# Lexicographic hyperbolic DEA

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

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

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

ARTÍCULO PUBLICADO EN LA REVISTA

*Journal of the Operational Research* Society (2019)

doi: 10.1080/01605682.2019.1599704

*Abstract*

The hyperbolic distance function (HDF) reduces all inputs and increases all outputs simultaneously and at the same rate. Although the corresponding data envelopment analysis (DEA) model is non-linear, for constant returns to scale it can be linearized and for variable returns to scale an efficient iterative approach based on the directional distance function (DDF) model can be used. However, HDF does not necessarily project onto an efficient target. To remedy this, lexicographic hyperbolic DEA (LexHDEA) is proposed in this paper. Thus, before solving the HDF model, the input or output dimensions that can be improved are determined. A reduced HDF model is then solved, looking for improvements only in these dimensions. If the corresponding target is efficient, then no further steps are necessary. Otherwise, a reduced HDF model that improves only those dimensions that can be further improved is solved. If this improved target is efficient the process stops. Otherwise the process is repeated until eventually the efficient frontier is reached. In addition to guaranteeing an efficient target the proposed approach also computes an efficiency measure that has indication of efficiency and units invariance. The proposed approach can be extended to handle a preference structure, non-discretionary variables and undesirable outputs.

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

Keywords: Data envelopment analysis; hyperbolic distance function; geometric distance function; lexicographic projection; lexicographic hyperbolic efficiency measure

## 1. Introduction

Data envelopment analysis (DEA) is an optimization-based mathematical approach commonly used to benchmark comparable operating units. The operating units being benchmarked are generally called decision making units (DMUs). DEA is a non-parametric tool that only uses data on the inputs consumed and the outputs produced by the operating units. The first step of the methodology is inferring the Production Possibility Set (PPS), i.e. the set of all feasible operating points. The PPS contains the observed data plus all linear combinations of them plus the operating points that they dominate. When Constant Returns to Scale (CRS) is assumed the operating points generated by upward and downward scaling of the linear combinations of the observations are also considered (Cooper et al. 2004, 2006).

The efficient frontier is the subset of non-dominated operating points within the PPS. Most DEA models consist in projecting the observed DMUs onto the efficient frontier. This is done maximizing some criterion that looks for improvements in the amount of inputs consumed and output produced by the DMU. If no such improvements are possible then the DMU is efficient. Otherwise an efficient target that dominates the DMU is computed and an efficiency score that measures the amount of improvement that can be achieved is provided. Some so-called least-distance DEA models (e.g. Aparicio et al. 2007, Aparicio 2016) aim at computing the efficient target that is closest to the DMU, something which can be desirable as it reduces the effort required to achieve efficiency. An alternative strategy is to compute a path to the efficient frontier that involves successive intermediate targets (e.g. Ghahraman and Prior 2016, Lozano and Calzada-Infante 2018).

There are different DEA models that differ in their orientation (input, output or non-oriented) as well as in the way the improvements along the different dimensions are sought (radial, non-radial, etc.). Radial DEA models, both oriented (e.g. Charnes et al. 1978, Banker et al. 1984) and non-oriented, like DDF (Chambers et al. 1996), can lead to weak efficient targets that do not exhaust all possible input and output improvements. The conventional way of dealing with this drawback is to carry out a phase II that maximizes the remaining input and output slacks. One minor problem of such approach is the possibility of alternative optima. An alternative approach is to carry out a lexicographic approach. Coelli (1998) was the first to propose such a multi-stage methodology using radial oriented models. More recently, Korhonen et al. (2018) and Lozano

and Soltani (2018a) have extended this multi-stage methodology to non-oriented radial and to DDF DEA models, labelling those approaches lexicographic radial and lexicographic DDF, respectively. The lexicographic Kalai-Smorodinsky and the lexicographic egalitarian DEA bargaining approaches in Lozano et al. (2018) can also be included in this family.

In this paper, the multi-stage methodology is used to enhance the hyperbolic distance function (HDF) proposed in Färe et al. (1985). Färe et al. (2016) enumerates a number of areas in which HDF has been used. One of the advantages of HDF over other approaches like DDF is that the projection path is based only on the inputs and outputs of the DMU being projected and therefore it does not require subjective information about the projection direction. Some other DEA approaches, like, for example, Slacks-based measure of efficiency (SBM, Tone 2001) are also based on the inputs and outputs of the DMU being projected and therefore do not require any exogenous information either, but they are non-radial and hence the projection direction is unknown a priori. Similarly, endogenous DDF approaches, like, for example, the largest improvement DDF approaches of Färe et al. (2013) and Hampf and Krüger (2015), do not require subjective indication of the projection direction but at the cost of the unpredictability of that projection direction. Hence, HDF has the desirable features of not requiring additional information and using a defined projection path. An additional argument in support of HDF is that it is the closest non-oriented relative of the mix-preserving radial oriented DEA models (Hasannasab et al. 2018). However, since HDF is in some sense a merger of the input and output radial oriented DEA models it has the same drawback as those models, i.e. the projection is not necessarily efficient. That is why it can benefit from a lexicographic approach that keep on improving on the input and output dimensions that still have slacks using, to that end, the hyperbolic criterion of using the same rate to improve every input and output. It is true that the proposed lexicographic approach introduces some deviation from the strict hyperbolic path but the proposed lexicographic continuation of this hyperbolic path uses the same hyperbolic improvement criterion in its movement along the weak efficient frontier and hence the projection path is still clearly defined. In other words, the proposed lexicographic hyperbolic projection coincides with the HDF projection if this is efficient but otherwise it computes a unique efficient target by resuming the hyperbolic trajectory along the subspace spanned by the input and output dimensions with remaining inefficiencies.

The structure of this paper is the following. First, in Section 2, the hyperbolic DEA model is presented. Section 3 presents the proposed lexicographic hyperbolic DEA (LexHDEA) approach which is illustrated, in Section 4, using a small dataset. Section 5 applies LHDEA to assess the efficiency of 26 organic cereal farms in Spain. Section 6 summarizes and concludes.

## 2. Relevant literature review

In this section the relevant literature, namely the hyperbolic distance function, the geometric distance function and the existing lexicographic approaches are reviewed. Färe et al. (1985) proposed the following model to compute the hyperbolic graph efficiency measure of DMU 0

$$
\begin{aligned}
&\theta_0^{HDF} = \text{Min} \quad \theta \\
&\text{s.t.} \\
&\sum_{j=1}^{n} \lambda_j x_{ij} \le \theta \cdot x_{i0} \quad \forall i \\
&\sum_{j=1}^{n} \lambda_j y_{kj} \ge \frac{1}{\theta} \cdot y_{k0} \quad \forall k \\
&\sum_{j=1}^{n} \lambda_j = 1 \\
&\lambda_j \ge 0 \ \forall j \quad \theta \text{ free}
\end{aligned}
\tag{1}
$$

where $x_{ij}$ and $y_{kj}$ are the amount of input i consumed and the amount of output k produced, respectively, by each DMU j and the vector $(\lambda_1, \lambda_2, ..., \lambda_n)$ represents the intensity variables used to form the linear combination of the observed DMUs.

This DEA model tries to simultaneously reduce the inputs and increase the outputs and does that using a uniform improvement factor for all. Although the efficiency score $\theta$ is declared as a free variable, it is clear that $0 \le \theta_0^{HDF} \le 1$. The target computed by this model is

$$\hat{x}_i = \sum_{j=1}^{n} \lambda_j x_{ij} \quad \forall i$$
$$\hat{y}_k = \sum_{j=1}^{n} \lambda_j y_{kj} \quad \forall k \tag{2}$$

Note that model (1) corresponds to the Variable Returns to Scale (VRS) case and that it is a non-linear program. Färe et al. (1985) showed that the corresponding CRS formulation (i.e. without the convexity constraint on the intensity variables) can be linearised and it is equivalent to the classical radial DEA model of Charnes et al. (1978). As regards model (1), Färe et al. (2016) have proposed an ingenious iterative solution approach that only needs to solve a finite sequence of linear programming (LP) models of DDF type. They present numerical results on real-world and simulated data confirming that the iterative DDF approach computes the exact HDF scores and it is rather fast, requiring most of the times at most two iterations to converge. An alternative approach has been recently proposed by Hasannasab et al. (2018) using a conic programming formulation to solve model (1) in an almost LP computation time. They also present a dual multiplier formulation and show the relationships between the corresponding shadow prices and the returns to scale and scale elasticity.

Another DEA approach that is related to the proposed lexicographic hyperbolic efficiency measure is the geometric distance function (GDF) of Portela and Thanassoulis (2007), which, for comparison, is formulated below.

<u>GDF DEA model</u>

$$\zeta_0^{GDF} = \text{Min} \; \frac{\left( \prod_{i=1}^{m} \alpha_i \right)^{1/m}}{\left( \prod_{k=1}^{s} \beta_k \right)^{1/s}}$$
$$\text{s.t.} \quad \sum_{j=1}^{n} \lambda_j x_{ij} \le \alpha_i \cdot x_{io} \qquad i = 1, \ldots, m$$
$$\sum_{j=1}^{n} \lambda_j y_{kj} \ge \beta_k \cdot y_{ko} \qquad k = 1, \ldots, s \tag{3}$$
$$\lambda_j \ge 0 \;\; \forall j \quad \alpha_i \le 1 \;\; \forall i \quad \beta_k \ge 1 \;\; \forall k$$

In the above model each input i is reduced by a specific factor $\alpha_i$ and each output k is increased by a specific factor $\beta_k$ and the efficiency score, i.e. the GDF, is computed as the ratio of the geometric mean of the inputs and of the output factors. As one of the reviewers rightly pointed out, the proposed LexhDEA and SBM, which is equivalent to the Enhanced Russell Graph Measure (ERM, Pastor et al. 1999) should be highly correlated as ERM minimizes the ratio of the arithmetic average of the input reductions to the arithmetic average of the output increases while the proposed LexHDEA approach minimizes, in a lexicographic way, the uniform input reductions and output increases. Actually, as it is shown in Section 3, due to the multiplicative character of the hyperbolic criterion, the proposed is even closer to the GDF approach (3) which minimizes the ratio of the geometric mean of the input reductions to the geometric mean of the output increases.

As regards the existing lexicographic DEA approaches, they have been applied to radial oriented models (Coelli 1998), radial non-oriented model (Korhonen et al. 2018), DDF model (Lozano and Soltani 2018a) and to the Kalai-Smorodinsky and egalitarian bargaining solutions (Lozano et al. 2018). What all these models have in common is that they do not project onto the efficient frontier and hence they may lead to weak efficient targets with input and output slacks. Hence, the lexicographic approach iterates solving a reduced version of the corresponding model that only keeps on improving the inputs and outputs that can be further improved.

Finally, let us comment on the comparability issue in DEA which comes from the fact that an inefficient DMU can be legitimately projected onto any efficient reference point that dominates it. And, being the efficiency score a measure of the distance to the efficient frontier, the efficiency score would be different depending on the specific target chosen. Actually, this has motivated the Multidirectional Efficiency Measure (MEM) of Lozano and Soltani (2018b). In any case, the closest-target and the largest improvement DEA approaches are two extreme projections while every other DEA approach, e.g. the proposed LexHDEA, DDF, SBM, etc, always lies somewhere in between, depending on the specific projection criterion chosen. Although this leads to a certain comparability issue it can be mitigated comparing the proposed approach with that which is conceptually closer, which, as indicated above, in our case is the GDF approach.

## 3. Proposed lexicographic hyperbolic DEA approach

The proposed lexicographic hyperbolic DEA (LexHDEA) approach carries out a sequence of steps. Since in each step of the lexicographic iterative process at least one variable reaches its limit and cannot improve further, the iterative process has a finite number of steps that cannot be greater than the total number of variables. In order to formulate them mathematically, let

$\varepsilon_i$ Desired precision level of input i (e.g. if input i has 2 decimal digits then $\varepsilon_i = 10^{-2}$)

$\hat{\varepsilon}_k$ Desired precision level of output k

$I^-_{t-1}$ Set of inputs that could still be improved in step t-1 (initially, it is assumed that all inputs can be improved, i.e. $I^-_0 = \{1, 2, ..., m\}$

$O^+_{t-1}$ Set of inputs that could still be improved in step t-1 (initially, it is assumed that all inputs can be improved, i.e. $O^+_0 = \{1, 2, ..., s\}$

$x_i^{t-1}$ Intermediate target value of input i after step t-1 (initially, $x_i^0 = x_{i0}$)

$y_k^{t-1}$ Intermediate target value of output k after step t-1 (initially, $y_k^0 = y_{k0}$)

In each step t, in order to compute $I^-_t$ and $O^+_t$, a mixed-integer linear program (MILP) is solved in the first place. The following binary variables indicate whether or not each input and output can be further improved

$u_i$ binary variable indicating whether or not input $i \in I^-_{t-1}$ can be reduced in iteration t

$v_k$ binary variable indicating whether or not output $k \in O^+_{t-1}$ can be increased in iteration t

The MILP model for determining which input and output dimensions can still be improved in step t can be formulated as

$$\text{Max} \quad \sum_{i \in I_{t-1}^-} u_i + \sum_{k \in O_{t-1}^+} v_k$$

s.t.

$$\begin{aligned}
&\sum_{j=1}^{n} \lambda_j x_{ij} \le x_i^{t-1} - \varepsilon_i \cdot u_i \quad \forall i \in I_{t-1}^- \\
&\sum_{j=1}^{n} \lambda_j x_{ij} \le x_i^{t-1} \qquad \forall i \notin I_{t-1}^- \\
&\sum_{j=1}^{n} \lambda_j y_{kj} \ge y_k^{t-1} + \hat{\varepsilon}_k \cdot v_k \quad \forall k \in O_{t-1}^+ \\
&\sum_{j=1}^{n} \lambda_j y_{kj} \ge y_k^{t-1} \qquad \forall k \notin O_{t-1}^+ \\
&\sum_{j=1}^{n} \lambda_j = 1 \\
&\lambda_j \ge 0 \ \forall j \quad u_i \in \{0,1\} \ \forall i \in I_{t-1}^- \quad v_k \in \{0,1\} \ \forall k \in O_{t-1}^+
\end{aligned} \tag{4}$$

This model detects all the input and outputs that can be improved in step t, i.e. that can be improved with respect to the step t-1 intermediate target $\left(x_i^{t-1}, y_k^{t-1}\right)$. Note that the inputs not in $I_{t-1}^-$ as well as the outputs not in $O_{t-1}^+$ could not be improved in step t-1 and, hence, they cannot be improved either in step t. Those inputs and outputs maintain the level they have reached, $x_i^{t-1}$ and $y_k^{t-1}$, respectively. For those inputs and outputs no $u_i$ or $v_k$ are needed because their value would be zero as they have reached their maximum possible improvement. After solving (8), the subsets of inputs and outputs that can be improved in step t are determined as:

$$I_t^- = \left\{ i \in I_{t-1}^- : u_i^* = 1 \right\} \qquad O_t^+ = \left\{ k \in O_{t-1}^+ : v_k^* = 1 \right\} \tag{5}$$

Once the sets $I_t^-$ and $O_t^+$ have been computed, the following reduced hyperbolic DEA model is solved

$$
\begin{aligned}
&\theta^t = \text{Min} \quad \theta \\
&\text{s.t.} \\
&\sum_{j=1}^{n} \lambda_j x_{ij} \le \theta \cdot x_{i0} \qquad \forall i \in I_t^- \\
&\sum_{j} \lambda_j x_{ij} \le x_i^{t-1} \qquad \forall i \notin I_t^- \\
&\sum_{j=1}^{n} \lambda_j y_{kj} \ge \frac{1}{\theta} \cdot y_{k0} \qquad \forall k \in O_t^+ \\
&\sum_{j} \lambda_j y_{kj} \ge y_k^{t-1} \qquad \forall k \notin O_t^+ \\
&\sum_{j=1}^{n} \lambda_j = 1 \\
&\lambda_j \ge 0 \ \forall j \quad \theta \text{ free}
\end{aligned}
\tag{6}
$$

The intermediate step t target computed by the above model is

$$
\begin{aligned}
&x_i^t = \theta^t \cdot x_{i0} \qquad \forall i \in I_t^- \\
&x_i^t = x_i^{t-1} \qquad \forall i \notin I_t^- \\
&y_k^t = \frac{1}{\theta^t} \cdot y_{k0} \qquad \forall k \in O_t^+ \\
&y_k^t = y_k^{t-1} \qquad \forall k \notin O_t^+
\end{aligned}
\tag{7}
$$

Note that in the above model the hyperbolic improvement is maximized but only in the $I_t^-$ and $O_t^+$ dimensions. Since, by construction, $I_t^- \subseteq I_{t-1}^-$ and $O_t^+ \subseteq O_{t-1}^+$, it follows that $\theta^t \le \theta^{t-1}$. Moreover, since in step t-1 at least one input $i \in I_{t-1}^-$ or one output in $k \in O_{t-1}^+$ limits the maximum reduction of $\theta$ in model (6), i.e. in the optimum at least one of the constraints involving $\theta$ must be binding, it follows that $I_t^- \subset I_{t-1}^-$ and/or $O_t^+ \subset O_{t-1}^+$. Moreover, the successive improvement factors $\theta^t$ form a strictly decreasing sequence. This is so because, by construction, all the input and output dimensions that remain in the sets $I_t^-$ and $O_t^+$ can be further improved with respect to

the step t-1 intermediate target $\left(x_i^{t-1}, y_k^{t-1}\right)$ and therefore in model (6) an additional reduction of $\theta^t$ over the reduction already obtained in the previous step ($\theta^{t-1}$) is feasible. Hence, $\theta^t < \theta^{t-1}$.

Since the cardinality of at least one of the sets $I_t^-$ and $O_t^+$ decreases in each iteration, it follows that after a certain number of steps $T \le m+s$ the process stops because solving model (4) for t=T+1 the optimal solution is zero, which leads to $I_{T+1}^- = O_{T+1}^+ = \varnothing$. That means that no input or output dimension can be further improved and hence the step T target $\left(x^T, y^T\right)$ is efficient. This is clear, since if the step T target $\left(x^T, y^T\right)$ were not efficient, then solving model (4) for t=T+1 would detect that some input and/or output dimensions could be further improved and hence $I_{T+1}^- \neq \varnothing$ or $O_{T+1}^+ \neq \varnothing$, contrary to what we have assumed.

Note that, as it can be seen in the numerical results shown in Sections 4 and 5, during the iterative process one of the sets $I_t^-$ and $O_t^+$ can become empty before the other does. That is no problem and can be perfectly handled by model (6). Actually, when that happens, it only means that, from that iteration on, and because the dimensions that can still be improved are all inputs or all outputs, the lexicographic continuation of the hyperbolic path implies following, for the last steps, a lexicographic radial oriented approach similar to the one originally proposed in Coelli (1998).

If we denote by $T(i)$ and $\hat{T}(k)$ the last step in which each input i and output k, respectively, improved, i.e.

$$T(i) = \max_t \left\{ t : i \in I_t^- \right\} \qquad \hat{T}(k) = \max_t \left\{ t : k \in O_t^+ \right\} \tag{8}$$

then the proposed LexHDEA target can be computed as

$$\begin{aligned} x_i^T &= x_i^t = x_i^{T(i)} = \theta^{T(i)} \cdot x_{i0} \qquad \forall i \; \forall T(i) \le t \le T \\ y_k^T &= y_k^t = y_k^{\hat{T}(k)} = \frac{1}{\theta^{\hat{T}(k)}} \cdot y_{k0} \qquad \forall i \; \forall \hat{T}(k) \le t \le T \end{aligned} \tag{9}$$

which means that $\theta_i = \theta^{T(i)}$ and $\theta_k = \theta^{\hat{T}(k)}$ represent the improvement factors, i.e. the efficiency score along each input and output dimension, respectively. The corresponding LexHDEA efficiency score of DMU 0 is defined as

$$\xi_0^{LexHDEA} = \frac{\left(\prod_{i=1}^{m} \frac{x_i^T}{x_{i0}}\right)^{\frac{1}{m}}}{\left(\prod_{k=1}^{s} \frac{y_k^T}{y_{k0}}\right)^{\frac{1}{s}}} = \frac{\left(\prod_{i=1}^{m} \theta^{T(i)}\right)^{\frac{1}{m}}}{\left(\prod_{k=1}^{s} \frac{1}{\theta^{\hat{T}(k)}}\right)^{\frac{1}{s}}} = \frac{\left(\prod_{i=1}^{m} \theta_i\right)^{\frac{1}{m}}}{\left(\prod_{k=1}^{s} \frac{1}{\theta_k}\right)^{\frac{1}{s}}} \tag{10}$$

Note that the above way of measuring the efficiency resembles the geometric distance function (GDF) metric where the numerator is the geometric mean of the input improvements and the numerator is the geometric mean of the output improvements. The difference between (10) and (3) is that in (3) the target is computed so as to minimize this metric while in (10) the target is computed using the lexicographic hyperbolic path and the metric is used to measure the distance to the efficient frontier, i.e. the level of improvement from the observed DMU to the target. It follows that $\zeta_0^{GDF} \leq \xi_0^{LexHDEA}$ and that $\zeta_0^{GDF} = \xi_0^{LexHDEA} = 1$ if and only if DMU 0 is efficient.

Note also that the hyperbolic graph efficiency measure $\theta_0^{HDF}$ corresponds to the first step of the proposed lexicographic hyperbolic DEA approach. More specifically,

$$\theta_0^{HDF} = \begin{cases} \theta^1 & \text{if } \left|I_1^-\right| + \left|O_1^+\right| = m + s \\ 1 & \text{otherwise} \end{cases} \tag{11}$$

Hence, since $\theta^{T(i)} \leq \theta_0^{HDF} \quad \forall i$ and $\theta^{\hat{T}(k)} \leq \theta_0^{HDF} \quad \forall k$, it follows that

$$\xi_0^{\text{LexHDEA}} = \frac{\left(\prod_{i=1}^{m} \theta^{T(i)}\right)^{\frac{1}{m}}}{\left(\prod_{k=1}^{s} \frac{1}{\theta^{\hat{T}(k)}}\right)^{\frac{1}{s}}} \leq \frac{\left(\prod_{i=1}^{m} \theta_0^{HDF}\right)^{\frac{1}{m}}}{\left(\prod_{k=1}^{s} \frac{1}{\theta_0^{HDF}}\right)^{\frac{1}{s}}} = \left(\theta_0^{HDF}\right)^2 \tag{12}$$

Finally, note that the reduced hyperbolic DEA model (6) is non-linear. However, the iterative DDF-based algorithm of Färe et al. (2016) can be appropriately extended to solve (6) using only LP models. Alternatively, a non-linear optimization solver can be used to directly solve (6).

## 4. Numerical example

In order to illustrate the proposed approach, consider the two-input, single-output dataset shown in Table 1. VRS has been assumed. Table 1 also shows the corresponding VRS hyperbolic graph efficiency measure $\theta_0^{HDF}$ and GDF score $\theta_0^{GDF}$. Five DMUs, namely A, B, F, G and H, are efficient. Of the five are inefficient, two are weak efficient (namely, DMUs E and J). This can be noted in that their HDF score is one but its GDF score is less than one.

========== Table 1 ==========

Table 2 shows the results of the application of the LexHDEA approach to DMU C. For the MILP model (8) $\varepsilon_1 = \varepsilon_2 = \hat{\varepsilon}_1 = 0.001$ have been used. For the successive reduced hyperbolic DEA model (6) both a non-linear solver and the iterative DDF algorithm of Färe et al. (2016) have been used. The steps of the lexicographic approach are the same. The only difference is that the non-linear solver computes the optimal solution of (6) directly while the iterative DDF algorithm sometimes requires more than one iteration, before finding the optimal value of $\theta^t$. For the sake of completeness, these internal iterations of the iterative DDF algorithm are shown in Table 2. Thus in the first and in the third steps only one iteration of DDF algorithm is required. Conversely, in step t=2, as the first iteration of step 2 lead to an infeasible target, a second iteration was required to find the optimal value $\theta^2 = 0.476$. The computing times of the LexHDEA approach for both the non-linear solver and the iterative DDF algorithm and for the GDF score are reported. It can

be noted that $\theta_C^{HDF} = \theta^1 = 0.667 > \theta^2 = 0.476 > \theta^3 = 0.333$. Note also that, for this DMU, the final target computed by LexHDEA coincides with that of GDF model (3) and, hence, $\zeta_C^{GDF} = \xi_C^{LexHDEA} = 0.224$. Note also that, as indicated in section 3, $\xi_C^{LexHDEA} = 0.224 \leq \left(\theta_C^{HDF}\right)^2 = (0.667)^2 = 0.444$.

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

The successive intermediate targets of the proposed LexHDEA approach for DMU C are graphically shown in Figure 1. Note that the internal iterations of the iterative DDF algorithm are not shown since they are not inherent to the LexHDEA approach; they are just auxiliary, intermediate computations of the iterative DDF algorithm.

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

The numerical results and the visualization of the successive steps of the LexHDEA path for the other inefficient DMUs are shown in the supplementary material file accompanying this paper.

## 5. Case study

In this section the proposed LexHDEA approach is applied to 26 organic cereal farms in Spain. The DMUs being benchmarked is 1 ha. of the corresponding farms. Three inputs are used, namely Fuel consumption (FC, measured in liters/ha), Total Carbon input (TCI, kg C/ha) and Total Nitrogen input (TNI, kg N/ha). The input TCI is considered non-discretionary. There are two outputs, namely Yield fresh matter (YFM, Mg/ha) and Net primary production (excluding fresh matter)(NPP, Mg dry matter/ha). One undesirable output, Total area-based greenhouse gases emissions (TAE, measured in kg CO2eq/ha) is also considered. Additional information on this dataset is provided in Gutiérrez et al. (2017).

Since this application involves VRS, a non-discretionary input and an undesirable output, the extensions of the proposed approach to this situation is used. This is presented in Appendix A. For completeness and for correctness checking, the extension of the iterative DDF-based algorithm of Färe et al. (2016) to this situation is also presented in Appendix B.

Table 3 shows the observed inputs and outputs as well as the corresponding GDF target and efficiency score. The GDF computing time is also reported. The target for input TCI is not shown because that input is non-discretionary and, therefore, the corresponding target coincides with the observed value. Note that 14 out of the 26 DMUs are efficient.

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

The proposed LexHDEA approach was applied to all DMUs. For the efficient ones the first iteration detected that $I_1^- = O_1^+ = \varnothing$ and the process stop. For the inefficient DMUs a lexicographic hyperbolic path was computed. Figure 2 shows the successive $\theta^t$ values for some DMUs, together with the variables that could not improve more once that point is reached. Thus, for example, for DMU O3, in the first step all discretionary variables improved according to $\theta^1 = 0.85372$ which means a 14.63% reduction for the inputs and the undesirable output and an increase of $\frac{1}{\theta^1} - 1 = 17.13\%$ for the outputs. After that first step input FC cannot improve more so the hyperbolic projection is continued but without that input. The corresponding variables improved a little bit up to $\theta^2 = 0.85393$ until the other discretionary input (TNI) cannot improve more. From that moment only the two outputs and the undesirable output can improve further and they do so until a value $\theta^3 = 0.84034$ is reached. From that moment only one output (YFM) can improve and does so until a final value $\theta^4 = 0.80622$ is attained. At that point no slacks remain and the efficient frontier has been reached. It can be noted that, in general, the variables with the smallest margin for improvement are two discretionary inputs (FC, TNI) while the output YFM is the one that can improve furthest.

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

Table 4 shows the intermediate targets computed by the LexHDEA approach for DMU O3. As it can be seen from the precision level of the targets a value of 0.01 has been used for all the variables. Note that, in every step, the iterative DDF has required one iteration only. Note also that the successive $\theta^t$ values are strictly decreasing, i.e. $\theta_{O3}^{HDF} = \theta^1 > \theta^2 > \theta^3 > \theta^4$ and that $\zeta_{O3}^{GDF} = 0.48464 \leq \xi_C^{LexHDEA} = 0.59029$. Note also that, as indicated in Appendix A,

$\xi_{O3}^{LexHDEA} = 0.59029 \leq \left(\theta_{O3}^{HDF}\right)^3 = (0.85372)^3 = 0.62222$. The above inequalities hold for all DMUs as it can be seen in Figure 3, which shows a plot of LexHDEA score $\xi_0^{LexHDEA}$ versus GDF score $\zeta_0^{GDF}$ and versus the third power of the HDF score $\left(\theta_0^{HDF}\right)^3$. Table 4 also reports the computing times for the LexHDEA approach using the non-linear solver and the iterative DDF algorithm. Similar tables for the other inefficient DMUs are included in the supplementary material file accompanying this paper.

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

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

## 6. Conclusions

In this paper, a lexicographic projection approach based on the hyperbolic DEA model is proposed. Its motivation is to guarantee the efficiency of the computed target. The proposed approach follows, as closely as possible, the hyperbolic graph path, deviating from it only when no further improvements along a certain dimension are feasible. This allows exhausting the input and output slacks that typically exist after the conventional hyperbolic graph projection and it is therefore an alternative to the common phase II that maximizes the sum of those slacks, an approach that does not guarantee a unique solution. A LexHDEA efficiency score has also been proposed. Same as HDF and GDF, with which it can be related, this LexHDEA score is units invariant. Moreover, as shown in the case study presented, it can be extended to take into account non-discretionary variables and undesirable outputs. Also, for the successive projection steps, the corresponding reduce hyperbolic DEA model can be directly solved using a non-linear solver or the iterative DDF algorithm proposed in Färe et al. (2016).

Possible continuations of this research include extending the approach to hyperbolic network DEA (HNDEA) models (e.g. Yu and Lee 2009). How to handle integer variables (e.g. Lozano and Villa 2006) can also be studied. Finally, another research endeavor would be applying the lexicographic approach to the directional hyperbolic distance function (e.g. Roshdi et al. 2018).

## Appendix A

In this appendix the modifications of the basic LexHDEA approach presented in section 3 for the case of some inputs being non-discretionary and some outputs undesirable is presented. Also, VRS is assumed. Let $I^D$ and $I^{ND}$ the set of discretionary and non-discretionary inputs, respectively and $B$ the set of undesirable outputs. Let also $z_{bj}$ be the value of the undesirable outputs produced by DMU j for each $b \in B$. We will assume joint weak disposability of desirable and undesirable outputs, which will be modelled using specific abatement factors for each DMU as per Kuosmanen (2005).

The first modification of the proposed approach to handle this situation is to exclude the non-discretionary variables from the set of dimensions that can be improved using the hyperbolic path. This can be done setting $I_0^- = I^D$ initially. The other two sets that indicate the desirable and undesirable outputs that can be improved are set initially to $O_0^+ = \{1,2,...,s\}$ and $B_0^- = B$, respectively.

Assuming that $\left(x^0, y^0, z^0\right) = \left(x_0, y_0, z_0\right)$, the model for identifying the dimensions that can be improved in each step t has to be modified also to include the undesirable outputs as follows

$$\text{Max} \quad \sum_{i \in I_{t-1}^-} u_i + \sum_{k \in O_{t-1}^+} v_k + \sum_{b \in B_{t-1}^-} w_b$$

s.t.

$$\begin{aligned}
&\sum_{j=1}^{n} \left(\lambda_j + \mu_j\right) \cdot x_{ij} \le x_i^{t-1} - \varepsilon_i \cdot u_i \quad \forall i \in I_{t-1}^- \\
&\sum_{j=1}^{n} \left(\lambda_j + \mu_j\right) \cdot x_{ij} \le x_i^{t-1} \qquad \forall i \notin I_{t-1}^- \\
&\sum_{j=1}^{n} \lambda_j y_{kj} \ge y_k^{t-1} + \hat{\varepsilon}_k \cdot v_k \quad \forall k \in O_{t-1}^+ \\
&\sum_{j=1}^{n} \lambda_j y_{kj} \ge y_k^{t-1} \qquad \forall k \notin O_{t-1}^+ \\
&\sum_{j=1}^{n} \lambda_j z_{bj} = z_b^{t-1} - \hat{\hat{\varepsilon}}_b \cdot w_b \quad \forall b \in B_{t-1}^- \\
&\sum_{j=1}^{n} \lambda_j z_{bj} = z_b^{t-1} \quad \forall b \notin B_{t-1}^- \\
&\sum_{j=1}^{n} \left(\lambda_j + \mu_j\right) = 1 \\
&\lambda_j, \mu_j \ge 0 \ \forall j \quad u_i \in \{0,1\} \ \forall i \in I_{t-1}^- \quad v_k \in \{0,1\} \ \forall k \in O_{t-1}^+ \quad w_b \in \{0,1\} \ \forall b \in B_{t-1}^-
\end{aligned} \tag{A.1}$$

The corresponding $I_t^-$, $O_t^+$ and $B_t^-$ sets are computed as

$$I_t^- = \left\{ i \in I_{t-1}^- : u_i^* = 1 \right\} \qquad O_t^+ = \left\{ k \in O_{t-1}^+ : v_k^* = 1 \right\} \qquad B_t^- = \left\{ b \in B_{t-1}^- : w_b^* = 1 \right\} \tag{A.2}$$

so that the stop condition is $I_{T+1}^- = O_{T+1}^+ = B_{T+1}^- = \varnothing$

The reduced VRS hyperbolic DEA model to be solved step t is

$$\theta^t = \text{Min} \quad \theta$$

s.t.

$$\sum_j \left(\lambda_j + \mu_j\right) \cdot x_{ij} \le \theta \cdot x_{i0} \qquad \forall i \in I_t^-$$

$$\sum_j \left(\lambda_j + \mu_j\right) \cdot x_{ij} \le x_i^{t-1} \qquad \forall i \notin I_t^-$$

$$\sum_{j=1}^{n} \lambda_j y_{kj} \ge \frac{1}{\theta} \cdot y_{k0} \qquad \forall k \in O_t^+$$

$$\sum_j \lambda_j y_{kj} \ge y_k^{t-1} \qquad \forall k \notin O_t^+ \qquad \text{(A.3)}$$

$$\sum_j \lambda_j z_{bj} = \theta \cdot z_{b0} \qquad \forall b \in B_t^-$$

$$\sum_j \lambda_j z_{bj} = z_b^{t-1} \qquad \forall b \notin B_t^-$$

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

$$\lambda_j, \mu_j \ge 0 \ \forall j \quad \theta \text{ free}$$

The corresponding step t intermediate target is computed as

$$\begin{aligned}
x_i^t &= \theta^t \cdot x_{i0} && \forall i \in I_t^- \\
x_i^t &= x_i^{t-1} && \forall i \notin I_t^- \\
y_k^t &= \frac{1}{\theta^t} \cdot y_{k0} && \forall k \in O_t^+ \\
y_k^t &= y_k^{t-1} && \forall k \notin O_t^+ \\
z_b^t &= \theta^t \cdot z_{b0} && \forall b \in B_t^- \\
z_b^t &= z_b^{t-1} && \forall b \notin B_t^-
\end{aligned} \qquad \text{(A.4)}$$

The final LexHDEA target can be computed as

$$
\begin{aligned}
x_i^T &= x_{i0} && \forall i \in I^{ND} \\
x_i^T &= \theta^{T(i)} \cdot x_{i0} && \forall i \in I^{D} \\
y_k^T &= \frac{1}{\theta^{\hat{T}(k)}} \cdot y_{k0} && \forall k \\
z_b^T &= \theta^{\hat{\hat{T}}(b)} \cdot z_{b0} && \forall b
\end{aligned}
\tag{A.5}
$$

where

$$
T(i) = \max_t \left\{ t : i \in I_t^- \right\} \qquad \hat{T}(k) = \max_t \left\{ t : k \in O_t^+ \right\} \qquad \hat{\hat{T}}(b) = \max_t \left\{ t : b \in B_t^- \right\} \tag{A.6}
$$

And, finally, the LexHDEA efficiency measure can be computed as

$$
\xi_0^{\text{LexHDEA}} = \frac{\left( \prod_{i \in I^D} \frac{x_i^T}{x_{i0}} \right)^{\frac{1}{|I^D|}} \cdot \left( \prod_{b=1}^{|B|} \frac{z_b^T}{z_{b0}} \right)^{\frac{1}{|B|}}}{\left( \prod_{k=1}^{s} \frac{y_k^T}{y_{k0}} \right)^{\frac{1}{s}}} = \frac{\left( \prod_{i \in I^D} \theta^{T(i)} \right)^{\frac{1}{|I^D|}} \cdot \left( \prod_{b=1}^{|B|} \theta^{\hat{\hat{T}}(b)} \right)^{\frac{1}{|B|}}}{\left( \prod_{k=1}^{s} \frac{1}{\theta^{\hat{T}(k)}} \right)^{\frac{1}{s}}} \tag{A.7}
$$

Note that $\zeta_0^{\text{GDF}} \leq \xi_0^{\text{LexHDEA}}$ is still valid, provided that the GDF DEA model (3) is modified to take into account the non-discretionary inputs and the undesirable outputs, i.e.

$$\zeta_0^{GDF} = \text{Min}\ \frac{\left(\prod_{i\in I^D}\alpha_i\right)^{1/|I^D|}\cdot\left(\prod_{b\in B}\gamma_b\right)^{1/|B|}}{\left(\prod_{k=1}^{s}\beta_k\right)^{1/s}}$$

$$\begin{aligned}
\text{s.t.}\quad & \sum_{j=1}^{n}\left(\lambda_j+\mu_j\right)\cdot x_{ij} \le \alpha_i\cdot x_{i0} && \forall i\in I^D \\
& \sum_{j=1}^{n}\left(\lambda_j+\mu_j\right)\cdot x_{ij} \le x_{i0} && \forall i\notin I^D \\
& \sum_{j=1}^{n}\lambda_j y_{kj} \ge \beta_k\cdot y_{k0} && k=1,\ldots,s \\
& \sum_{j=1}^{n}\lambda_j z_{bj} = \gamma_b\cdot z_{b0} && \forall b\in B \\
& \sum_{j=1}^{n}\left(\lambda_j+\mu_j\right)=1 \\
& \lambda_j,\mu_j\ge 0\ \ \forall j \quad \alpha_i\le 1\ \ \forall i\in I^D \quad \beta_k\ge 1\ \ \forall k \quad \gamma_b\le 1\ \ \forall b\in B
\end{aligned} \tag{A.8}$$

However, (12) must be modified so that, in this case,

$$\xi_0^{LexHDEA} = \frac{\left(\prod_{i\in I^D}\theta^{T(i)}\right)^{\frac{1}{|I^D|}}\cdot\left(\prod_{b=1}^{|B|}\theta^{\hat{\hat{T}}(b)}\right)^{\frac{1}{|B|}}}{\left(\prod_{k=1}^{s}\frac{1}{\theta^{\hat{T}(k)}}\right)^{\frac{1}{s}}} \le \frac{\left(\prod_{i\in I^D}\theta_0^{HDF}\right)^{\frac{1}{|I^D|}}\cdot\left(\prod_{b=1}^{|B|}\theta_0^{HDF}\right)^{\frac{1}{|B|}}}{\left(\prod_{k=1}^{s}\frac{1}{\theta_0^{HDF}}\right)^{\frac{1}{s}}} = \left(\theta_0^{HDF}\right)^3 \tag{A.9}$$

## Appendix B

In this appendix the extension of the iterative DDF-based algorithm of Färe et al. (2016) for the case of some inputs being non-discretionary and some outputs undesirable is presented. VRS is assumed. Recall that the objective is to solve the non-linear model (A.3) using an iterative

algorithm that solves a DDF model with a specific directional vector $g=\left(g^x, g^y, g^z\right)$ in each iteration. Following Färe et al. (2016), the linear approximation of the non-linear model (A.3) is

$$
\begin{aligned}
&\theta^t = \text{Min} \quad \theta \\
&\text{s.t.} \\
&\sum_j \left(\lambda_j + \mu_j\right) \cdot x_{ij} \le x_{i0} - \beta \cdot g_i^x \qquad \forall i \in I_t^- \\
&\sum_j \left(\lambda_j + \mu_j\right) \cdot x_{ij} \le x_i^{t-1} \qquad \forall i \notin I_t^- \\
&\sum_{j=1}^n \lambda_j y_{kj} \ge y_{k0} + \beta \cdot g_k^y \qquad \forall k \in O_t^+ \\
&\sum_j \lambda_j y_{kj} \ge y_k^{t-1} \qquad \forall k \notin O_t^+ \\
&\sum_j \lambda_j z_{bj} = z_{b0} - \beta \cdot g_b^z \qquad \forall b \in B_t^- \\
&\sum_j \lambda_j z_{bj} = z_b^{t-1} \qquad \forall b \notin B_t^- \\
&\sum_{j=1}^n \left(\lambda_j + \mu_j\right) = 1 \\
&\lambda_j, \mu_j \ge 0 \; \forall j \quad \theta \text{ free}
\end{aligned}
\tag{B.1}
$$

where, in the first iteration $g=\left(x_0, y_0, z_0\right)$.

The dual of the above model is

$$
\begin{aligned}
\text{Max} \quad & \sum_{k\in O_t^+} v_k y_{k0} + \sum_{k\notin O_t^+} v_k y_k^{t-1} - \sum_{i\in I_t^-} u_i x_{i0} - \sum_{i\notin I_t^-} u_i x_i^{t-1} - \sum_{b\in B_t^-} w_b z_{b0} - \sum_{b\notin B_t^-} w_b z_b^{t-1} - q_0 \\
\text{s.t.} \quad & \sum_{k=1}^{s} v_k y_{kj} - \sum_{i=1}^{m} u_i x_{ij} - \sum_{b\in B} w_b z_{bj} - q_0 \le 0 \qquad \forall j \\
& \sum_{i=1}^{m} u_i x_{ij} + q_0 \ge 0 \qquad \forall j \\
& \sum_{i\in I_t^-} u_i g_i^x + \sum_{k\in O_t^+} v_k g_k^y + \sum_{b\in B_t^-} w_b g_b^z = 1 \\
& u_i \ge 0 \ \forall i \quad v_k \ge 0 \ \forall k \quad w_b \text{ free } \forall b \quad q_0 \text{ free}
\end{aligned}
\tag{B.2}
$$

The optimal solution of this dual model $\left(u^*, v^*, w^*, q_0^*\right)$ provides a supporting hyperplane whose intersection with the reduced hyperbolic path

$$
\begin{aligned}
x_i &= \theta \cdot x_{i0} \quad \forall i \in I_t^- \\
x_i &= x_i^{t-1} \quad \forall i \notin I_t^- \\
y_k &= \frac{1}{\theta} \cdot y_{k0} \quad \forall k \in O_t^+ \\
y_k &= y_k^{t-1} \quad \forall k \notin O_t^+ \\
z_b &= \theta \cdot z_{b0} \quad \forall b \in B_t^- \\
z_b &= z_b^{t-1} \quad \forall b \notin B_t^-
\end{aligned}
\tag{B.3}
$$

occurs at

$$
\theta^* = \begin{cases}
-\dfrac{B}{A} & \text{if } C = 0 \ \& \ A \neq 0 \\
-\dfrac{C}{B} & \text{if } A = 0 \ \& \ B \neq 0 \\
-\dfrac{B + \sqrt{B^2 - 4 \cdot A \cdot C}}{2 \cdot A} & \text{otherwise}
\end{cases}
\tag{B.4}
$$

where

$$A = -\sum_{i \in I_t^-} u_i^* x_{i0} - \sum_{b \in B_t^-} w_b^* z_{b0}$$

$$B = \sum_{k \notin O_t^+} v_k^* y_k^{t-1} - \sum_{i \notin I_t^-} u_i^* x_i^{t-1} - \sum_{b \notin B_t^-} w_b^* z_b^{t-1} - q_0^* \qquad \text{(B.5)}$$

$$C = \sum_{k \in O_t^+} v_k y_{k0}$$

If this intersection is feasible, then it is the optimal solution $\theta^t$ of the non-linear model (A.3). Otherwise, the DDF model (B.1) is solved again but this time with an updated directional vector $g = \left( x_0 - \theta^* \cdot x_0, \frac{1}{\theta^*} \cdot y_0 - y_0, z_0 - \theta^* \cdot z_0 \right)$.

As regards the feasibility checking mentioned above, it is only necessary to check the consistency of the following linear system, which can be done solving the corresponding LP with any linear (e.g. constant) objective function

$$\begin{aligned}
&\sum_j \left( \lambda_j + \mu_j \right) \cdot x_{ij} \le \theta^* \cdot x_{i0} && \forall i \in I_t^- \\
&\sum_j \left( \lambda_j + \mu_j \right) \cdot x_{ij} \le x_i^{t-1} && \forall i \notin I_t^- \\
&\sum_{j=1}^n \lambda_j y_{kj} \ge \frac{1}{\theta^*} \cdot y_{k0} && \forall k \in O_t^+ \\
&\sum_j \lambda_j y_{kj} \ge y_k^{t-1} && \forall k \notin O_t^+ \\
&\sum_j \lambda_j z_{bj} = \theta^* \cdot z_{b0} && \forall b \in B_t^- \\
&\sum_j \lambda_j z_{bj} = z_b^{t-1} && \forall b \notin B_t^- \\
&\sum_{j=1}^n \left( \lambda_j + \mu_j \right) = 1 \\
&\lambda_j, \mu_j \ge 0 \ \forall j
\end{aligned} \qquad \text{(B.6)}$$

As indicated in Färe et al (2016) the above system can be reduced setting $\lambda_j = \mu_j = 0$ for all DMUs j not in the optimal supporting hyperplane $\left(u^*, v^*, w^*, q_0^*\right)$, i.e. those DMUs j for which

$$\sum_{k=1}^{m} v_k y_{kj} - \sum_{i=1}^{m} u_i^* x_{ij} - \sum_{b \in B} w_b^* z_{bj} - q_0^* < 0 \qquad \text{(B.7)}$$

**Acknowledgements**

This research was carried out with the financial support of the Spanish Ministry of Economy, Industry and Competitiveness, grant DPI2017-85343-P.

**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., Ruíz, J.L. and Sirvent, I., "Closest targets and minimum distance to the Pareto-efficient frontier in DEA", *Journal of Productivity Analysis*, 28 (2007) 209-218

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

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

Coelli, T., "A multi-stage methodology for the solution of orientated DEA models*", Operations Research Letters*, 23 (1998) 143-149

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

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

Färe, R., Grosskopf, S. and Lovell, C.A.K., *The measurement of efficiency of production*, Kluwer-Nijhoff, Boston, 1985

Färe, R., Grosskopf, S. and Whittaker, G., "Directional output distance functions: Endogenous constraints based on exogenous normalization constraints", *Journal of Productivity Analysis*, 40 (2013) 267-269

Färe, R., Margaritis, D., Rouse, P. and Roshdi, I., "Estimating the hyperbolic distance function: A directional distance function approach", *European Journal of Operational Research*, 254 (2016) 312-319

Ghahraman, A. and Prior, D., "A learning ladder toward efficiency: Proposing network-based stepwise benchmark selection", *Omega*, 63 (2016) 83-93

Gutiérrez, E., Aguilera, E., Lozano, S. and Guzmán, G.I., "A two-stage DEA approach for quantifying and analysing the inefficiency of conventional and organic rain-fed cereals in Spain", *Journal of Cleaner Production*, 149 (2017) 335-348

Hampf, B. and Krüger, J. J., "Optimal directions for directional distance functions: An exploration of potential reductions of greenhouse gases", *American Journal of Agricultural Economics*, 97, 3 (2015) 920-938

Hasannasab, M., Margaritis, D., Roshdi, I. and Rouse, P., "Hyperbolic efficiency measurement: A conic programming approach", *European Journal of Operational Research*, (2018) (doi: 10.1016/j.ejor.2018.12.005)

Korhonen, P.J., Dehnokhalaji, A. and Nasrabadi, N., "A lexicographic radial projection onto the efficient frontier in Data Envelopment Analysis", *European Journal of Operational Research*, 265 (2018) 1005-1012

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*, 81 (2018) 195-207

Lozano, S., Hinojosa, M.A. and Mármol, A.M., "Extending the bargaining approach to DEA target setting", *Omega*, (2018) (doi: 10.1016/j.omega.2018.05.015)

Lozano, S. and Soltani, N., “DEA target setting using lexicographic and endogenous Directional Distance Function approaches”, *Journal of Productivity Analysis*, 50, 1-2 (2018a) 55-70

Lozano, S. and Soltani, N., “Efficiency assessment using a multidirectional DDF approach”, *International Transactions in Operational Research*, (2018b) (doi: 10.1111/itor.12617)

Lozano, S. and Villa, G., “Determining a sequence of targets in DEA”, *Journal of the Operational Research Society*, 56 (2005) 1439-1447

Lozano, S. and Villa, G., “Data envelopment analysis of integer-valued inputs and outputs”, *Computers and Operations Research*, 33, 10 (2006) 3004-3014

Pastor, J.T., Aparicio, J., Alcaraz, J., Vidal, F. and Pastor, D., “The Reverse Directional Distance Function”, in *Advances in Efficiency and Productivity*, J. Aparicio et al. (eds.), International Series in Operations Research & Management Science, vol. 249 (2016) 15-55

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

Roshdi, I., Hasannasab, M., Margaritis, D. and Rouse P., “Generalised weak disposability and efficiency measurement in environmental technologies”, *European Journal of Operational Research*, 266 (2018) 1000-1012

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)

Yu, M.M. and Lee, B.C.Y., “Efficiency and effectiveness of service business: Evidence from international tourist hotels in Taiwan", *Tourism Management*, 30, 4 (2009) 571-580

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 Productivity Analysis*, 40 (2013) 257-266

**List of figure and table captions**

Table 1. Numerical example: inputs/output data and HDF and GDF scores

Table 2. LexHDEA and GDF results for DMU C

Table 3. Observed data and GDF results for case study

Table 4. LexHDEA results for DMU O3

Figure 1. Graphical visualization of LexHDEA path for DMU C

Figure 2. Evolution of successive $\theta^t$ values along LexHDEA path for some inefficient DMUs

Figure 3. $\xi_0^{\text{LexHDEA}}$ versus $\zeta_0^{\text{GDF}}$ (left panel) and $\xi_0^{\text{LexHDEA}}$ versus $\left(\theta_0^{\text{HDF}}\right)^3$ (right panel)

Table 1. Numerical example: inputs/output data and HDF and GDF scores

| **DMU** | $x_1$ | $x_2$ | y | $\theta_0^{HDF}$ | $\zeta_0^{GDF}$ |
|---|---|---|---|---|---|
| A | 10 | 18 | 10 | 1.000 | 1.000 |
| B | 10 | 20 | 21 | 1.000 | 1.000 |
| C | 15 | 60 | 10 | 0.667 | 0.224 |
| D | 25 | 30 | 10 | 0.571 | 0.235 |
| E | 40 | 50 | 33 | 1.000 | 0.568 |
| F | 19 | 34 | 33 | 1.000 | 1.000 |
| G | 26 | 13 | 29 | 1.000 | 1.000 |
| H | 33 | 21 | 31 | 1.000 | 1.000 |
| I | 35 | 36 | 15 | 0.550 | 0.268 |
| J | 35 | 13 | 15 | 1.000 | 0.446 |

Table 2. LexHDEA and GDF results for DMU C

| LexHDEA projection of DMU C | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| t | $I_t^-$ | $O_t^+$ | DDF iter. | $\theta^t$ | $x_1^t$ | $x_2^t$ | $y^t$ | DDF iter. status |
| 0 | $x_1$,$x_2$ | y | - | - | 15 | 60 | 10 | - |
| 1 | $x_1$,$x_2$ | y | 1 | 0.667 | 10.000 | 40.000 | 15.000 | Feasible |
| 2 | $x_2$ | y | 1 | 0.355 | - | 21.303 | 28.165 | Infeasible |
| | | | 2 | 0.476 | - | 28.571 | 21.000 | Feasible |
| 3 | $x_2$ | - | 1 | 0.333 | - | 20.000 | - | Feasible |
| 4 | ∅ | ∅ | - | - | - | - | - | - |
| $\xi_C^{LexHDEA} = 0.224$ | | | LexHDEA target | | 10.000 | 20.000 | 21.000 | - |
| LexHDEA computing time (non-linear solver): 0:00:03.231 | | | | | | | | |
| LexHDEA computing time (iterative DDF algorithm): 0:00:02.422 | | | | | | | | |
| | | | | | | | | |
| GDF projection of DMU C | | | | | | | | |
| $\zeta_C^{GDF} = 0.224$ | | | GDF target | | 10.000 | 20.000 | 21.000 | - |
| GDF computing time (non-linear solver): 0:00:00.412 | | | | | | | | |

Note: computing time format HH:MM:SS:mmm

Table 3. Observed data and GDF results for case study

| DMU | Observed Data | | | | | | GDF | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | $\zeta_0^{GDF}$ | Target | | | | | Time |
| | FC | TCI | TNI | YFM | NPP | TAE | | FC | TNI | YFM | NPP | TAE | |
| O1 | 39.30 | 472.18 | 46.40 | 1,880 | 3,109.96 | 414.36 | 1.00000 | - | - | - | - | - | 0:00:02.024 |
| O2 | 52.48 | 307.35 | 36.96 | 2,000 | 3,252.72 | 488.50 | 0.70500 | 52.02 | 36.96 | 2,208.75 | 3,252.72 | 363.51 | 0:00:01.494 |
| O3 | 49.66 | 324.24 | 38.98 | 1,400 | 2,538.91 | 457.68 | 0.48464 | 47.30 | 38.98 | 1,796.55 | 2,538.91 | 257.47 | 0:00:17.072 |
| O4 | 47.75 | 526.26 | 48.31 | 1,200 | 2,104.91 | 271.49 | 0.38664 | 47.75 | 48.08 | 1,451.77 | 2,104.91 | 115.73 | 0:00:01.640 |
| O5 | 49.91 | 118.84 | 25.87 | 1,500 | 2,657.88 | 497.30 | 1.00000 | - | - | - | - | - | 0:00:01.261 |
| O6 | 48.86 | 104.24 | 23.34 | 1,200 | 2,300.97 | 473.18 | 1.00000 | - | - | - | - | - | 0:00:15.326 |
| O7 | 44.69 | 373.53 | 40.64 | 1,000 | 2,063.03 | 462.66 | 0.31630 | 44.69 | 40.50 | 1,469.64 | 2,063.03 | 177.72 | 0:00:01.509 |
| O8 | 45.20 | 485.83 | 46.60 | 2,000 | 3,252.72 | 444.70 | 0.62541 | 44.03 | 46.60 | 2,281.73 | 3,252.72 | 301.00 | 0:00:01.369 |
| O9 | 46.73 | 472.98 | 45.90 | 1,880 | 3,109.96 | 456.10 | 0.56441 | 43.94 | 45.90 | 2,185.58 | 3,109.96 | 286.25 | 0:00:01.628 |
| O10 | 55.65 | 384.70 | 40.60 | 3,000 | 4,442.42 | 494.81 | 1.00000 | - | - | - | - | - | 0:00:01.216 |
| O11 | 56.52 | 1,297.82 | 94.23 | 3,000 | 4,442.42 | 88.18 | 1.00000 | - | - | - | - | - | 0:00:20.524 |
| O12 | 55.49 | 228.24 | 32.23 | 2,500 | 3,847.57 | 523.94 | 1.00000 | - | - | - | - | - | 0:00:16.737 |
| O13 | 51.07 | 318.49 | 36.26 | 2,000 | 3,460.38 | 498.89 | 0.75127 | 51.07 | 36.26 | 2,268.91 | 3,460.38 | 399.20 | 0:00:02.059 |
| O14 | 47.27 | 516.38 | 47.88 | 1,000 | 1,899.65 | 291.75 | 0.28984 | 47.27 | 45.85 | 1,287.88 | 1,899.65 | 98.07 | 0:00:01.603 |
| O15 | 43.14 | 141.08 | 45.24 | 1,800 | 3,201.68 | 701.83 | 1.00000 | - | - | - | - | - | 0:00:17.261 |
| O16 | 45.20 | 500.09 | 46.18 | 2,000 | 3,460.38 | 454.86 | 0.65948 | 45.20 | 46.18 | 2,385.10 | 3,460.38 | 327.58 | 0:00:09.955 |
| O17 | 43.52 | 209.54 | 29.90 | 2,000 | 3,460.38 | 480.56 | 1.00000 | - | - | - | - | - | 0:00:02.159 |
| O18 | 56.36 | 1,119.50 | 48.44 | 1,500 | 2,412.81 | 154.74 | 0.59091 | 41.31 | 48.44 | 1,500.00 | 2,513.34 | 109.02 | 0:00:01.301 |
| O19 | 53.11 | 477.46 | 59.86 | 2,500 | 3,439.12 | 373.85 | 0.69762 | 41.78 | 49.93 | 2,500.00 | 3,452.98 | 322.63 | 0:00:01.559 |
| O20 | 55.85 | 469.99 | 43.26 | 2,600 | 3,541.75 | 348.11 | 1.00000 | - | - | - | - | - | 0:00:02.121 |
| O21 | 42.67 | 372.38 | 50.23 | 2,800 | 3,747.01 | 384.58 | 1.00000 | - | - | - | - | - | 0:00:01.129 |
| O22 | 48.91 | 123.70 | 25.55 | 1,600 | 2,776.85 | 557.43 | 1.00000 | - | - | - | - | - | 0:00:02.139 |
| O23 | 49.61 | 247.59 | 39.08 | 1,125 | 2,211.74 | 448.77 | 0.41974 | 45.25 | 37.30 | 1,591.32 | 2,211.74 | 240.15 | 0:00:07.439 |
| O24 | 43.34 | 1,431.67 | 73.89 | 2,000 | 3,252.72 | 78.04 | 0.75328 | 41.44 | 69.50 | 2,000.00 | 3,356.55 | 62.97 | 0:00:16.723 |
| O25 | 33.90 | 1,252.08 | 57.13 | 1,500 | 2,813.62 | 50.36 | 1.00000 | - | - | - | - | - | 0:00:22.287 |
| O26 | 41.79 | 1,093.36 | 49.60 | 1,200 | 2,425.56 | 164.42 | 0.42096 | 38.38 | 49.60 | 1,343.31 | 2,425.56 | 76.41 | 0:00:02.116 |

Note: computing time format HH:MM:SS:mmm

Table 4. LexHDEA results for DMU O3

| t | $I_t^-$ | $O_t^+$ | DDF iter. | $\theta^t$ | Step t target | | | | | DDF iter. status |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | FC | TNI | YFM | NPP | TAE | |
| 0 | FC,TNI | YFM,NPP,TAE | - | - | 49.66 | 38.98 | 1,400.00 | 2,538.91 | 457.68 | - |
| 1 | FC,TNI | YFM,NPP,TAE | 1 | 0.85372 | 42.40 | 33.28 | 1,639.87 | 2,973.93 | 390.74 | Feasible |
| 2 | TNI | YFM,NPP,TAE | 1 | 0.85293 | - | 33.25 | 1,641.40 | 2,976.70 | 390.37 | Feasible |
| 3 | ∅ | YFM,NPP,TAE | 1 | 0.84034 | - | - | 1,665.98 | 3,021.27 | 384.61 | Feasible |
| 4 | ∅ | YFM | 1 | 0.80622 | - | - | 1,736.50 | - | - | Feasible |
| 5 | ∅ | ∅ | - | - | - | - | - | - | - | - |
| $\xi_{O3}^{LexHDEA} = 0.59029$ | | | LexHDEA target | | 42.40 | 33.25 | 1,736.50 | 3,021.27 | 384.61 | - |
| LexHDEA computing time (non-linear solver): 0:00:03.653 | | | | | | | | | | |
| LexHDEA computing time (iterative DDF algorithm): 0:00:07.524 | | | | | | | | | | |

Note: computing time format HH:MM:SS:mmm

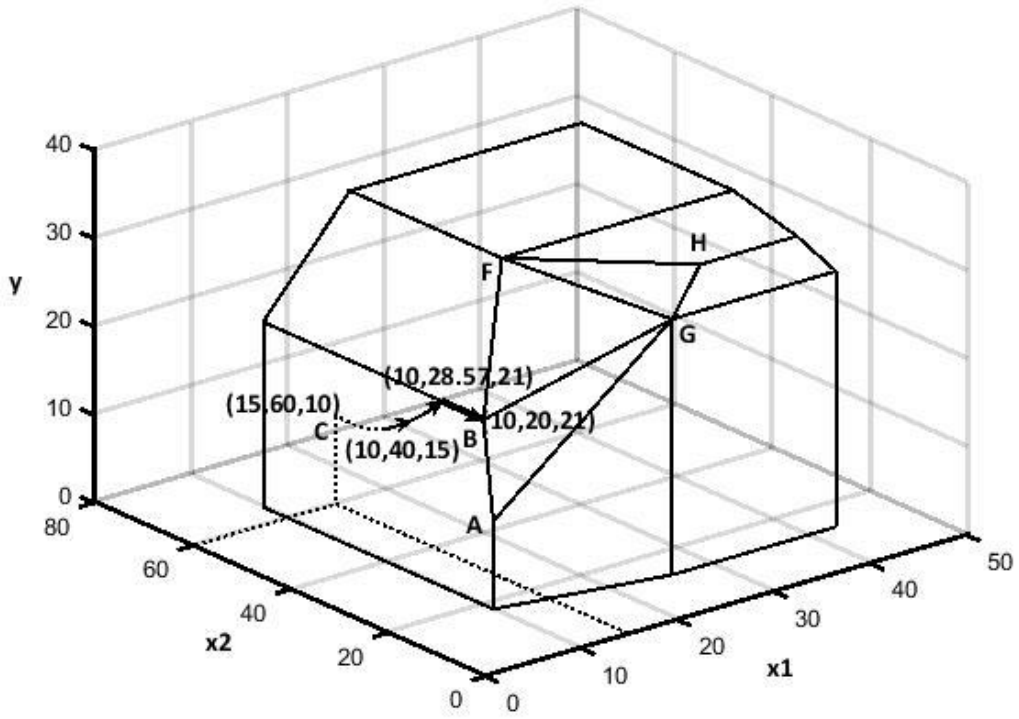


Figure 1. Graphical visualization of LexHDEA path for DMU C

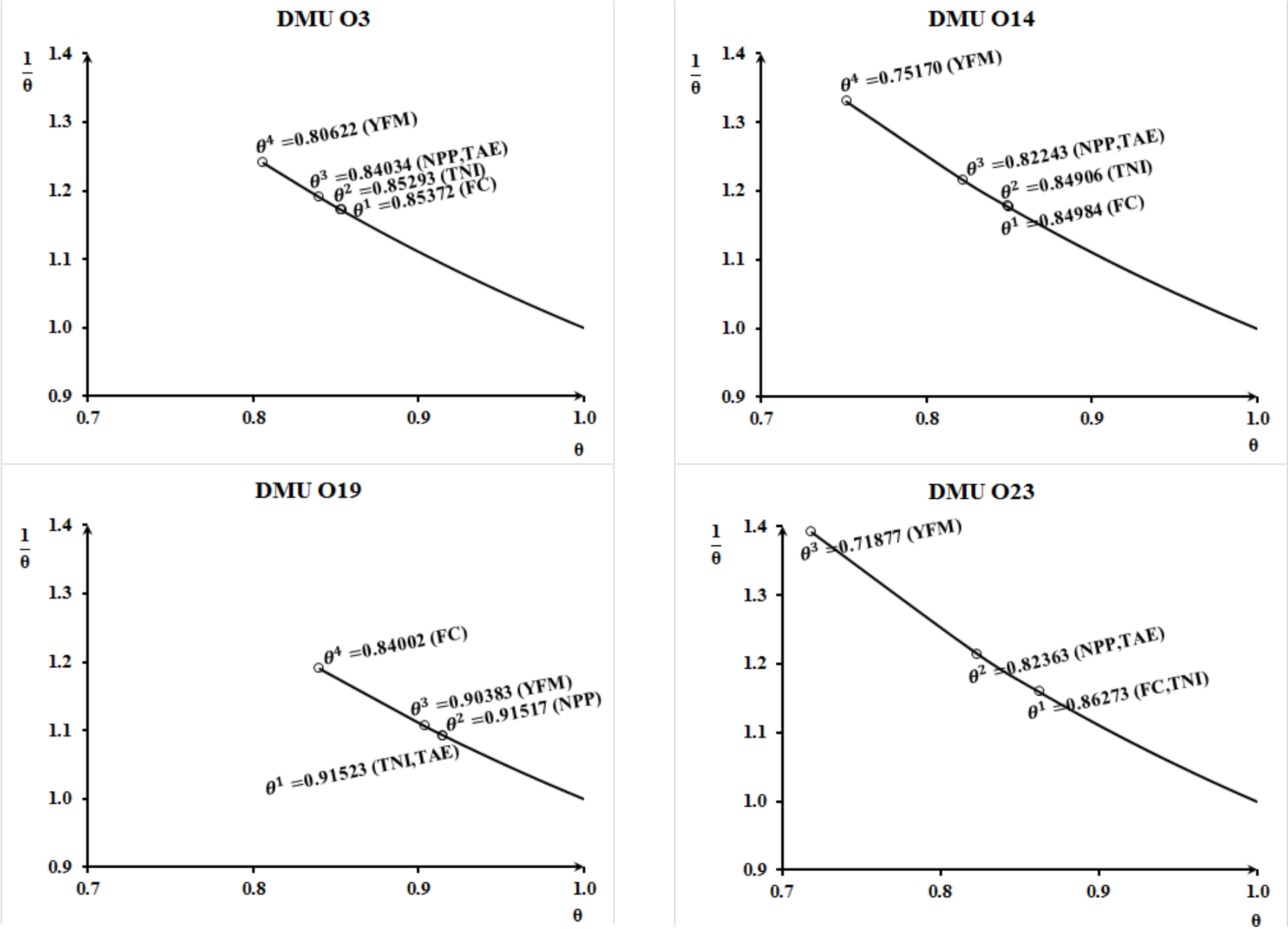


Figure 2. Evolution of successive $\theta^t$ values along LexHDEA path for some inefficient DMUs

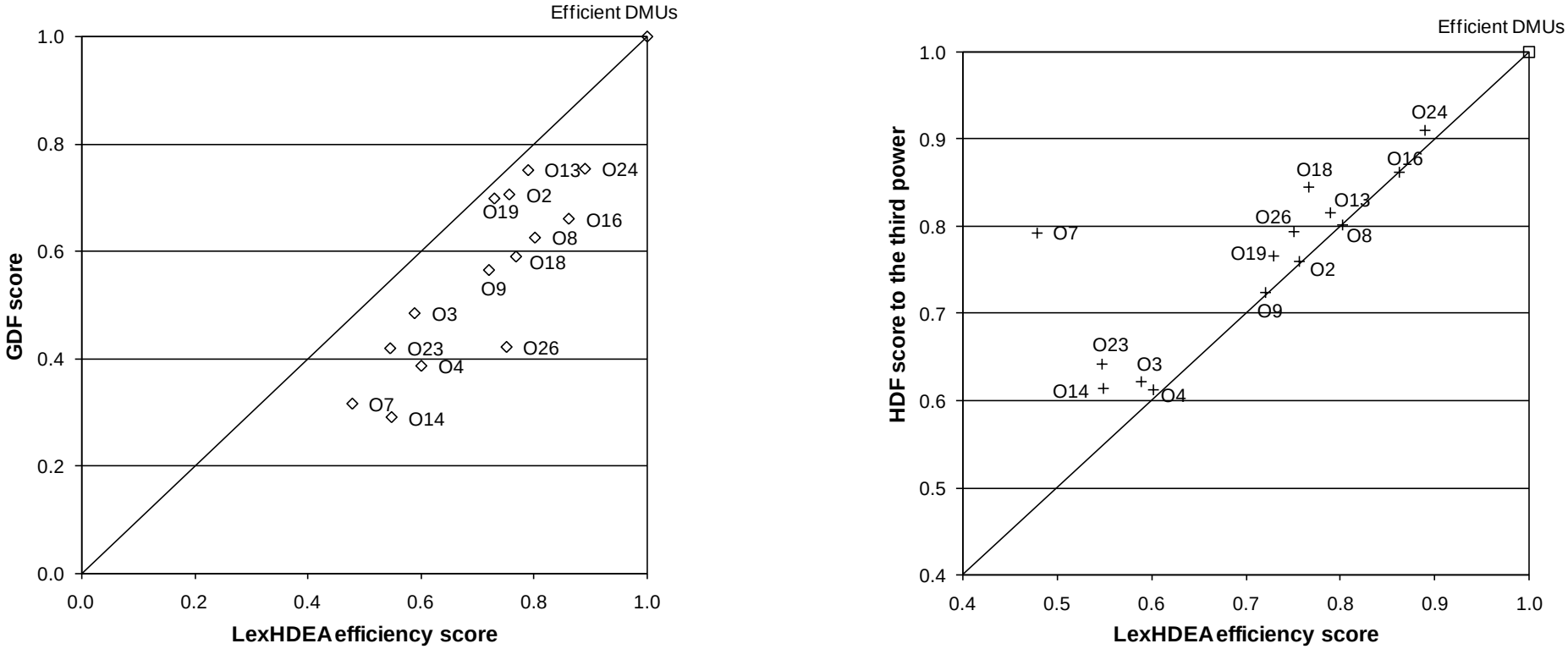


Figure 3. $\xi_0^{\text{LexHDEA}}$ versus $\zeta_0^{\text{GDF}}$ (left panel) and $\xi_0^{\text{LexHDEA}}$ versus $\left(\theta_0^{\text{HDF}}\right)^3$ (right panel)