Potential-Based Efficiency Assessment and Target Setting

Narges Soltani[a] and Sebastián Lozano[b§]

[a] Department of Mathematics, Kharazmi University, Tehran, Iran

[b] Dept. of Industrial Management, University of Seville, Spain

§   Corresponding author:
Escuela Superior de Ingenieros
Camino de los Descubrimientos, s/n, 41092 Sevilla, Spain
(Phone: +34-954487208   E-mail: slozano@us.es)

Potential-Based Efficiency Assessment and Target Setting

Abstract

In this paper, a new approach to the problem of efficiency assessment is presented, one that comes from applying a Physics analogy so that each feasible operating point is assigned an Efficiency Potential, i.e. a real value that decreases if the inputs are reduced and/or the outputs are increased. The gradient of that efficiency potential at any operating point points to the less input, more output region. For each Decision Making Unit (DMU) the minimum efficiency potential in that region is computed. The proposed Potential-Based Measure of efficiency (PBM) is computed as the ratio of that minimum efficiency potential and that of the DMU. A minimum efficiency potential target setting model is also formulated to identify a closest minimum efficiency potential target. The proposed approach can take into account undesirable outputs, nondiscretionary variables and preference structures. The method has been applied to a container shipping lines dataset.

Keywords: DEA; efficiency potential; potential-based efficiency measure; closest target; container shipping lines

1. Introduction

Data Envelopment Analysis (DEA) is a well-known non-parametric method for assessing the relative efficiency of a set of homogeneous organizational units (usually termed Decision Making Units, DMUs). These DMUs are assumed to consume different inputs and produce different outputs. DEA only uses the information about the amount of each input consumed by each DMU and the amount of each output produced by each DMU. With the observed data and some standard assumptions (like envelopment of observed data, free disposability and convexity) an appropriate Production Possibility Set (PPS) is inferred (for more details see Cooper et al. 2006). This PPS contains all the operating points that are deemed feasible. The non-dominated subset within the PPS is called its efficient frontier, also called strong-efficient frontier.

For each DMU, an appropriate DEA model is used to project it onto the efficient frontier and compute an efficiency score. There are different DEA models leading to different efficiency/inefficiency measures. Table 1 lists some commonly used efficiency measures,

together with their main features. The list includes the oriented radial efficiency measures computed by the CCR (Charnes et al. 1978) and BCC (Banker et al. 1984) DEA models, the oriented non-radial Russell (NRRM, Färe and Lovell 1978), the hyperbolic efficiency measure (Färe et al. 1985), the Russell graph measure of efficiency (RGM, Färe et al. 1985), the enhanced Russell graph efficiency measure (ERGM, Pastor et al. 1999) and its equivalent slacks-based measure of efficiency (SBM, Tone 2001), the directional distance function (Chambers et al. 1996), the measure of efficiency proportions (MEP, Cooper and Tone 1997) and its equivalent measure of efficiency dominance (MED, Bardhan et al. 1996), the range adjusted measure of efficiency (RAM, Cooper et al. 1999), the geometric distance function (GDF, Portela and Thanassoulis 2007), slacks-based inefficiency measure (SBI, Fukuyama and Weber 2009), the range directional measure (RDM) and the slack-free multidirectional efficiency analysis measure of efficiency (MEA) (Asmild and Pastor 2010) and the epsilon-based measure of efficiency (EBM, Tone 2010).

As can be seen in Table 1, most efficiency measures are non-radial and result from linear optimization models. Practically all the efficiency measures considered are units invariant although only a few are also translation invariant. Most of these efficiency measures are comprehensive except the radial CCR and BCC, the oriented NRRM, the hyperbolic efficiency and the DDF, all of which do not exhaust all input and output slacks. Not all the efficiency measures are reference-set independent, usually because they involve statistics covering the whole dataset (like the input and output ranges or the ideal point, for example). In the case of DDF and SBI, this property, as well as others, depends on the specific direction vector considered (see Wang et al. 2017 for a review of the different methods of selecting the direction vector in DDF). Almost all efficiency measures are normalized scores but only a few, namely RGM, ERGM/SBM and GDF, are strongly monotonic in inputs and outputs.

The Potential-Based Measure of efficiency (PBM) proposed in this paper is based on the efficiency potential, which translates the Potential Field concept, common in Physics, to the PPS context. This analogy allows defining Efficiency Filed Vectors, Efficiency Field Lines, isopotential surfaces, etc. (see Lozano and Calzada-Infante, 2018) and, in particular, allows computing an efficiency score for each point in the PPS. Thus, the gradient of the efficiency potential at a given DMUo points to the less input, more output region. All the points in that region have a lower efficiency potential than DMUo. The larger the input reduction and the output increase, the larger the reduction in the efficiency potential achieved. It seems natural to compute the minimum efficiency potential in that region and to measure the efficiency as the corresponding relative reduction of the efficiency potential from DMUo. Such a PBM efficiency score is non-radial, non-linear, units invariant, comprehensive, reference-set dependent, normalized and strongly monotonic in inputs and outputs. Also, since the minimum efficiency potential (MEffP) optimization model may have alternative optima, an approach similar to Zofio et al. (2013) to select the closest (using weighted Euclidean distance) MEffP operation point is proposed for target setting.

Interestingly, although having its origin in a Physics analogy, the proposed potential-based measure of efficiency has a functional form that relates it to two existing DEA efficiency measures, namely the enhanced Russell graph measure of efficiency (ERGM) (a.k.a. slacks-based measure of efficiency, SBM) and the geometric distance function (GDF). Thus, while ERGM (and equivalently SBM) maximizes the ratio of the average relative input reduction to the average relative output increase and GDF maximizes the ratio of the geometric mean of the relative input reductions to the geometric mean of the relative output increases, PBM maximizes the ratio of the product of the relative input reductions to the product of the relative output increases. In all cases, larger relative input reductions and relative output increases (i.e. larger margin for input and output improvement) lead to lower efficiency scores.

The structure of the paper is the following. Section 2 introduces the efficiency potential and presents the proposed PBM efficiency score, illustrating it with a simple example. Section 3 extends the PBM approach to situations in which non-discretionary inputs or outputs,

undesirable outputs or a preference structure is present. Section 4 presents the target setting model for selecting the closest MEffP operating point. Section 5 applies the proposed approach to assess the efficiency of 18 container shipping lines. Finally, Section 6 summarizes and concludes.

2. Proposed PBM efficiency score

The efficiency score proposed in this paper is based on the concept of efficiency potential, which is a novel concept inspired by the field potential concept commonly used in Physics. Same as an electric field potential can be assigned to each point of the Euclidean space, in this analogy an efficiency potential is assigned to each point of the Production Possibility Set. This allows defining equi-potential surfaces and efficiency field vectors perpendicular to the latter. Since the efficiency potential is monotonic in inputs and outputs, i.e. it increases if any input is increased or any output is decreased, all the operating points in the region that dominates a given DMU have less efficiency potential than that DMU. Among these, the operating points with the minimum efficiency potential are guaranteed to be efficient and therefore potential targets for the DMU being assessed. It seems natural also, and that is the potential-based efficiency score, to measure the efficiency of the DMU in terms of the corresponding reduction in the efficiency potential. This is because if the inputs of the DMU can be significantly reduced or its outputs significantly increased then the minimum efficiency potential would be much lower than that of the DMU and its potential-based efficiency score will be low. And the opposite, if the inputs of the DMU can only be slightly reduced and its outputs only slightly increased then the reduction of the efficiency potential will also be small and hence its potential-based efficiency score will be high.

In order to make the efficiency potential dimensionless we have normalized the input and output variables. This is equivalent to defining the efficiency potential on the Production Possibility Set defined with these normalized inputs and outputs. Although other normalization constants could have been used, we used the average inputs and outputs, which, according to the definition of the efficiency potential given by equation (4), corresponds to assigning an efficiency potential of one to the average DMU. This seems a sensible choice.

2.1. Efficiency potential function

In this section, we introduce the efficiency potential function. Assume we have data about $\mathrm{DMU}_j (j \in \{1,...,n\})$ that produces output vector $y_j = (y_{kj})_{k=1}^{s}$ consuming input vector $x_j = (x_{ij})_{i=1}^{m}$. We assume that $\hat{x}_{ij} > 0$, $\hat{y}_{kj} > 0$ for all i, j and k. This assumption guarantees that the efficiency potential function, defined below, is always a positive number. The Production Possibility Set (PPS) constructed based on envelopment, convexity and Variable Returns to Scale (VRS) assumptions is

$$T^{VRS} = \{(x,y) \in R^{m+s} \mid x \geq X\lambda, y \leq Y\lambda, e^{\mathrm{t}}\lambda = 1, \lambda \geq 0\} \tag{1}$$

where e refers to unit vector. Note that, although in this paper we assume VRS, the proposed approach can be applied, with the corresponding changes, to any other returns to scale assumption. In particular, for Constant Returns to Scale (CRS) it is sufficient to drop the convexity constraint $e^{t}\lambda = 1$. We assume VRS so that the extension of the proposed approach to undesirable outputs (presented in section 3.2) considers that case, which is more difficult to model than the CRS case.

Using the average input and output values x_i^{aver} and y_k^{aver} we may define the dimensionless variables

$$\hat{x}_{ij} = \frac{x_{ij}}{x_i^{\mathrm{aver}}} \quad \forall i \forall j \qquad \hat{y}_{kj} = \frac{y_{kj}}{y_k^{\mathrm{aver}}} \quad \forall k \forall j \tag{2}$$

which leads to the following dimensionless PPS

$$\hat{T}^{VRS} = \{(\hat{x},\hat{y}) \mid \hat{x} \geq \hat{X}\lambda, \hat{y} \leq \hat{Y}\lambda, e^{t}\lambda = 1, \lambda \geq 0\} \tag{3}$$

in which $\hat{x} = (\hat{x}_i)_{i=1}^m$ and $\hat{y} = (\hat{y}_k)_{k=1}^s$. Note that there is a one-to-one correspondence between T^{VRS} and $\hat{T}^{VRS}$. The following potential function is defined on every point of the PPS $(x, y) \in T^{VRS} \cap R_{++}^{m+s}$ [1] as follows

$$P(x, y) = \left(\frac{\displaystyle\prod_{i=1}^m \frac{x_i}{x_i^{aver}}}{\displaystyle\prod_{k=1}^s \frac{y_k}{y_k^{aver}}} \right)^{1/(m+s)} = \left(\frac{\displaystyle\prod_{i=1}^m \hat{x}_i}{\displaystyle\prod_{k=1}^s \hat{y}_k} \right)^{1/(m+s)} \tag{4}$$

Note that this function is units invariant. It assigns to each feasible operating point a value so that: if any input is decreased, the efficiency potential decreases and if any output is increased, the efficiency potential decreases. Although this is not essential, note that the specific normalization used in (4) implies that the efficiency potential of the average DMU is 1.

As regards taking the $m+s$ root in the efficiency potential definition (4), the idea is to compute the joint geometric mean of the normalized inputs and outputs. We use a joint geometric mean for symmetry. Otherwise, as it happens with GDF, which computes geometric means for the inputs and outputs separately (see Section 2.3), the treatment of the inputs and output variables is not symmetric as it depends on the number of each of these types of variables.

2.2. MEffP DEA model and PBM efficiency

Let DMUo be the DMU whose efficiency is to be assessed. Its corresponding efficiency potential is $P_o = P(x_o, y_o)$. Let us consider the region of the PPS that dominates DMUo. We can compute the minimum efficiency potential (MEffP) within that region solving

---

[1] $R_{++}^{m+s} = \{(x, y) \in R^{m+s} : (x, y) > 0\}$

MEffP DEA model

$$\left(P_o^*\right)^{m+s} = min \ \frac{\displaystyle\prod_{i=1}^{m} \hat{x}_i}{\displaystyle\prod_{k=1}^{s} \hat{y}_k}$$

$$s.t. \ \sum_{j=1}^{n} \lambda_j \hat{x}_{ij} \leq \hat{x}_i \leq \hat{x}_{io} \qquad i = 1,...,m,$$

$$\sum_{j=1}^{n} \lambda_j \hat{y}_{kj} \geq \hat{y}_k \geq \hat{y}_{ko} \qquad k = 1,...,s,$$

$$\sum_{j=1}^{n} \lambda_j = 1,$$

$$\lambda_j \geq 0 \ \ \forall j$$

(5)

Note that the objective function of model (5) is not the efficiency potential but its m+s power. Since the power function is monotonic it is equivalent to minimize the efficiency potential or the efficiency potential raised to its m+s power. Even with this simplification, the above model is still a linearly constrained nonlinear problem that can be solved using common global optimization packages like LINGO or GAMS. In particular, for the case study presented in Section 5, the solver that we chose for GAMS to solve this model was "Convex Over and Under ENvelopes for Nonlinear Estimation (Couenne)" which is an open source library for solving global optimization problems.

Let $(\hat{x}^*, \hat{y}^*)$ be an optimal solution of model (5), which corresponds to the operating point $x_i^* = \hat{x}_i^* \cdot x_i^{\text{aver}} \ \forall i$, $y_k^* = \hat{y}_k^* \cdot y_k^{\text{aver}} \ \forall k$. The MEffP among all points that dominate DMUo is $P_o^* = P(x^*, y^*)$. The proposed Potential-Based Measure of efficiency (PBM) of DMUo is computed as:

$$\pi_o = \frac{P_o^*}{P_o} = \frac{P(x^*, y^*)}{P(x_o, y_o)} = \frac{\left(\dfrac{\prod_{i=1}^{m} \dfrac{x_i^*}{x_i^{\text{aver}}}}{\prod_{k=1}^{s} \dfrac{y_k^*}{y_k^{\text{aver}}}}\right)^{1/(m+s)}}{\left(\dfrac{\prod_{i=1}^{m} \dfrac{x_{io}}{x_i^{\text{aver}}}}{\prod_{k=1}^{s} \dfrac{y_{ko}}{y_k^{\text{aver}}}}\right)^{1/(m+s)}} = \frac{\left(\dfrac{\prod_{i=1}^{m} \hat{x}_i^*}{\prod_{k=1}^{s} \hat{y}_k^*}\right)^{1/(m+s)}}{\left(\dfrac{\prod_{i=1}^{m} \hat{x}_{io}}{\prod_{k=1}^{s} \hat{y}_{ko}}\right)^{1/(m+s)}} \tag{6}$$

2.3 Alternative PBM formulation and relation with GDF and ERGM

Instead of computing the PBM score in two steps, first solving (4) and then applying (6), we can formulate the computation of π_o directly. Since $\hat{x}_{io} > 0 \ \forall i$, $\hat{y}_{ko} > 0 \ \forall k$ we can define variables $\theta_i \leq 1$, $\varphi_k \geq 1$ so that $\hat{x}_i = \theta_i \cdot \hat{x}_{io} \ \forall i$, $\hat{y}_k = \varphi_k \cdot \hat{y}_{ko} \ \forall k$. Thus

$$\frac{\prod_{i=1}^{m} \hat{x}_i}{\prod_{k=1}^{s} \hat{y}_k} = \frac{\prod_{i=1}^{m} \theta_i \cdot \hat{x}_{io}}{\prod_{k=1}^{s} \varphi_k \cdot \hat{y}_{ko}} = \frac{\prod_{i=1}^{m} \theta_i}{\prod_{k=1}^{s} \varphi_k} \cdot \frac{\prod_{i=1}^{m} \hat{x}_{io}}{\prod_{k=1}^{s} \hat{y}_{ko}} \tag{7}$$

Hence, the reduction in the efficiency potential of point (x, y) with respect to (x_o, y_o) is

$$\frac{P(x, y)}{P(x_o, y_o)} = \frac{\left(\dfrac{\prod_{i=1}^{m} \hat{x}_i}{\prod_{k=1}^{s} \hat{y}_k}\right)^{1/(m+s)}}{\left(\dfrac{\prod_{i=1}^{m} \hat{x}_{io}}{\prod_{k=1}^{s} \hat{y}_{ko}}\right)^{1/(m+s)}} = \frac{\left(\dfrac{\prod_{i=1}^{m} \theta_i}{\prod_{k=1}^{s} \varphi_k} \cdot \dfrac{\prod_{i=1}^{m} \hat{x}_{io}}{\prod_{k=1}^{s} \hat{y}_{ko}}\right)^{1/(m+s)}}{\left(\dfrac{\prod_{i=1}^{m} \hat{x}_{io}}{\prod_{k=1}^{s} \hat{y}_{ko}}\right)^{1/(m+s)}} = \left(\frac{\prod_{i=1}^{m} \theta_i}{\prod_{k=1}^{s} \varphi_k}\right)^{1/(m+s)} \tag{8}$$

Therefore, since the PBM efficiency score π_o corresponds to minimizing the ratio (8), we can solve model (9) instead of model (5) to obtain the PBM efficiency score directly. This requires the following change of variables $\hat{x}_i = \theta_i \cdot \hat{x}_{io} \ \forall i, \hat{y}_k = \varphi_k \cdot \hat{y}_{ko} \ \forall k$ (so that $\theta_i \leq 1, \varphi_k \geq 1$) as well as multiplying the input and output constraints by x_i^{aver} and y_k^{aver} respectively.

PBM DEA model

$$(\pi_o)^{m+s} = min \ \frac{\displaystyle\prod_{i=1}^{m} \theta_i}{\displaystyle\prod_{k=1}^{s} \varphi_k}$$

$$s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta_i \cdot x_{io} \qquad i = 1,...,m,$$

$$\sum_{j=1}^{n} \lambda_j y_{kj} \geq \varphi_k \cdot y_{ko} \qquad k = 1,...,s,$$

$$\sum_{j=1}^{n} \lambda_j = 1,$$

$$\lambda_j \geq 0 \ \forall j \quad \theta_i \leq 1 \ \forall i \quad \varphi_k \geq 1 \ \forall k$$

(9)

Same as (5) the above formulation has a non-linear objective function. The advantages of using it are that it produces the efficiency score directly, it shows that the PBM efficiency score is independent of the normalization coefficients used in the efficiency potential (4) and it facilitates the comparison with other DEA models like GDF and ERGM (see below). Also, upon the request of one of the reviewers, in the Appendix we derive the dual of model (9).

Note that the proposed PBM efficiency score has the following properties (see Table 1 for comparison with other efficiency measures):

✓ It is a normalized score, i.e. $0 \leq \pi_o \leq 1$. Thus, since the constraints of model (5) impose $0 \leq \hat{x}_i \leq \hat{x}_{io} \ \forall i, \hat{y}_k \geq \hat{y}_{ko} \geq 0 \ \forall k$, it follows that $0 \leq P(x^*, y^*) \leq P(x_o, y_o)$ and therefore $0 \leq \pi_o \leq 1$. Analogously, the constraints in model (9) guarantee that

$0 \le \theta_i \le 1 \ \forall i, \ \varphi_k \ge 1 \ \forall k$ and therefore the corresponding optimal objective function value $0 \le (\pi_o)^{m+s} \le 1$.

- ✓ $\pi_o = 1$ if and only if (x_o, y_o) is technically efficient. By definition, DMU$_O$ is efficient if and only if $\pi_o = 1$, which is equivalent to $\theta_i^* = 1 \ \forall i, \varphi_k^* = 1 \ \forall k$ in model (9), i.e. no input or output can improve, and hence DMU$_O$ is technical efficient.

- ✓ It is a comprehensive efficiency measure. Thus, the target computed by model (5) or model (9) is efficient and cannot improve any of its inputs or outputs any further. Otherwise, its efficiency potential could be reduced and hence it would not be optimal.

- ✓ It is units invariant. Thus, equation (6) shows that the efficiency potential depends only on the normalized input and output values, which do not depend on their units of measurement.

- ✓ It is strongly monotonic. A reduction in any input reduces the efficiency potential of the DMU and therefore increases its PBM efficiency score. Similarly, increasing any output reduces its efficiency potential and therefore increases its PBM efficiency score. Therefore, π_o is strictly increasing in output quantities and decreasing in input quantities. This can also be seen in model (9), whose objective function is strictly increasing in θ_i and strictly decreasing in φ_k.

- ✓ It is reference-set dependent. An efficiency measure is reference-set dependent if it is determined only by the reference set of the DMU concerned (Tone 2011). Thus, the PBM efficiency score of a DMU only depends on its efficiency potential and that of its corresponding minimum efficiency potential target. Although model (5) is sensitive to the normalization constants used to compute the dimensionless inputs and outputs, the PBM efficiency score (6) is not, as model (9) shows. In other words, although the efficiency potential uses the average input and output values, the PBM efficiency score does not depend on the value of those normalization constants.

Note that the PBM DEA model above has a resemblance to the GDF DEA model (Portela and Thanassoulis 2007), which, for comparison, is formulated below.

GDF DEA model

$$\text{GDF}(x_o, y_o) = Min \frac{\left(\prod_{i=1}^{m} \theta_i\right)^{1/m}}{\left(\prod_{k=1}^{s} \varphi_k\right)^{1/s}}$$

$$s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta_i \cdot x_{io} \qquad i = 1,...,m, \qquad (10)$$

$$\sum_{j=1}^{n} \lambda_j y_{kj} \geq \varphi_k \cdot y_{ko} \qquad k = 1,...,s,$$

$$\sum_{j=1}^{n} \lambda_j = 1,$$

$$\lambda_j \geq 0 \ \forall j \quad \theta_i \leq 1 \ \forall i \quad \varphi_k \geq 1 \ \forall k$$

It is easy to check that $\text{GDF}(x_o, y_o) \leq \pi_o$. That is so because models (9) and (10) have the same feasibility region and therefore the optimal solution to model (9) is feasible in model (10). Denoting θ_i^* and φ_k^* the optimal values of those variables in model (9) and taking into account

that $\theta_i^* \leq 1 \ \forall i$ and $\varphi_k^* \geq 1 \ \forall k$, it follows that $\text{GDF}(x_o, y_o) \leq \dfrac{\left(\prod_{i=1}^{m} \theta_i^*\right)^{1/m}}{\left(\prod_{k=1}^{s} \varphi_k^*\right)^{1/s}} \leq \left(\dfrac{\prod_{i=1}^{m} \theta_i^*}{\prod_{k=1}^{s} \varphi_k^*}\right)^{1/(m+s)} = \pi_o.$

It is also easy to check that $\pi_o = 1 \Leftrightarrow \text{GDF}(x_o, y_o) = 1$. Note that Portela and Thanassoulis (2007) did not propose solving the above GDF model to compute the corresponding efficiency measure. Instead they used the GDF efficiency score to decompose profit efficiency.

The PBM DEA model also resembles, although somewhat less, the ERGM DEA model (Pastor et al. 1999), which, for comparison, is also formulated below. Note that although, in principle, ERGM is a non-linear optimization model, unlike the PBM and GDF models, it can be linearized (Pastor et al. 1999).

ERGM DEA model

$$\text{ERGM}(x_o, y_o) = \min \quad \frac{\dfrac{1}{m}\displaystyle\sum_{i=1}^{m}\theta_i}{\dfrac{1}{s}\displaystyle\sum_{k=1}^{s}\varphi_k}$$

$$s.t. \quad \sum_{j=1}^{n}\lambda_j x_{ij} \leq \theta_i \cdot x_{io} \qquad i = 1,...,m,$$

$$\sum_{j=1}^{n}\lambda_j y_{kj} \geq \varphi_k \cdot y_{ko} \qquad k = 1,...,s,$$

$$\sum_{j=1}^{n}\lambda_j = 1,$$

$$\lambda_j \geq 0 \;\; \forall j \quad \theta_i \leq 1 \;\; \forall i \quad \varphi_k \geq 1 \;\; \forall k$$

(11)

2.4. Illustrative examples

Consider the 1YY (i.e. two outputs and a single, constant input) dataset shown in Table 2. The table also shows the input and output vector of the average DMU and the efficiency potentials of all the DMUs. The isopotential curves $P(x, y) = \alpha$ for the 1YY case are $y_1 y_2 = \dfrac{y_1^{aver} \cdot y_2^{aver}}{\alpha^3}$.

These are hyperbolas whose efficiency potential decreases as they move in the NE direction. Figure 1 shows two of those curves, one passing through DMU C and the other through DMU B.

=========================== Table 2 ===========================

=========================== Figure 1 ===========================

As can be seen, the only inefficient DMU is C and its efficiency potential is $P(x_C, y_C) = 1.298$. The MEffP computed by model (5) is $P(x^*, y^*) = 0.900$, which corresponds to the efficiency potential of DMU B. Therefore, its PBM efficiency score is $\pi_C = \dfrac{P(x^*, y^*)}{P(x_C, y_C)} = \dfrac{0.900}{1.298} = 0.693$.

Figure 1 helps to illustrate the proposed approach. Thus, it can be seen that the isopotential curves correspond to lower values of efficiency potential as they move north-east and that all the points in the region that dominates DMU C have lower efficiency potential. Actually, the more to the right and up an operating point is, the lower its efficiency potential. The MEffP corresponds to the isopotential curve that is tangent to the PPS, which occurs for DMU B. The output improvements from DMU C to DMU B lead, therefore, to the maximum reduction of the efficiency potential. And the PBM efficiency score, which reflects that reduction, measures the extent of those output improvements.

Consider now the XY (i.e. single input, single output) dataset shown in Table 3 and plotted in Figure 2. The table also shows the input and output vector of the average DMU and the efficiency potentials of all the DMUs. For the XY case, the isopotential curves $P(x, y) = \alpha$ are straight lines $y = \dfrac{y^{aver} / x^{aver}}{\alpha^2} \cdot x$. These correspond to rays starting at the origin and whose slope increases as the efficiency potential decreases. Figure 2 shows two of those curves, one passing through DMU C and the other passing through DMUs A and B. The efficiency potential of DMU C is $P(x_C, y_C) = 1.515$. The MEffP computed by model (4) is $P(x^*, y^*) = 0.787$, which corresponds to the efficiency potential of both DMU A and DMU B. Note that, on purpose, we have considered the case in which there are multiple MEffP operating points. In fact, all the points in the segment $C'B$ are MEffP and any of them can be the projection computed by models (5) or (9). The existence of alternative optima is dealt with in Section 4 and does not affect the PBM efficiency score, which is well defined. In particular,

$$\pi_C = \frac{P(x^*, y^*)}{P(x_C, y_C)} = \frac{0.787}{1.515} = 0.519 .$$

========================== Table 3 ==========================

========================== Figure 2 ==========================

In Figure 2, similar comments as in Figure 1 can be made. In this case, the isopotential curves, which are now straight lines passing through the origin, correspond to lower values of efficiency potential as they move north-west (i.e. in the direction of less input and more output). It is

apparent that the points in the region that dominates DMU C have lower efficiency potential than DMU C. As before, the MEffP corresponds to the isopotential curve that is tangent to the PPS, which occurs for the facet defined by DMUs A and B, which defines the Most Productive Scale Size section of the efficient frontier (Banker 1984). As indicated above, any of the operating points in the segment C'B can be chosen as all of them are MEffP. It may make sense, from a managerial point of view, to choose that point that is closest to the DMU being projected as this will minimize the effort required to attain efficiency. That will be dealt with in Section 4.

3. Extensions of the proposed approach

In this Section we will show how the proposed PBM can be extended to different situations. Although we will deal with each of these situations separately, in an actual application they can occur simultaneously.

3.1. Non discretionary variables

Let the set of inputs be formed by two subsets: discretionary (I^{D}) and non-discretionary (I^{ND}). Similarly, let O^{D} and O^{ND} be the corresponding discretionary and non-discretionary outputs. The efficiency potential is the same as before, i.e.

$$P(x, y) = \left(\frac{\prod\limits_{i \in I^{\mathrm{D}}} \hat{x}_i \prod\limits_{i \in I^{\mathrm{ND}}} \hat{x}_i}{\prod\limits_{k \in I^{\mathrm{D}}} \hat{y}_k \prod\limits_{k \in I^{\mathrm{ND}}} \hat{y}_k} \right)^{1/(m+s)} \tag{12}$$

Handling non-discretionary variables as per Banker and Morey (1986), the corresponding MEffP DEA model is

$$\left(P_o^*\right)^{m+s} = min \ \frac{\prod\limits_{i=1}^{m} \hat{x}_i}{\prod\limits_{k=1}^{s} \hat{y}_k}$$

$$s.t. \ \sum_{j=1}^{n} \lambda_j \hat{x}_{ij} \le \hat{x}_i \le \hat{x}_{io} \qquad \forall i \in I^{\mathrm{D}},$$

$$\sum_{j=1}^{n} \lambda_j \hat{x}_{ij} \le \hat{x}_i = \hat{x}_{io} \qquad \forall i \in I^{\mathrm{ND}},$$

$$\sum_{j=1}^{n} \lambda_j \hat{y}_{kj} \ge \hat{y}_k \ge \hat{y}_{ko} \qquad \forall k \in O^{\mathrm{D}},$$

$$\sum_{j=1}^{n} \lambda_j \hat{y}_{kj} \ge \hat{y}_k = \hat{y}_{ko} \qquad \forall k \in O^{\mathrm{ND}},$$

$$\sum_{j=1}^{n} \lambda_j = 1,$$

$$\lambda_j \ge 0 \ \ \forall j$$

(13)

Note, in the above formulation, the existence of some equality constraints. These are used to indicate that, for the non-discretionary inputs and outputs, the corresponding target is the observed value and hence it is not properly a decision variable.

Using (6), the PBM efficiency score is then

$$\pi_o = \frac{P_o^*}{P_o} = \frac{P(x^*, y^*)}{P(x_o, y_o)} = \left(\frac{\dfrac{\prod\limits_{i \in I^{\mathrm{D}}} \hat{x}_i^* \prod\limits_{i \in I^{\mathrm{ND}}} \hat{x}_i^*}{\prod\limits_{k \in O^{\mathrm{D}}} \hat{y}_k^* \prod\limits_{k \in O^{\mathrm{ND}}} \hat{y}_k^*}}{\dfrac{\prod\limits_{i \in I^{\mathrm{D}}} \hat{x}_{io} \prod\limits_{i \in I^{\mathrm{ND}}} \hat{x}_{io}}{\prod\limits_{k \in O^{\mathrm{D}}} \hat{y}_{ko} \prod\limits_{k \in O^{\mathrm{ND}}} \hat{y}_{ko}}} \right)^{1/(m+s)} = \left(\frac{\dfrac{\prod\limits_{i \in I^{\mathrm{D}}} \hat{x}_i^*}{\prod\limits_{k \in O^{\mathrm{D}}} \hat{y}_k^*}}{\dfrac{\prod\limits_{i \in I^{\mathrm{D}}} \hat{x}_{io}}{\prod\limits_{k \in O^{\mathrm{D}}} \hat{y}_{ko}}} \right)^{1/(m+s)}$$

(14)

The corresponding PBM DEA model is

$$\left(\pi_o\right)^{m+s} = \min \ \frac{\prod\limits_{i \in I^{\mathrm{D}}} \theta_i}{\prod\limits_{k \in O^{\mathrm{D}}} \varphi_k}$$

$$s.t. \ \sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta_i \cdot x_{io} \qquad \forall i \in I^{\mathrm{D}},$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{io} \qquad \forall i \in I^{\mathrm{ND}},$$

$$\sum_{j=1}^{n} \lambda_j y_{kj} \geq \varphi_k \cdot y_{ko} \qquad \forall k \in O^{\mathrm{D}},$$

$$\sum_{j=1}^{n} \lambda_j y_{kj} \geq y_{ko} \qquad \forall k \in O^{\mathrm{ND}},$$

$$\sum_{j=1}^{n} \lambda_j = 1,$$

$$\lambda_j \geq 0 \ \forall j \quad \theta_i \leq 1 \ \forall i \in I^{\mathrm{D}} \quad \varphi_k \geq 1 \ \forall k \in O^{\mathrm{D}}$$

(15)

3.2. Undesirable outputs

Let $z_j = (z_{bj})_{b=1}^{q} > 0$ the undesirable outputs produced by each DMU j. To construct the PPS in this case, we assume joint weak disposability of the desirable and the undesirable outputs and use Kuosmanen (2005) approach. Assuming non-uniform abatement factors for the different observed DMUs, as Kuosmanen (2005) proposes, is not only reasonable but also directly leads to a linear programming optimization model. Thus, we arrive at the following PPS

$$T^{\mathrm{VRS}} = \{(x,y,z) \mid x \geq \sum_{j=1}^{n} (\lambda_j + \mu_j) x_j, y \leq \sum_{j=1}^{n} \lambda_j y_j, z = \sum_{j=1}^{n} \lambda_j z_j, \sum_{j=1}^{n} (\lambda_j + \mu_j) = 1, \lambda_j, \mu_j \geq 0\} \quad (16)$$

An associated dimensionless PPS $\hat{T}^{\mathrm{VRS}}$ can analogously be defined. We can define the efficiency potential as the following function, defined on $T^{\mathrm{VRS}} \cap R_{++}^{m+s+q}$

$$P(x,y,z) = \left(\frac{\prod\limits_{i=1}^{m} \dfrac{x_i}{x_i^{aver}} \prod\limits_{b=1}^{q} \dfrac{z_b}{z_b^{aver}}}{\prod\limits_{k=1}^{s} \dfrac{y_k}{y_k^{aver}}} \right)^{1/(m+s+q)} = \left(\frac{\prod\limits_{i=1}^{m} \hat{x}_i \prod\limits_{b=1}^{q} \hat{z}_b}{\prod\limits_{k=1}^{s} \hat{y}_k} \right)^{1/(m+s+q)} \tag{17}$$

The above expression is the way the proposed PBM can be extended to handle undesirable outputs. Thus, in (17) the undesirable outputs are normalized, same as the inputs and the outputs. Also, they appear in the numerator because, same as the inputs, lower values of undesirable outputs are preferable and therefore lead to lower efficiency potential. Finally, the number of undesirable outputs is taken into account when computing the joint geometric mean of all the normalized variables, i.e. its $m+s+q$ root. Therefore, (17) is a consistent and sound extension of the efficiency potential definition (4).

The corresponding MEffP and PBM DEA models are

$$\left(P_o^* \right)^{m+s+q} = min \ \frac{\prod\limits_{i=1}^{m} \hat{x}_i \prod\limits_{b=1}^{q} \hat{z}_b}{\prod\limits_{k=1}^{s} \hat{y}_k}$$

$$s.t. \ \sum\limits_{j=1}^{n} (\lambda_j + \mu_j) \cdot \hat{x}_{ij} \leq \hat{x}_i \leq \hat{x}_{io} \qquad i = 1,...,m,$$

$$\sum\limits_{j=1}^{n} \lambda_j \hat{y}_{kj} \geq \hat{y}_k \geq \hat{y}_{ko} \qquad\qquad k = 1,...,s, \tag{18}$$

$$\sum\limits_{j=1}^{n} \lambda_j \hat{z}_{bj} = \hat{z}_b \leq \hat{z}_{bo} \qquad\qquad b = 1,...,q,$$

$$\sum\limits_{j=1}^{n} (\lambda_j + \mu_j) = 1$$

$$\lambda_j, \mu_j \geq 0 \ \ \forall j$$

$$(\pi_o)^{m+s+q} = min \ \frac{\prod\limits_{i=1}^{m}\theta_i \prod\limits_{b=1}^{q}\phi_b}{\prod\limits_{k=1}^{s}\varphi_k}$$

$$
s.t. \ \sum_{j=1}^{n}(\lambda_j + \mu_j)\cdot x_{ij} \leq \theta_i \cdot x_{io} \qquad i = 1,...,m,
$$

$$
\sum_{j=1}^{n}\lambda_j y_{kj} \geq \varphi_k \cdot y_{ko} \qquad\qquad k = 1,...,s,
$$

$$
\sum_{j=1}^{n}\lambda_j z_{bj} = \phi_b \cdot z_{bo} \qquad\qquad b = 1,...,q,
$$

$$
\sum_{j=1}^{n}(\lambda_j + \mu_j) = 1
$$

$$
\lambda_j, \mu_j \geq 0 \ \forall j \quad \theta_i \leq 1 \ \forall i \quad \varphi_k \geq 1 \ \forall k \quad \phi_b \leq 1 \ \forall b
$$

(19)

and the PBM efficiency score

$$
\pi_o = \frac{P_o^*}{P_o} = \frac{P(x^*,y^*,z^*)}{P(x_o,y_o,z_o)} = \frac{\left(\dfrac{\prod\limits_{i=1}^{m}\hat{x}_i^* \prod\limits_{b=1}^{q}\hat{z}_b^*}{\prod\limits_{k=1}^{s}\hat{y}_k^*}\right)^{1/(m+s+q)}}{\left(\dfrac{\prod\limits_{i=1}^{m}\hat{x}_{io} \prod\limits_{b=1}^{q}\hat{z}_{bo}}{\prod\limits_{k=1}^{s}\hat{y}_{ko}}\right)^{1/(m+s+q)}} = \left(\dfrac{\prod\limits_{i=1}^{m}\theta_i^* \prod\limits_{b=1}^{q}\phi_b^*}{\prod\limits_{k=1}^{s}\varphi_k^*}\right)^{1/(m+s+q)}
$$

(20)

3.3. Preference structure

Let assume that the Decision Maker has a preference structure that assigns more weights to certain input and outputs. This may be given by a set of weights strictly positive weights (w_i^x, w_k^y) such that $\sum\limits_{i=1}^{m} w_i^x + \sum\limits_{k=1}^{s} w_k^y = 1$. The corresponding efficiency potential can be defined as

$$P(x, y) = \frac{\prod_{i=1}^{m} (\hat{x}_i)^{w_i^x}}{\prod_{k=1}^{s} (\hat{y}_k)^{w_k^y}} \tag{21}$$

Note that the efficiency potential (4) is a special case of (21) corresponding to $w_i^x = w_k^y = \dfrac{1}{m+s}$ $\forall i \forall k$.

In this case it is better to compute the PBM efficiency score directly using the following model

$$\pi_o = min \frac{\prod_{i=1}^{m} \theta_i^{w_i^x}}{\prod_{k=1}^{s} \varphi_k^{w_k^y}}$$

$$s.t. \ \sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta_i x_{io} \qquad i = 1,...,m, \tag{22}$$

$$\sum_{j=1}^{n} \lambda_j y_{kj} \geq \varphi_k \, y_{ko} \qquad k = 1,...,s,$$

$$\sum_{j=1}^{n} \lambda_j = 1,$$

$$\lambda_j \geq 0 \ \forall j \quad \theta_i \leq 1 \ \forall i \quad \varphi_k \geq 1 \ \forall k$$

Note that model (22) is also a linearly constrained nonlinear problem and in principle should not be harder to solve to optimality than model (9). Note also that the GDF DEA model (10) is equal to the square root of the efficiency score (22) for the special case $w_i^x = \dfrac{1}{2m} \forall i, w_k^y = \dfrac{1}{2s} \forall k$, in which

$\sum_{i=1}^{m} w_i^x + \sum_{k=1}^{s} w_k^y = 1$. Therefore, the weighted version of the proposed approach can be said to include the

GDF efficiency measure as a special case, one that gives the same importance to each input and the same importance to each output (but not the same importance to each input and output as the unweighted PBM efficiency measure does). In other words, GDF treats inputs symmetrically between themselves and the same with the outputs, but it does not give the same importance to an input as to an output. When there is a preference structure this (and any other weighting option) is valid but in the absence of a preference

structure we believe that it is more neutral and objective to give the same importance to each input and output, which the unweighted PBM efficiency measure does.

4. MEffP target setting model

As was seen in one of the illustrations of Section 2.4., the MEffP DEA model may have alternative optima which means that, although it can be used as is for computing the MEffP P_o^* and the corresponding PBM efficiency score π_o, it requires some refining if it is to be used for target setting purposes. The idea is to use a secondary goal to select one among the MEffP operating points. This secondary criterion may be to minimize the distance to the DMU being projected. The literature on DEA approaches to compute closest efficient targets is abundant, dating back to Frei and Harker (1999). The reader is referred to Aparicio (2016) and Aparicio et al. (2017a, 2017b) for recent developments in the field as well as an up-to-date review of the literature on this topic. In particular what we propose is similar to what Zofio et al. (2013) use to select the closest target among the maximum profit operating points. Zofio et al. (2013) used Euclidean distance because they considered normalized direction vectors. Instead, in order to normalize the distance measure, we will use a weighted Euclidean distance, with weights $1/x_{io}$ for ith input and $1/y_{ko}$ for kth output. Thus, the distance between a DMUo and its corresponding target (x_o^*, y_o^*) can be computed as

$$
\ell_2 = \left\| (x_o^*, y_o^*) - (x_o, y_o) \right\|_{w,2} = \sqrt{\sum_{i=1}^{m} \frac{1}{x_{io}^2}(x_{io} - x_{io}^*)^2 + \sum_{k=1}^{s} \frac{1}{y_{ko}^2}(y_{ko}^* - y_{ko})^2} =
$$
$$
= \sqrt{\sum_{i=1}^{m}(1 - \theta_i^*)^2 + \sum_{k=1}^{s}(\varphi_k^* - 1)^2}
\tag{23}
$$

In order to compute the closest MEffP we can use a lexicographic approach in which we first solve model (9) and then, in a second phase, we minimize the weighted Euclidean distance objective function, imposing that optimal value as a constraint, i.e.

$$min \sum_{i=1}^{m}(1-\theta_i)^2 + \sum_{k=1}^{s}(\varphi_k-1)^2$$

$$s.t. \quad \sum_{j=1}^{n}\lambda_j x_{ij} \leq \tilde{x}_i = \theta_i \cdot x_{io} \qquad i=1,...,m,$$

$$\sum_{j=1}^{n}\lambda_j y_{kj} \geq \tilde{y}_k = \varphi_k \cdot y_{ko} \qquad k=1,...,s,$$

$$\sum_{j=1}^{n}\lambda_j = 1,$$

$$\frac{\prod_{i=1}^{m}\theta_i}{\prod_{k=1}^{s}\varphi_k} = \left(\pi_o\right)^{m+s}$$

$$\lambda_j \geq 0 \ \forall j \quad \theta_i \leq 1 \ \forall i \quad \varphi_k \geq 1 \ \forall k$$

$$(24)$$

This model has a quadratic objective function and one non-linear constraint. Note that the right-hand side of the non-linear constraint is a constant, as π_o has been previously computed using (9). Model (24) can be solved with GAMS and with solver Couenne as before.

Note that instead of using the average values in equation (2), the median or any other typical value could be used to define the efficiency potential. However, this does not affect neither the PBM efficiency score computed with model (9) nor the MEffP targets computed with model (24).

In Figure 3 we revisit the XY illustrative example of Table 3. Recall that all the points along the segment $C'B$ are MEffP. Using (24) or (25) we can find the closest among those benchmarks. Specifically, for DMU C, its nearest MEffP operating point is $(x_C^*, y_C^*)=(1.598, 3.196)$ and its weighted l_2 distance is equal to 0.705. Therefore, using (x_C^*, y_C^*) as target for DMU C means reducing its input from 5 to 1.598 (a 68% reduction) and increasing its output from 2.7 to 3.196 (a 18.4% increase). This translates into a reduction of the efficiency potential from 1.515 to 0.787, leading to a PBM efficiency score of $\frac{0.787}{1.515}=0.519$. Actually, among the efficient MEffP

operating points that dominate DMU C, (x_C^*, y_C^*) is the one that is closest and hence it should be easiest to reach as a target.

=========================== Figure 3 ===========================

5. Case study

In this section we apply the proposed approach to the 18 container shipping lines (CSLs) studied in Gutiérrez et al. (2014). This application considers three inputs (Labor, number of ships and fleet capacity, measured in Twenty-foot equivalent units, TEU) and two outputs: number of containers carried (in TEU) and turnover (in million USD). The dataset, together with the corresponding average DMU and efficiency potentials, are shown in Table 4.

=========================== Table 4 ===========================

Gutiérrez et al (2014) showed that eight of the DMUs (namely Maersk, CSCL, OOCL, NYK, K_Line, HMM, Wan_Hai and Delmas) are technical efficient. The proposed PBM approach confirms that the MEffP targets of these 8 DMUs are themselves and therefore $P(x^*, y^*) = P(x_o, y_o) \Rightarrow \pi_o = 1$. For the inefficient DMUs, Table 5 shows the corresponding MEffP and PBM score computed using (5) and (6) and the optimal θ_i and φ_k variables and l_2 distance computed by model (24). The ranking of the inefficient DMUs derived from the PBM efficiency scores are also shown within parentheses. Thus, for example, MSC, which is the second largest CSL in the sample, can reduce its efficiency potential from 1.293 to a MEffP value of 1,059, which means a PBM efficiency score of $\frac{1.059}{1.293} = 0.819$. In terms of the input and output improvements computed by the proposed approach it seems that most of the inefficiencies affect the inputs, with reductions of 23.6%, 26.6% and 30.5% for labor, number of ships and fleet capacity, respectively. A small increase of 5.6% in one of the outputs, namely Turnover, is also estimated. Even these improvements are relatively large, even larger improvements are computed for other DMUs, thus making MSC rank 4th among the inefficient DMUs.

=========================== Table 5 ===========================

For comparison, the GDF and ERGM scores and input and output improvements computed using (10) and (11) are also shown in Tables 6 and 7, respectively. Note that, as indicated in Section 2, the PBM efficiency score is always greater than (or at least equal to) the GDF efficiency score. Thus, for example, for MSC GDF is 0.704, which is lower than the PBM score (0.819). The ERGM score is slightly lower (0.691) and all three methods rank MSC 4th among the inefficient DMUs. The targets computed for MSC by the three methods are rather different, however. Thus, GDF estimates smaller input reductions (7.5%, 23% and 18.3% for labor, number of ships and fleet capacity, respectively) but a much larger increase in Turnover (40.7%). Something similar happens with the target computed by ERGM: the estimated input reductions are relatively small (1%, 20% and 15% for labor, number of ships and fleet capacity, respectively) and the estimated Turnover increase is rather large (54.7%). Interestingly, for MSC the three methods estimate a 0% increase in the Containers carried output. In fact, all three methods coincide in that this output only has margin for improvement for two of the DMUs only, namely CSAV and Hamburg Süd.

=========================== Table 6 ===========================

=========================== Table 7 ===========================

The ranking of the inefficient DMUs computed by PBM and GDF coincide, differing slightly with that of ERGM (Spearman rank order correlation coefficient = 0.964, significant at 0.01 level, two-tailed test). Note also that all three methods identify more inefficiencies (and hence more potential for improvement) in the inputs than in the outputs. This can also be seen in Figure 4, which shows, using parallel coordinates, the observed and corresponding target values for each inefficient DMUs. As can be seen, the efficiency score θ_i for the employees input for the COSCON DMU is very low in all three methods (0.047 for PBM, 0.076 for GDF and 0.105 for ERGM). This may be due to a mistake in the observed data for that DMU, which at 71,584 employees is clearly an outlier. In any case, here we report the results for the dataset listed in Gutiérrez et al. (2014) and shown in Table 4.

=========================== Figure 4 ===========================

24

Finally, Table 8 shows, for each inefficient DMU, the intensity variables λ_j that determine the target computed by the proposed PBM approach. Thus, for example, the target computed by the proposed approach for MSC corresponds to combining, with a weight of practically 50% each, the largest DMU (Maersk) and a medium-sized efficient DMU (CSCL). Actually Maersk is only a reference for MSC and CMA_CGM, which are the second and third largest companies in the sample. Correspondingly, the smaller efficient companies (like HMM and Wan Hai) are used as benchmarks mainly for the smaller inefficient DMUs (like Hamburg Süd or YML). Some efficient DMUs, like Delmas and OOCL, are never used as benchmarks, indicating that they have singular operating points that do not serve as reference for the other DMUs. Note, finally, that the peer group of the inefficient DMUs is always formed by at most three benchmarks.

=========================== Table 8 ===========================

6. Conclusions

In this paper a new efficiency measure has been introduced. Although the mathematical formulation of the proposed PBM efficiency measure has similarities with some existing methods (like GDF or ERGM) its motivation is completely different. It consists in applying a Physics analogy so that each feasible operating point is assigned an Efficiency Potential, i.e. a positive value that is lower, the lower the input consumption and the higher the output production. The larger the input reduction and the output increase, the larger the Efficiency Potential reduction achieved. The proposed approach computes the MEffP associated to each DMU and the proposed PBM efficiency score is the ratio of the MEffP to the efficiency potential of the observed DMU. PBM can also be applied in non-discretionary variables, undesirable outputs and preference structure scenarios. An appropriate target setting DEA model has been also proposed for the case that there are multiple MEffP operating points.

Note that we do not claim that the PBM efficiency score proposed in this paper has some advantages or is superior to existing, well-established efficiency scores like, for example, ERGM/SBM or GDF. The PBM efficiency score has, it is true, a number of interesting features such as normalization, efficiency indication, units-invariance, reference-set dependence and strong monotonicity, but so have also other efficiency measures. And, PBM has the drawback of

requiring solving a non-linear optimization model. Even so, the fact that PBM makes use of the novel concept of efficiency potential, which is based on a Physics analogy and presents a new perspective on efficiency assessment, makes it, in our opinion, a worthy proposal. The fact that the efficiency potential concept can be used with non-discretionary variables, and undesirable outputs, as shown in the paper, and may also be extended to fuzzy data, network DEA, etc. shows the flexibility and usefulness of this approach, which we believe is a promising research topic.

The proposed approach has, nevertheless, some limitations, such as not being able to handle zero or negative data, nor integer variables. Some of these issues may also be topics for further research.

Appendix

In this section, the dual of model (9) is formulated. Thus, following Bazaraa et al. (1993),given a primal model:

$$
\begin{aligned}
&min \quad f(z) \\
&s.t. \quad g(z) \le 0 \\
&\qquad h(z) = 0 \\
&\qquad z \in Z
\end{aligned}
$$

its Lagrangean dual problem is

$$
\begin{aligned}
&Max \quad L^*(u,u_0) \\
&s.t. \quad u \ge 0, \; u_0 \text{ free}
\end{aligned}
$$

where $L^*(u,u_0) = \inf_{z \in Z} L(z;u,u_0) = \inf_{z \in Z} \{f(z) + ug(z) + u_0h(z)\}$ and u and u_0 are the dual variables corresponding to the inequality and equality constraints, respectively.

The PBM model (9), where the dual variables associated to each constraint are indicated to its right, is reproduced below

$$\min \ f\left(\lambda,\theta,\varphi\right) = \frac{\prod\limits_{i=1}^{m}\theta_i}{\prod\limits_{k=1}^{s}\varphi_k}$$

$$s.t. \quad \sum_{j=1}^{n}\lambda_j x_{ij} - \theta_i \cdot x_{io} \leq 0 \qquad i = 1,...,m, \qquad\qquad \left(v_i \geq 0\right)$$

$$\varphi_k \cdot y_{ko} - \sum_{j=1}^{n}\lambda_j y_{kj} \leq 0 \qquad k = 1,...,s, \qquad\qquad \left(u_k \geq 0\right)$$

$$\theta_i - 1 \leq 0 \qquad\qquad i = 1,...,m, \qquad\qquad \left(\hat{v}_i \geq 0\right)$$

$$1 - \varphi_k \leq 0 \qquad\qquad k = 1,...,s, \qquad\qquad \left(\hat{u}_k \geq 0\right)$$

$$\sum_{j=1}^{n}\lambda_j - 1 = 0, \qquad\qquad\qquad\qquad \left(u_0 \text{ free}\right)$$

$$\lambda_j \geq 0 \ \ \forall j$$

The corresponding Lagrangean function is

$$L(\lambda,\theta,\varphi; v,\hat{v},u,\hat{u},u_0) = \frac{\prod\limits_{i=1}^{m}\theta_i}{\prod\limits_{k=1}^{s}\varphi_k} + \sum_{i=1}^{m}v_i \cdot \left(\sum_{j=1}^{n}\lambda_j x_{ij} - \theta_i x_{io}\right) + \sum_{k=1}^{s}u_k \cdot \left(\varphi_k y_{ko} - \sum_{j=1}^{n}\lambda_j y_{kj}\right)$$

$$+ \sum_{i=1}^{m}\hat{v}_i \cdot (\theta_i - 1) + \sum_{k=1}^{s}\hat{u}_k \cdot \left(1 - \varphi_k\right) + u_0 \cdot \left(\sum_{j=1}^{n}\lambda_j - 1\right)$$

$$= \frac{\prod\limits_{i=1}^{m}\theta_i}{\prod\limits_{k=1}^{s}\varphi_k} + \sum_{j=1}^{n}\lambda_j \cdot \left(\sum_{i=1}^{m}v_i x_{ij} - \sum_{k=1}^{s}u_k y_{kj} + u_0\right) +$$

$$+ \sum_{i=1}^{m}\theta_i \cdot \left(\hat{v}_{\boldsymbol{i}} - v_i x_{io}\right) + \sum_{k=1}^{s}\varphi_k \cdot \left(u_k y_{ko} - \hat{u}_k\right) + \sum_{k=1}^{s}\hat{u}_k - \sum_{i=1}^{m}\hat{v}_i - u_0$$

In order to compute $L^*(v,\hat{v},u,\hat{u},u_0) = \inf\limits_{\lambda \geq 0,(\theta,\varphi)\in R^{m+s}} L(\lambda,\theta,\varphi,\lambda;v,\hat{v},u,\hat{u},u_0)$, note that

$L^*(v,\hat{v},u,\hat{u},u_0) = -\infty$ if $\sum\limits_{i=1}^{m} v_i x_{ij} - \sum\limits_{k=1}^{s} u_k y_{kj} + u_0 < 0$ for any DMU j as the corresponding $\lambda_j = \infty$.

If, on the contrary, $\sum\limits_{i=1}^{m} v_i x_{ij} - \sum\limits_{k=1}^{s} u_k y_{kj} + u_0 \geq 0 \quad \forall j$, then the terms

$\lambda_j \cdot \left(\sum\limits_{i=1}^{m} v_i x_{ij} - \sum\limits_{k=1}^{s} u_k y_{kj} + u_0 \right) = 0 \quad \forall j$ and hence do not contribute to $L^*(v,\hat{v},u,\hat{u},u_0)$.

As regards the θ_i and φ_k variables,

$$\frac{\partial L}{\partial \theta_i} = \frac{\prod\limits_{i'\neq i}\theta_{i'}}{\prod\limits_{k=1}^{s}\varphi_k} + \hat{v}_i - v_i x_{io} = \frac{\prod\limits_{i'=1}^{m}\theta_{i'}}{\prod\limits_{k=1}^{s}\varphi_k} \cdot \frac{1}{\theta_i} + \hat{v}_i - v_i x_{io} = \frac{f(\lambda,\theta,\varphi)}{\theta_i} + \hat{v}_i - v_i x_{io}$$

$$\frac{\partial L}{\partial \varphi_k} = -\frac{\prod\limits_{i=1}^{m}\theta_i}{\prod\limits_{k'\neq k}\varphi_{k'}} \cdot \frac{1}{(\varphi_k)^2} + u_k y_{ko} - \hat{u}_k = -\frac{\prod\limits_{i=1}^{m}\theta_i}{\prod\limits_{k'=1}^{s}\varphi_{k'}} \cdot \frac{1}{\varphi_k} + u_k y_{ko} - \hat{u}_k = -\frac{f(\lambda,\theta,\varphi)}{\varphi_k} + u_k y_{ko} - \hat{u}_k$$

$$\frac{\partial L}{\partial \theta_i} = 0 \implies \frac{f(\lambda,\theta,\varphi)}{\theta_i} + \hat{v}_i - v_i x_{io} = 0 \implies \begin{cases} \theta_i = \dfrac{f(\lambda,\theta,\varphi)}{v_i x_{io} - \hat{v}_i} \quad \forall i \\[2mm] v_i x_{io} - \hat{v}_i = \dfrac{f(\lambda,\theta,\varphi)}{\theta_i} \quad \forall i \end{cases}$$

$$\frac{\partial L}{\partial \varphi_k} = 0 \implies -\frac{f(\lambda,\theta,\varphi)}{\varphi_k} + u_k y_{ko} - \hat{u}_k = 0 \implies \begin{cases} \varphi_k = \dfrac{f(\lambda,\theta,\varphi)}{u_k y_{ko} - \hat{u}_k} \quad \forall k \\[2mm] u_k y_{ko} - \hat{u}_k = \dfrac{f(\lambda,\theta,\varphi)}{\varphi_k} \quad \forall k \end{cases}$$

Hence

$$\theta_i = \frac{f(\lambda,\theta,\varphi)}{v_i x_{io} - \hat{v}_i} \le 1 \quad \forall i \quad \Rightarrow \quad v_i x_{io} - \hat{v}_i \ge f(\lambda,\theta,\varphi) \quad \forall i$$

$$\varphi_k = \frac{f(\lambda,\theta,\varphi)}{u_k y_{ko} - \hat{u}_k} \ge 1 \quad \forall k \quad \Rightarrow \quad u_k y_{ko} - \hat{u}_k \le f(\lambda,\theta,\varphi) \quad \forall k$$

$$f(\lambda,\theta,\varphi) = \frac{\prod_{i=1}^{m}\theta_i}{\prod_{k=1}^{s}\varphi_k} = \frac{f(\lambda,\theta,\varphi)^m}{f(\lambda,\theta,\varphi)^s} \cdot \frac{\prod_{k=1}^{s}(u_k y_{ko} - \hat{u}_k)}{\prod_{i=1}^{m}(v_i x_{io} - \hat{v}_i)} \quad \Rightarrow$$

$$\Rightarrow \quad f(\lambda,\theta,\varphi)^{(1+s-m)} = \frac{\prod_{k=1}^{s}(u_k y_{ko} - \hat{u}_k)}{\prod_{i=1}^{m}(v_i x_{io} - \hat{v}_i)} \Rightarrow
\begin{cases}
f(\lambda,\theta,\varphi) = \left[\dfrac{\prod_{k=1}^{s}(u_k y_{ko} - \hat{u}_k)}{\prod_{i=1}^{m}(v_i x_{io} - \hat{v}_i)}\right]^{\frac{1}{1+s-m}} & \text{if } 1+s-m \ne 0 \\[3em]
\dfrac{\prod_{k=1}^{s}(u_k y_{ko} - \hat{u}_k)}{\prod_{i=1}^{m}(v_i x_{io} - \hat{v}_i)} = 1 & \text{if } 1+s-m = 0
\end{cases}$$

$$v_i x_{io} - \hat{v}_i \ge f(\lambda,\theta,\varphi) \quad \forall i \quad \Rightarrow \quad v_i x_{io} - \hat{v}_i \ge \left(\frac{\prod_{k'=1}^{s}(u_{k'} y_{k'o} - \hat{u}_{k'})}{\prod_{i'=1}^{m}(v_{i'} x_{i'o} - \hat{v}_{i'})}\right)^{\frac{1}{1+s-m}} \quad \forall i$$

$$\Rightarrow \quad \left(v_i x_{io} - \hat{v}_i\right)^{\frac{2+s-m}{1+s-m}} \ge \left(\frac{\prod_{k'=1}^{s}(u_{k'} y_{k'o} - \hat{u}_{k'})}{\prod_{i'\ne i}(v_{i'} x_{i'o} - \hat{v}_{i'})}\right)^{\frac{1}{1+s-m}} \quad \forall i$$

$$\Rightarrow \quad \left(v_i x_{io} - \hat{v}_i\right)^{2+s-m} \ge \frac{\prod_{k'=1}^{s}(u_{k'} y_{k'o} - \hat{u}_{k'})}{\prod_{i'\ne i}(v_{i'} x_{i'o} - \hat{v}_{i'})} \quad \forall i$$

$$u_k\, y_{ko} - \hat{u}_k \le f\left(\lambda,\theta,\varphi\right) \quad \forall k \quad \Rightarrow \quad u_k\, y_{ko} - \hat{u}_k \le \left(\frac{\displaystyle\prod_{k'=1}^{s} \left(u_{k'}\, y_{k'o} - \hat{u}_{k'}\right)}{\displaystyle\prod_{i'=1}^{m} \left(v_{i'}\, x_{i'o} - \hat{v}_{i'}\right)} \right)^{\frac{1}{1+s-m}} \quad \forall k$$

$$\Rightarrow \quad \left(u_k\, y_{ko} - \hat{u}_k\right)^{\frac{s-m}{1+s-m}} \le \left(\frac{\displaystyle\prod_{k'\neq k} \left(u_{k'}\, y_{k'o} - \hat{u}_{k'}\right)}{\displaystyle\prod_{i'=1}^{m} \left(v_{i'}\, x_{i'o} - \hat{v}_{i'}\right)} \right)^{\frac{1}{1+s-m}} \quad \forall k$$

$$\Rightarrow \quad \left(u_k\, y_{ko} - \hat{u}_k\right)^{s-m} \le \frac{\displaystyle\prod_{k'\neq k} \left(u_{k'}\, y_{k'o} - \hat{u}_{k'}\right)}{\displaystyle\prod_{i'=1}^{m} \left(v_{i'}\, x_{i'o} - \hat{v}_{i'}\right)} \quad \forall k$$

Therefore

$$L*(v,\hat{v},u,\hat{u},u_0) = \begin{cases} -\infty & \text{if } \exists j : \displaystyle\sum_{i=1}^{m} v_i x_{ij} - \sum_{k=1}^{s} u_k y_{kj} + u_0 < 0 \\[4ex] = \dfrac{\displaystyle\prod_{i=1}^{m}\theta_i}{\displaystyle\prod_{k=1}^{s}\varphi_k} + \displaystyle\sum_{i=1}^{m}\theta_i\cdot\left(\hat{v}_i - v_i x_{io}\right) + \sum_{k=1}^{s}\varphi_k\cdot\left(u_k y_{ko} - \hat{u}_k\right) + \sum_{k=1}^{s}\hat{u}_k - \sum_{i=1}^{m}\hat{v}_i - u_0 = \\[4ex] = f\left(\lambda,\theta,\varphi\right) - \displaystyle\sum_{i=1}^{m} f\left(\lambda,\theta,\varphi\right) + \sum_{k=1}^{s} f\left(\lambda,\theta,\varphi\right) + \sum_{k=1}^{s}\hat{u}_k - \sum_{i=1}^{m}\hat{v}_i - u_0 = \\[4ex] = \left(1+s-m\right)\cdot f\left(\lambda,\theta,\varphi\right) + \displaystyle\sum_{k=1}^{s}\hat{u}_k - \sum_{i=1}^{m}\hat{v}_i - u_0 = \\[4ex] = \left(1+s-m\right)\cdot \left(\dfrac{\displaystyle\prod_{k=1}^{s} \left(u_k y_{ko} - \hat{u}_k\right)}{\displaystyle\prod_{i=1}^{m} \left(v_i x_{io} - \hat{v}_i\right)} \right)^{\frac{1}{1+s-m}} + \displaystyle\sum_{k=1}^{s}\hat{u}_k - \sum_{i=1}^{m}\hat{v}_i - u_0 & \text{otherwise} \end{cases}$$

and the dual of model (9) is

$$max \quad (1+s-m) \cdot \left[\frac{\prod\limits_{k=1}^{s}(u_k y_{ko} - \hat{u}_k)}{\prod\limits_{i=1}^{m}(v_i x_{io} - \hat{v}_i)} \right]^{\frac{1}{1+s-m}} + \sum_{k=1}^{s} \hat{u}_k - \sum_{i=1}^{m} \hat{v}_i - u_0$$

$$s.t. \quad \sum_{i=1}^{m} v_i x_{ij} - \sum_{k=1}^{s} u_k y_{kj} + u_0 \geq 0 \quad \forall j$$

$$\left(v_i x_{io} - \hat{v}_i \right)^{2+s-m} \geq \frac{\prod\limits_{k'=1}^{s}(u_{k'} y_{k'o} - \hat{u}_{k'})}{\prod\limits_{i' \neq i}(v_{i'} x_{i'o} - \hat{v}_{i'})} \quad \forall i$$

$$\left(u_k y_{ko} - \hat{u}_k \right)^{s-m} \leq \frac{\prod\limits_{k' \neq k}(u_{k'} y_{k'o} - \hat{u}_{k'})}{\prod\limits_{i'=1}^{m}(v_{i'} x_{i'o} - \hat{v}_{i'})} \quad \forall k$$

$$u, v, \hat{u}, \hat{v} \geq 0 \; , u_0 \text{ free}$$

Solving this dual model seems to be harder than solving the primal model. In the special case $1+s-m=0$, the objective function is linear and the dual of model (9) reduces to

$$max \quad \sum_{k=1}^{s} \hat{u}_k - \sum_{i=1}^{m} \hat{v}_i - u_0$$

$$s.t. \quad \sum_{i=1}^{m} v_i x_{ij} - \sum_{k=1}^{s} u_k y_{kj} + u_0 \geq 0 \quad \forall j$$

$$\frac{\prod\limits_{k=1}^{s}(u_k y_{ko} - \hat{u}_k)}{\prod\limits_{i=1}^{m}(v_i x_{io} - \hat{v}_i)} = 1$$

$$u, v, \hat{u}, \hat{v} \geq 0 \; , u_0 \text{ free}$$

As it is customary in DEA, the dual models above provide shadow prices for the inputs and outputs as well as the coefficients of the supporting hyperplane on which the efficient projection

lies. It is easy to see that in the optimum at least one of the constraints

$$\sum_{i=1}^{m} v_i x_{ij} - \sum_{k=1}^{s} u_k y_{kj} + u_0 \geq 0 \ \forall j$$ must be binding. Otherwise it would be feasible to decrease the

intercept variable u_0 thus increasing the objective function, which would be a contradiction with the assumed optimality. Moreover, due to complementary slackness, we know that

$$\sum_{i=1}^{m} v_i^* x_{ij} - \sum_{k=1}^{s} u_k^* y_{kj} + u_0^* = 0 \ \forall j : \lambda_j^* > 0$$. Similarly, $\hat{v}_i^* = 0 \ \forall i : \theta_i^* < 1$ and $\hat{u}_k^* = 0 \ \forall k : \varphi_k^* > 1$.

Note also that if DMU 0 is efficient, the optimal solution of model (9) is $\theta_i^* = 1 \ \forall i, \varphi_k^* = 1 \ \forall k, \lambda_0^* = 1, \lambda_j^* = 0 \ \forall j \neq 0$, and, in the optimum of the dual

$$\sum_{i=1}^{m} v_i^* x_{i0} - \sum_{k=1}^{s} u_k^* y_{k0} + u_o^* = 0 \quad \Rightarrow \quad u_o^* = -\sum_{i=1}^{m} v_i^* x_{i0} + \sum_{k=1}^{s} u_k^* y_{k0}$$

$$u_k^* y_{ko} - \hat{u}_k^* = \frac{f\left(\lambda^*, \theta^*, \varphi^*\right)}{\varphi_k^*} = 1 \ \forall k \quad \Rightarrow \quad \sum_{k=1}^{s} \hat{u}_k^* = \sum_{k=1}^{s} \left(u_k^* y_{ko} - 1\right) = \sum_{k=1}^{s} u_k^* y_{ko} - s$$

$$v_i^* x_{io} - \hat{v}_i^* = \frac{f\left(\lambda^*, \theta^*, \varphi^*\right)}{\theta_i^*} = 1 \ \forall i \quad \Rightarrow \quad \sum_{i=1}^{m} \hat{v}_i^* = \sum_{i=1}^{m} \left(v_i^* x_{io} - 1\right) = \sum_{i=1}^{m} v_i^* x_{io} - m$$

and hence the corresponding optimal dual objective function value is

$$\left(1 + s - m\right) \cdot \left[\frac{\displaystyle\prod_{k=1}^{s}\left(u_k^* y_{ko} - \hat{u}_k^*\right)}{\displaystyle\prod_{i=1}^{m}\left(v_i^* x_{io} - \hat{v}_i^*\right)} \right]^{\frac{1}{1+s-m}} + \sum_{k=1}^{s} \hat{u}_k^* - \sum_{i=1}^{m} \hat{v}_i^* - u_o^* =$$

$$= \left(1 + s - m\right) + \sum_{k=1}^{s} u_k^* y_{ko} - s - \left(\sum_{i=1}^{m} v_i^* x_{io} - m\right) - \left(-\sum_{i=1}^{m} v_i^* x_{i0} + \sum_{k=1}^{s} u_k^* y_{k0}\right) = 1$$

References

Aparicio, J., "A survey on measuring efficiency through the determination of the least distance in data envelopment analysis", *Journal of Centrum Cathedra*, 9, 2 (2016) 143–167

Aparicio, J., Cordero, J.M. and Pastor, J.T., "The determination of the least distance to the strongly efficient frontier in Data Envelopment Analysis oriented models: Modelling and computational aspects", *Omega*, 71 (2017a) 1–10

Aparicio, J., Pastor, J.T., Sainz-Pardo, J.L. and Vidal, F., "Estimating and decomposing overall inefficiency by determining the least distance to the strongly efficient frontier in data envelopment analysis", *Operational Research International Journal*, (2017b) DOI 10.1007/s12351-017-0339-0

Asmild, M. and Pastor, J.T., Slack free MEA and RDM with comprehensive efficiency measures, *Omega*, 38 (2010) 475–483

Banker, R.D., Estimating most productive scale size using data envelopment analysis, *European Journal of Operational Research*, 17 (1984) 35–44

Banker, R.D., Charnes, A. and Cooper, W.W., "Some models for estimating technical and scale inefficiencies in data envelopment analysis", *Management Science*, 30, 9 (1984) 1078–1092

Banker, R.D. and Morey, R., "Efficiency analysis for exogenously fixed inputs and outputs", *Operations Research*, 34 (1986) 513–521

Bardhan, I., Bowlin, W.F., Cooper, W.W. and Sueyoshi, T., "Models and Measures for Efficiency Dominance in DEA. Part I: Additive Models and MED Measures", *Journal of the Operations Research Society of Japan*, 39, 3 (1996) 322–332

Bazaraa, M.S., Sherali, H.D. and Shetty, C.M., *Nonlinear programming: theory and algorithms*, John Wiley and Sons, New York, 1993

Chambers, R.G., Chung, Y. and Färe, R., "Benefit and Distance Functions", *Journal of Economic Theory*, 70 (1996) 407–419

Charnes, A., Cooper, W.W. and Rhodes, E., "Measuring the efficiency of decision making units", *European Journal of Operational Research*, 2, 6 (1978) 429–444

Cooper, W.W. and Tone, K., "Measures of inefficiency in data envelopment analysis and stochastic frontier estimation", *European Journal of Operational Research*, 99 (1997) 72–88

Cooper, W.W., Park, K.S. and Pastor, J.T., "RAM: A Range Adjusted Measure of Inefficiency for Use with Additive Models, and Relations to Other Models and Measures in DEA", *Journal of Productivity Analysis*, 11, 1 (1999) 5–42

Cooper, W.W., Seiford, L.M. and Zhu, J., *Handbook on Data Envelopment Analysis*, Springer, New York, 2004

Cooper, W.W., Seiford, L.M. and Tone, K., *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, 2nd edition, Springer, New York, 2006

Färe, R., Grosskopf, S. and Lovell, C.A.K., *The Measurement of Efficiency of Production*. Kluwer-Nijhoff Publishing, Dordrecht, 1985

Färe, R. and Lovell, C.A.K., "Measuring the technical efficiency of production", *Journal of Economic Theory*, 19, 1 (1978) 150–162

Frei, F.X. and Harker, P.T., "Projections onto efficient frontiers: theoretical and computational extensions to DEA", *Journal of Productivity Analysis*, 11 (1999) 275–300

Gutiérrez, E., Lozano, S. and Furió, S., "Evaluating efficiency of international container shipping lines: A bootstrap DEA approach", *Maritime Economics & Logistics*, 16, 1 (2014) 55–71

Kuosmanen, T., "Weak disposability in nonparametric production analysis with undesirable outputs", *American Journal of Agricultural Economics*, 87, 4 (2005) 1077–1082

Lozano, S. and Calzada-Infante, L., "Computing gradient-based stepwise benchmarking paths", *Omega*, (2017) (DOI: 10.1016/j.omega.2017.11.002)

Pastor, J.T., Ruiz, J.L. and Sirvent, I., "An enhanced DEA Russell graph efficiency measure", *European Journal of Operational Research*, 115 (1999) 596–607

Portela, M.C.A.S. and Thanassoulis, E., "Developing a decomposable measure of profit efficiency using DEA", *Journal of the Operational Research Society*, 58 (2007) 481–490

Tone, K., "A slacks-based measure of efficiency in data envelopment analysis", *European Journal of Operational Research*, 130, 3 (2001) 498–509

Wang, K., Xian, Y., Lee, C.Y., Wei, Y.M. and Huang, Z., "On selecting directions for directional distance functions in a non-parametric framework: a review", *Annals of Operations Research*, (2017) DOI: 10.1007/s10479-017-2423-5

Zofio, J.L., Pastor, J.T. and Aparicio, J., "The directional profit efficiency measure: on why profit inefficiency is either technical or allocative", *Journal of Productive Analysis*, 40 (2013) 257–266

List of figures and table captions

Table 1. Summary of commonly-used efficiency measures

Table 2. Data for 1YY illustrative example

Table 3. Data for XY illustrative example

Table 4. CSL dataset, average DMU and efficiency potentials

Table 5. PBM approach results for case study

Table 6. GDF approach results for case study

Table 7. ERGM approach results for case study

Table 8. Intensity variables and benchmarks for the inefficient DMUs

Figure 1. Isopotential curves and Minimum efficiency potential of DMU C of 1YY illustrative example

Figure 2. Isopotential curves and minimum efficiency potential of DMU C of XY illustrative example

Figure 3. Minimum efficiency potential target for DMU C of XY illustrative example

Figure 4. Observed and target values for the inefficient DMUs

Table 1. Summary of existing efficiency measures

Properties	Efficiency measures													
	CCR	BCC	NRRM	Hyperbolic	RGM	ERGM	DDF	MED	RAM	GDF	SBI	RDM	MEA	EBM
Radial	✓	✓	×	×	×	×	×	×	×	×	×	×	×	×
Non-radial	×	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Linear	✓	✓	✓	×	×	✓	✓	×	✓	×	✓	✓	✓	✓
Nonlinear	×	×	×	✓	✓	×	×	✓	×	✓	×	×	×	×
Translation Invariance[1]	×	✓[2]	✓[2]	×	×	✓	✓[3]	×	✓	×	✓[3]	✓	✓	✓
Units Invariance	✓	✓	✓	✓	✓	✓	✓[3]	✓	✓	✓	✓[3]	✓	✓	✓
Comprehensive measure[4]	×	×	×	×	✓	✓	×	✓	✓	✓	✓	✓	✓	✓
Reference-set dependence[5]	✓	✓	✓	✓	✓	✓	✓[3]	✓	×	✓	✓[3]	×	×	×
Normalized Score[6]	✓	✓	✓	✓	✓	✓	✓[3]	✓	✓	✓	×	✓	✓	✓
Strong monononicity[7]	×	×	×	×	✓	✓	×	×	×	✓	✓[3]	×	×	✓

Notes:

1: For this property we assume VRS, except in the case of CCR.

2: Input-oriented (output-oriented) is invariant w.r.t output (input) translation.

3: It depends on the selected direction vector.

4: The inputs or outputs cannot improve more.

5: An efficiency measure is reference-set dependent if it is determined only by the reference set of the DMU concerned. It should not be influenced by statistics covering the whole data set.

6: Normalized score: The measure is between zero and one.

7: A technical efficient measure should be strictly increasing in output quantities and decreasing in input quantities.

Table 2. Data for 1YY illustrative example

DMU	A	B	C	D	Aver. DMU
x	1	1	1	1	1.000
y_1	2	3	1.5	3.5	2.500
y_2	4	3	2	1.5	2.625
$P(x, y)$	0.936	0.900	1.298	1.077	1.000

Table 3. Data for XY illustrative example

DMU	A	B	C	D	E	Aver.
x	1	2	5	3	4.5	3.10
y	2	4	2.7	5	5.5	3.84
$P(x, y)$	0.787	0.787	1.515	0.862	1.007	1.000

Table 4. CSL dataset, average DMU and efficiency potentials

DMU	Inputs			Outputs		$P(x, y)$
	Employees	# ships	Fleet capacity (TEU)	Containers carried (TEU)	Turnover	
Maersk	24,500	430	1,753,996	13,800,000	19,962.05	1.190
MSC	19,000	399	1,516,381	10,290,000	11,000.00	1.293
CMA_CGM	17,000	284	944,514	7,882,000	10,600.48	1.142
Hapag_Lloyd	6,670	112	460,241	4,637,000	6,194.31	0.843
COSCON	71,584	142	490,836	5,200,000	4,306.71	1.513
Evergreen_Line	4,141	167	593,443	5,815,000	2,704.17	0.985
APL	4,980	129	528,515	4,930,000	5,485.00	0.851
CSCL	4,311	121	457,648	6,700,000	3,090.31	0.837
OOCL	7,748	64	297,367	4,159,000	4,350.20	0.780
CSAV	6,972	65	194,010	1,790,500	3,028.00	0.896
MOL	10,012	90	345,218	3,030,000	4,900.00	0.943
NYK	1,619	77	357,686	3,550,000	4,955.44	0.618
Hamburg_Süd	4,791	90	288,297	2,300,000	4,463.36	0.845
K_Line	7,119	92	334,741	3,081,000	10,982.96	0.745
YML	4,197	82	325,828	2,780,000	2,933.76	0.867
HMM	2,038	52	255,643	2,510,000	5,256.00	0.592
Wan_Hai	769	63	122,069	2,685,166	1,595.08	0.547
Delmas	727	63	90,978	692,000	1,765.82	0.656
Average DMU	11,010	140	519,856	4,768,426	5,976.31	1.000

Table 5. Results of proposed PBM approach for case study

DMU	PBM approach							
	$P(x^*, y^*)$	PBM eff. score	θ (Employees)	θ (ships)	θ (Fleet capacity)	φ (Containers carried)	φ (Turnover)	ℓ_2 distance
MSC	1.059	0.819 (4)	0.764	0.734	0.695	1.000	1.056	0.223
CMA CGM	0.965	0.846 (3)	0.671	0.875	0.736	1.000	1.000	0.193
Hapag Lloyd	0.735	0.872 (1)	0.672	0.846	0.886	1.000	1.000	0.144
COSCON	0.740	0.489 (10)	0.047	0.834	0.713	1.000	1.000	1.018
Evergreen Line	0.775	0.787 (5)	0.858	0.724	0.651	1.000	1.337	0.332
APL	0.740	0.870 (2)	0.839	0.756	0.785	1.000	1.000	0.131
CSAV	0.566	0.632 (9)	0.182	0.899	0.903	1.461	1.000	0.902
MOL	0.606	0.642 (8)	0.176	0.863	0.721	1.000	1.000	0.776
Hamburg Süd	0.584	0.691 (6)	0.368	0.786	0.604	1.108	1.000	0.613
YML	0.572	0.660 (7)	0.282	0.578	0.765	1.000	1.000	0.748

Table 6. Results of GDF approach for case study

DMU	GDF approach					
	GDF eff. score	θ (Employees)	θ (ships)	θ (Fleet capacity)	φ (Containers carried)	φ (Turnover)
MSC	0.704 (4)	0.925	0.770	0.817	1.000	1.407
CMA CGM	0.756 (3)	0.671	0.736	0.875	1.000	1.000
Hapag Lloyd	0.796 (1)	0.672	0.886	0.846	1.000	1.000
COSCON	0.300 (10)	0.076	0.767	0.829	1.000	1.477
Evergreen Line	0.629 (5)	1.000	0.663	0.722	1.000	1.548
APL	0.793 (2)	0.839	0.785	0.756	1.000	1.000
CSAV	0.435 (9)	0.208	0.878	1.000	1.447	1.178
MOL	0.478 (8)	0.183	0.717	0.888	1.000	1.042
Hamburg Süd	0.531 (6)	0.425	0.578	0.887	1.091	1.178
YML	0.495 (7)	0.460	0.713	0.866	1.000	1.765

Table 7. Results of proposed ERGM approach for case study

DMU	ERGM approach					
	ERGM eff. score	θ (Employees)	θ (ships)	θ (Fleet capacity)	φ (Containers carried)	φ (Turnover)
MSC	0.691 (4)	0.990	0.800	0.850	1.000	1.547
CMA CGM	0.761 (3)	0.671	0.736	0.875	1.000	1.000
Hapag Lloyd	0.797 (1)	0.886	0.933	0.842	1.000	1.225
COSCON	0.437 (10)	0.105	0.909	1.000	1.000	2.076
Evergreen Line	0.624 (5)	1.000	0.663	0.722	1.000	1.548
APL	0.794 (2)	0.839	0.785	0.756	1.000	1.000
CSAV	0.530 (8)	0.208	0.878	1.000	1.447	1.178
MOL	0.552 (7)	0.666	0.983	0.949	1.000	2.137
Hamburg Süd	0.555 (6)	0.425	0.578	0.887	1.091	1.178
YML	0.492 (9)	0.460	0.713	0.866	1.000	1.765

Table 8. Intensity variables and benchmarks for the inefficient DMUs

DMU	Benchmarks							
	Maersk	**CSCL**	**OOCL**	**NYK**	**K Line**	**HMM**	**Wan Hai**	**Delmas**
MSC	0.506	0.494						
CMA CGM	0.312	0.403			0.285			
Hapag Lloyd		0.394		0.279	0.327			
COSCON		0.532		0.411	0.057			
Evergreen Line		0.719		0.281				
APL		0.473		0.293	0.234			
CSAV						0.391	0.609	
MOL				0.491		0.452	0.057	
Hamburg Süd						0.784	0.216	
YML				0.155		0.224	0.622	

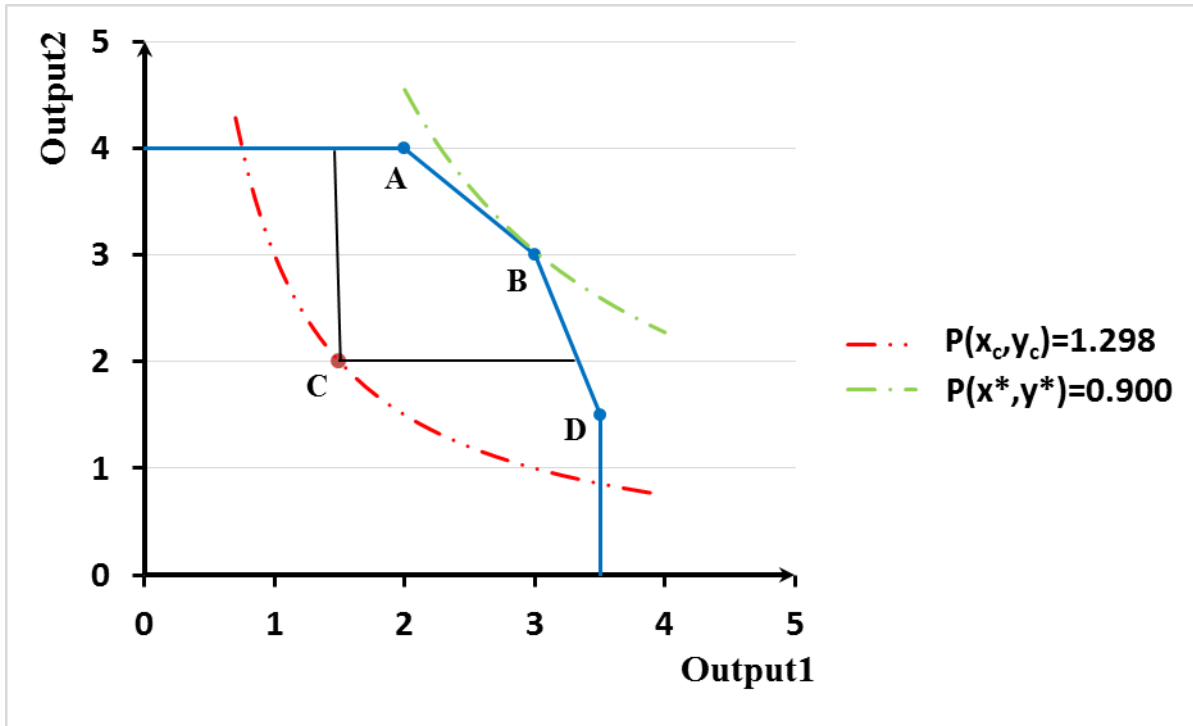Figure 1. Minimum efficiency potential of DMU C of 1YY illustrative example

Figure 2. Minimum efficiency potential of DMU C of XY illustrative example
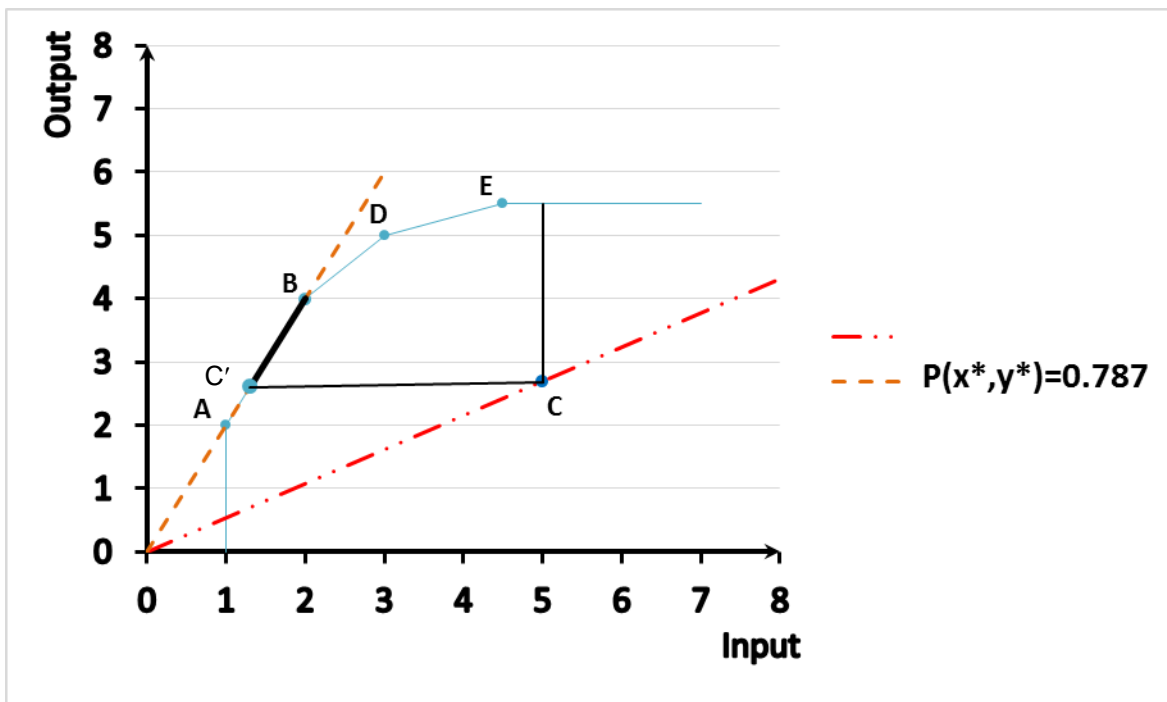
Figure 3. Minimum efficiency potential target for DMU C of XY illustrative example
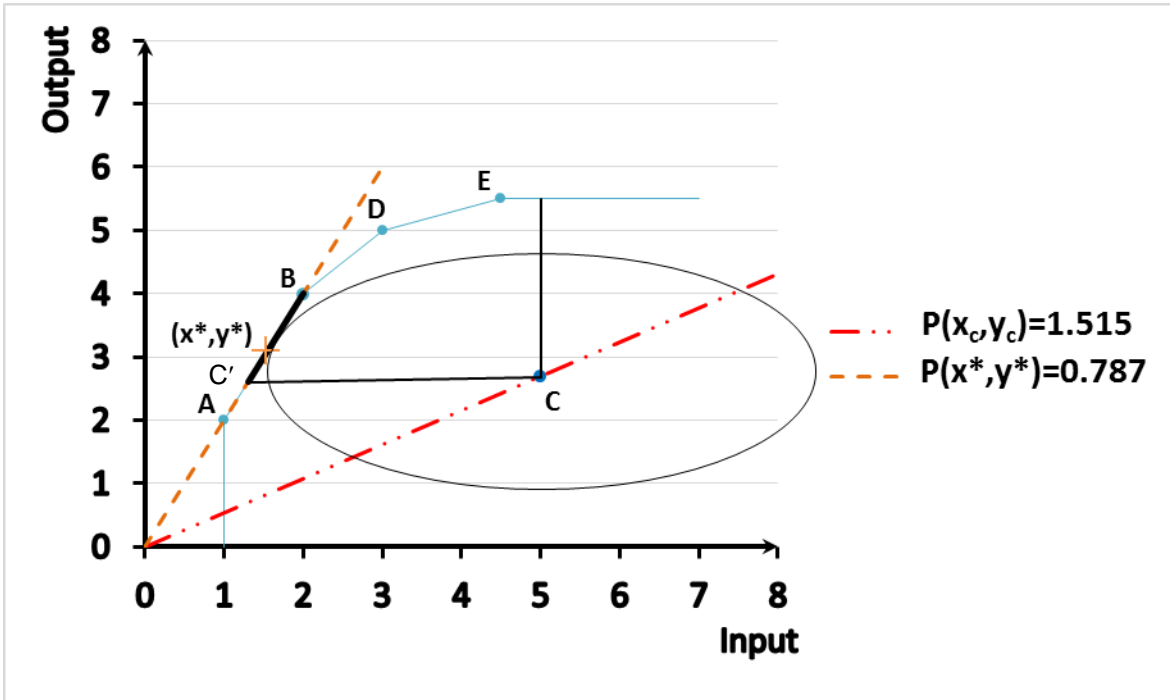
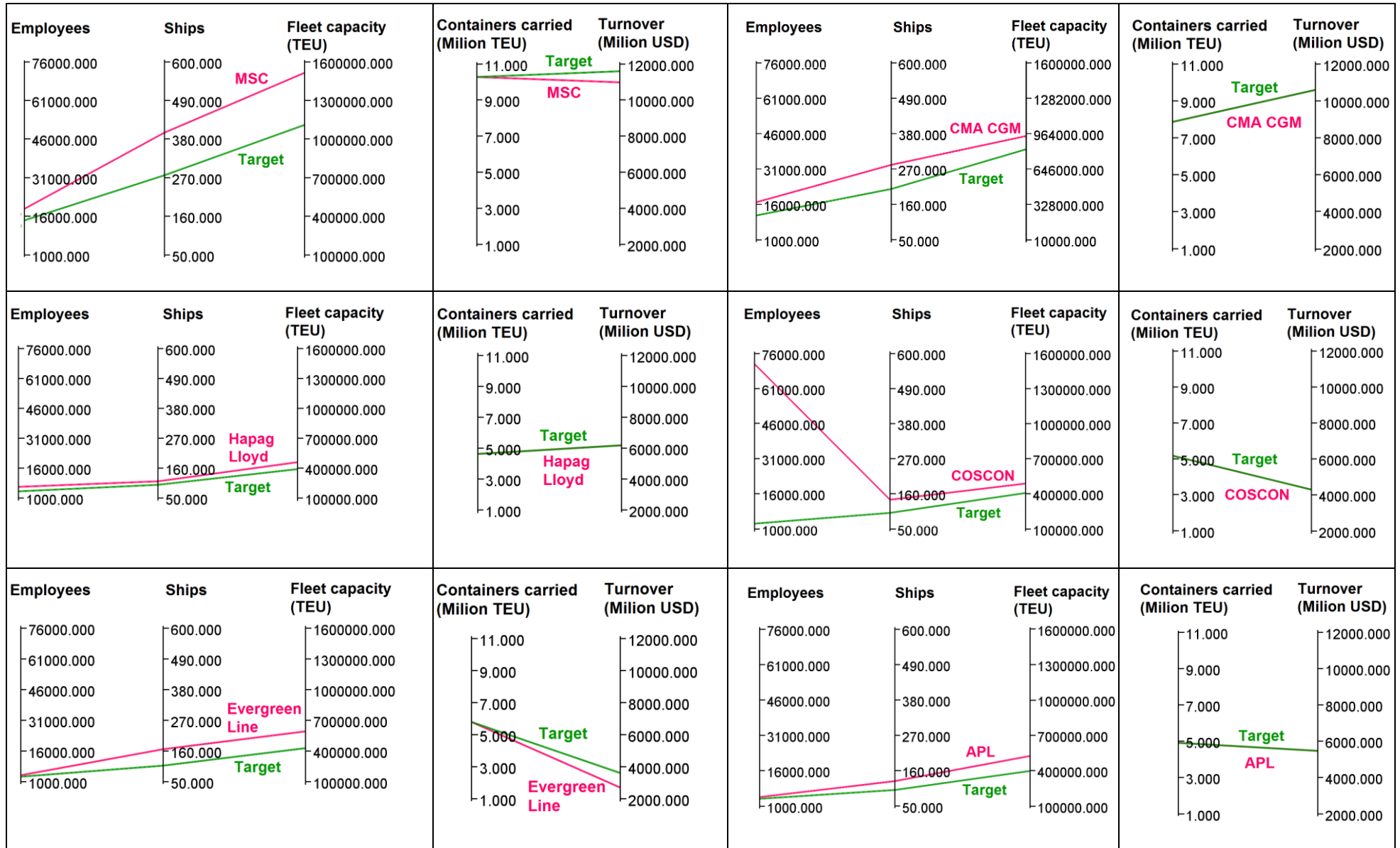Figure 4. Observed and target values for the inefficient DMUs

Figure 4. Observed and target values for the inefficient DMUs (cont.)