# The Importance of Discriminant Metrics in Efficiency Analysis

KELLY PATRICIA MURILLO
Department of Mathematics,
University of Aveiro,
Campus Universitário de Santiago, 3810-193 Aveiro,
PORTUGAL

*Abstract:* - In this article, a non-parametric deterministic method that combines efficiency models with mathematical techniques to examine decision units is applied. In order to better understand the calculated efficiencies, characterize and identify possible improvements in less efficient units, four discrimination metrics are proposed. The metrics are determined by how the efficiency index is calculated. The metric that best represents data and allows for more detailed analysis of results is taken as a reference to build a new metric with a more complete structure. The latter allows a general characterization of the decision unit in the context studied. The methodology presented in this study is discussed through an empirical application, which allows examining the efficiency of European countries in production sectors.

*Key-Words:* - Metrics analysis, efficiency models, Data envelope analysis, Production sectors.

Received: April 26, 2023. Revised: October 6, 2023. Accepted: October 18, 2023. Published: October 27, 2023.

## 1 Introduction

In studies involving data analysis, additional mathematical tools are often applied to analysis models with the aim of differentiating the database more strictly. In efficiency analysis, there are many aspects to consider for establishing a strong structure that allows a complete and reliable study. Part of the data analysis involves identifying units with highly differentiated characteristics, which may interfere in some way with the results obtained.

This is because units, whose input/output values correspond precisely to the minimum/maximum in which the entire data set varies, may be worse/better positioned relative to the majority. Regarding variables, their classification beyond being inputs or outputs should be considered. The analysis should establish whether the variables are stationary, desirable, undesirable, intermediate, etc. On the other hand, the efficiency models are determined according to the context, the available database, and the objectives. This determines whether a linear or nonlinear, parametric, static, or dynamic model will be used, etc.

However, it is very common in this type of analysis to emphasize the most appropriate selection of the data, the correct variables, and more adjusted models moving to the background, crucial aspects such as the metrics used in the study, forgetting that these can decisively influence the results obtained as well as the other aspects mentioned.

In the literature, efficiency measurement has been based on parametric and nonparametric border analyses. The two most commonly used methods are the regression analysis approach, which leads to the use of econometric methods, and the Data Envelopment Analysis DEA which uses linear programming, [1], [2]. DEA allows solving problems of simultaneous maximization of products or simultaneous minimization of inputs, building an optimal production frontier, and comparing each observation unit against the expected optimum. The DEA model is based on radial contractions at all undesirable inputs and outputs and radial expansions at all desirable outputs.

Recently, in some studies to evaluate efficiency where it is required to establish inefficiency indices in each variable individually, the Multidirectional Efficiency Analysis MEA, initially proposed in, [3], has been used. This model allows the reduction of inputs and expansion of outputs, looking for a separate potential improvement in each input variable and each output variable.

The use of the most suitable non-parametric model for the intended analysis depends on the objectives of each particular study. DEA and MEA are the most important models currently used to measure efficiency. [4] compares a set of public and private schools using DEA. In, [5], DEA is used to evaluate the financial performance of the textile industry in Haryana, located in the northern part of India. The authors in, [6], build on the existing studies in Dynamic Network Data Envelopment Analysis (DNDEA) and proposing a sequential structure incorporating dual-role characteristics of the production factors.

In, [7], DEA techniques were applied to examine the comparative efficiency of higher education institutions. Other interesting studies with DEA are, [8], [9], [10], [11].

Regarding MEA, in, [12], the use of MEA allows investigating how railway reforms affect the inefficiencies of specific cost factors. In, [13], MEA is used to assess the level of energy and environmental efficiency and the trend of China's transport sector. In, [14], both integrated MEA efficiency levels and efficiency standards were detected, which are represented by the specific MEA efficiency of the variable according to each type of emission or discharge of industrial pollutants from major cities in China. In, [15], the efficiency of European countries in the context of Circular Economy, is examined, considering the sector of plastic. Other interesting studies with applications in MEA are, [16], [17].

The metrics we present correspond to the need to establish a well-structured and complete methodology that allows evaluating efficiency from various angles and clearly establishing which factors must be improved. Metrics allow us to examine the factors that influence the behavior of decision units in any context, such as public health, business efficiency, educational quality, energy efficiency, and the circular economy, among others.

In this paper, four metrics are distributed in four approaches, which allow for different visualizations of the efficiency index against the given database. The index is calculated from a non-parametric and deterministic model to measure the technical efficiency of decision units. After selecting the metric that best represents the data and allows a more detailed analysis of the results, a fifth metric is defined from that. This new metric has a more complete structure, which allows a general characterization of the decision unit in the studied context.

To visualize the effects of the metrics studied, particularly in this work, we examined the efficiency of European countries in production sectors compared to their waste management capacity.

The remainder of this document is presented as follows: In the next section, the model for calculating the efficiency index is presented, and, based on this, the four proposed approaches are discussed. In Section 3, the efficiency of 26 European countries in three productive sectors is examined. In Section 4, the general comments and the final observations are established.

## 2 Methodology and Characterization of Metrics

In the literature, we find different models to examine the efficiency of decision units, which basically depend on the characteristics of the database and the objective of the study. This can be oriented to the input, oriented to the output, with a constant scale, or with a variant scale, among other features.

It should be noted that, for the purpose of this study, the efficiency indicator of each decision unit can be obtained using either of the two non-parametric and deterministic methods mentioned above: the traditional DEA, [18], or the last-model MEA, [19]. Specifically, about the DEA. The most commonly used DEA models are the DEA-CCR model, introduced in, [1], and the DEA-BCC model, introduced in, [2]. The DEA-CCR model assumes constant scale returns, and the DEA-BCC model, on the other hand, allows variable scale returns. The MEA model, on the other hand, can be adjusted to use VRS (a model with variable scale returns) or CRS (a model with constant scale returns), according to the objectives of the problem in context.

In order to facilitate reading this article, unify the concepts, and structure the metrics to be studied, we introduce an efficiency index, the one determined by the DEA-CCR model. Note, however, that another model (such as DEA-BCC or MEA) could be selected. This is because the model to use to calculate the index depends on the problem to be studied, the orientation, the scale, and the type of study to be done. The selected model has been the one that best corresponds to the characteristics of the problem to be studied, allowing us to concretely visualize the differences in the glass, paper, and plastic production sectors discussed in the numerical simulation.

Consider $k$ a decision unit. Suppose that any unit $k \in N$ produces $y_j(k), j \in [J]$ outputs, using $x_i(k)$, $i \in [I]$ inputs.

**Definition 2.1** Let $\mathrm{Z} = \{z(k)\}_{\mathrm{N}}$ be a given database with set of values $z(k) = (x_i(k), y_j(k))$. The technical efficiency index for a specific observation, $z(\bar{k}) = (x_i(\bar{k}), y_j(\bar{k}))$, $\bar{k} \in [K]$, is the optimal solution $h_k^*$ of the linear problem $P(x_i(\bar{k}), y_j(\bar{k}))$:

$$\min h_{\bar{k}} \quad \sum_{i=1}^{n} v_i x_{i\bar{k}} \tag{1}$$

such that

$$\sum_{r=1}^{m} u_r y_{rj} - \sum_{i=1}^{n} v_i x_{ij} \leq 0 \tag{2}$$

$$\sum_{r=1}^{m} u_r y_{r\bar{k}} = 1 \tag{3}$$

$$u_r, v_i \geq 0 \tag{4}$$

where u are the weights of the outputs; v are the weights of the inputs and, $r \in [M]$.

The efficiency index obtained in Definition 2.1 is a value $0 \leq h_k^* \leq 1$. The linear program (1)-(4) is executed for the $N$ decision units of the study and an index is obtained for each. This index allows us to compare the performance of each unit, according to the variables considered in the study.

## 2.1 Effective Metrics Characterization

In this section, different metrics will be distributed in approaches that will allow different visualizations of the efficiency index compared to the given database. The index is calculated from a non-parametric and deterministic model to measure the technical efficiency of the decision units.

Based on the efficiency model described in the previous section, five metrics are considered in this study: M1, M2, M3, M4, and M5. The first four (M1–M4) are related to the four approaches E1, E2, E3, and E4, which are determined by how efficiency indices $h_k^*$ are calculated:

(E1) Efficiency calculation, where metric M1 considers all study variables;

(E2) Efficiency calculation with the M2 metric, in which only a representative subset of the database is considered, obtained through Principal Component Analysis (PCA);

(E3) Calculation of efficiency, considering a composition metric. The M3 metric is defined in combination with the results of the E1 and E2 approaches;

(E4) Calculation of the efficiency index, with the metric M4, defined as a metric of composition by ranges.

A fifth metric, M5, is defined based on the best results obtained in previous approaches. The M5 metric presents a composition structure and is divided by ranges (Section 2.2). This new metric has a more complete structure, which allows a general characterization of the decision unit in the studied context.

To continue, we will describe the metrics and approaches specifically.

Let $Z = \{z(k)\}_N$ be a database given with $z(k) = (x_i(k), y_j(k))$, where $x_i(k), i \in [I]$ represents the inputs and $y_j(k), j \in [J]$ represents the products. Consider $\bar{Z} = \{\bar{z}(k)\}_N$ a subset of the database with $\bar{z}(k) = (\bar{x}_i(k), \bar{y}_j(k))$, where $\bar{x}_i(k), i \in [I]$ y $\bar{y}_j(k), j \in [J]$ represent the inputs and outputs selected by the PCA.

**(E1)** Efficiency calculation, considering all study variables.

After debugging the database on which the study will be conducted, it is important to take a first look at the results on all the selected data, applying the M1 metric (Definition 2.1.1). This involves calculating the efficiency index with all variables, analyzing the results, and examining whether they make sense within the context being studied.

**Definition 2.1.1:** Let $Z = \{z(k)\}_N$ be a database with $z(k) = (x_i(k), y_j(k))$. The efficiency index of each DMU $k \in N$, is defined as the value

$$h_k^{1*}: \text{optimal solution of } P\left(x_i(k), y_j(k)\right) \tag{5}$$

**(E2)** Efficiency calculation, considering a subset of the initial database.

For this approach, the M2 metric (Definition 2.1.2) is used. Thus, the relevance of each variable in the model must be established first. An efficiency study consistent with the results depends largely on the relevance of the variables considered in the study.

The most representative variables $\bar{Z} = \{\bar{z}(k)\}_N$ can be determined by statistical techniques such as Principal Component Analysis (PCA). We selected the most relevant variables for this study, using the PCA, accompanied by a dimensionality test called *test-dim*. The PCA analysis, proposed in, [20], transforms a series of correlated variables into a series of uncorrelated variables, [21]. Once the PCA is complete, an attenuation test is performed. This allows testing the number of axes in a multivariate analysis. The procedure is based on the calculation of the RV coefficient, [22].

**Definition 2.1.2:** Let $\bar{Z} = \{\bar{z}(k)\}_N$ be a subset of $Z = \{z(k)\}_N$ with $\bar{z}(k) = (\bar{x}_i(k), \bar{y}_j(k))$. The efficiency index of each DMU $k \in \mathrm{N}$, is defined as the value,

$\mathrm{h}_{\mathrm{k}}^{2*}$: optimal solution of $P\left(\bar{x}_i(k), \bar{y}_j(k)\right)$ (6)

**(E3)** Efficiency calculation, considering a composition metric.

The M3 metric (Definition 2.1.3) used in this approach is defined in combination with the results of the E1 and E2 approaches.

***Definition 2.1.3:*** Let $Z = \{z(k)\}_N$ be a database with $z(k) = (x_i(k), y_j(k))$, and $\bar{Z} = \{\bar{z}(k)\}_N$ a subset of the database with $\bar{z}(k) = (\bar{x}_i(k), \bar{y}_j(k))$. The efficiency index of each DMU $k \in N$, is defined as the value

$$\mathrm{h}_{\mathrm{k}}^{3*} = (\mathrm{h}_{\mathrm{k}}^{1*} + \mathrm{h}_{\mathrm{k}}^{2*})/2 \qquad (7)$$

where $\mathrm{h}_{\mathrm{k}}^{1*}$ is the optimal solution of $P(x_i(k), y_j(k))$ *with* $Z = \{z(k)\}_N$, and $\mathrm{h}_{\mathrm{k}}^{2*}$ is the optimal solution of $P(\bar{x}_i(k), \bar{y}_j(k))$ with $\bar{Z} = \{\bar{z}(k)\}_N$. Then $\mathrm{h}_{\mathrm{k}}^{1*}$ and $\mathrm{h}_{\mathrm{k}}^{2*}$ are the values in (5) and (6), respectively.

**(E4)** Calculation of the efficiency index, with a metric of composition by ranges.

The M4 metric (Definition 2.1.4) in this approach is defined in combination with the results of the E1 and E2 approaches, differentiating between units that are fully efficient ($\mathrm{h}_{\mathrm{k}}^{1*} = 1$) and the other remaining units ($\mathrm{h}_{\mathrm{k}}^{1*} \leq 0{,}99$). The latter correspond to efficient but not fully efficient units and inefficient units.

**Definition 2.1.4:** Let $Z = \{z(k)\}_N$ be a database given with $z(k) = (x_i(k), y_j(k))$, and $\bar{Z} = \{\bar{z}(k)\}_N$ a subset of the database with $\bar{z}(k) = (\bar{x}_i(k), \bar{y}_j(k))$. The efficiency index of each DMU $k \in N$, is defined as the value

$$h_k^{4*} = \begin{cases} h_k^{1*}/2, & h_k^{1*} \leq 0{,}99; \\ (h_k^{1*} + h_k^{2*})/2, & h_k^{1*} = 1, \end{cases} \qquad (8)$$

where $h_k^{1*}$ is the optimal solution of $P(x_i(k), y_j(k))$ with $Z = \{z(k)\}_N$; and $h_k^{2*}$ is the optimal solution of $P(\bar{x}_i(k), \bar{y}_j(k))$ with $\bar{Z} = \{\bar{z}(k)\}_N$. Therefore $h_k^{1*}$ and $h_k^{2*}$ are the values in (5) and (6), respectively.

Note that when all variables are relevant to the study, we have to $h_k^{1*} = h_k^{2*}$. Then the metrics M1 and M2 are equal, and therefore the focus E1 and E2 become only one. On the other hand, if all units are fully efficient ($h_k^{*} = 1$), we have $h_k^{4*} = h_k^{3*}$; the metrics M3 and M4 are the same, and then the approaches E3 and E4 are the same. Clearly, the latter case would lose interest, not analysis, and it would be necessary to establish other metrics that would allow us to differentiate and characterize the database more strictly.

## 2.2 Composite Metric Characterization

Once it has been decided which is the metric that best represents the data and allows a better analysis of results, a more complete metric is proposed that allows a characterization of each decision unit in the context studied (considering all areas of study).

Consider that all units k, from the previous section, present the following structure:

$k \equiv k(s, p, t) \in S \mathrm{x} P \mathrm{x} T$, a decision unit, identifying a sector $s \in S$, a country $p \in P$ and a year $t \in T$. Suppose that any unit $k \in N$ produces $y_j(k), j \in [J]$ outputs, using $x_i(k), i \in [I]$ inputs.

Next, we will define the M5 metric (Definition 2.2.1) with weights for each sector from the M4 metric.

***Definition 2.2.1:*** Let $Z = \{z(k)\}_N$ be a database and $\bar{Z} = \{\bar{z}(k)\}_N$ a subset $Z$ with $\bar{z}(k) = (\bar{x}_i(k), \bar{y}_j(k))$, the efficiency index of each DMU $k \in N$, is defined as the value

$$h_k^{5*} = a_1 h_{kS_1}^{4*} + a_2 h_{kS_2}^{4*} + a_3 h_{kS_3}^{4*} + \cdots + a_n h_{kS_4}^{4*} \qquad (9)$$

with

$$h_{kS_1}^{4*} = \begin{cases} h_{kS_i}^{1*}/2, & h_{kS_i}^{1*} \leq 0{,}99; \\ (h_{kS_i}^{1*} + h_{kS_i}^{2*})/2, & h_{kS_i}^{1*} = 1, \end{cases} \qquad (10)$$

where the values $a_i$ correspond to the weights that are attributed to each sector $Si$; $h_k^{1*}$ is the optimal solution of $P(xi(k), yj(k))$ with $Z = \{z(k)\}_N$ in $Si$; $h_k^{2*}$ is the optimal solution of $P(\bar{x}_i(k), \bar{y}_j(k))$ with $\bar{Z} = \{\bar{z}(k)\}_N$ in $Si$. Therefore, the indices $h_{kS_1}^{4*}$ correspond to the value $h_k^{4*}$ in (8). Considering the sector $Si$; $h_{kS_i}^{1*}$ and $h_{kS_i}^{2*}$ are the values in (5) and (6), respectively.

# 3 Numerical Application

To visualize the effects of the metrics studied, particularly in this work, we will examine the efficiency of European countries in production sectors compared to their waste management capacity. The metrics defined in the previous section are applied in three production sectors: (S1) glass production, (S2) paper production, and (S3) plastic production.

Twenty-six European countries are considered as decision-making units for the analysis: Austria (AT), Belgium (BE), Bulgaria (BG), Cyprus (CY), Czech Republic (CZ), Denmark (DK), Estonia (USA), Finland (FI), France (FR), Germany (DE), Greece (EL), Hungary (HU), Ireland (IE), Italy (IT), Latvia (LV), Luxembourg (LU), Malta (MT), Norway (NO), the Netherlands (NL), Poland (PL), Portugal (PT), Romania (RO), Slovakia (SK), Slovenia (SI), Spain (ES) and the United Kingdom (UK). Acronyms in brackets represent the nomenclature of each country, used throughout the article.

The analysis was done over a period of eleven years (2006–2016). Input and output variables are selected for waste generation, recovery, and recycling in order to examine the performance of countries in relation to the circular economy of the three production sectors (Table 1).

Table 1. Inputs/outputs variables

| Inputs | | Outputs | |
|---|---|---|---|
| I1 | Labor | O1 | Gross Domestic Product |
| I2 | Capital | O2 | CO2 Emission |
| I3 | Energy consumed | O3 | Waste |
| | | O4 | Recycling |
| | | O5 | Recovery |

If we consider the study in terms of a circular economy, CO2 emissions and waste, would be undesirable variables. Therefore, it becomes necessary to use complement variables instead. These variables are defined as the maximum value of the variable in a complete database minus the value of the variable for the unit under consideration, [23]. In order to obtain the efficiency indices, the other variables are used in their normal form and complement the CO2 and waste emissions.

The data is organized according to each approach. A software package developed by Python called pyDEA, [24], is used for data processing and the calculation of efficiency indices. Specifically, for the E2 approach, the database was reduced to labor and energy consumed inputs and to waste, recycling, and recovery outputs. Calculations are made for all countries and sectors throughout the study period. In Table 2, the results obtained in each metric for the years 2006 and 2016 in sector S1 are reflected.

We rank efficiency indices in four ranges:

$$R_1: 0{,}9 \leq h_k^{1*} \leq 1$$
$$R_2: 0{,}7 \leq h_k^{1*} < 0{,}9 \qquad (11)$$
$$R_3: 0{,}5 \leq h_k^{1*} < 0{,}7$$
$$R_4: 0{,}0 \leq h_k^{1*} < 0{,}5$$

The differences in the metrics applied are noticeable. In fact, in 2006, on metric M1, 61,5% of the countries studied had an efficiency index in the range $R_1$; 30,7% in $R_2$ and 7,6% in $R_3$. None of the countries obtained with this metric an index below 0.63. According to M2, the result was 42,3% in $R_1$, 15,3% in $R_2$, 3,8% in $R_3$ and 38,4% in $R_4$. The indices with this metric were the lowest of the four approaches, with a score of 0,18.

About the metric M3, 46,1% of countries have an index in $R_1$; 15,3% in $R_2$; 26,9% in $R_3$ and 11,5% in $R_4$. According to M4, 46,1% have in $R_1$; 7,6% in $R_2$; 3,8% in $R_3$ and 42,3% in $R_4$. The second-highest percentage occurred in the last rank.

Considering decision units with indexes in the ranges $R_1$or $R_2$, efficient, and index units in the ranges $R_3$ or $R_4$ as inefficient. With M1, 92,2% of the units in 2006 are efficient; with M2, 57,6%; with M3, 61,4% and with M4 only 53,7%. This leads to the conclusion that M4 metric is stricter in unit discrimination.

Table 2. Comparison of metrics in S1, 2006 and 2016

| | M1 | | M2 | | M3 | | M4 | |
|---|---|---|---|---|---|---|---|---|
| | 2006 | 2016 | 2006 | 2016 | 2006 | 2016 | 2006 | 2016 |
| AT | 1 | 0,948489 | 0,912849 | 0,791136 | 0,956425 | 0,869812 | 0,956425 | 0,474244 |
| BE | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| BG | 1 | 0,963084 | 0,308296 | 0,294299 | 0,654148 | 0,628691 | 0,654148 | 0,481542 |
| CY | 0,822949 | 1 | 0,451486 | 0,754327 | 0,637218 | 0,877163 | 0,411474 | 0,877163 |
| CZ | 0,803267 | 0,664682 | 0,416285 | 0,399762 | 0,609776 | 0,532222 | 0,401633 | 0,332341 |
| DK | 1 | 1 | 1 | 0,917238 | 1 | 0,958619 | 1 | 0,958619 |
| EE | 0,78036 | 1 | 0,504981 | 1 | 0,642671 | 1 | 0,39018 | 1 |
| FI | 0,887468 | 1 | 0,22232 | 0,591797 | 0,554894 | 0,795898 | 0,443734 | 0,795898 |
| FR | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| DE | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| EL | 0,809087 | 1 | 0,187884 | 0,223689 | 0,498486 | 0,611845 | 0,404544 | 0,611845 |
| HU | 0,75695 | 1 | 0,193342 | 0,422808 | 0,475146 | 0,711404 | 0,378475 | 0,711404 |
| IE | 1 | 1 | 0,926412 | 1 | 0,963206 | 1 | 0,963206 | 1 |
| IT | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| LV | 1 | 1 | 0,702587 | 0,816731 | 0,851294 | 0,908365 | 0,851294 | 0,908365 |
| LU | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| MT | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| NL | 0,97489 | 1 | 0,819753 | 0,684073 | 0,897322 | 0,842036 | 0,487445 | 0,842036 |
| NO | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| PL | 1 | 1 | 0,424622 | 0,638928 | 0,712311 | 0,819464 | 0,712311 | 0,819464 |
| PT | 1 | 1 | 1 | 0,904042 | 1 | 0,952021 | 1 | 0,952021 |
| RO | 0,821275 | 0,70385 | 0,25121 | 0,412551 | 0,536243 | 0,5582 | 0,410638 | 0,351925 |
| SK | 0,65961 | 0,728363 | 0,229803 | 0,356795 | 0,444706 | 0,542579 | 0,329805 | 0,364182 |
| SI | 0,635076 | 1 | 0,411571 | 1 | 0,523324 | 1 | 0,317538 | 1 |
| ES | 0,864859 | 0,875553 | 0,851896 | 0,80423 | 0,858377 | 0,839892 | 0,432429 | 0,437777 |
| UK | 1 | 1 | 0,855523 | 0,80975 | 0,927761 | 0,904875 | 0,927761 | 0,904875 |

## 3.1 Relative Position

Undoubtedly, another interesting aspect is to analyze how both change the position in an efficiency ranking according to the metric applied. These differences can be examined in Table 3 for 2006, sector S1.

BG is located in the 3rd position when the M1 metric is used, in the 21º position when the M2 metric is used, in the 17º position when the M3 metric is used, and in the 15º position when the M4 metric is used. LV is located in the 10th position when metric M1 is used, in position 15º when metric M2 is used, in position 15º when metric M3 is used, and in position 13º when metric M4 is used.

The case of PL is relevant to this analysis and clearly shows the objective of this study. PL is placed in 13th position when the M1 metric is used, 18º when the M2 metric is used, 16º when the M3 metric is used, and 14º when the M4 metric is used. It is therefore fully efficient in the M1 metric, inefficient in the M2 metric, and efficient in the M3 and M4 metrics.

Table 3. Countries' relative positions M1-M4, 2006

| No. | M1 | | M2 | | M3 | | M4 | |
|---|---|---|---|---|---|---|---|---|
| 1 | AT | 1 | BE | 1 | BE | 1 | BE | 1 |
| 2 | BE | 1 | DE | 1 | DE | 1 | DE | 1 |
| 3 | BG | 1 | DK | 1 | DK | 1 | DK | 1 |
| 4 | DE | 1 | FR | 1 | FR | 1 | FR | 1 |
| 5 | DK | 1 | IT | 1 | IT | 1 | IT | 1 |
| 6 | FR | 1 | LU | 1 | LU | 1 | LU | 1 |
| 7 | IE | 1 | MT | 1 | MT | 1 | MT | 1 |
| 8 | IT | 1 | NO | 1 | NO | 1 | NO | 1 |
| 9 | LU | 1 | PT | 1 | PT | 1 | PT | 1 |
| 10 | LV | 1 | IE | 0,926412 | IE | 0,963206 | IE | 0,963206 |
| 11 | MT | 1 | AT | 0,912849 | AT | 0,956425 | AT | 0,956425 |
| 12 | NO | 1 | UK | 0,855523 | UK | 0,927761 | UK | 0,927761 |
| 13 | PL | 1 | ES | 0,851896 | NL | 0,897322 | LV | 0,851294 |
| 14 | PT | 1 | NL | 0,819753 | ES | 0,858377 | PL | 0,712311 |
| 15 | UK | 1 | LV | 0,702587 | LV | 0,851294 | BG | 0,654148 |
| 16 | NL | 0,97489 | EE | 0,504981 | PL | 0,712311 | NL | 0,487445 |
| 17 | FI | 0,887468 | CY | 0,451486 | BG | 0,654148 | FI | 0,443734 |
| 18 | ES | 0,864859 | PL | 0,424622 | EE | 0,642671 | ES | 0,432429 |
| 19 | CY | 0,822949 | CZ | 0,416285 | CY | 0,637218 | CY | 0,411474 |
| 20 | RO | 0,821275 | SI | 0,411571 | CZ | 0,609776 | RO | 0,410638 |
| 21 | EL | 0,809087 | BG | 0,308206 | FI | 0,554894 | EL | 0,404544 |
| 22 | CZ | 0,803267 | RO | 0,25121 | RO | 0,536243 | CZ | 0,401633 |
| 23 | EE | 0,78036 | SK | 0,229808 | SI | 0,523324 | EE | 0,39018 |
| 24 | HU | 0,75695 | FI | 0,22232 | EL | 0,498486 | HU | 0,378475 |
| 25 | SK | 0,65961 | HU | 0,193342 | HU | 0,475146 | SK | 0,329805 |
| 26 | SI | 0,635076 | EL | 0,187884 | SK | 0,444706 | SI | 0,317538 |

On the other hand, it should be noted that the relative position of decision units does not necessarily provide all the information on the index. For example, the countries NL, FI, ES, CY, RO, EL, CZ, EE, HU, SK, and SI have the same positions (16 to 26, respectively) in M1 and M4 metrics. However, the scores are very different $0,63 < h_k^{1*} < 0,97$ for M1, and $0,31 < h_k^{3*} < 0,48$ for M4. Countries BE, DE, DK, FR, IT, LU, MT, NO, PT, IE, AT and UK, have the same positions (1-12) in the metrics M2, M3, and M4. However, scores are only different for M2 ($0,85 < h_k^{1*} < 1,00$). The indices for these 12 countries calculated with M3 ($0,92 < h_k^{3*} < 1,00$) and M4 ($0,92 < h_k^{4*} < 1,00$) are equal.

### 3.2 Evolution over Time

Figure 1, Figure 2 and Figure 3 show the efficiency results per year (2006-2016), according to each approach (E1-E4) and sectors S1-S3, respectively.

Each year is represented by a color, and each country corresponds to a column within the graph. The lines represent the variability of the results obtained for each country in each year. When the lines are very close in a certain country, it means that throughout the study period, their efficiency values were very similar, such as the cases of LV (Figure 1), PT (Figure 2), and IE (Figure 3). On the contrary, if the lines are very separated in a certain country, it means that there are great differences in efficiency results throughout the study period, such as BG, HU, and RO (Figure 1, Figure 2 and Figure 3). Looking at the results simultaneously, we can see that E4 is the approach with the greatest variability in the three sectors over the years. Therefore, the M4 metric allows for greater discrimination among decision units.

In Figure 4, Figure 5 and Figure 6, the country indices calculated in the M4 metric for sectors S1–S3 are represented, respectively. Each year is represented by a colored line, and each country corresponds to a section within the graph. Each of the 10 circles in the graph represents a value on the scale from 0 to 1. The closer each colored line is to the center, the lower the efficiency value obtained in that year. For example, if we analyze the LV and RO countries, we can see that the score changes from year to year in greater proportion in some sectors. However, we also find countries like LU where efficiency is always 1, independent of the approach.

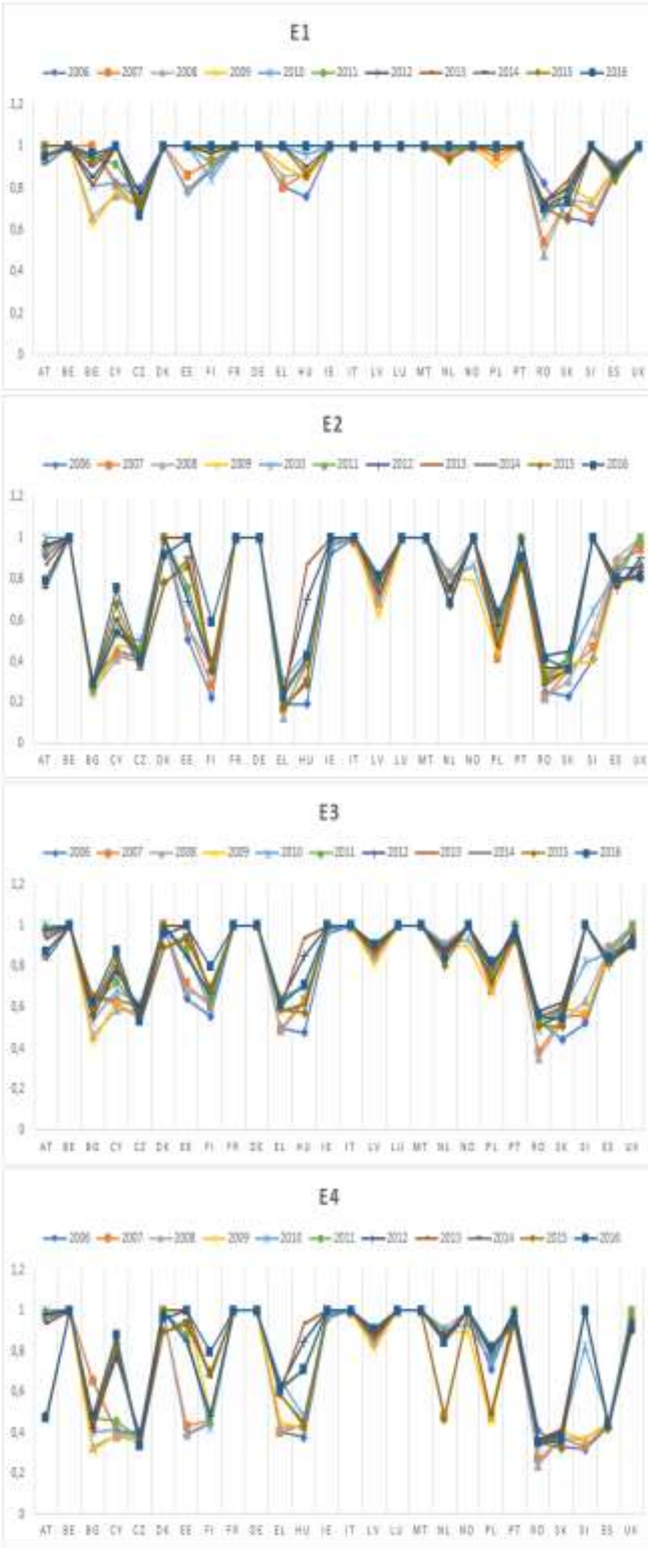


Fig. 1: Decision units in S1, E1-E4 approaches

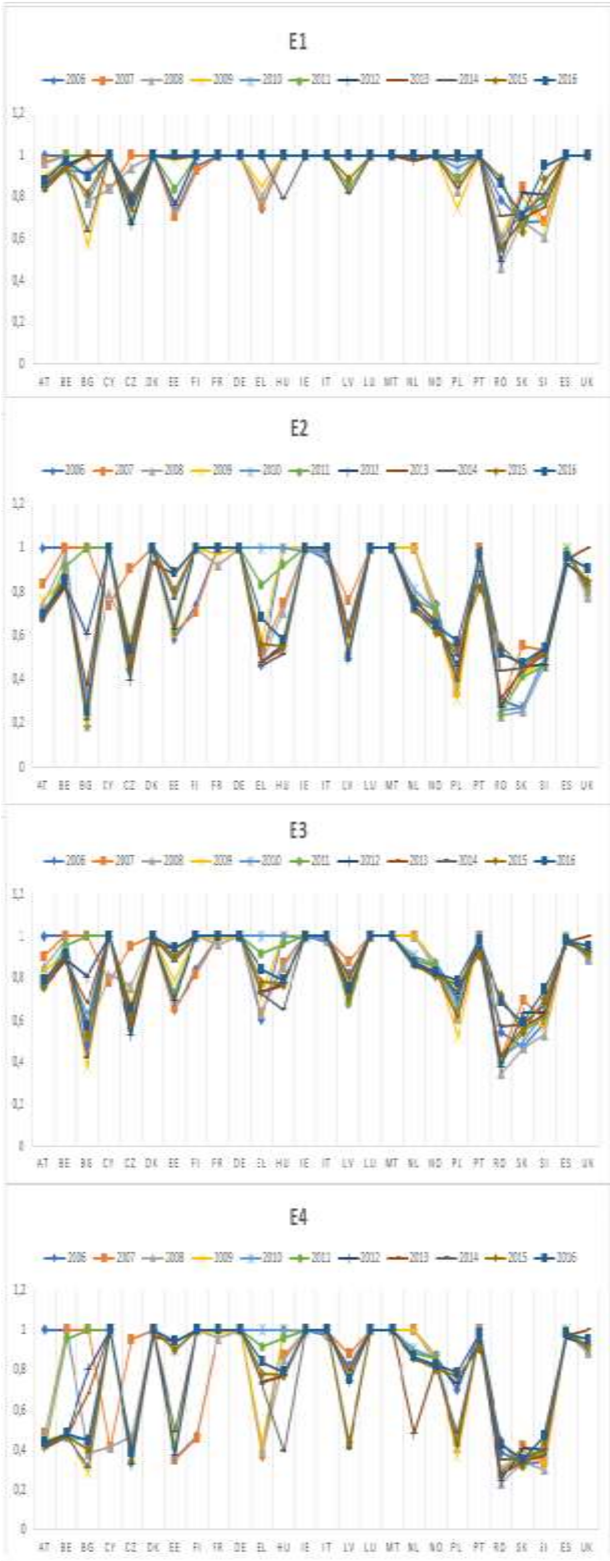


Fig. 2: Decision units in S2, E1-E4 approaches

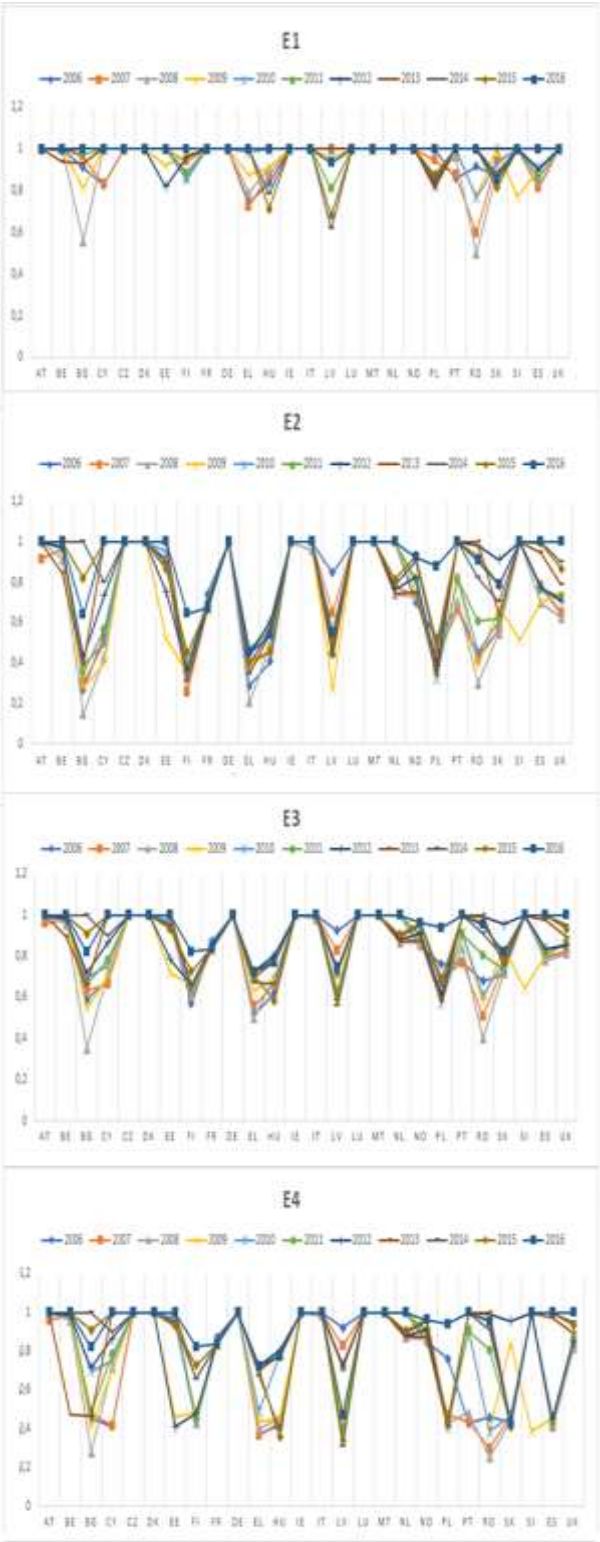

Fig. 3: Decision units in S3, E1-E4 approaches

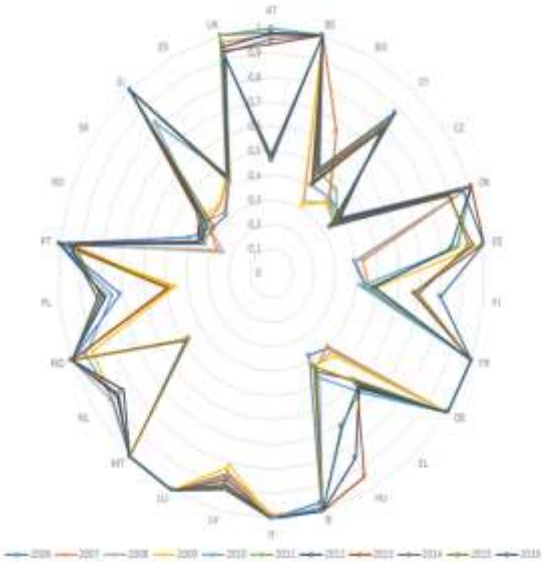
Fig. 4: Efficiency in S1, M4

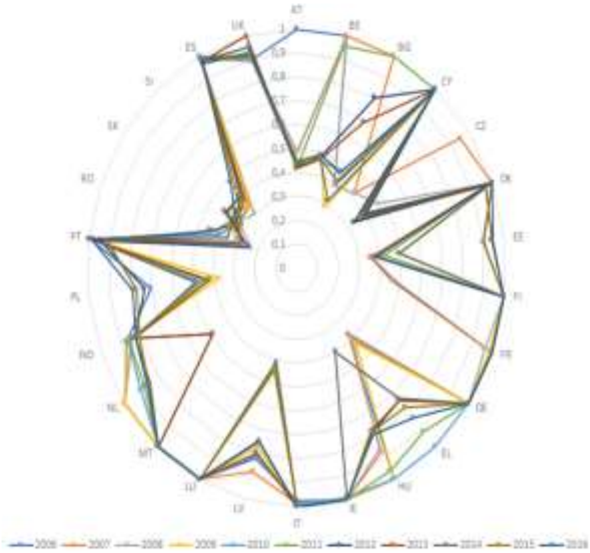
Fig. 5: Efficiency in S2, M4

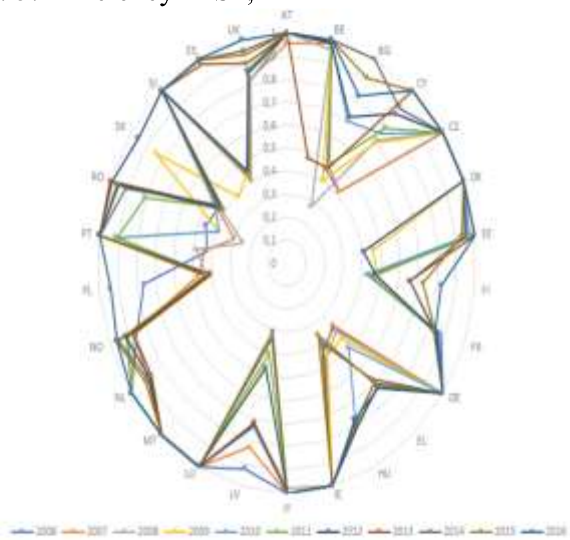
Fig. 6: Efficiency in S3, M4

Below, we introduce the calculations with the M5 metric. For this, it is necessary to define the values $ai$, corresponding to the weights, that are attributed to each sector $Si$. In this study, we will give the same importance in the calculation of efficiency to the three sectors (glass, plastic, and paper); however, the values $ai$ can be assigned differently.

Consider $a_1 = a_2 = a_3 = \frac{1}{3}$ . Then the index to be calculated is

$$h_k^* = \frac{1}{3} h_{kS_1}^{4*} + \frac{1}{3} h_{kS_2}^{4*} + \frac{1}{3} h_{kS_3}^{4*} \quad (12)$$

Table 4 shows the results obtained in (12) with the M5 metric. For this study, all units (26 countries) and all sectors (S1–S3) are considered.

Table 4. Global performance (%), M5 metric

| | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $R_1$ | 38,5 | 34,6 | 42,3 | 38,5 | 42,3 | 46,2 | 38,5 | 46,2 | 42,3 | 46,2 | 50,0 |
| $R_2$ | 19,2 | 26,9 | 11,5 | 11,5 | 30,8 | 23,1 | 26,9 | 26,9 | 23,1 | 19,2 | 30,8 |
| $R_3$ | 23,1 | 15,4 | 30,8 | 30,8 | 19,2 | 19,2 | 34,6 | 19,2 | 26,9 | 30,8 | 15,4 |
| $R_4$ | 19,2 | 23,1 | 15,4 | 19,2 | 7,7 | 11,5 | 0,0 | 7,7 | 7,7 | 3,8 | 3,8 |

Performance percentages by rank in each metric are shown in Table 5.

Table 5. Metrics by rank (%), year 2006

| | Efficients DMU | | Inefficients DMU | | | |
|---|---|---|---|---|---|---|
| Metrics | $R_1$ | $R_2$ | $R_3$ | $R_4$ | Min | Max |
| M1 | 61,5 | 30,7 | 7,6 | 0 | 0,63 | 1 |
| M2 | 42,3 | 15,3 | 3,8 | 38,4 | 0,18 | 1 |
| M3 | 46,1 | 15,3 | 26,9 | 11,5 | 0,44 | 1 |
| M4 | 46,1 | 7,6 | 3,8 | 42,3 | 0,31 | 1 |
| M5 | 38,5 | 19,2 | 23,1 | 19,2 | 0,36 | 1 |

The differences between the metrics are noticeable, and the importance of the M5 metric in relation to each of the previous ones is reflected. For example, with the M1 metric, 92,2% of the sample is efficient; with the M2 metric, 57,6%; with the M3 metric, 61,4%; with the M4 metric, only 53,7%; and with the M5 metric, 57,7%.

# 4 Conclusions

The study is based on the DEA-CCR model and shows the importance of selecting the most appropriate metric in each context for measuring efficiency.

In this work, different metrics are considered, determined by how efficiency indices are calculated. Specifically, five metrics are considered: M1, M2, M3, M4, and M5. The first four (M1–M4) are related to four approaches: E1, E2, E3, and E4, respectively. Next, it is proposed to build the fifth metric, M5, which it considers the strictest of the four previous metrics.

In numerical simulation, three sectors are considered: (S1) glass production, (S2) paper production, and (S3) plastic production. The study involved 26 European countries for eleven years (2006–2016). The variables were selected for waste generation, recovery, and recycling in the three sectors.

The differences between the five metrics and the three sectors are very noticeable. The lowest efficiency levels are in the M2 metric, followed by the M4 metric, then M3, and the highest levels in the M1 metric. The last metric defined in the M4 approach is the one with the most data variability, allowing for more detailed analysis. With this metric, it is possible to more strictly differentiate efficient units, establishing a ranking even when the units are all on the border.

The results show that a metric not analyzed in depth can give the wrong idea that the decision units studied are mostly (or entirely) efficient. It is important to consider that in addition to calculating the efficiency index, in this type of study, it is possible to make additional estimates and calculations to obtain the maximum amount of information from the database.

*References:*

[1] Charnes, A., Cooper, W. W., & Rhodes, E., Measuring the efficiency of decision-making units. Eur J Oper Res 2, 1978, pp. 429–444.

[2] Banker, R. D., Charnes, A., & Cooper, W., Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management Science, 30 (9), 1984, pp. 1078-1092.

[3] Bogetoft, P., & Hougaard, J. L., Efficiency Evaluations Based on Potential (Non-proportional) Improvements. Journal of Productivity Analysis, 12(3), 1999, pp. 233–47.

[4] Mancebon, M. J., & Muniz, M. A., Private versus public high schools in Spain: disentangling managerial and programme efficiencies, Journal of the Operational Research Society, 2007, pp. 1–10.

[5] Darji, I. S., & Dahiya, S., Financial performance of the textile industry in India: the case of Haryana. Research Journal of Textile and Apparel, ahead-of-print, 2021.

[6] Fukuyama, H., Mike Tsionas, M., Tan, Y., Dynamic network data envelopment analysis with a sequential structure and behavioural-causal analysis: Application to the Chinese banking industry, European Journal of Operational Research, Volume 307, 2023, pp. 1360-1373.

[7] Cunha, M., & Rocha, V., On the Efficiency of Public Higher Education Institutions in Portugal: An Exploratory Study, FEP working papers, No. 468, 2012.

[8] Ahn, T., Charnes, A., & Cooper, W., Some statistical and DEA evaluations of relative efficiencies of public and private institutions of higher education, Socio-Economic Planning Sci, 22, 1998, pp. 259–269.

[9] Bradley, S., Johnes G., & Millington, J., The effect of competition on the efficiency of secondary schools in England, European Journal Operational Research, 135, 2001, pp. 545–568.

[10] Mizala, A., Romaguera, P., & Farren, D., The technical efficiency of schools in Chile, Applied Economics, 34, 2002, pp. 1533–1552.

[11] Barbetta, G. P., & Turati, G. (2003). Efficiency of junior high schools and the role of proprietary structure, Annals of Public and Cooperative Economics 74, 2003, pp. 529–551.

[12] Asmild, M., Holvad, T., Hougaard, J., & Kronborg, D., Railway reforms: do they influence operating efficiency? Transportation, 36(5), 2009, pp. 617-638.

[13] Gongbing, B., Pingchun, W., Feng, Y., & Liang, L., Energy and Environmental Efficiency of China's Transportation Sector: A Multidirectional Analysis Approach. Mathematical Problems in Engineering, 2014, pp. 1-12.

[14] Wang, K., Yu, S., Li, M. J., & Wei, Y., Multi-directional efficiency analysis-based regional industrial environmental performance evaluation of China. Natural Hazards, 2015.

[15] Robaina, M., Murillo, K., Rocha, E., & Villar, J., Circular economy in plastic waste - Efficiency analysis of European countries, Science of the Total Environment, Volume 730, 2020.

[16] Hougaard JL, Kronborg D, Overgard, C. Improvement Potential in Danish Elderly Care. Health Care Management Science, 7(3), 2004, pp. 225–235.

[17] Asmild M, Balezentis T, Hougaard J., Multi-directional Productivity Change: MEA Malmquist, Journal of Productivity Analysis, 46, 2016, pp.109–119.

[18] Murillo K., & Rocha E., Sustainable development in education: a non-parametric analysis. INTED2021 Proceedings, 2021, pp. 4536-4544.

[19] Murillo, K. P., Rocha, E. M., & Ramalho, J. S., About the Efficiency Behavior of the Portuguese Manufacturing Firms During the Financial Crisis. Libertas Mathematica (new series), 38 (1), 2018, pp. 1–27.

[20] Pearson K. On Lines and Planes of Closest Fit to Systems of Points in Space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2(11), 1901, pp. 559–72.

[21] Abdi, H., & Williams, L. J., Component Analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 2 (4), 2010, pp. 433–459.

[22] Dray, S., On the number of principal components: a test of dimensionality based on measurements of similarity between matrices. Computational Statistics & Data Analysis, 52 (4), 2008, pp. 2228–2237.

[23] Jahanshahloo, G.R., Lotfi, F. Hosseinzadeh, Shoja, N., Tohidi, G., Razavyan, S., Undesirable inputs and outputs in DEA models. Appl. Math. Comput. 169 (2), 2005, pp. 917–925.

[24] Raith, A., Fauzi, F., & Perederieieva, O., PyDEA documentation, 2016.

**Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)**

Kelly Patricia Murillo has written, reviewed, and actively participated in all the publication stages of this manuscript.

**Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself**

This work is supported by the Center for Research and Development in Mathematics and Applications (CIDMA) through the Portuguese Foundation for Science and Technology (FCT - Fundação para a Ciência e a Tecnologia), UIDB/04106/2020, UIDP/04106/2020, and by national funds (OE), through FCT, I.P., in the scope of the framework contract foreseen in the numbers 4, 5 and 6 of the article 23, of the Decree-Law 57/2016, of August 29, changed by Law 57/2017, of July 19.

**Conflict of Interest**

The author declares no conflict of interest.