	0.886	0.909	1	1	1	1	0.779	0.812
Dominican Republic	0.876	0.878	0.853	0.87	1	1	1	1	0.767	0.771
Ecuador	0.931	1	0.906	0.918	1	1	0.892	1	0.812	0.823
El Salvador	1	0.915	0.813	0.881	1	1	0.926	0.929	1	0.651
Guatemala	0.849	0.868	0.794	1	1	1	0.838	0.863	0.648	0.677
Haiti	1	0.794	0.538	0.515	1	0.814	0.707	0.69	0.615	0.602
Honduras	1	0.907	0.856	0.86	1	1	1	1	0.83	0.84
Jamaica	0.955	1	1	0.998	1	1	1	1	1	0.925
Mexico	0.788	0.452	1	0.524	1	0.737	1	0.56	0.664	0.301
Nicaragua	1	1	0.897	1	1	1	1	1	0.887	1
Panama	0.91	1	1	1	1	1	1	1	0.724	0.752
Paraguay	0.87	1	1	1	1	1	1	0.892	0.768	0.797
Peru	0.803	0.774	0.738	0.707	1	1	0.948	0.916	0.713	0.681
Uruguay	0.911	0.939	0.859	0.891	1	1	1	1	0.813	1

B.4. $C = 0.7$

Country Name	2012		2014		2016		2018		2020	
	Classic	Modified	Classic	Modified	Classic	Modified	Classic	Modified	Classic	Modified
Argentina	1	0.629	0.858	0.545	1	0.921	1	0.65	1	0.576
Belize	1	1	1	1	1	1	1	1	1	1
Bolivia	1	1	0.93	0.943	1	1	1	1	0.838	0.856
Brazil	1	1	1	0.729	1	0.637	1	0.91	1	0.99
Chile	1	0.794	1	0.979	1	1	0.957	0.811	1	0.78
Colombia	1	0.771	1	0.584	1	1	1	1	0.9	0.566
Costa Rica	1	1	1	1	1	1	1	1	0.981	1
Cuba	1	1	1	1	1	1	1	1	0.866	1
Dominican Republic	1	0.961	1	1	1	1	1	1	1	1
Ecuador	0.996	1	0.971	1	1	1	0.957	1	0.876	1
El Salvador	1	1	0.864	0.885	1	1	1	1	1	0.731
Guatemala	0.932	0.956	1	1	1	1	1	1	1	0.765
Haiti	1	1	1	0.572	1	1	1	1	1	0.667
Honduras	1	1	1	1	1	1	1	1	0.901	1
Jamaica	1	1	1	1	1	1	1	1	1	1
Mexico	0.853	0.725	0.927	0.797	1	0.98	1	0.797	0.729	0.532
Nicaragua	1	1	1	0.968	1	1	1	1	1	1
Panama	0.987	1	1	1	1	1	1	1	1	0.826
Paraguay	1	1	1	0.794	1	1	1	0.962	0.861	0.863
Peru	1	1	1	0.689	1	1	1	0.979	1	0.652
Uruguay	1	1	1	0.989	1	1	1	1	0.894	0.939

B.5. $C = 0.85$

Country Name	2012		2014		2016		2018		2020	
	Classic	Modified	Classic	Modified	Classic	Modified	Classic	Modified	Classic	Modified
Argentina	1	1	1	1	1	1	1	1	1	0.911
Belize	1	1	1	1	1	1	1	1	1	1
Bolivia	1	1	1	1	1	1	1	1	1	1
Brazil	1	1	1	1	1	1	1	1	1	1
Chile	1	1	1	1	1	1	1	1	1	1
Colombia	1	1	1	1	1	1	1	1	1	1
Costa Rica	1	1	1	1	1	1	1	1	0.933	0.933
Cuba	1	1	0.988	1	1	1	1	1	1	1
Dominican Republic	1	1	1	1	1	1	1	1	1	1
Ecuador	1	1	1	1	1	1	1	1	1	1
El Salvador	1	1	1	1	1	1	1	0.998	1	1
Guatemala	1	1	1	1	1	1	1	1	1	1
Haiti	1	1	1	1	1	1	1	1	1	1
Honduras	1	1	1	1	1	1	1	1	1	0.915
Jamaica	1	1	1	1	1	1	1	1	0.989	1
Mexico	1	1	1	1	1	1	1	1	0.792	1
Nicaragua	1	1	1	1	1	1	1	1	1	1
Panama	1	1	1	1	1	1	1	1	1	1
Paraguay	1	1	1	1	1	1	1	1	1	1
Peru	1	1	1	0.896	1	1	1	1	1	1
Uruguay	1	1	0.95	1	1	1	1	1	1	1

References

[1] A.P. Afanasiev, V.E. Krivonozhko, A.V. Lychev and O.V. Sukhoroslov, Multidimensional frontier visualization based on optimization methods using parallel computations. *J. Global Optim.* **76** (2020) 563–574.

[2] R. Banker, Stochastic data envelopment analysis. Working Paper. Carnegie Mellon University (1988).

[3] R.D. Banker, Maximum likelihood, consistency and data envelopment analysis: a statistical foundation. *Manage. Sci.* **39** (1993) 1265–1273.

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

[5] R.D. Banker, S.M. Datar and C.F. Kemerer, A model to evaluate variables impacting the productivity of software maintenance projects. *Manage. Sci.* **37** (1991) 1–18.

[6] R. Banker, K. Kotarac and L. Neralić, Sensitivity and stability in stochastic data envelopment analysis. *J. Oper. Res. Soc.* **66** (2015) 134–147.

[7] A. Charnes and W.W. Cooper, The Non-Archimedean CCR Ratio for Efficiency Analysis: A Rejoinder to Boyd and Färe. Texas Univ at Austin Center for Cybernetic Studies (1984).

[8] A. Charnes, W.W. Cooper and E. Rhodes, Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* **2** (1978) 429–444.

[9] A. Chames, W. Cooper and E. Rhodes, Short communication: measuring the efficiency of decision making units. *Eur. J. Oper. Res.* **3** (1979) 339.

[10] S. Daneshvar, G. Izbirak and A. Javadi, Sensitivity analysis on modified variable returns to scale model in Data Envelopment Analysis using facet analysis. *Comput. Ind. Eng.* **76** (2014) 32–39.

[11] M. Davtalab-Olyaie, I. Roshdi, V. Partovi Nia and M. Asgharian, On characterizing full dimensional weak facets in DEA with variable returns to scale technology. *Optimization* **64** (2015) 2455–2476.

[12] EPI, Environmental Performance Index (EPI) (2021, 16th May). https://epi.yale.edu/about-epi.

[13] M.K. Epstein and J.C. Henderson, Data envelopment analysis for managerial control and diagnosis. *Decis. Sci.* **20** (1989) 90–119.

[14] L.M. Fonseca, J.P. Domingues and A.M. Dima, Mapping the sustainable development goals relationships. *Sustainability* **12** (2020) 3359.

[15] Z. Huang and S.X. Li, Dominance stochastic models in data envelopment analysis. *Eur. J. Oper. Res.* **95** (1996) 390–403.

[16] M.D. Ibrahim and A.A. Alola, Integrated analysis of energy-economic development-environmental sustainability nexus: case study of MENA countries. *Sci. Total Environ.* **737** (2020) 139768.

[17] M. Ibrahim, S. Daneshvar, H. Güden and B. Vizvari, Target setting in data envelopment analysis: efficiency improvement models with predefined inputs/outputs. *OPSEARCH* **57** (2020) 1319–1336.

[18] S. Jradi and J. Ruggiero, Stochastic data envelopment analysis: a quantile regression approach to estimate the production frontier. *Eur. J. Oper. Res.* **278** (2019) 385–393.

[19] S. Jradi, C.F. Parmeter and J. Ruggiero, Quantile estimation of the stochastic frontier model. *Econ. Lett.* **182** (2019) 15–18.

[20] S. Kazemi, M. Tavana, M. Toloo and N.A. Zenkevich, A common weights model for investigating efficiency-based leadership in the russian banking industry. *RAIRO: Oper. Res.* **55** (2021) 213–229.

[21] E. Kazemzadeh, J.A. Fuinhas, M. Koengkan, F. Osmani and N. Silva, Do energy efficiency and export quality affect the ecological footprint in emerging countries? A two-step approach using the SBM–DEA model and panel quantile regression. *Environ. Syst. Decis.* (2022) 1–18. DOI: 10.1007/s10669-022-09846-2.

[22] E. Koçak, H. Kınacıand K. Shehzad, Environmental efficiency of disaggregated energy R&D expenditures in OECD: a bootstrap DEA approach. *Environ. Sci. Pollut. Res.* **28** (2021) 19381–19390.

[23] M. Koengkan, J.A. Fuinhas, E. Kazemzadeh, F. Osmani, N.K. Alavijeh, A. Auza and M. Teixeira, Measuring the economic efficiency performance in Latin American and Caribbean countries: an empirical evidence from stochastic production frontier and data envelopment analysis. *Int. Econ.* **169** (2022) 43–54.

[24] V. Moutinho, J.A. Fuinhas, A.C. Marques and R. Santiago, Assessing eco-efficiency through the DEA analysis and decoupling index in the Latin America countries. *J. Cleaner Prod.* **205** (2018) 512–524.

[25] O.B. Olesen and N.C. Petersen, Facet analysis in data envelopment analysis. In: Data Envelopment Analysis, edited by J. Zhu. Springer US (2015).

[26] O.B. Olesen and N.C. Petersen, Stochastic data envelopment analysis – a review. *Eur. J. Oper. Res.* **251** (2016) 2–21.

[27] J. Ondrich and J. Ruggiero, Efficiency measurement in the stochastic frontier model. *Eur. J. Oper. Res.* **129** (2001) 434–442.

[28] S. Sarkar, Performance measurement using a novel directional distance function based super efficiency model and neighbourhood theory. *RAIRO: Oper. Res.* **55** (2021) 3617–3638.

[29] J.K. Sengupta, Theory of DEA models. In: Dynamics of Data Envelopment Analysis. Springer (1995) 1–37.

[30] A. Takeda and H. Nishino, On measuring the inefficiency with the inner-product norm in data envelopment analysis. *Eur. J. Oper. Res.* **133** (2001) 377–393.

[31] H. Yang and M. Pollitt, The necessity of distinguishing weak and strong disposability among undesirable outputs in DEA: environmental performance of Chinese coal-fired power plants. *Energy Policy* **38** (2010) 4440–4444.

[32] Z. Zhou, W. Sun, H. Xiao, Q. Jin and W. Liu, Stochastic leader–follower DEA models for two-stage systems. *J. Manage. Sci. Eng.* **6** (2021) 413–434.

[33] Q. Zhu, J. Aparicio, F. Li, J. Wu and G. Kou, Determining closest targets on the extended facet production possibility set in data envelopment analysis: modeling and computational aspects. *Eur. J. Oper. Res.* **296** (2022) 927–939.