Manuscript

Integer interval DEA: an axiomatic derivation of the technology and an additive, slacks-based model

Manuel Arana-Jiménez[a], M. Carmen Sánchez-Gil [a,1], Atefeh Younesi[a], Sebastián Lozano[b]

[a]*Department of Statistics and Operational Research, University of Cádiz, Spain*
[b]*Department of Industrial Management University of Seville, Spain*

Abstract

The present paper studies the efficiency assessment of Decision-Making Units (DMUs) when their inputs and outputs are described under uncertainty of the type of integer interval data. An axiomatic derivation of the production possibility set (PPS) is presented. An additive, slacks-based data envelopment analysis (DEA) model is formulated, consisting of two phases. This has required the use of adequate arithmetic and LU-partial orders for integer intervals. This novel integer interval DEA approach is the first step towards DEA models under fuzzy integer intervals, with the extension of the corresponding arithmetic and LU-partial orders to fuzzy integer intervals. The proposed method is applied on a dataset, taken from the literature, that involves both continuous and integer interval variables.

Keywords: `Efficiency; DEA; integer interval data; production possibility set; slacks-based model`
2018 MSC:

*mcarmen.sanchez@uca.es
Email addresses: `manuel.arana@uca.es` (Manuel Arana-Jiménez), `mcarmen.sanchez@uca.es` (M. Carmen Sánchez-Gil), `atefehyounesi1363@gmail.com` (Atefeh Younesi), `slozano@us.es` (Sebastián Lozano)

1. Introduction

Data envelopment analysis (DEA) is a non-parametric, data-driven methodology for assessing the efficiency of a set of comparable organizational units commonly termed Decision Making Units (DMUs) (Zhu [44], Cooper et al. [11]). DMUs are assumed to consume inputs (i.e. resources) in order to produce outputs. DEA only requires data about the input consumption and the output production of the DMUs. From these, and using some basic axioms (like free disposability and convexity), a Production Possibility Set (PPS) is inferred. The PPS, also known as the DEA technology, contains all the operating points that are deemed feasible. The non-dominated subset of the PPS is the efficient frontier (EF). DMUs that belong to EF are labelled efficient while the DMUs that do not belong to EF are labelled inefficient and can be projected onto EF. The projection of a DMU onto EF is called its target and the distance from the DMU to the target, which is a measure of the potential improvements that the DMU can achieved, is used to compute a quantitative efficiency score.

There are different ways of carrying out the projection onto the efficient frontier and computing the corresponding efficiency scores, e.g. using efficiency potential (Lozano and Calzada-Infante [29], Soltani and Lozano [36]), multi-directional approaches (Lozano and Soltani [30]) or lexicographic approaches (Lozano and Soltani [31], [32]), among others. There are, however, two types of DEA approaches that are of relevance in this research. One is integer DEA, i.e. DEA models that can handle integer data, and the other is imprecise or interval DEA models, in which some or all the inputs and outputs are given as interval data.

As regards integer DEA, it was first addressed in Lozano and Villa [33] and subsequently studied in Kuosmanen and Kazemi Matin [27] and Kazemi Matin and Kuosmanen [21]. Advanced integer DEA models involve Directional Distance Function (DDF) (e.g. Tan et al. [38]), super-efficiency (e.g. Du et al. [13], Chen et al. [9]), flexible measures (e.g. Kordrostami et al. [25]), two-stage systems (e.g. Ajirlo et al. [1]) or congestion (e.g. Khoveyni et al. [23]). A related problem is that of handling variables that can only take certain discrete values (e.g. Amirteimoori and Kordrostami [2]). Integer DEA has been applied, for example, to hotel performance (Wu et al.[41]), sports (e.g. Wu et al. [42], Chen et al. [10]) and transportation (e.g. Lozano et al. [34], Yu and Hsu [43]).

As regards interval DEA, there have been also many developments, most of them involving radial multiplier formulations (e.g. Despotis and Smirlis [12], Zhu [45]), although there are also additive imprecise DEA approaches (e.g. Lee et al. [28]), FDH interval DEA models (e.g. Jahanshaloo et al. [19]), non-radial, non-oriented imprecise DEA approaches (e.g. Azizi et al. [5]), Ideal point approaches (e.g. Jahanshahloo et al. [17]), inverted DEA approaches (e.g. Inuiguchi and Mizoshita [16]), interval DEA with negative data (e.g. Hatami-Marbini et al. [15]), flexible measure interval DEA approaches (e.g. Kordrostami and Jahani Sayyad Noveiri [26]) and common weights imprecise DEA approaches (e.g. Hatami-Marbini et al. [14]). Applications include manufacturing industry (e.g. Wang et al. [39]), banks and bank branches (e.g. Jahanshaloo et al [18], Inuiguchi and Mizoshita [16], Hatami-Marbini et al. [15]), power plants (e.g. Khalili-Damghani et al. [22]), etc.

This paper studies the situation when we have inputs and outputs that are both integer and interval-valued, as a mathematical modelling of the uncertainty on integer data. To the best of our knowledge, the closest existing DEA approach is the fuzzy integer DEA model of Kordrostami et al. [24], which extend the integer DEA model of Jie et al. [20]. The approach proposed in this paper has numerous differences with respect to Kordrostami et al. [24]. Thus, while [24] considers fuzzy integer data in our case the uncertainty is modelled with interval integer data. While [24] uses a fuzzy ranking approach, what derives a defuzzification of the data instead of fully keeping the fuzzy information given by the original data, in the present approach we establish the order relation between the elements of the PPS using interval orders, together with interval arithmetic. Also, while [24] uses a radial oriented approach, we use an additive, non-oriented approach. While [24] computes a crisp target, we compute an integer interval target. More important, while [24] uses the integer PPS of Kuosmanen and Kazemi Matin [27], we carry out an axiomatic derivation of a new integer interval PPS. This PPS can be used as a base to derive, in a continuation of this research, a fuzzy integer interval PPS, and a corresponding DEA model with fuzzy integer data using partial orders and arithmetic on fuzzy sets.

The structure of the paper is the following. In Section 2 the basic concepts of the DEA methodology as well as an slacks-based DEA model for efficiency assessment are reviewed. In Section 3 additional concepts on integer intervals are introduced, in particular, arithmetic operations and partial orders. Those concepts are later used, in Section 4, to define a new integer interval DEA technology and a new slacks-based DEA approach involving two phases. Numerical experiments are presented and discussed in Section 5, comparing the proposed approach with other approaches from the literature. Finally, in Section 6, conclusions are drawn.

2. Crisp production possibility set and slack-based measure

Let us consider a set of n DMUs. For $j \in J = \{1, \ldots, n\}$, each DMU_j has m inputs $X_j = (x_{1j}, \ldots, x_{mj}) \in \mathbb{R}^m$, produces s outputs $Y_j = (y_{1j}, \ldots, y_{sj}) \in \mathbb{R}^s$. In the classic Charnes et al. [8] DEA model, the production possibility set (PPS) or technology, denoted by T, satisfies the following axioms:

(A1) Envelopment: $(X_j, Y_j) \in T$, for all $j \in J$.

(A2) Free disposability: $(x, y) \in T$, $(x', y') \in \mathbb{R}^{m+s}$, $x' \geqq x$, $y' \leqq y \Rightarrow (x', y') \in T$.

(A3) Convexity: $(x, y), (x', y') \in T$, then $\lambda(x, y) + (1 - \lambda)(x', y') \in T$, for all $\lambda \in [0, 1]$.

(A4) Scalability: $(x, y) \in T \Rightarrow (\lambda x, \lambda y) \in T$, for all $\lambda \in \mathbb{R}_+$.

Following the minimum extrapolation principle (see [6]), the DEA PPS, which contains all the feasible input-output bundles, is the intersection of all the sets that satisfy axioms (A1)-(A4) and can be expressed as

$$T_{DEA} = \left\{ (x,y) \in \mathbb{R}_+^{m+s} : x \geq \sum_{j=1}^{n} \lambda_j X_j, y \leq \sum_{j=1}^{n} \lambda_j Y_j, \lambda_j \geq 0 \right\}.$$

Let us recall that a DMU p is said to be efficient if and only if for any $(x,y) \in T_{DEA}$ such that $x \leqq X_p$ and $y \geqq Y_p$, then $(x,y) = (X_p, Y_p)$. This can be determined solving the following normalized slacks-based DEA model

$$\begin{aligned}
\text{(DEA)}\ \ I(X_p,Y_p) = \quad \text{Max} \quad & \sum_{i=1}^{M} \frac{s_i^x}{x_{ip}} + \sum_{r=1}^{S} \frac{s_r^y}{y_{rp}} \qquad (1)\\
\text{s.t.} \quad & \sum_{j=1}^{N} \lambda_j x_{ij} \leq x_{ip} - s_i^x, \quad i = 1,\dots,M,\\
& \sum_{j=1}^{N} \lambda_j y_{rj} \geq y_{rp} + s_r^y, \quad r = 1,\dots,S,\\
& \lambda_j \geq 0, \quad j = 1,\dots,N,\\
& s_i^x, s_r^y \geq 0, \quad i = 1,\dots,M, \ \ r = 1,\dots,S.
\end{aligned}$$

where λ_j, $j = 1,\dots,n$, are the intensity variables used for defining the corresponding efficient target of DMU_p. The inefficiency measure $I(X_p,Y_p)$ is units invariant and non-negative. Moreover, a DMU_p is efficient if and only if $I(X_p,Y_p) = 0$.

3. Notation and preliminaries

In this paper, uncertainty on the production possibility set is presented by modeling the corresponding inequality relationships using partial orders on integer intervals. This requires introducing first the following notation and results.

Let $\mathbb{R}$ be the real number set. We denote by $\mathcal{K}_C = \left\{ \left[\underline{a}, \overline{a}\right] \mid \underline{a}, \overline{a} \in \mathbb{R} \text{ and } \underline{a} \leq \overline{a} \right\}$ the family of all bounded closed intervals in $\mathbb{R}$.

Definition 1. *Let* $A = [\underline{a}, \overline{a}] \in \mathcal{K}_C$, $B = [\underline{b}, \overline{b}] \in \mathcal{K}_C$

- *Addition:* $A + B := \{a + b \mid a \in A, b \in B\} = [\underline{a} + \underline{b}, \overline{a} + \overline{b}]$,
- *Opposite value:* $-A = \{-a : a \in A\} = [-\overline{a}, -\underline{a}]$,
- *Multiplication:* $A \cdot B := \{a \cdot b \mid a \in A, b \in B\} = [min(A \cdot B), max(A \cdot B)]$, *where* $A \cdot B = \{\underline{a} \cdot \underline{b}, \underline{a} \cdot \overline{b}, \overline{a} \cdot \underline{b}, \overline{a} \cdot \overline{b}\}$.
- *Multiplication by scalar: for any* λ,

$$\lambda \cdot A := \begin{cases} [\lambda \cdot \underline{a}, \lambda \cdot \overline{a}] & \lambda \geq 0 \\ [\lambda \cdot \overline{a}, \lambda \cdot \underline{a}] & \lambda < 0 \end{cases}$$

Example 1. *Consider the following examples of the defined operations for continuous intervals. Note that, when applied to continuous intervals, all these operations produce continuous interval domains.* $[-5, 2] + [-4, -1] = [-9, 1]$, $-[2, 7] = [-7, -2]$, $[2, 4] \cdot [4, 6] = [8, 24]$, $3 \cdot [2, 4] = [6, 12]$, $-3 \cdot [2, 4] = [-12, -6]$.

3.1. Integer Set Arithmetic

Apt and Zoeteweij [3] have defined the following arithmetic operations on integer intervals A and B:

- Addition: $A + B := \{a + b \mid a \in A, b \in B\}$,
- Subtraction: $A - B := \{a - b \mid a \in A, b \in B\}$,
- Multiplication: $A * B := \{a * b \mid a \in A, b \in B\}$,
- Multiplication by scalar: for any integer λ,

$$\lambda * A := \begin{cases} \lambda * a & \lambda \geq 0 \\ -\lambda * a & \lambda < 0 \end{cases}$$

Example 2. *To illustrate the previous arithmetic operations between integer intervals, consider the following examples. For the case of sum and subtraction,* $\{3, 4, 5\} + \{2, 3, 4\} = \{5, 6, 7, 8, 9\}$, $\{3, 4, 5, 6\} - \{2, 3, 4, 5, 6, 7\} = \{-4, , -3, -2, -1, 0, 1, 2, 3, 4\}$; *and for the case of multiplication,* $\{2, 3, 4\} * \{4, 5, 6\} = \{8, 10, 12, 12, 15, 16, 18, 20, 24\}$,

$3 * \{2, 3, 4\} = \{6, 9, 12\}$. *Note, from the last example, that* $\{2, 3, 4\} * \{4, 5, 6\}$ *does not contains all integer numbers from* 8 *to* 24, *and also* $3 * \{2, 3, 4\}$ *does not contains all integer numbers from* 6 *to* 12.

Therefore, for A, B integer intervals and a λ an integer the following holds:

- $A + B, A - B$ are integer intervals.
- $A * B$ does not correspond to an integer interval, in general. And the same for $\lambda * A$.

To deal with this problem, it is necessary to introduce a new multiplication operation for the multiplication between two integer interval to be an integer interval also. Let $\mathbb{Z}$ be the integer set. We denote by $\mathcal{K}_{\mathbb{Z}} = \left\{ \left[\underline{a}, \overline{a}\right]_{\mathbb{Z}} \mid \underline{a}, \overline{a} \in \mathbb{Z} \text{ and } \underline{a} \leq \overline{a} \right\}$ a closed integer interval in $\mathbb{Z}$.

Definition 2. *Let* $A = [\underline{a}, \overline{a}] \in \mathcal{K}_{\mathbb{Z}}$, $B = [\underline{b}, \overline{b}] \in \mathcal{K}_{\mathbb{Z}}$

- *Addition:* $[\underline{a}, \overline{a}]_{\mathbb{Z}} + [\underline{b}, \overline{b}]_{\mathbb{Z}} = [\underline{a} + \underline{b}, \overline{a} + \overline{b}]_{\mathbb{Z}}$
- *Subtraction:* $[\underline{a}, \overline{a}]_{\mathbb{Z}} - [\underline{b}, \overline{b}]_{\mathbb{Z}} = [\underline{a} - \overline{b}, \overline{a} - \underline{b}]_{\mathbb{Z}}$
- *Multiplication:* $[\underline{a}, \overline{a}]_{\mathbb{Z}} \cdot [\underline{b}, \overline{b}]_{\mathbb{Z}} = [min(A \cdot B), max(A \cdot B)]_{\mathbb{Z}}$,

  where $A \cdot B = \{\underline{a} \cdot \underline{b}, \underline{a} \cdot \overline{b}, \overline{a} \cdot \underline{b}, \overline{a} \cdot \overline{b}\}$.
- *Multiplication by scalar: for any integer* λ,

$$\lambda \cdot A := \begin{cases} [\lambda \cdot \underline{a}, \lambda \cdot \overline{a}]_{\mathbb{Z}} & \lambda \geq 0 \\ [\lambda \cdot \overline{a}, \lambda \cdot \underline{a}]_{\mathbb{Z}} & \lambda < 0 \end{cases}$$

Example 3. *Consider the following examples of the above operations for integer intervals.* $[4, 5]_{\mathbb{Z}} + [-1, 2]_{\mathbb{Z}} = [3, 7]_{\mathbb{Z}}$, $[-4, 5]_{\mathbb{Z}} - [-1, 2]_{\mathbb{Z}} = [-6, 4]_{\mathbb{Z}}$, $[2, 4]_{\mathbb{Z}} \cdot [4, 6]_{\mathbb{Z}} = [8, 24]_{\mathbb{Z}}$, $3 \cdot [2, 4]_{\mathbb{Z}} = [6, 12]_{\mathbb{Z}}$. *It can be seen that the arithmetic operations for integer intervals defined above always produce integer intervals.*

It is also useful to define the continuous extension of an integer interval $[\underline{a}, \overline{a}]_{\mathbb{Z}}$ as $C([\underline{a}, \overline{a}]_{\mathbb{Z}}) = [\underline{a}, \overline{a}]$. Conversely, given $\underline{a} \leq \overline{a}$ with $\underline{a}, \overline{a} \in \mathbb{Z}$, we define the integer projection of $[\underline{a}, \overline{a}] \in \mathcal{K}_C$ as $\mathbb{Z}([\underline{a}, \overline{a}]) = [\underline{a}, \overline{a}]_{\mathbb{Z}} \in \mathcal{K}_{\mathbb{Z}}$; and in this case, it is said that $[\underline{a}, \overline{a}] \in \mathcal{K}_{C\to\mathbb{Z}}$. In other words, $\mathcal{K}_{C\to\mathbb{Z}}$ is the set of intervals whose endpoints are integer. Note also that $\mathbb{Z}(C([\underline{a}, \overline{a}]_{\mathbb{Z}})) = [\underline{a}, \overline{a}]_{\mathbb{Z}}$.

It is necessary also to define a partial order relationship for integer intervals. To this aim, we will use an adaptation of LU-fuzzy partial orders on intervals, which are well known in the literature, (see, e.g., [40, 37] and the references therein), to integer intervals. .

Definition 3. *Given two intervals $A = [\underline{a}, \overline{a}], B = [\underline{b}, \overline{b}] \in \mathcal{K}_C$, we say that:*

(i) $[\underline{a}, \overline{a}] \leqq [\underline{b}, \overline{b}]$ *if and only if* $\underline{a} \leq \underline{b}$ *and* $\overline{a} \leq \overline{b}$.

(ii) $[\underline{a}, \overline{a}] \prec [\underline{b}, \overline{b}]$ *if and only if* $\underline{a} < \underline{b}$ *and* $\overline{a} < \overline{b}$.

Definition 4. *Given two integer intervals $A = [\underline{a}, \overline{a}]_{\mathbb{Z}}, B = [\underline{b}, \overline{b}]_{\mathbb{Z}} \in \mathcal{K}_{\mathbb{Z}}$, we say that:*

(i) $[\underline{a}, \overline{a}]_{\mathbb{Z}} \leqq [\underline{b}, \overline{b}]_{\mathbb{Z}}$ *if and only if* $\underline{a} \leq \underline{b}$ *and* $\overline{a} \leq \overline{b}$.

(ii) $[\underline{a}, \overline{a}]_{\mathbb{Z}} \prec [\underline{b}, \overline{b}]_{\mathbb{Z}}$ *if and only if* $\underline{a} < \underline{b}$ *and* $\overline{a} < \overline{b}$.

In a similar manner, we define the relationships $A \geqq B$ and $A \succ B$ for intervals and integer intervals, which really means $B \leqq A$ and $B \prec A$, respectively. Note that, for the sake of simplicity, we use the same symbols of partial orders to compare intervals in $\mathcal{K}_C$ as to compare integer intervals in $\mathcal{K}_{\mathbb{Z}}$.

In the next section, to define the corresponding DEA technology, we will need to relate intervals and integer intervals. We will use the property that an integer interval in $\mathcal{K}_{\mathbb{Z}}$ is contained in Z and within the interval in $\mathcal{K}_C$ whose endpoints are the same, that is, $[\underline{a}, \overline{a}]_{\mathbb{Z}} \subseteq [\underline{a}, \overline{a}] \cap \mathbb{Z}$ for all $\underline{a} \leq \overline{a}$ with $\underline{a}, \overline{a} \in \mathbb{Z}$. Furthermore, it is derived that, given $\underline{a} \leq \overline{a}, \underline{b} \leq \overline{b}$ with $\underline{a}, \overline{a}, \underline{b}, \overline{b} \in \mathbb{Z}$, then $[\underline{a}, \overline{a}]_{\mathbb{Z}} \leqq (\prec)[\underline{b}, \overline{b}]_{\mathbb{Z}}$ if and only if $C([\underline{a}, \overline{a}]_{\mathbb{Z}}) = [\underline{a}, \overline{a}] \leqq (\prec)[\underline{b}, \overline{b}] = C([\underline{b}, \overline{b}]_{\mathbb{Z}})$.

4. Proposed integer interval PPS and slack-based measure of inefficiency

Let us consider a set of N DMUs. Each DMU_j, with $j \in J = \{1, \ldots, N\}$, consumes M inputs given by $X_j = (x_{1j}, \ldots, x_{Mj}) \in (\mathcal{K}_{\mathbb{Z}+})^M$, with $x_{ij} = [\underline{x_{ij}}, \overline{x_{ij}}]_{\mathbb{Z}} \in$

$\mathcal{K}_{\mathbb{Z}+}$ for $i \in \{1, \ldots, M\}$. Each DMU_j also produces S outputs given by $Y_j = (y_{1j}, \ldots, y_{Sj}) \in (\mathcal{K}_{\mathbb{Z}+})^S$, with $y_{rj} = [\underline{y_{rj}}, \overline{y_{rj}}]_{\mathbb{Z}} \in \mathcal{K}_{\mathbb{Z}+}$ for $r \in \{1, \ldots, S\}$. Their continuous extensions are $C(X_j) = \big(C(x_{1j}), \ldots, C(x_{Mj})\big)$ and $C(Y_j) = \big(C(y_{1j}), \ldots, C(y_{Sj})\big)$, with $C(x_{ij}) = [\underline{x_{ij}}, \overline{x_{ij}}] \in \mathcal{K}_C$, and $C(y_{rj}) = [\underline{y_{rj}}, \overline{y_{rj}}] \in \mathcal{K}_C$, respectively.

Let us consider the following axioms, which are analogous to (A1)-(A4) in Section 2 but considering integer interval inputs and outputs and using the corresponding partial order introduced in Definitions 3 and 4:

(B1) *Envelopment:* $(X_j, Y_j) \in T$, *for all* $j \in J$.

(B2) *Free disposability:* $(x, y) \in T$, $(x', y') \in (\mathcal{K}_{\mathbb{Z}}+)^{M+S}$, *such that* $x' \geqq x$, $y' \leqq y \;\; \Rightarrow (x', y') \in T$.

(B3) *Convexity:* $(x, y), (x', y') \in T$, $\alpha \in [0, 1]$, *such that* $\alpha(C(x), C(y)) + (1 - \alpha)(C(x'), C(y')) \in (\mathcal{K}_{C\to\mathbb{Z}})^{M+S} \Rightarrow (x'', y'') = \mathbb{Z}(\alpha(C(x), C(y)) + (1-\alpha)(C(x'), C(y'))) \in T$.

(B4) *Scalability:* $(x, y) \in T$, $\alpha \geq 0$, *and* $\alpha(C(x), C(y)) \in (\mathcal{K}_{C\to\mathbb{Z}})^{M+S} \Rightarrow (x'', y'') = \mathbb{Z}(\alpha(C(x), C(y))) \in T$.

Theorem 1. *Under axioms (B1), (B2), (B3) and (B4), the interval production possibility set that results from the minimum extrapolation principle is*

$$T_{IIDEA} = \left\{ (x, y) \in (\mathcal{K}_{\mathbb{Z}}+)^{M+S} : C(x) \geqq \sum_{j=1}^{N} \lambda_j C(X_j), C(y) \leqq \sum_{j=1}^{N} \lambda_j C(Y_j), \lambda_j \geq 0, \forall j \right\}$$

Proof. See Appendix A. □

After the characterization result for the T_{IIDEA} given in Theorem 1, we can formulate the following integer interval DEA (IIDEA) model, which is a slacks-based measure of inefficiency,

$$
\begin{aligned}
\text{(IIDEA)}\ \ I(X_p, Y_p) = \quad \text{Max} \quad & \sum_{i=1}^{M} \frac{\underline{s_i^x} + \overline{s_i^x}}{\underline{x_{ip}} + \overline{x_{ip}}} + \sum_{r=1}^{S} \frac{\underline{s_r^y} + \overline{s_r^y}}{\underline{y_{rp}} + \overline{y_{rp}}} \\
\text{s.t.} \quad & \sum_{j=1}^{N} \lambda_j C(x_{ij}) \leqq C(x_{ip}) - C(s_i^x), \quad i = 1, \ldots, M, \\
& \sum_{j=1}^{N} \lambda_j C(y_{rj}) \geqq C(y_{rp}) + C(s_r^y), \quad r = 1, \ldots, S, \\
& \lambda_j \geq 0, \quad j = 1, \ldots, N, \\
& s_i^x, s_r^y \in \mathcal{K}_{\mathbb{Z}+}, \quad i = 1, \ldots, M, \ \ r = 1, \ldots, S.
\end{aligned}
\tag{2}
$$

where it is assumed that all inputs $x_{ij} = [\underline{x_{ij}}, \overline{x_{ij}}]_{\mathbb{Z}}$, and outputs $y_{rj} = [\underline{y_{rj}}, \overline{y_{rj}}]_{\mathbb{Z}}$ are non-negative integer intervals and belong to $\mathcal{K}_{\mathbb{Z}+}$, $\forall i, j, r$.

Let us denote a feasible solution for (IIDEA) as $(\boldsymbol{s}^{x*}, \boldsymbol{s}^{y*}, \boldsymbol{\lambda}^*)$, where $\boldsymbol{s}^{x*} = (s_1^{x*}, \ldots, s_M^{x*}) \in (\mathcal{K}_{\mathbb{Z}+})^M$, $\boldsymbol{s}^{y*} = (s_1^{y*}, \ldots, s_S^{y*}) \in (\mathcal{K}_{\mathbb{Z}+})^S$, and $\boldsymbol{\lambda}^* = (\lambda_1^*, \ldots, \lambda_N^*) \in \mathbb{R}^N$. We will deal directly with (IIDEA) model, without any ranking function. Note that its objective function is a real number, i.e. $I(X_p, Y_p) \in \mathbb{R}$.

Definition 5. *A DMU p is said to be efficient if and only if $(x, y) \in T_{IIDEA}$, $x \leqq X_p$ and $y \geqq Y_p$ implies $(x, y) = (X_p, Y_p)$.*

Given the above integer-interval (IIDEA) model (2), efficient DMUs have a null inefficiency measure, i.e.

Theorem 2. *If DMU_p is efficient, then $I(X_p, Y_p) = 0$.*

Proof. Suppose that $I(X_p, Y_p) > 0$, with $(\boldsymbol{s}^{x*}, \boldsymbol{s}^{y*}, \boldsymbol{\lambda}^*)$ an optimal solution for (IIDEA). Let $x^* = (x_1^*, \ldots, x_M^*) \in (\mathcal{K}_{\mathbb{Z}_+})^M$, where $x_i^* = x_{ip} - s_i^{x*} = [\underline{x_{ip}} - \overline{s_i^{x*}}, \overline{x_{ip}} - \underline{s_i^{x*}}]_{\mathbb{Z}}$ for each $i = 1, \ldots, M$. And let $y^* = (y_1^*, \ldots, y_S^*) \in (\mathcal{K}_{\mathbb{Z}}+)^S$, defined as $y_r^* = y_{rp} + s_r^{y*} = [\underline{y_{rp}} + \underline{s_r^{y*}}, \overline{y_{rp}} + \overline{s_r^{y*}}]_{\mathbb{Z}}$ for $r = 1, \ldots, S$. By the model constraints,

$$
C(x^*) \geqq \sum_{j=1}^{N} \lambda_j^* C(X_j) \quad \text{and} \quad C(y^*) \leqq \sum_{j=1}^{N} \lambda_j^* C(Y_j)
$$

and hence, $(x^*, y^*) \in T_{IIDEA}$. It is clear also that $x^* \leqq X_p$ and $y^* \geqq Y_p$.

If $I(X_p, Y_p) > 0$, then $(\boldsymbol{s}^{x*}, \boldsymbol{s}^{y*}) \neq 0$, i.e., $\boldsymbol{s}^{x*} \geqq 0$, with $s^{x*}_{i_0} \neq 0$ for some i_0, or/and $\boldsymbol{s}^{y*} \geqq 0$, with $s^{y*}_{r_0} \neq 0$ for some r_0. In the first case, it must happen that $\overline{s^{x*}_{i_0}} > 0$ and therefore $x^* \leqq X_p$, with $x^* \neq X_p$. This means that $(x^*, y^*) \in T_{IIDEA}$, $x^* \leqq X_p, x^* \neq X_p$, and $y^* \geqq Y_p$, which implies that DMU p is not efficient, reaching a contradiction. Analogously, we also reach a contradiction for the second case. □

To solve (IIDEA) model at its current stage (2), we take into account the arithmetic operations (Definition 2 and order relations (Definition 4) defined in the previous section. Therefore, the Integer Interval Data Envelopment Analysis problem (IIDEA) can be reformulated or parameterized as

$$
\begin{aligned}
\text{(PIIDEA)}\ \ I(X_p, Y_p) = \quad & \text{Max} \quad \sum_{i=1}^{M} \frac{\underline{s_i^x} + \overline{s_i^x}}{\underline{x_{ip}} + \overline{x_{ip}}} + \sum_{r=1}^{S} \frac{\underline{s_r^y} + \overline{s_r^y}}{\underline{y_{rp}} + \overline{y_{rp}}} && (3)\\
& \text{s.t.} \quad \sum_{j=1}^{N} \lambda_j \underline{x_{ij}} \leq \underline{x_{ip}} - \overline{s_i^x}, \quad i = 1, \ldots, M,\\
& \qquad\ \ \sum_{j=1}^{N} \lambda_j \overline{x_{ij}} \leq \overline{x_{ip}} - \underline{s_i^x}, \quad i = 1, \ldots, M,\\
& \qquad\ \ \sum_{j=1}^{N} \lambda_j \underline{y_{rj}} \geq \underline{y_{rp}} + \underline{s_r^y}, \quad r = 1, \ldots, S,\\
& \qquad\ \ \sum_{j=1}^{N} \lambda_j \overline{y_{rj}} \geq \overline{y_{rp}} + \overline{s_r^y}, \quad r = 1, \ldots, S,\\
& \qquad\ \ \underline{s_i^x} \leq \overline{s_i^x}, \quad i = 1, \ldots, M,\\
& \qquad\ \ \underline{s_r^y} \leq \overline{s_r^y}, \quad r = 1, \ldots, S,\\
& \qquad\ \ \lambda_j \geq 0, \quad j = 1, \ldots, N,\\
& \qquad\ \ \underline{s_i^x}, \overline{s_i^x}, \underline{s_r^y}, \overline{s_r^y} \in \mathbb{Z}_+, \quad i = 1, \ldots, M, \ \ r = 1, \ldots, S.
\end{aligned}
$$

The first four sets of constraints are just the corresponding transformation of the inputs/outputs constraints from model (2), given the order relation for integer intervals, Definition 4. The fifth and sixth set of constraints ensure the slacks $s_i^x = [\underline{s_i^x}, \overline{s_i^x}]_{\mathbb{Z}_+}$ and $s_r^y = [\underline{s_r^y}, \overline{s_r^y}]_{\mathbb{Z}_+}$ are integer intervals $\mathcal{K}_{\mathbb{Z}_+}$.

The relationship between the (IIDEA) and (PIIDEA) solutions is demonstrated in the following proposition.

Proposition 1. *$(\boldsymbol{s}^{x*}, \boldsymbol{s}^{y*}, \boldsymbol{\lambda}^*)$ with $\boldsymbol{s}^{x*} \in (\mathcal{K}_{\mathbb{Z}_+})^M$, $\boldsymbol{s}^{y*} \in (\mathcal{K}_{\mathbb{Z}_+})^S$ and $\boldsymbol{\lambda}^* \in \mathbb{R}_+^N$ is an optimal solution of (IIDEA) if and only if its corresponding components or parameterization $(\underline{s_1^{x*}}, \overline{s_1^{x*}}, \ldots, \underline{s_M^{x*}}, \overline{s_M^{x*}}, \underline{s_1^{y*}}, \overline{s_1^{y*}}, \ldots, \underline{s_S^{y*}}, \overline{s_S^{y*}} \lambda_1^*, \ldots, \lambda_N^*)$, with $\lambda_j^* \in \mathbb{R}_+$, $j = 1, \ldots, N$, $\underline{s_i^{x*}}, \overline{s_i^{x*}} \in \mathbb{Z}_+$, $i = 1, \ldots, M$, and $\underline{s_r^{y*}}, \overline{s_r^{y*}} \in \mathbb{Z}_+$ for $r = 1, \ldots, S$, is an optimal solution of (PIIDEA).*

Proof. The constraint in (IIDEA) (2) are equivalent to the constraint conditions in (PIIDEA) (3), given Definitions 2 and 4. The rest of the proof is straightforward. □

Although Theorem 2 establishes it as a necessary condition, a null inefficiency measure, i.e. $I(X_p, Y_p) = 0$, is not sufficient to guarantee the efficiency of DMU_j in the integer intervals case, as it happens in the crisp model (1). This can be seen in the following example.

Example 4. *Consider six DMUs that consume two different inputs and produce a constant amount of output. Figure 1 shows the inputs of these DMUs, that produce a single and constant output. Therefore, by decreasing each input we move towards the efficiency frontier, represented with a thick grey line and delimited by DMUs 1, 2, and 6. As data are integer intervals, the inputs of each DMU are the set of integer points within such integer intervals,shown in the Figure with different shaped symbols (filled points are used for the efficient DMUs). In this small example, we can observe the different classes of DMUs in terms of their efficiency characterization. Note that DMU_3 and DMU_5 have non-zero slacks for both inputs and thus $I(X_3, Y_3) > 0$ and $I(X_5, Y_5) > 0$. According to Theorem 2, they are inefficient. On the contrary, DMU_1, DMU_2, DMU_4 and DMU_6 have zero slacks for both inputs, and hence $I(X_1, Y_1) = 0$, $I(X_2, Y_2) = 0$, $I(X_4, Y_4) = 0$ and $I(X_6, Y_6) = 0$. But this does not imply that these DMUs are efficient in the integer intervals framework. According to Definition 5, it is clear that DMU_1,*

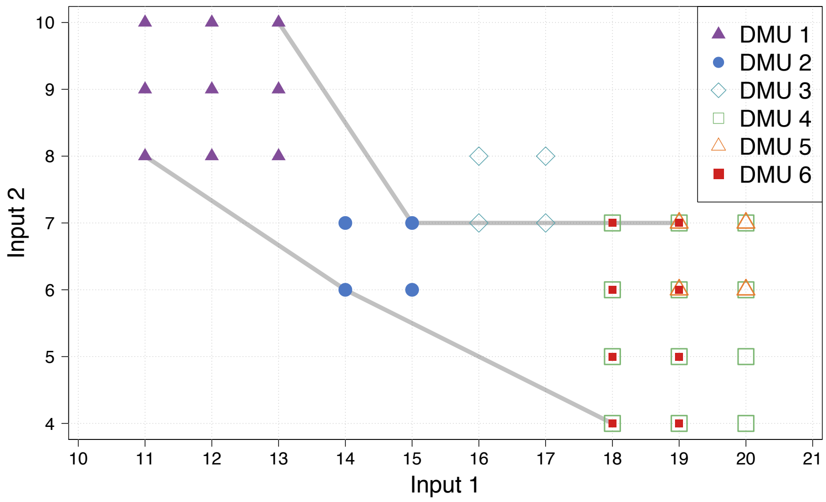

Figure 1: Consider these six *DMUs* that consume two inputs and produce a single constant output (see Example 4). The data are integer intervals and the set of integer points corresponding to each DMU is represented using different shaped symbols (see legend). The thick grey line represents the efficiency frontier. In this small example we can observe the different classes of DMUs in terms of their efficiency characterization. According to Definition 5, DMU_1, DMU_2 and DMU_6 (plotted with filled points) are efficient while the rest of DMUs are not efficient. While the inefficiency scores of the former are null, $I(X_3, Y_3) > 0$ and $I(X_5, Y_5) > 0$. Note that although DMU_4 is not efficient ($X_{1,6} \leqq X_{1,4}$, $X_{1,6} \neq X_{1,4}$) its inefficiency score $I(X_4, Y_4) = 0$. It is an example of weakly efficient DMU (see Definition 6), and this is why it is necessary a second phase for a correct efficiency characterization.

DMU_2 and DMU_6 are efficient. However, DMU_4 is not efficient, since $X_{1,6} \leqq X_{1,4}$, $X_{2,6} = X_{2,4}$ and $Y_6 = Y_4$, as it can be observed in Figure 1. Therefore, in order to exhaust all possible input and output slacks, a phase II is required to determine the efficiency character of the DMUs with null inefficiency measure $I(X_p, Y_p)$. This is performed by model (4) below, which uses additional integer-valued specific left and right slack variables, L^x, R^x, L^y and R^y. These variables allows us to detect if there still exists some remaining slack, for any input or output, that can be removed. The optimal solution of model (4) for DMU_4 has a non-zero objective function value $H(X_4, Y_4) > 0$,

which tells us that DMU_4 is not efficient but weakly efficient (see Definition 6).

Therefore, given an optimal solution for (3) $(\boldsymbol{s}^{x*}, \boldsymbol{s}^{y*}, \boldsymbol{\lambda}^*)$, we can formulate the following Phase II model to exhaust all remaining input and output slacks.

$$
\begin{aligned}
(PIIDEA)_2 \;\; H(X_p, Y_p) = \quad \text{Max} \quad & \sum_{i=1}^{M}(L_i^x + R_i^x) + \sum_{r=1}^{S}(L_r^y + R_r^y) \\
\text{s.t.} \quad & \sum_{j=1}^{N} \lambda_j \underline{x_{ij}} \le \underline{x_{ip}} - \overline{s_i^{x*}} - R_i^x, \quad i = 1, \ldots, M, \\
& \sum_{j=1}^{N} \lambda_j \overline{x_{ij}} \le \overline{x_{ip}} - \underline{s_i^{x*}} - L_i^x, \quad i = 1, \ldots, M, \\
& \sum_{j=1}^{N} \lambda_j \underline{y_{rj}} \ge \underline{y_{rp}} + \underline{s_r^{y*}} + L_r^y, \quad r = 1, \ldots, S, \\
& \sum_{j=1}^{N} \lambda_j \overline{y_{rj}} \ge \overline{y_{rp}} + \overline{s_r^{y*}} + R_r^y, \quad r = 1, \ldots, S, \\
& \lambda_j \ge 0, \quad j = 1, \ldots, N, \\
& L_i^x, R_i^x, L_r^y, R_r^y \in \mathbb{Z}_+, \quad i = 1, \ldots, M, \;\; r = 1, \ldots, S.
\end{aligned}
\tag{4}
$$

Theorem 3. *Given a DMU_p with $I(X_p, Y_p) = 0$, then $H(X_p, Y_p) = 0$ if and only if DMU_p is efficient.*

Proof. If $I(X_p, Y_p) = 0$, for a maximizing problem with non-negative variables, it is clear that $\underline{s_i^x} = \overline{s_i^x} = \underline{s_r^y} = \overline{s_r^y} = 0$, $\forall i$ and $\forall r$. Moreover, if $H(X_p, Y_p) = 0$ as well, owing to similar reasoning, this implies that all the variables $L_i^x, R_i^x, L_r^y, R_r^y$ are equal to zero, for all $i = 1, \ldots, M$, and $r = 1, \ldots, S$. Now let us assume that DMU_p is not efficient. This means that there exist $(x^*, y^*) \in T_{IIDEA}$ such that $x^* \leqq X_p$ and $y^* \geqq Y_p$, with $(x^*, y^*) \neq (X_p, Y_p)$. I.e., $x^*_{i_0} \leqq x_{i_0 p}, x^*_{i_0} \neq x_{i_0 p}$ for some $i_0 \in \{1, \ldots, M\}$, or $y^*_{r_0} \geqq y_{r_0 p}, y^*_{r_0} \neq y_{r_0 p}$ for some $r_0 \in \{1, \ldots, S\}$. In the first case, by Definition 4, as either $\underline{x^*_{i_0}} < \underline{x_{i_0 p}}$ or $\overline{x^*_{i_0}} < \overline{x_{i_0 p}}$, we can compute a new feasible solution for $(PIIDEA)_2$, such that $R^{x^*}_{i_0} = \underline{x_{i_0 p}} - \underline{x^*_{i_0}} > 0$ or $L^{x^*}_{i_0} = \overline{x_{i_0 p}} - \overline{x^*_{i_0}} > 0$. Its feasibility holds since $(x^*, y^*) \in T_{IIDEA}$, i.e.

$$\begin{aligned}
\sum_{j=1}^{N} \lambda_j \underline{x_{i_0 j}} &\leq \underline{x_{i_0 p}} - \overline{s^{x*}_{i_0}} - R^{x^*}_{i_0} = \underline{x^*_{i_0}} \\
\sum_{j=1}^{N} \lambda_j \overline{x_{ij}} &\leq \overline{x_{i_0 p}} - \underline{s^{x*}_{i_0}} - L^{x^*}_{i_0} = \overline{x^*_{i_0}}.
\end{aligned}$$

In this way we have reached a contradiction, since we have found a feasible solution with an objective function value larger than the supposed optimal value $H(X_p, Y_p) = 0$. For the second case, we also reach a contradiction with a similar reasoning, just defining a new solution with $L^{y*}_{r_0} = \underline{y_{r_0 p}} - \underline{y^*_{r_0}} > 0$ or $R^{y*}_{r_0} = \overline{y_{r_0 p}} - \overline{y^*_{r_0}} > 0$. Therefore, if $I(X_p, Y_p) = 0$, and $H(X_p, Y_p) = 0$ then DMU_p is efficient.

Finally, to proof that the efficiency of a DMU_p implies both $I(X_p, Y_p) = 0$ and $H(X_p, Y_p) = 0$, we only need to proof the latter since the necessary condition $I(X_p, Y_p) = 0$ was established in Theorem 2. Now let us suppose the opposite, $H(X_p, Y_p) > 0$. Then we can compute $(x^*, y^*) \in T_{IIDEA}$ such that $x^* \leqq X_p$ and $y^* \geqq Y_p$, with $(x^*, y^*) \neq (X_p, Y_p)$, as follows. We have four possibilities, $L^x_{i_0} > 0$, or $R^x_{i_0} > 0$ for some $i_0 \in \{1, \ldots, M\}$, or, $L^y_{r_0} > 0$, or $R^{y*}_{r_0} > 0$ for some $r_0 \in \{1, \ldots, S\}$. For the two first cases, let $y^* = Y_p$, and $x^*_i = x_{ip}$ for all $i \in \{1, \ldots, M\}$, with $i \neq i_0$. And $\underline{x^*_{i_0}} = \underline{x_{i_0 p}} - \overline{s^{x*}_{i_0}} - R^x_{i_0}$, $\overline{x^*_{i_0}} = \overline{x_{ip}} - \underline{s^{x*}_{i}} - L^x_{i_0}$. Then, $x^* \leqq X_p$ and $y^* \geqq Y_p$, with $(x^*, y^*) \neq (X_p, Y_p)$, which is a contradiction to the fact that DMU_p is efficient. Analogously, for the other two cases, let $x^* = X_p$, and $y^*_r = y_{rp}$ for all $r \in \{1, \ldots, S\}$, with $r \neq r_0$. And $\underline{y^*_{r_0}} = \underline{y_{r_0 p}} - \overline{s^{y*}_{r_0}} - L^y_{r_0}$, $\overline{y^*_{r_0}} = \overline{y_{r_0 p}} - \underline{s^{y*}_{r_0}} - R^y_{r_0}$. Again, $x^* \leqq X_p$ and $y^* \geqq Y_p$, with $(x^*, y^*) \neq (X_p, Y_p)$, which is a contradiction to the fact that DMU_p is efficient. $\square$

Let $(\boldsymbol{s}^{x*}, \boldsymbol{s}^{y*}, \boldsymbol{\lambda}^*)$ be the optimal solution for (3) and let $(\boldsymbol{L}^{x*}, \boldsymbol{R}^{x*}, \boldsymbol{L}^{y*}, \boldsymbol{R}^{y*}, \boldsymbol{\lambda}^{**})$ the optimal solution for (4) for a given DMU_p, we can compute its input and output targets X_p^{target} and Y_p^{target} as

$$\underline{x_{ip}^{target}} = \underline{x_{ip}} - \overline{s_i^{x*}} - R_i^{x*}, \qquad \overline{x_{ip}^{target}} = \overline{x_{ip}} - \underline{s_i^{x*}} - L_i^{x*}, \qquad i = 1, \ldots, M, \tag{5}$$

$$\underline{y_{rp}^{target}} = \underline{y_{rp}} + \underline{s_r^{y*}} + L_r^{y*}, \qquad \overline{y_{rp}^{target}} = \overline{y_{rp}} + \overline{s_r^{y*}} + R_r^{y*}, \qquad r = 1, \ldots, S. \tag{6}$$

Theorem 4. *$(X_p^{target}, Y_p^{target})$ is efficient.*

Proof. By the constraints of (4), it follows that $(X_p^{target}, Y_p^{target}) \in T_{IIDEA}$. Suppose that $(X_p^{target}, Y_p^{target})$ is not efficient. Then, there must exist $(x', y') \in T_{IIDEA}$ such that $x' \leqq X_p^{target}$ and $y' \geqq Y_p^{target}$, with $(x', y') \neq (X_p^{target}, Y_p^{target})$. This implies that for some $\lambda' \geq 0$,

$$C(x') \geqq \sum_{j=1}^{N} \lambda'_j C(X_j), \quad C(y') \leqq \sum_{j=1}^{N} \lambda'_j C(Y_j),$$

which is equivalent to

$$\underline{x'_i} \geq \sum_{j=1}^{N} \lambda'_j \underline{x_{ij}}, \qquad \overline{x'_i} \geq \sum_{j=1}^{N} \lambda'_j \overline{x_{ij}}, \qquad i = 1, \ldots, M,$$

$$\underline{y'_r} \leq \sum_{j=1}^{N} \lambda'_j \underline{y_{rj}}, \qquad \overline{y'_r} \leq \sum_{j=1}^{N} \lambda'_j \overline{y_{rj}}, \qquad r = 1, \ldots, S.$$

Besides,

$$\underline{x'_i} \leq \underline{x_{ip}^{target}} \qquad \overline{x'_i} \leq \overline{x_{ip}^{target}} \qquad i = 1, \ldots, M,$$

$$\underline{y'_r} \geq \underline{y_{rp}^{target}} \qquad \overline{y'_r} \geq \overline{y_{rp}^{target}} \qquad r = 1, \ldots, S.$$

where at least one of these inequalities is strict for some $i_0 \in \{1, \ldots, M\}$ or $r_0 \in \{1, \ldots, S\}$, since $(x', y') \neq (X_p^{target}, Y_p^{target})$.

Combining the above constraints, it follows that

$$\sum_{j=1}^{N} \lambda'_j \underline{x_{ij}} \leq \underline{x_{ip}} - \overline{s_i^{x*}} - R_i^{x*}, \qquad \sum_{j=1}^{N} \lambda'_j \overline{x_{ij}} \leq \overline{x_{ip}} - \underline{s_i^{x*}} - L_i^{x*}, \quad i = 1, \ldots, M,$$

$$\sum_{j=1}^{N} \lambda'_j \underline{y_{rj}} \geq \underline{y_{rp}} + \underline{s_r^{y*}} + L_r^{y*}, \qquad \sum_{j=1}^{N} \lambda'_j \overline{y_{rj}} \geq \overline{y_{rp}} + \overline{s_r^{y*}} + R_r^{y*}, \quad r = 1, \ldots, S,$$

where at least one of these inequalities is strict for some $i_0 \in \{1, \ldots, M\}$ or $r_0 \in \{1, \ldots, S\}$. Therefore, there exists some $\delta^L_{i_0}, \delta^R_{i_0}, \epsilon^L_{r_0}, \epsilon^R_{r_0} \in \mathbb{Z}_+$, where at least one of them is non-zero, such that

$$\sum_{j=1}^{N} \lambda'_j \underline{x_{i_0 j}} \leq \underline{x_{i_0 p}} - \overline{s^{x*}_{i_0}} - R^{x*}_{i_0} - \delta^R_{i_0}, \qquad \sum_{j=1}^{N} \lambda'_j \overline{x_{i_0 j}} \leq \overline{x_{i_0 p}} - \underline{s^{x*}_{i_0}} - L^{x*}_{i_0} - \delta^L_{i_0},$$

$$\sum_{j=1}^{N} \lambda'_j \underline{y_{r_0 j}} \geq \underline{y_{r_0 p}} + \underline{s^{y*}_{r_0}} + L^{y*}_{r_0} + \epsilon^L_{r_0}, \qquad \sum_{j=1}^{N} \lambda'_j \overline{y_{r_0 j}} \geq \overline{y_{r_0 p}} + \overline{s^{y*}_{r_0}} + R^{y*}_{r_0} + \epsilon^R_{r_0},$$

If we define the new variables for the corresponding sharp constraints, as

$$L^{x**}_{i_0} = L^{x*}_{i_0} + \delta^L_{i_0}, \quad R^{x**}_{i_0} = R^{x*}_{i_0} + \delta^R_{i_0}\,; \quad L^{x**}_{i} = L^{x*}_{i}, \quad R^{x**}_{i} = R^{x*}_{i} \qquad i = 1, \ldots, M, \;\; i \neq i_0$$

$$L^{y**}_{r_0} = L^{y*}_{r_0} + \epsilon^L_{r_0}, \quad R^{y**}_{r_0} = R^{y*}_{r_0} + \epsilon^R_{r_0}\,; \quad L^{y**}_{r} = L^{y*}_{r}, \quad R^{y**}_{r} = R^{y*}_{r} \qquad r = 1, \ldots, S, \;\; r \neq r_0$$

then $(\boldsymbol{L}^{x**}, \boldsymbol{R}^{x**}, \boldsymbol{L}^{y**}, \boldsymbol{R}^{y**}, \boldsymbol{\lambda}')$ would be a feasible solution in (4) with a larger objective function value than the supposed optimum, which implies a contradiction.

□

Definition 6. *For each DMU p, $p \in \{1, \ldots, N\}$, consider the inefficiency measurements $I(X_p, Y_p)$ computed in the (IIDEA), and $H(X_p, Y_p)$ obtained in Phase II, $(PIIDEA)_2$. We say that the DMU p is*

(i) efficient if $I(X_p, Y_p) = 0$ and $H(X_p, Y_p) = 0$,

(ii) weakly efficient if $I(X_p, Y_p) = 0$ and $H(X_p, Y_p) > 0$,

(iii) inefficient if $I(X_p, Y_p) > 0$.

5. Numerical experiments

5.1. Small illustrative case

Let us go back to the small dataset of Example 4 again to illustrate the proposed approach step by step, as well as the need for Phase II for the efficiency characterization and the computation of the targets. Recall that there are six

Table 1: Data for Case 5.1

DMU (j)	1	2	3	4	5	6
x_{1j}	$(11, 13)$	$(14, 15)$	$(16, 17)$	$(18, 20)$	$(19, 20)$	$(18, 19)$
x_{2j}	$(8, 10)$	$(6, 7)$	$(7, 8)$	$(4, 7)$	$(6, 7)$	$(4, 7)$
y_{1j}	$(10, 10)$	$(10, 10)$	$(10, 10)$	$(10, 10)$	$(10, 10)$	$(10, 10)$

Table 2: Results for Phases I & II, and DMU efficiency status classification for Case 5.1.

	DMU	1	2	3	4	5	6
Phase I	$I(X_p, Y_P)$	0.00	0.00	0.25	0.00	0.26	0.00
	s_1^x	$(0, 0)$	$(0, 0)$	$(2, 2)$	$(0, 0)$	$(5, 5)$	$(0, 0)$
	s_2^x	$(0, 0)$	$(0, 0)$	$(1, 1)$	$(0, 0)$	$(0, 0)$	$(0, 0)$
	s_1^y	$(0, 0)$	$(0, 0)$	$(0, 0)$	$(0, 0)$	$(0, 0)$	$(0, 0)$
Phase II	$H(X_p, Y_P)$	0	0	0	1	0	0
	L_1^x	0	0	0	1	0	0
	R_1^x	0	0	0	0	0	0
	L_2^x	0	0	0	0	0	0
	R_2^x	0	0	0	0	0	0
	L_1^y	0	0	0	0	0	0
	R_1^y	0	0	0	0	0	0
	X_1^{target}	$(11, 13)$	$(14, 15)$	$(14, 15)$	$(18, 19)$	$(14, 15)$	$(18, 19)$
	X_2^{target}	$(8, 10)$	$(6, 7)$	$(6, 7)$	$(4, 7)$	$(6, 7)$	$(4, 7)$
	Y_1^{target}	$(10, 10)$	$(10, 10)$	$(10, 10)$	$(10, 10)$	$(10, 10)$	$(10, 10)$
	Eff. Status	efficient	efficient	inefficient	weakly efficient	inefficient	efficient

DMUs, with two inputs and a single constant output (see Table 1). All the variables are assumed to be integer intervals.

Among these six DMUs, there are three classified as efficient, two inefficient, and one weakly efficient case, as established in Definition 6. Below we show, using DMU 1 as an example, the model solved and the results of the phases of the proposed approach.

Phase I: The corresponding (PIIDEA) (3) problem for DMU_1 is

$$
\begin{array}{llll}
I(X_1,Y_1) = & Max & \dfrac{\underline{s_1^x}+\overline{s_1^x}}{11+13}+\dfrac{\underline{s_2^x}+\overline{s_2^x}}{8+10}+\dfrac{\underline{s_1^y}+\overline{s_1^y}}{10+10} & \\
 & \text{s.t.} & \left.\begin{array}{lll} 11\lambda_1+14\lambda_2+16\lambda_3+18\lambda_4+19\lambda_5+18\lambda_6 & \leq & 11-\overline{s_1^x} \\ 13\lambda_1+15\lambda_2+17\lambda_3+20\lambda_4+20\lambda_5+19\lambda_6 & \leq & 13-\underline{s_1^x} \end{array}\right\} & i=1 \\
 & & \left.\begin{array}{lll} 8\lambda_1+6\lambda_2+7\lambda_3+4\lambda_4+6\lambda_5+4\lambda_6 & \leq & 8-\overline{s_2^x} \\ 10\lambda_1+7\lambda_2+8\lambda_3+7\lambda_4+7\lambda_5+7\lambda_6 & \leq & 10-\underline{s_2^x} \end{array}\right\} & i=2 \\
 & & \left.\begin{array}{lll} 10\lambda_1+10\lambda_2+10\lambda_3+10\lambda_4+10\lambda_5+10\lambda_6 & \geq & 10+\overline{s_1^y} \\ 10\lambda_1+10\lambda_2+10\lambda_3+10\lambda_4+10\lambda_5+10\lambda_6 & \geq & 10+\underline{s_1^y} \end{array}\right\} & r=1 \\
 & & \underline{s_i^x} \leq \overline{s_i^x} \quad i=1,2, & \\
 & & \underline{s_1^y} \leq \overline{s_1^y} & \\
 & & \lambda_j \geq 0, \quad j=1,\ldots,6, & \\
 & & \underline{s_i^x}, \overline{s_i^x}, \underline{s_1^y}, \overline{s_1^y} \in \mathbb{Z}_+ \quad i=1,2 &
\end{array}
$$

The optimal solution of the above Linear Program (LP) is $(\boldsymbol{s}^{x*}, \boldsymbol{s}^{y*}, \boldsymbol{\lambda}^*) = (\underline{s_1^{x*}} = 0, \overline{s_1^{x*}} = 0,\ \underline{s_2^{x*}} = 0,\ \overline{s_2^{x*}} = 0,\ \underline{s_1^{y*}} = 0,\ \overline{s_1^{y*}} = 0,\ \lambda_1^* = 1,\ \lambda_2^* = 0,\ \lambda_3^* = 0,\ \lambda_4^* = 0,\ \lambda_5^* = 0, \lambda_6^* = 0)$ As $I(X_1, Y_1) = 0$, it is a candidate to be an efficient DMU, but we cannot be sure yet. To confirm its efficiency status we need to solve the Phase II model below.

Phase II: Given the solution obtained in Phase I for DMU_1, specifically the slacks $(\boldsymbol{s}^{x*}, \boldsymbol{s}^{y*})$, the corresponding $(PIIDEA)_2$ model (4) is formulated as

$$
\begin{aligned}
H(X_1, Y_1) = \quad & \text{Max} \quad L_1^x + R_1^x + L_2^x + R_2^x + L_1^y + R_1^y \\
& \text{s.t.} \quad \left.\begin{aligned} 11\lambda_1 + 14\lambda_2 + 16\lambda_3 + 18\lambda_4 + 19\lambda_5 + 18\lambda_6 &\leq 11 - R_1^x \\ 13\lambda_1 + 15\lambda_2 + 17\lambda_3 + 20\lambda_4 + 20\lambda_5 + 19\lambda_6 &\leq 13 - L_1^x \end{aligned}\right\} \; i = 1 \\
& \qquad\;\; \left.\begin{aligned} 8\lambda_1 + 6\lambda_2 + 7\lambda_3 + 4\lambda_4 + 6\lambda_5 + 4\lambda_6 &\leq 8 - R_2^x \\ 10\lambda_1 + 7\lambda_2 + 8\lambda_3 + 7\lambda_4 + 7\lambda_5 + 7\lambda_6 &\leq 10 - L_2^x \end{aligned}\right\} \; i = 2 \\
& \qquad\;\; \left.\begin{aligned} 10\lambda_1 + 10\lambda_2 + 10\lambda_3 + 10\lambda_4 + 10\lambda_5 + 10\lambda_6 &\geq 10 + R_1^y \\ 10\lambda_1 + 10\lambda_2 + 10\lambda_3 + 10\lambda_4 + 10\lambda_5 + 10\lambda_6 &\geq 10 + L_1^y \end{aligned}\right\} \; r = 1 \\
& \qquad \lambda_j \geq 0, \quad j = 1, \ldots, 6, \\
& \qquad L_i^x, R_i^x, L_r^y, R_r^y \in \mathbb{Z}_+, \quad i = 1, 2, \;\; r = 1.
\end{aligned}
$$

The optimal solution of the above LP problem is $(\boldsymbol{L}^{x*}, \boldsymbol{R}^{x*}, \boldsymbol{L}^{y*}, \boldsymbol{R}^{y*}, \boldsymbol{\lambda}^{**}) = (L_1^{x*} = 0, R_1^{x*} = 0, L_2^{x*} = 0, R_2^{x*} = 0, L_1^{y*} = 0, R_1^{y*} = 0, \lambda_1^{**} = 1, \lambda_2^{**} = 0, \lambda_3^{**} = 0, \lambda_4^{**} = 0, \lambda_5^{**} = 0, \lambda_6^{**} = 0)$. The left and right slack variables $L_i^x, R_i^x, L_r^y, R_r^y$ represent the potential improvements that may remain and correspond to moving, if possible, towards the efficiency frontier. Only for efficient DMUs these variables are all null, as it happens for DMU_1.

In this case, the corresponding input and output targets, as per (5) and (6), coincide with those of the observed DMU, i.e.

$$
\begin{aligned}
&\underline{x_{11}^{target}} = \underline{x_{11}} - \overline{s_1^{x*}} - R_1^{x*} = 11 - 0 - 0 = 11, &&\overline{x_{11}^{target}} = \overline{x_{11}} - \underline{s_1^{x*}} - L_1^{x*} = 13 - 0 - 0 = 13, \\
&\underline{x_{21}^{target}} = \underline{x_{21}} - \overline{s_2^{x*}} - R_2^{x*} = 8 - 0 - 0 = 8, &&\overline{x_{21}^{target}} = \overline{x_{21}} - \underline{s_2^{x*}} - L_2^{x*} = 10 - 0 - 0 = 10, \\
&\underline{y_{11}^{target}} = \underline{y_{11}} + \underline{s_1^{y*}} + L_1^{y*} = 10 + 0 + 0 = 10, &&\overline{y_{11}^{target}} = \overline{y_{11}} + \overline{s_1^{y*}} + R_1^{y*} = 10 + 0 + 0 = 10.
\end{aligned}
$$

As it can be seen in Table 2, in the case of DMU_4, the solution of the Phase I is $I(X_4, Y_4) = 0$, similar to what happens for DMU_1, DMU_2 and DMU_6. Unlike them, however, for DMU_4, the Phase II solution $L_1^x = 1$ and $H(X_4, Y_4) = 1$ indicates that the upper limit of the first input of DMU_4 can be feasibly reduced by one unit and hence DMU_4 is not efficient.

Table 3: Phase I results for Case 5.2. This is a hybrid problem. The second input and both outputs are integer, whereas the other three inputs are continuous.

p	$I(X_p, Y_p)$	Input slacks intervals				Output slacks intervals	
		s_1^x	$\mathbf{s_2^x}$	s_3^x	s_4^x	$\mathbf{s_1^y}$	$\mathbf{s_2^y}$
1	0.00	(0.00 , 0.00)	(0 , 0)	(0.00 , 0.00)	(0.00 , 0.00)	(0 , 0)	(0 , 0)
2	2.29	(0.44 , 0.44)	(0 , 0)	(6.70 , 6.70)	(0.00 , 0.00)	(0 , 0)	(75 , 75)
3	1.21	(0.00 , 0.00)	(1 , 1)	(40.09 , 40.09)	(0.00 , 0.00)	(0 , 6)	(59 , 62)
4	1.37	(0.00 , 0.00)	(53 , 53)	(0.00 , 0.00)	(12.06 , 12.06)	(0 , 2)	(78 , 79)
5	0.56	(169.13 , 169.13)	(0 , 0)	(2.37 , 2.37)	(0.10 , 0.10)	(0 , 10)	(33 , 37)
6	0.00	(0.00 , 0.00)	(0 , 0)	(0.00 , 0.00)	(0.00 , 0.00)	(0 , 0)	(0 , 0)
7	0.00	(0.00 , 0.00)	(0 , 0)	(0.00 , 0.00)	(0.00 , 0.00)	(0 , 0)	(0 , 0)
8	0.00	(0.00 , 0.00)	(0 , 0)	(0.00 , 0.00)	(0.00 , 0.00)	(0 , 0)	(0 , 0)
9	0.00	(0.00 , 0.00)	(0 , 0)	(0.00 , 0.00)	(0.00 , 0.00)	(0 , 0)	(0 , 0)
10	1.91	(0.05 , 0.05)	(39 , 39)	(7.19 , 7.19)	(0.08 , 0.08)	(0 , 0)	(143, 143)
11	1.03	(18.42 , 18.42)	(0 , 0)	(21.92 , 21.92)	(26.45 , 26.45)	(0 , 0)	(0 , 0)
12	1.14	(0.00 , 0.00)	(21 , 21)	(0.19, 0.19)	(16.16, 16.16)	(0 , 7)	(67 , 70)
13	1.37	(142.78 , 142.78)	(60 , 60)	(0.00 , 0.00)	(15.82 , 15.82)	(0 , 4)	(58 , 59)
14	2.31	(0.00 , 0.00)	(24 , 24)	(5.98 , 5.98)	(52.49 , 52.49)	(0 , 0)	(152 , 152)
15	3.61	(0.00 , 0.00)	(110 , 110)	(10.36 , 10.36)	(36.90 , 36.90)	(0 , 0)	(158 , 158)
16	0.00	(0.00 , 0.00)	(0 , 0)	(0.00 , 0.00)	(0.00 , 0.00)	(0 , 0)	(0 , 0)
17	0.00	(0.00 , 0.00)	(0 , 0)	(0.00 , 0.00)	(0.00 , 0.00)	(0 , 0)	(0 , 0)
18	6.94	(0.00 , 0.00)	(141 , 141)	(13.97 , 13.97)	(27.17 , 27.17)	(0 , 13)	(311 , 315)
19	0.00	(0.00 , 0.00)	(0 , 0)	(0.00 , 0.00)	(0.00 , 0.00)	(0 , 0)	(0 , 0)
20	2.79	(0.02 , 0.02)	(92 , 92)	(36.33 , 36.33)	(03.16 , 3.16)	(20 , 24)	(200 , 202)
21	1.96	(0.00 , 0.00)	(56 , 56)	(28.37 , 28.37)	(38.25, 38.25)	(0 , 0)	(74 , 74)
22	0.00	(0.00 , 0.00)	(0 , 0)	(0.00 , 0.00)	(0.00 , 0.00)	(0 , 0)	(0 , 0)
23	2.14	(0.00 , 0.00)	(56 , 56)	(13.03 , 13.03)	(24.06, 24.06)	(0 , 16)	(210 , 216)
24	0.95	(0.04 , 0.04)	(62 , 62)	(19.74 , 19.74)	(0.09 , 0.09)	(0 , 0)	(38 , 38)
25	3.90	(29.81 , 29.81)	(91, 91)	(0.03 , 0.03)	(2.52 , 2.52)	(36 , 45)	(252 , 255)
26	1.35	(0.06,0.06)	(58,58)	(11.52,11.52)	(18.33,18.33)	(0,4)	(50,52)

5.2. *Larger real-world application*

In this section, we take a real problem, which not only is bigger but also includes both integer and continuous variables. This case may be found more often than the pure integer one in the real world. The extension of the proposed approach to the hybrid scenario is not difficult and has been included in Appendix B.

The dataset considered comes from Majid Azadi et al. [4]. The original data are given as triangular fuzzy numbers. To adapt them as intervals we have considered the corresponding zero $\alpha-$levels. The DMUs correspond to 26 suppliers of raw materials with four crisp inputs and two integer interval outputs. The inputs are the economic criteria given by the total cost of shipments (TC), and the number of shipments per month (NS) and the social criteria given by the eco-design cost (ED) and the cost of work safety and labor health (CS). Except for the NS input, the rest of the inputs are continuous variables. The two outputs are the number of shipments to arrive on time (NOT) and the number of bills received from the supplier without errors (NB). Both outputs are integer interval variables.

The results from the Phase I, model (B.2) (see Appendix B), are shown in Table 3. The results of the Phase II model (B.3), as well as the input and output targets, which are interval variables, and the corresponding efficiency status are given in Table 4. As we can see in the table, all DMUs are classified as either efficient or inefficient, i.e., there are no weakly efficient DMUs in this case.

Table 5 compares the results of the proposed approach with the inefficiency scores and the corresponding targets when the integrality of the integer variables is ignored. These results correspond to relaxing the integrality of the corresponding input and output slacks in models (B.2) and (B.3), which are the hybrid equivalent of models (3) and (4). Because they are relaxations of the original models, they can compute slightly higher inefficiency scores. However, we claim that those results are not valid because they correspond to targets that, as shown in Table 5, do not always respect the integer character of some of the variables (the second input, and the two outputs in the current instance).

On the contrary, the proposed approach considers both the integer and the interval-valued character of those variables.

For the sake of comparison, Table 6 also includes the results when other existing approaches are applied, in particular the model (3.8) from Kordrostami et al. [24], which also considers a hybrid case of integer and continuous variables. As already discussed in Section 1, these authors consider fuzzy data, whereas we consider that the uncertainty is given in terms of interval data. To apply their models we consider interval data as a particular case of trapezoidal fuzzy data (a, b, c, d), when $a = b$ and $c = d$. We do not include the results from their alternative model (3.9), since they are the same in the case of interval data.

Among the main differences between our approach and Kordrostami et al. [24], already discussed in Section 1, we have that Kordrostami et al. [24] use a fuzzy ranking approach and get crisp targets (see last columns of Table 6), while we use integer interval arithmetic and compute integer interval targets. In addition, they use a radial oriented approach (while we apply an additive, non-oriented approach) and they use the integer PPS of Kuosmanen & Matin [27] (while we use a specific integer interval PPS).

In spite of these differences, analysing the results of both approaches, we can check that they are in good agreement. In particular, the corresponding efficient characterisations coincide. Thus, the efficient DMUs identified by the proposed approach, with both inefficient null values $I(X_p, Y_p) = H(X_p, Y_p) = 0$, have also an efficiency score of 1 with the Kordrostami et al. approach. And those inefficient DMUS, with $I(X_p, Y_p) > 0$, have an efficiency score less than the unity. Also, the Spearman rank-order correlation coefficient between both approaches is $\rho = -0.91$. Finally, regarding the targets for the efficient DMUs for the second input variables, which are not interval since the original input data was not interval, they coincide with Kordrostami et al.'s targets. Moreover, the integer (non-interval) output targets from Kordrostami et al.'s model are contained within the integer interval targets computed from the proposed approach.

Table 4: Results for Phase II, including targets and DMU efficiency status classification for Case 5.2. Second input and both outputs are integer, whereas the other three inputs are continuous.

			Phase II slack variables												TARGETS						
p	$I(X_p, Y_p)$	$H(X_p, Y_p)$	Input s.v.								Output s.v.				TC	**NS**	ED	CS	**NOT**	**NB**	Efficiency
			L_1^x	R_1^x	$\mathbf{L_2^x}$	$\mathbf{R_2^x}$	L_3^x	R_3^x	L_4^x	R_4^x	$\mathbf{L_1^y}$	$\mathbf{R_1^y}$	$\mathbf{L_2^y}$	$\mathbf{R_2^y}$	X_1^{target}	$\mathbf{X_2^{target}}$	X_3^{target}	X_4^{target}	$\mathbf{Y_1^{target}}$	$\mathbf{Y_2^{target}}$	Status
1	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	(316.00 , 316.00)	(251 , 251)	(61.00 , 61.00)	(18.00 , 18.00)	(199 , 239)	(76 , 90)	Eff.
2	2.29	4	0	0	0	0	0	0	0	0	3	0	1	0	(280.56 , 280.56)	(164 , 164)	(38.30 , 38.30)	(21.00 , 21.00)	(156 , 193)	(105 , 117)	Ineff.
3	1.21	0	0	0	0	0	0	0	0	0	0	0	0	0	(309.00 , 309.00)	(197 , 197)	(42.91 , 42.91)	(40.00 , 40.00)	(203 , 249)	(137 , 154)	Ineff.
4	1.37	0	0	0	0	0	0	0	0	0	0	0	0	0	(291.00 , 291.00)	(165 , 165)	(37.00 , 37.00)	(32.94 , 32.94)	(167, 209)	(163 , 178)	Ineff.
5	0.56	0	0	0	0	0	0	0	0	0	0	0	0	0	(427.87 , 427.87)	(178 , 178)	(49.63 , 49.63)	(28.90 , 28.90)	(197 , 247)	(196 , 214)	Ineff.
6	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	(341.00 , 341.00)	(142 , 142)	(19.00 , 19.00)	(33.00 , 33.00)	(129 , 169)	(129 , 143)	Eff.
7	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	(475.00 , 475.00)	(149 , 149)	(74.00 , 74.00)	(18.00 , 18.00)	(193 , 233)	(111 , 125)	Eff.
8	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	(254.00 , 254.00)	(172 , 172)	(53.00 , 53.00)	(35.00 , 35.00)	(134 , 174)	(250 , 264)	Eff.
9	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	(328.00 ,328.00)	(135 , 135)	(83.00 , 83.00)	(47.00 , 47.00)	(184 , 224)	(58 , 72)	Eff.
10	1.91	1	0	0	0	0	0	0	0	0	1	0	0	0	(309.95 , 309.95)	(134 , 134)	(33.81 , 33.81)	(15.92 , 15.92)	(114 , 153)	(231 , 245)	Ineff.
11	1.03	6	0	0	0	0	0	0	0	0	5	0	1	0	(302.58 , 302.58)	(121 , 121)	(35.08 , 35.08)	(18.55 , 18.55)	(130 , 165)	(154 , 167)	Ineff.
12	1.14	0	0	0	0	0	0	0	0	0	0	0	0	0	(329.00 , 329.00)	(183 , 183)	(37.81 , 37.81)	(36.84 , 36.84)	(195 , 242)	(157 , 174)	Ineff.
13	1.37	0	0	0	0	0	0	0	0	0	0	0	0	0	(332.22 , 332.22)	(152 , 152)	(32.00 , 32.00)	(26.18 , 26.18)	(156 , 200)	(197 , 212)	Ineff.
14	2.31	0	0	0	0	0	0	0	0	0	0	0	0	0	(259.00 , 259.00)	(165 , 165)	(50.02 , 50.02)	(32.51 , 32.51)	(129 , 169)	(249 , 263)	Ineff.
15	3.61	8	0	0	0	0	0	0	0	0	6	0	2	0	(274.00 , 274.00)	(107 , 107)	(27.64 , 27.64)	(14.10 , 14.10)	(91 , 125)	(228 , 240)	Ineff.
16	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	(264.00 , 264.00)	(158 , 158)	(25.00 , 25.00)	(35.00 , 35.00)	(193 , 233)	(45 , 59)	Eff.
17	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	(327.00 , 327.00)	(124 , 124)	(32.00 , 32.00)	(16.00, 16.00)	(107 , 147)	(271 , 285)	Eff.
18	6.94	0	0	0	0	0	0	0	0	0	0	0	0	0	(429.00 , 429.00)	(166 , 166)	(43.03 , 43.03)	(21.83 , 21.83)	(142 , 195)	(357 , 375)	Ineff.
19	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	(262.00 , 262.00)	(138 , 138)	(25.00 , 25.00)	(31.00 , 31.00)	(122 , 162)	(173 , 187)	Eff.
20	2.79	0	0	0	0	0	0	0	0	0	0	0	0	0	(384.98 , 384.98)	(146 , 146)	(37.67 , 37.67)	(18.84 , 18.84)	(126 , 173)	(319 , 335)	Ineff.
21	1.96	1	0	0	0	0	0	0	0	0	1	0	0	0	(249.00 , 249.00)	(161 , 161)	(40.63 , 40.63)	(33.75 , 33.75)	(151 , 190)	(165 , 178)	Ineff.
22	0.00	0	0	0	0	0	0	0	0	0	0	0	0	0	(337.00 , 337.00)	(203 , 203)	(27.00 , 27.00)	(33.00 , 33.00)	(104 , 144)	(271 , 285)	Eff.
23	2.05	0	0	0	0	0	0	0	0	0	0	0	0	0	(365.00 , 365.00)	(236 , 236)	(71.97 , 71.97)	(49.94 , 49.94)	(185 , 241)	(353 , 373)	Ineff.
24	0.95	2	0	0	0	0	0	0	0	0	2	0	0	0	(295.96 , 295.96)	(123 , 123)	(29.26 , 29.26)	(17.91 , 17.91)	(114 , 152)	(216 , 229)	Ineff.
25	3.90	0	0	0	0	0	0	0	0	0	0	0	0	0	(398.19 , 398.19)	(151 , 151)	(38.97 , 38.97)	(19.48 , 19.48)	(130 , 179)	(330 ,347)	Ineff.
26	1.35	0	0	0	0	0	0	0	0	0	0	0	0	0	(326.94,326.94)	(160,160)	(31.48,31.48)	(29.67,29.67)	(173,217)	(163,179)	Ineff.

Table 5: Comparison of the results for Case 5.2, without integrality constraints.

	Hybrid case		Continuous case		TARGETS for continous case						
p	$I(X_p, Y_p)$	$H(X_p, Y_p)$	$I(X_p, Y_p)$	$H(X_p, Y_p)$	TC X_1^{target}	NS X_2^{target}	ED X_3^{target}	CS X_4^{target}	NOT Y_1^{target}	NB Y_2^{target}	Efficiency Status
1	0.00	0	0.00	0.00	(316.00, 316.00)	(251.00, 251.00)	(61.00, 61.00)	(18.00, 18.00)	(199.00, 239.00)	(76.00, 90.00)	Eff.
2	2.29	4	2.31	5.22	(281.00, 281.00)	(164.00, 164.00)	(38.31, 38.31)	(21.00, 21.00)	(156.86, 193.00)	(105.12, 117.76)	Ineff.
3	1.21	0	1.22	0.00	(309.00, 309.00)	(196.90, 196.90)	(42.91, 42.91)	(40.00, 40.00)	(203.00, 249.94)	(137.58, 154.01)	Ineff.
4	1.37	0	1.37	0.00	(291.00, 291.00)	(164.62, 164.62)	(37.00, 37.00)	(32.94, 32.94)	(167.00, 209.46)	(163.13, 177.99)	Ineff.
5	0.56	0	0.56	0.00	(426.42, 426.42)	(178.00, 178.00)	(49.32, 49.32)	(29.00, 29.00)	(197.00, 247.29)	(195.67, 213.27)	Ineff.
6	0.00	0	0.00	0.00	(341.00, 341.00)	(142.00, 142.00)	(19.00, 19.00)	(33.00, 33.00)	(129.00, 169.00)	(129.00, 143.00)	Eff.
7	0.00	0	0.00	0.00	(475.00, 475.00)	(149.00, 149.00)	(74.00, 74.00)	(18.00, 18.00)	(193.00, 233.00)	(111.00, 125.00)	Eff.
8	0.00	0	0.00	0.00	(254.00, 254.00)	(172.00, 172.00)	(53.00, 53.00)	(35.00, 35.00)	(134.00, 174.00)	(250.00, 264.00)	Eff.
9	0.00	0	0.00	0.00	(328.00, 328.00)	(135.00, 135.00)	(83.00, 83.00)	(47.00, 47.00)	(184.00, 224.00)	(58.00, 72.00)	Eff.
10	1.91	1	1.91	2.35	(310.00, 310.00)	(133.61, 133.61)	(33.67, 33.67)	(16.00, 16.00)	(114.74, 153.00)	(231.88, 245.27)	Ineff.
11	1.03	6	1.03	6.78	(302.58, 302.58)	(121.00, 121.00)	(35.08, 35.08)	(18.55, 18.55)	(130.02, 165.00)	(154.76, 167.00)	Ineff.
12	1.14	0	1.14	0.00	(329.00, 329.00)	(183.11, 183.11)	(38.00, 38.00)	(36.92, 36.92)	(195.00, 242.75)	(158.11, 174.83)	Ineff.
13	1.37	0	1.38	0.00	(330.82, 330.82)	(151.38, 151.38)	(32.00, 32.00)	(26.08, 26.08)	(156.00, 199.92)	(196.33, 211.70)	Ineff.
14	2.31	0	2.32	0.63	(259.00, 259.00)	(164.73, 164.73)	(50.08, 50.08)	(32.51, 32.51)	(129.47, 169.00)	(249.36, 263.19)	Ineff.
15	3.61	8	3.61	8.38	(274.00, 274.00)	(106.21, 106.21)	(27.67, 27.67)	(14.10, 14.10)	(91.21, 125.00)	(228.28, 240.11)	Ineff.
16	0.00	0	0.00	0.00	(264.00, 264.00)	(158.00, 158.00)	(25.00, 25.00)	(35.00, 35.00)	(193.00, 233.00)	(45.00, 59.00)	Eff.
17	0.00	0	0.00	0.00	(327.00, 327.00)	(124.00, 124.00)	(32.00, 32.00)	(16.00, 16.00)	(107.00, 147.00)	(271.00, 285.00)	Eff.
18	7.20	0	7.21	0.00	(429.00, 429.00)	(165.09, 165.09)	(42.88, 42.88)	(21.71, 21.71)	(142.00, 194.76)	(356.79, 375.26)	Ineff.
19	0.00	0	0.00	0.00	(262.00, 262.00)	(138.00, 138.00)	(25.00, 25.00)	(31.00, 31.00)	(122.00, 162.00)	(173.00, 187.00)	Eff.
20	2.79	0	2.79	0.00	(385.00, 385.00)	(145.99, 145.99)	(37.68, 37.68)	(18.84, 18.84)	(125.98, 173.07)	(319.07, 335.55)	Ineff.
21	1.96	1	1.97	1.85	(249.00, 249.00)	(160.88, 160.88)	(40.75, 40.75)	(33.80, 33.80)	(151.37, 190.00)	(165.05, 178.57)	Ineff.
22	0.00	0	0.00	0.00	(337.00, 337.00)	(203.00, 203.00)	(27.00, 27.00)	(33.00, 33.00)	(104.00, 144.00)	(271.00, 285.00)	Eff.
23	2.05	0	2.14	0.00	(365.00, 365.00)	(235.92, 235.92)	(71.98, 71.98)	(46.94, 46.94)	(185.00, 241.15)	(353.38, 373.04)	Ineff.
24	0.95	2	0.95	3.46	(296.00, 296.00)	(122.55, 122.55)	(29.10, 29.10)	(18.00, 18.00)	(114.56, 152.00)	(216.19, 229.29)	Ineff.
25	3.90	0	3.90	0.00	(398.53, 398.53)	(151.12, 151.12)	(39.00, 39.00)	(19.50, 19.50)	(130.41, 179.16)	(330.28, 347.34)	Ineff.
26	1.35	0	1.36	0.00	(327.00, 327.00)	(159.84, 159.84)	(31.48, 31.48)	(29.67, 29.67)	(173.00, 217.77)	(163.42, 179.09)	Ineff.

Table 6: Inefficiency measurements (Phases I and II) for Case 5.2, compared to the Efficiency score and integer input and output targets from Kordrostami et al. [24], see their model (3.8).

	proposed approach					Kordrostami et al. [24]			
p	$I(X_p, Y_p)$	$H(X_p, Y_p)$	X_2^{target}	Y_1^{target}	Y_2^{target}	Efficiency	X_2^{target}	Y_1^{target}	Y_2^{target}
1	0.00	0	$(251, 251)$	$(199, 239)$	$(76, 90)$	1.00	251	219	83
2	2.29	4	$(164, 164)$	$(156, 193)$	$(105, 117)$	0.95	155	173	64
3	1.21	0	$(197, 197)$	$(203, 249)$	$(137, 154)$	0.93	177	223	85
4	1.37	0	$(165, 165)$	$(167, 209)$	$(163, 178)$	0.84	151	187	92
5	0.56	0	$(178, 178)$	$(197, 247)$	$(196, 214)$	0.95	169	217	170
6	0.00	0	$(142, 142)$	$(129, 169)$	$(129, 143)$	1.00	142	149	136
7	0.00	0	$(149, 149)$	$(193, 233)$	$(111, 125)$	1.00	149	213	118
8	0.00	0	$(172, 172)$	$(134, 174)$	$(250, 264)$	1.00	172	154	257
9	0.00	0	$(135, 135)$	$(184, 224)$	$(58, 72)$	1.00	135	204	65
10	1.91	1	$(134, 134)$	$(114, 153)$	$(231, 245)$	0.83	143	133	95
11	1.03	6	$(121, 121)$	$(130, 165)$	$(154, 167)$	0.96	116	145	160
12	1.14	0	$(183, 183)$	$(195, 242)$	$(157, 174)$	0.85	172	215	97
13	1.37	0	$(152, 152)$	$(156, 200)$	$(197, 212)$	0.78	154	176	146
14	2.31	0	$(165, 165)$	$(129, 169)$	$(249, 263)$	0.79	129	149	104
15	3.61	8	$(107, 107)$	$(91, 125)$	$(228, 240)$	0.53	91	105	75
16	0.00	0	$(158, 158)$	$(193, 233)$	$(45, 59)$	1.00	158	213	52
17	0.00	0	$(124, 124)$	$(107, 147)$	$(271, 285)$	1.00	124	127	278
18	7.20	0	$(166, 166)$	$(142, 195)$	$(357, 375)$	0.49	137	162	53
19	0.00	0	$(138, 138)$	$(122, 162)$	$(173, 187)$	1.00	138	142	180
20	2.79	0	$(146, 146)$	$(126, 173)$	$(319, 335)$	0.60	140	128	126
21	1.96	1	$(161, 161)$	$(151, 190)$	$(165, 178)$	0.91	141	170	97
22	0.00	0	$(203, 203)$	$(104, 144)$	$(271, 285)$	1.00	203	124	278
23	2.05	0	$(236, 236)$	$(185, 241)$	$(353, 373)$	0.77	179	205	150
24	0.95	2	$(123, 123)$	$(114, 152)$	$(216, 229)$	0.89	141	132	184
25	3.90	0	$(151, 151)$	$(130, 179)$	$(330, 347)$	0.63	110	114	85
26	1.35	0	$(160, 160)$	$(173, 217)$	$(163, 179)$	0.80	163	193	120

6. Conclusions

This paper presents a new integer and interval-valued DEA approach and associated slacks-based measure of inefficiency. It requires solving two crisp linear optimization models that allow the computation of the corresponding input and output targets, as well as determining the efficiency status of each DMU. Computational experiments have been presented to validate the proposed approach.

It has been shown that a null value of the Phase I inefficiency score is a necessary but not sufficient condition for efficiency, i.e. the Phase I model cannot discriminate between efficient and weakly efficient DMUs. This is analogous to what happens with radial DEA models in crisp case although it does not happen in the slacks-based case. This highlights the differences between crisp and interval data scenarios. Hence the need for the Phase II model, which also provides efficient input and output targets.

The proposed approach can handle data that are simultaneously uncertain and integer. Existing interval DEA approaches do not consider integer data and, conversely, integer DEA approaches assume crisp data. Although at the cost of requiring interval arithmetic and relational operators, with a higher number of constraints in its parameterization form, the proposed approach is able to address the joint integer interval scenario. It does so in a rigorous way, defining the corresponding integer interval PPS, its corresponding efficient subset and finally, formulating the models that compute the inefficiency scores and the efficient targets.

As regards potential research directions, we envisage extending the proposed integer interval arithmetic and LU-partial order approach to the data case with fuzzy integer intervals. The approach should be non-oriented and guarantee efficient (i.e. non-dominated) fuzzy targets. As a first step, the fuzzy integer interval DEA technology needs to be axiomatically derived. Another interesting line of research, often neglected in the fuzzy DEA literature, is that of applying this type of approaches to real-world situations, e.g. manufacturing, healthcare or transportation, in which there may be uncertainty in the input and output data.

7. Acknowledgements

The first and second authors are partially supported by grant PID2019-105824GB-I00. The fourth author acknowledges the financial support of the Spanish Ministry of Science, Innovation and Universities, grant PGC2018-095786-B-I00.

8. References

[1] Ajirlo, S.F., Kordrostami, S. and Amirteimoori, A. Two-stage additive integer-valued data envelopment analysis models: case of Iranian power industry. *Journal of modelling in Management*, 14, 1 (2019) 199-213

[2] Amirteimoori, A. and Kordrostami, S. , Data envelopment analysis with discrete-valued inputs and outputs, *Expert Systems*, 31 (2014) 335-342

[3] Apt, K.R. and Zoeteweij, P.A. Comparative Study of Arithmetic Constraints on Integer Intervals. *International workshop on constraint solving and constraint logic programming*, 3010 (2004) 1-24

[4] Azadi, M., Jafarian, M., Farzipoor Saen, R. and Mirhedayatian, S.M. A new fuzzy DEA model for evaluation of efficiency and effectiveness of suppliers in sustainable supply chain management context. *Computers and Operation Research*, 54 (2015) 274-285

[5] Azizi, H., Kordrostami, S. and Amirteimoori, A. Slack-based measure of efficiency in imprecise data envelopment analysis: An approach based on data envelopment analysis with double frontier. *Computers and Industrial Engineering*, 79(2015) 42-51

[6] Banker, R.J., 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

[7] Charnes, A., Cooper, W. W., Golany, B. and Seiford, L. Foundations of data envelopment analysis for pareto-koopmans efficient empirical production functions. *Journal of Econometrics*, 30 (1985) 91-107

[8] Charnes, A., Cooper, W. W., and Rhodes, E. Measuring the efficiencies of DMUs.*European Journal of Operational Research*, 2, 6 (1978) 429-444

[9] Chen, Y.C., Chiu, Y.H., Huang, C.W. and Tu, C.H., The analysis of bank business performance and market risk-applying fuzzy DEA. *Economic Modelling*, 32, 1 (2013) 225-232

[10] Chen, Y., Gong, Y. and Li, X. Evaluating NBA player performance using bounded integer data envelopment analysis. *INFOR: Information Systems and Operational Research*, 55, 1 (2017) 38-51

[11] Cooper, W.W., Huang, Z. and Li, S.X., Chance Constrained DEA, in *Handbook on Data Envelopment Analysis*, Cooper, W.W., Seiford, L.M. and Zhu, J. (eds.), 2004, pp 229-264

[12] Despotis, D.K. and Smirlis, Y.G. Data envelopment analysis with imprecise data. *European Journal of Operational Research*, 140 (2002) 24-36

[13] Du, J., Chen, C.M., Chen, Y., Cook, W.D. and Zhu, J. Addititive super-efficiency in integer-valued data envelopment analysis. *European Journal of Operational Research*, 218 (2012) 186-192

[14] Hatami-Marbini, A., Beigi, Z.G., Fukuyama, H. and Gholami, K. Modeling Centralized Resources Allocation and Target Setting in Imprecise Data Envelopment Analysis. *International Journal of Information Technology & Decision Making*, 14, 6 (2015) 1189-1213

[15] Hatami-Marbini, A., Ebrahimnejad, A. and Agrell, P.J. Interval data without sign restrictions in DEA. *Applied Mathematical Modelling* , 38 (2014) 2028-2036

[16] Inuiguchi, M. and Mizoshita, F. Qualitative and quantitative data envelopment analysis with interval data. *Annals of Operations Research*, 195 (2012) 189-220

[17] Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., Rezaie, V. and Khanmohammadi, M. Ranking DMUs by ideal points with interval data in DEA. *Applied Mathematical Modelling*, 35, (2011) 218-229

[18] Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., Rostamy Malkhalifeh, M. and Ahadzadeh Namin, M. A generalized model for data envelopment analysis with interval data. *Applied Mathematical Modelling*, 33, (2009) 3237-3244

[19] Jahanshahloo, G.R., Hosseinzadeh Lotfi, F. and Moradi, M. Sensitivity and stability analysis in DEA with interval data. *Applied Mathematics and Computation*, 156, 2 (2004) 463-477

[20] Jie, T., Yan, Q. and Xu, W. A technical note on " A note on integer-valued radial model in DEA". *Computers and Industrial Engineering*, 87 (2015) 308-310

[21] Kazemi Matin, R. and Kuosmanen, T. Theory of integer-valued data envelopment analysis under alternative returns to scale axioms. *Omega*, 37,5 (2009) 988-995

[22] Khalili-Damghani, K., Tavana, M. and Haji-Saami, E. A data envelopment analysis model with interval data and undesirable output for combined cycle power plant performance aseeessment. *Expert Systems with Applications*, 42 (2015) 760-773

[23] Khoveyni, M., Eslami, R., Fukuyama, H., Yang, G.L. and Sahoo, B.K. Integer data in DEA: Illustrating the drawbacks and recognizing congestion. *Computers and Industrial Engineering*, 135 (2019) 675-688

[24] Kordrostami, S., Amirteimoori, A. and Jahani Sayyad Noveiri, M. Fuzzy integer-valued Data Envelopment Analysis. *RAIRO-Operations Research*, 52 (2018) 1429-1444

[25] Kordrostami, S., Amirteimoori, A. and Jahani Sayyad Noveiri, M. Inputs and outputs classification in integer-valued data envelopment analysis. *Measurement*, 139 (2019) 317-325

[26] Kordrostami, S. and Jahani Sayyad Noveiri, M. Evaluating the performance and classifying the interval data in data envelopment analysis. *International Journal of Management Science and Engineering Management*, 9, 4 (2014) 243-248

[27] Kuosmanen, T. and Kazemi Matin, R. Theory of integer-valued data envelopment analysis. *European Journal of Opertional Research*, 192, 2 (2009) 658-667

[28] Lee, Y.K., Park, K.S. and Kim, S.H. Identification of inefficiencies in an additive model based IDEA (imprecise data envelopment analysis). *Computers and Operations Research*, 29 (2002) 1661-1676

[29] Lozano, S. and Calzada-Infante, L. Computing gradient-based stepwise benchmarking paths. *Omega*, 81 (2018) 195-207

[30] Lozano, S. and Soltani, N. Efficiency asssessment using a multidirectional DDF approach. *International Transactions in Operational Research*, 27 (2000a) 2064-2080

[31] Lozano, S. and Soltani, N. Lexicographic hyperbolic DEA. *Journal of Operational Research Society*, 71, 6 (2020b) 979-990

[32] Lozano, S. and Soltani, N. A modified discrete Raiffa approach for efficiency assessment and target setting. *Annals of Operations Research*, 292 (2020) 71-95

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

[34] Lozano, S., Villa, G. and Canca, D. Appliction of centralised DEA approach to capital budgeting in Spanish ports. *Computers and Industrial Engineering*, 60, (2011) 455-465

[35] 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

[36] Soltani, N. and Lozano, S. Potential-Based Efficiency Assessment and Target Setting. *Computers and Industrial Engineering*, 126 (2018) 611-624 arithmetic. *Fuzzy Sets and Systems*, 161 (2010) 1564-1584

[37] Stefanini, L. and Arana-Jiménez, M. Karush-–Kuhn—Tucker conditions for interval and fuzzy optimization in several variables under total and directional generalized differentiability. *Fuzzy Sets and Systems*, 362 (2019) 1-34

[38] Tan, Y., Shetty, U., Diabat, A. and Pakkala, T.P.M. Aggregate directional distance formulation of DEA with integer variables. *Annals of Operations Research*, 325 (2015) 741-756

[39] Wang, Y.M., Greatbanks, R. and Yang, J.B. Interval efficiency assessment using data envelopment analysis. *Fuzzy Sets and Systems*, 153 (2005) 347-370

[40] Wu, H.C. The optimality conditions for optimization problems with convex constraints and multiple fuzzy-valued objective functions. *Fuzzy Optimization and Decision Making*, 8 (2009) 295-321

[41] Wu, J., Liang, L. and Song, H. Measuring hotel performance using the integer DEA model. *Tourism Economics*, 16, 4 (2010) 867-882

[42] Wu, J., Liang, L. and Zhou, Z.X. Measuring the performance of nations at Beijing summer Olympics using integer-valued DEA model. *Journal of Sports Economics*, 11 (2010) 549-566

[43] Yu, S.H. and Hsu, C.W. A unified extension of super-efficiency in additive data envelopment analysis with integer-valued inputs and outputs: an application to a municipal bus system. *Annals of Operations Research*, 287 (2020) 515-535

[44] Zhu, J. Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets and DEA Excel Solver. *Kluwer Academic Publishers*, Boston, 2002

[45] Zhu, J. Imprecise data envelopment analysis (IDEA): a review and improvement with an application *European Journal of Operational Research*, 144 (2003) 513-529

Appendix A. Proof of Theorem 1

Proof. Denote by T_{true} the result of the minimum extrapolation principle axioms (B1), (B2), (B3) and (B4). To prove the theorem it is necessary to show that $T_{FDEA} = T_{true}$. To this end, let us divide the proof into two parts.

(i) $T_{true} \subseteq T_{IIDEA}$.

It is sufficient to prove that T_{IIDEA} satisfies (B1), (B2), (B3) and (B4), since this implies that T_{IIDEA} contains the intersection of all sets that satisfies the previous axioms, and consequently contains T_{true}. Therefore, let us check the axioms (B1), (B2), (B3) and (B4) by T_{IIDEA}.

- Check (B1). It is clear since, given $j \in J$, then (X_j, Y_j), with $\lambda_j = 1$ and $\lambda_{j'} = 0$, for all $j' \neq j$, satisfies conditions in T_{IIDEA}.
- Check (B2). Given $(x, y) \in T_{IIDEA}$, $x' \geqq x$, $y' \leqq y$, $(x', y') \in (\mathcal{K}_{\mathbb{Z}}+)^{m+s}$, we have to prove that $(x', y') \in T_{IIDEA}$. By hypothesis, there exists $\lambda \geqq 0$ such

that

$$C(x) \geqq \sum_{j=1}^{n} \lambda_j C(X_j), \quad C(y) \leqq \sum_{j=1}^{n} \lambda_j C(Y_j). \tag{A.1}$$

Combining (A.1) with $x' \geqq x$, $y' \leqq y$, it follows that

$$C(x') \geqq C(x) \geqq \sum_{j=1}^{n} \lambda_j C(X_j), \quad C(y') \leqq C(y) \leqq \sum_{j=1}^{n} \lambda_j C(Y_j). \tag{A.2}$$

Therefore, $(x', y') \in T_{IIDEA}$.

- Check (B3). Let us consider $(x, y), (x', y') \in T_{IIDEA}$, and $\alpha \geq 0$, what means that there exist $\lambda, \lambda' \geqq 0$ such that

$$C(x) \geqq \sum_{j=1}^{n} \lambda_j C(X_j), \quad C(x') \geqq \sum_{j=1}^{n} \lambda'_j C(X_j), \tag{A.3}$$

$$C(y) \leqq \sum_{j=1}^{n} \lambda_j C(Y_j), \quad C(y') \leqq \sum_{j=1}^{n} \lambda'_j C(Y_j). \tag{A.4}$$

Multiplying by α each side in the first interval inequality in (A.3), by $(1-\alpha)$ each side in the second interval inequality in (A.3), and then combining the interval inequalities, we get

$$\alpha C(x) + (1-\alpha) C(x') \geqq \sum_{j=1}^{n} (\alpha \lambda_j + (1-\alpha)\lambda'_j) C(X_j), \tag{A.5}$$

Proceeding in a similar way with y and y' and inequalities (A.4), we have

$$\alpha C(y) + (1-\alpha) C(y') \leqq \sum_{j=1}^{n} (\alpha \lambda_j + (1-\alpha)\lambda'_j) C(Y_j), \tag{A.6}$$

We can see that $(\alpha C(x) + (1-\alpha)C(x'), \alpha C(y) + (1-\alpha)C(y')) = \alpha(C(x), C(y)) + (1-\alpha)(C(x'), C(y'))$. By hypothesis, $\alpha(C(x), C(y)) + (1-\alpha)(C(x'), C(y')) \in (\mathcal{K}_{C\to\mathbb{Z}})^{m+s}$. Define $\lambda'' = (\lambda''_1, \ldots, \lambda''_n)$, with $\lambda''_j = \alpha\lambda_j + (1-\alpha)\lambda'_j \geq 0$, for all $j = 1, \ldots, n$, and substitute them in expressions (A.5) and (A.6). Then, it follows that $(x'', y'') = \mathbb{Z}(\alpha(C(x), C(y)) + (1-\alpha)(C(x'), C(y'))) \in T_{IIDEA}$.

- Check (B4). Given $(x, y) \in T_{IIDEA}$, there exists $\lambda \geqq 0$ such that (A.1) holds. Given $\alpha \in \mathbb{R}_+$, and $(\alpha C(x), \alpha C(y)) \in (\mathcal{K}_{C\to\mathbb{Z}})^{m+s}$, it follows that there exists $\mathbb{Z}((\alpha C(x), \alpha C(y))) = (\alpha x, \alpha y) \in (\mathcal{K}_{\mathbb{Z}}+)^{m+s}$. Define $\bar{\lambda} = \alpha\lambda = (\alpha\lambda_1, \ldots, \alpha\lambda_n) \geqq 0$. Then, multiplying by α each side in the inequalities in (A.1),

$$C(\alpha x) \geqq \sum_{j=1}^{n} \alpha\lambda_j C(X_j) = \sum_{j=1}^{n} \bar{\lambda}_j C(X_j), \quad C(\alpha y) \leqq \sum_{j=1}^{n} \alpha\lambda_j C(Y_j) = \sum_{j=1}^{n} \bar{\lambda}_j C(Y_j).$$

Therefore, $(\alpha x, \alpha y) \in T_{IIDEA}$

(ii) $T_{IIDEA} \subseteq T_{true}$.

We need to prove that every element of T_{IIDEA} belongs to T_{true}. To this purpose, consider $(x, y) \in T_{IIDEA}$, which means that there exists $\lambda \geqq 0$, $\lambda \in \mathbb{R}^n$, such that

$$C(x) \geqq \sum_{j=1}^{n} \lambda_j C(X_j), \quad C(y) \leqq \sum_{j=1}^{n} \lambda_j C(Y_j), \tag{A.7}$$

what is equivalent to say

$$[\underline{x_i}, \overline{x_i}] \geqq \sum_{j=1}^{n} \lambda_j [\underline{x_{ij}}, \overline{x_{ij}}] = \Big[\sum_{j=1}^{n} \lambda_j \underline{x_{ij}}, \sum_{j=1}^{n} \lambda_j \overline{x_{ij}}\Big], \quad i = 1, \ldots, m, \tag{A.8}$$

$$[\underline{y_r}, \overline{y_r}] \leqq \sum_{j=1}^{n} \lambda_j [\underline{y_{ij}}, \overline{y_{ij}}] = \Big[\sum_{j=1}^{n} \lambda_j \underline{y_{ij}}, \sum_{j=1}^{n} \lambda_j \overline{y_{ij}}\Big], \quad r = 1, \ldots, s. \tag{A.9}$$

The relationships (A.8) and (A.9) imply

$$\underline{x_i} \geq \sum_{j=1}^{n} \lambda_j \underline{x_{ij}}, \quad \overline{x_i} \geq \sum_{j=1}^{n} \lambda_j \overline{x_{ij}}, \quad i = 1, \ldots, m, \tag{A.10}$$

$$\underline{y_r} \leq \sum_{j=1}^{n} \lambda_j \underline{y_{rj}}, \quad \overline{y_r} \leq \sum_{j=1}^{n} \lambda_j \overline{y_{rj}}, \quad r = 1, \ldots, s. \tag{A.11}$$

Taking into account the inequalities given by (A.10) and (A.11), we consider the following two cases:

- Suppose that there exists some index and some inequality, among those given by (A.10) and (A.11), such that the inequality becomes equality. For the sake of simplicity, suppose that the equality is verified for an inequality in the first group of (A.10), that is, there exists $i \in \{1, \ldots, m\}$ such that $\underline{x_i} = \sum_{j=1}^{n} \lambda_j \underline{x_{ij}} \in \mathbb{Z}_+$. The latter implies that $\lambda_j \in \mathbb{Q}_+$, for all j, with $\mathbb{Q} \subseteq \mathbb{R}$ the subset of rational numbers. Then, there exist $u_j, v_j \in \mathbb{Z}_+, v_j \neq 0$, with u_j a pair number, such that $\lambda_j = \frac{u_j}{v_j}$, for all j. Define $v = \prod_{j=1}^{n} v_j$, and $n_j = v\lambda_j \in \mathbb{Z}_+$. We point out that n_j is a pair number, that is, $0.5n_j \in \mathbb{Z}_+$, what is used in a next step in this proof. If we multiply each side of the interval inequalities given in (A.7), then if follows

$$vC(x) \geqq \sum_{j=1}^{n} n_j C(X_j), \quad vC(y) \leqq \sum_{j=1}^{n} n_j C(Y_j). \tag{A.12}$$

We have that $(X_j, Y_j) \in T_{true}$ by (B1), for all $j \in J$. Since $(n_j X_j, n_j Y_j)$ and $(0.5n_j X_j, 0.5n_j Y_j) \in (\mathcal{K}_{\mathbb{Z}_+})^{m+s}$, then, by (B4), it follows that $(n_j X_j, n_j Y_j)$ and $(0.5n_j X_j, 0.5n_j Y_j) \in T_{true}$, for all $j \in J$, and the relationships (A.12) can be written as

$$vx \geqq \sum_{j=1}^{n} n_j X_j, \quad vy \leqq \sum_{j=1}^{n} n_j Y_j. \tag{A.13}$$

Following, and reasoning by induction, let us prove that

$$\Big(\sum_{j=1}^{k} n_j X_j, \sum_{j=1}^{k} n_j Y_j\Big) \in T_{true}, \quad k = 1, \ldots, n. \tag{A.14}$$

To this matter, first we check that in the case $k = 1$ it holds, such as it has been proved before. Following, we check that if cases $k \leq t$ are true, this implies that the case $k = t + 1$ is also true. We can write $(\sum_{j=1}^{t+1} n_j X_j, \sum_{j=1}^{t+1} n_j Y_j)$ as the convex sum of two elements of T_{true}, multiplied by a scalar. Define $\alpha = 0.5$ and $\alpha' = 2$, then:

$$\begin{aligned}
\Big(\sum_{j=1}^{t+1} n_j X_j, \sum_{j=1}^{t+1} n_j Y_j\Big) &= \Big(\sum_{j=1}^{t} n_j X_j, \sum_{j=1}^{t} n_j Y_j\Big) + \Big(n_{t+1} X_{t+1}, n_{t+1} Y_{t+1}\Big) \\
&= \mathbb{Z}\Big(\alpha'\Big(\alpha\Big(\sum_{j=1}^{t} n_j C(X_j), \sum_{j=1}^{t} n_j C(Y_j)\Big) + \\
&\quad + \ (1-\alpha)\Big(n_{t+1} C(X_{t+1}), n_{t+1} C(Y_{t+1})\Big)\Big)\Big).
\end{aligned}$$

Since

$$\alpha\Big(\sum_{j=1}^{t} n_j C(X_j), \sum_{j=1}^{t} n_j C(Y_j)\Big) + (1-\alpha)\Big(n_{t+1}C(X_{t+1}), n_{t+1}C(Y_{t+1})\Big) =$$

$$\Big(\sum_{j=1}^{t+1} 0.5 n_j C(X_j), \sum_{j=1}^{t+1} 0.5 n_j C(Y_j)\Big) \in (\mathcal{K}_{C\to\mathbb{Z}})^{m+s},$$

then, by (B3), it follows that

$$\mathbb{Z}\Big(\alpha\Big(\sum_{j=1}^{t} n_j C(X_j), \sum_{j=1}^{t} n_j C(Y_j)\Big) + (1-\alpha)\Big(n_{t+1}C(X_{t+1}), n_{t+1}C(Y_{t+1})\Big)\Big) \in T_{true}.$$

If it is multiplied by $\alpha' = 2$, then, by (B4), it follows that

$$\alpha'\Big(\mathbb{Z}\Big(\alpha\Big(\sum_{j=1}^{t} n_j C(X_j), \sum_{j=1}^{t} n_j C(Y_j)\Big) + (1-\alpha)\Big(n_{t+1}C(X_{t+1}), n_{t+1}C(Y_{t+1})\Big)\Big)\Big) \in T_{true}.$$

Thus, $\left(\sum_{j=1}^{t+1} n_j X_j, \sum_{j=1}^{t+1} n_j Y_j\right) \in T_{true}$, and therefore (A.14) holds. As a consequence of (A.14), we have that $\left(\sum_{j=1}^{n} n_j X_j, \sum_{j=1}^{n} n_j Y_j\right) \in T_{true}$. Since (vx, vy) satisfies (A.13), then it also satisfies (A.12). Then, by (B2), we have that $(vx, vy) \in T_{true}$. And since $\frac{1}{v}(vx, vy) = (x, y) \in (\mathcal{K}_{\mathbb{Z}_+})^{m+s}$, then, by (B2), it follows that $(x, y) \in T_{true}$.

- Suppose that there exists no index and inequality, among those given by (A.10) and (A.11), such that the inequality becomes equality. In such a case, all inequalities are sharp, and it is not difficult to see that there exists $\delta > 0$, small enough, such that

$$\underline{x_i} > \sum_{j=1}^{n} (\lambda_j + \delta)\underline{x_{ij}}, \quad \overline{x_i} > \sum_{j=1}^{n} (\lambda_j + \delta)\overline{x_{ij}}, \quad i = 1, \ldots, m, \tag{A.15}$$

$$\underline{y_r} < \sum_{j=1}^{n} (\lambda_j + \delta)\underline{y_{rj}}, \quad \overline{y_r} < \sum_{j=1}^{n} (\lambda_j + \delta)\overline{y_{rj}}, \quad r = 1, \ldots, s. \tag{A.16}$$

We choose $\lambda_j' \in (\lambda_j, \lambda_j + \delta) \cap \mathbb{Q}_+ \neq \emptyset$, for $j \in \{1, \ldots, n\}$. Then, from (A.10), (A.11), (A.15) and (A.16), it follows

$$\underline{x_i} > \sum_{j=1}^{n} \lambda_j' \underline{x_{ij}}, \quad \overline{x_i} > \sum_{j=1}^{n} \lambda_j' \overline{x_{ij}}, \quad i = 1, \ldots, m, \tag{A.17}$$

$$\underline{y_r} < \sum_{j=1}^{n} \lambda'_j \underline{y_{rj}}, \quad \overline{y_r} < \sum_{j=1}^{n} \lambda'_j \overline{y_{rj}}, \quad r = 1, \ldots, s. \tag{A.18}$$

In particular,

$$\underline{x_i} \geq \sum_{j=1}^{n} \lambda'_j \underline{x_{ij}}, \quad \overline{x_i} \geq \sum_{j=1}^{n} \lambda'_j \overline{x_{ij}}, \quad i = 1, \ldots, m, \tag{A.19}$$

$$\underline{y_r} \leq \sum_{j=1}^{n} \lambda'_j \underline{y_{rj}}, \quad \overline{y_r} \leq \sum_{j=1}^{n} \lambda'_j \overline{y_{rj}}, \quad r = 1, \ldots, s. \tag{A.20}$$

Reasoning as above, we conclude that $(x, y) \in T_{true}$. Therefore, $T_{FDEA} \subseteq T_{true}$ and the proof is complete. □

Appendix B. Extension to the hybrid data scenario

Consider the hybrid scenario in which, in addition to integer interval data, there exist some inputs or outputs that are given as continuous intervals. Then, and following Lozano and Villa [33], we can partition each index set into two subsets; one for continuous variables, and another for integer variables. In this manner, for input and output variables, we have $O^X = O^{XI} \cup O^{XNI}$, $O^Y = O^{YI} \cup O^{YNI}$, respectively, where $O^X = \{1, \ldots, M\}$ and $O^Y = \{1, \ldots, S\}$. So, inputs $x_{ij} = [\underline{x_{ij}}, \overline{x_{ij}}]_{\mathbb{Z}} \in \mathcal{K}_{\mathbb{Z}_+}$, for all $i \in O^{XI}$, and $x_{ij} = [\underline{x_{ij}}, \overline{x_{ij}}] \in \mathcal{K}_{C_+}$, for all $i \in O^{XNI}$; and outputs $y_{rj} = [\underline{y_{rj}}, \overline{y_{rj}}]_{\mathbb{Z}} \in \mathcal{K}_{\mathbb{Z}_+}$ for all $r \in O^{YI}$, and $y_{rj} = [\underline{y_{rj}}, \overline{y_{rj}}] \in \mathcal{K}_{C_+}$ for all $r \in O^{YNI}$.

The model (IIDEA) becomes as follows, under the consideration of a hybrid interval DEA.

$$
\begin{aligned}
\text{(HIDEA)}\;\; I(X_p, Y_p) = \quad \text{Max} \quad & \sum_{i=1}^{M} \frac{\underline{s_i^x} + \overline{s_i^x}}{\underline{x_{ip}} + \overline{x_{ip}}} + \sum_{r=1}^{S} \frac{\underline{s_r^y} + \overline{s_r^y}}{\underline{y_{rp}} + \overline{y_{rp}}} && \text{(B.1)} \\
\text{s.t.} \quad & \sum_{j=1}^{N} \lambda_j C(x_{ij}) \leqq C(x_{ip}) - C(s_i^x), \quad i \in O^{XI}, \\
& \sum_{j=1}^{N} \lambda_j x_{ij} \leqq x_{ip} - s_i^x, \quad i \in O^{XNI}, \\
& \sum_{j=1}^{N} \lambda_j C(y_{rj}) \geqq C(y_{rp}) + C(s_r^y), \quad r \in O^{YI}, \\
& \sum_{j=1}^{N} \lambda_j y_{rj} \geqq y_{rp} + s_r^y, \quad r \in O^{YNI}, \\
& \lambda_j \geq 0, \quad j = 1, \ldots, N, \\
& s_i^x, s_r^y \in \mathcal{K}_{\mathbb{Z}+}, \quad i \in O^{XI}, \;\; r \in O^{YI}. \\
& s_i^x, s_r^y \in \mathcal{K}_{C+}, \quad i \in O^{XNI}, \;\; r \in O^{YNI}.
\end{aligned}
$$

To solve (HIDEA) model (B.1), we consider its following parameterization, which can be considered as the Phase I of the solution method.

$$
\begin{aligned}
\text{(PIHIDEA)}\ \ I(X_p, Y_p) = \quad \text{Max} \quad & \sum_{i=1}^{M} \frac{\underline{s_i^x} + \overline{s_i^x}}{\underline{x_{ip}} + \overline{x_{ip}}} + \sum_{r=1}^{S} \frac{\underline{s_r^y} + \overline{s_r^y}}{\underline{y_{rp}} + \overline{y_{rp}}} \qquad \text{(B.2)} \\
\text{s.t.} \quad & \sum_{j=1}^{N} \lambda_j \underline{x_{ij}} \leq \underline{x_{ip}} - \overline{s_i^x}, \quad i \in O^X, \\
& \sum_{j=1}^{N} \lambda_j \overline{x_{ij}} \leq \overline{x_{ip}} - \underline{s_i^x}, \quad i \in O^X, \\
& \sum_{j=1}^{N} \lambda_j \underline{y_{rj}} \geq \underline{y_{rp}} + \underline{s_r^y}, \quad r \in O^Y, \\
& \sum_{j=1}^{N} \lambda_j \overline{y_{rj}} \geq \overline{y_{rp}} + \overline{s_r^y}, \quad r \in O^Y, \\
& \underline{s_i^x} \leq \overline{s_i^x}, \quad i \in O^X, \\
& \underline{s_r^y} \leq \overline{s_r^y}, \quad r \in O^Y, \\
& \lambda_j \geq 0, \quad j = 1, \ldots, N, \\
& \underline{s_i^x}, \overline{s_i^x}, \underline{s_r^y}, \overline{s_r^y} \in \mathbb{Z}_+, \quad i \in O^{XI},\ \ r \in O^{YI}, \\
& \underline{s_i^x}, \overline{s_i^x}, \underline{s_r^y}, \overline{s_r^y} \geq 0, \quad i \in O^{XNI},\ \ r \in O^{YNI}.
\end{aligned}
$$

As it can be seen, the only difference with respect the corresponding model (3) is the that only the slacks of the integer inputs and outputs are forced to be integer. The slacks of the other inputs and outputs are considered continuous variables.

Given $(\boldsymbol{s}^{x*}, \boldsymbol{s}^{y*}, \boldsymbol{\lambda}^*)$, optimal solution for (B.2), we proceed with the phase II of the method.

$$
\begin{aligned}
(PHIDEA)_2 \;\; H(X_p, Y_p) = \quad \text{Max} \quad & \sum_{i=1}^{M} L_i^x + R_i^x + \sum_{r=1}^{S} L_r^y + R_r^y && \text{(B.3)}\\
\text{s.t.} \quad & \sum_{j=1}^{N} \lambda_j \underline{x_{ij}} \leq \underline{x_{ip}} - \overline{s_i^{x*}} - R_i^x, \quad i \in O^X, \\
& \sum_{j=1}^{N} \lambda_j \overline{x_{ij}} \leq \overline{x_{ip}} - \underline{s_i^{x*}} - L_i^x, \quad i \in O^X, \\
& \sum_{j=1}^{N} \lambda_j \underline{y_{rj}} \geq \underline{y_{rp}} + \underline{s_r^{y*}} + L_r^y, \quad r \in O^Y, \\
& \sum_{j=1}^{N} \lambda_j \overline{y_{rj}} \geq \overline{y_{rp}} + \overline{s_r^{y*}} + R_r^y, \quad r \in O^Y, \\
& \lambda_j \geq 0, \quad j = 1, \ldots, N, \\
& L_i^x, R_i^x, L_r^y, R_r^y \in \mathbb{Z}_+, \quad i \in O^{XI}, \;\; r \in O^{YI}, \\
& L_i^x, R_i^x, L_r^y, R_r^y \geq 0, \quad i \in O^{XNI}, \;\; r \in O^{YNI}.
\end{aligned}
$$

Given a DMU_p with $I(X_p, Y_p) = 0$, then $H(X_p, Y_p) = 0$ if and only if DMU_p is efficient. In other words, a DMU_p is efficient if and only if both $I(X_p, Y_p) = 0$ and $H(X_p, Y_p) = 0$.

Let $(\boldsymbol{s}^{x*}, \boldsymbol{s}^{y*}, \boldsymbol{\lambda}^*)$ be the optimal solution of (B.2), and $(\boldsymbol{L}^{x*}, \boldsymbol{R}^{x*}, \boldsymbol{L}^{y*}, \boldsymbol{R}^{y*}, \boldsymbol{\lambda}^{**})$ the optimal solution of (B.3) for a given DMU_p, then we can compute its input and output targets X_p^{target} and Y_p^{target} as

$$
\underline{x_{ip}^{target}} = \underline{x_{ip}} - \overline{s_i^{x*}} - R_i^{x*}, \qquad \overline{x_{ip}^{target}} = \overline{x_{ip}} - \underline{s_i^{x*}} - L_i^{x*}, \qquad i \in O^X, \tag{B.4}
$$

$$
\underline{y_{rp}^{target}} = \underline{y_{rp}} + \underline{s_r^{y*}} + L_r^{y*}, \qquad \overline{y_{rp}^{target}} = \overline{y_{rp}} + \overline{s_r^{y*}} + R_r^{y*}, \qquad r \in O^Y. \tag{B.5}
$$