



# Exact and stochastic methods for robustness analysis in the context of Imprecise Data Envelopment Analysis

Anna Labijak-Kowalska<sup>1</sup> · Miłosz Kadziński<sup>1</sup>

Received: 7 September 2021 / Revised: 6 October 2022 / Accepted: 30 October 2022 /  
Published online: 2 March 2023  
© The Author(s) 2023

## Abstract

We consider the problem of measuring the efficiency of decision-making units with a ratio-based model. In this perspective, we introduce a framework for robustness analysis that admits both interval and ordinal performances on inputs and outputs. The proposed methodology exploits the uncertainty related to the imprecise data and all feasible input/output weight vectors delimited through linear constraints. We offer methods for verifying the robustness of three types of outcomes: efficiency scores, efficiency preference relations, and efficiency ranks. On the one hand, we formulate mathematical programming models to compute the extreme, necessary, and possible results. On the other hand, we incorporate the stochastic analysis driven by the Monte Carlo simulations to derive the probability distribution of different outcomes. The framework is implemented in R and made available on open-source software. Its use is illustrated in two case studies concerning Chinese ports or industrial robots.

**Keywords** Data Envelopment Analysis · Imprecise performances · Robustness analysis · Monte Carlo simulation · Open-source software

## 1 Introduction

Data Envelopment Analysis (DEA) measures the relative efficiency of Decision Making Units (DMUs) (Cooper et al. 2014). The standard Charnes-Cooper-Rhodes (CCR) model used in DEA generalizes the single output/input productivity measure (Farrell 1957) by transforming the characterization of each DMU in terms of

---

✉ Miłosz Kadziński  
milosz.kadzinski@cs.put.poznan.pl

Anna Labijak-Kowalska  
anna.labijak@cs.put.poznan.pl

<sup>1</sup> Faculty of Computing and Telecommunications, Poznań University of Technology, Piotrowo 2, 60-965 Poznań, Poland

multiple desired outputs and multiple input factors (Charnes et al. 1978). Specifically, the efficiency is quantified as a ratio between a single virtual output and a single virtual input (Salo and Punkka 2011). When evaluating the efficiency of each DMU, the weights involved in the definition of efficiency measure are selected to identify the most advantageous scenario. This means that an efficiency score of a given DMU is maximized subject to both the constraint that all DMUs can have scores lesser or equal to the unity and the feasibility of input/output weights. As a result, DEA generates relative efficiency measures, which depend on the set of analyzed DMUs, leading to the identification of the so-called efficient frontier (Charnes et al. 1994). The units that lie on the frontier attain the score of one. In contrast, the units with a score lesser than one are below the efficient frontier, hence being classified as inefficient.

The main advantages of DEA derive from its following features (Charnes et al. 1994). First, DEA conducts a detailed analysis of performance measures for each DMU instead of focusing on the population averages. This allows for understanding the status of efficiency for individual observations. Moreover, in the case of inefficiency, one could identify its sources and point out the desired modifications of inputs and/or outputs for projecting the DMU onto the efficient frontier (see Aparicio et al. 2007; Chen and Wang 2020; Wu et al. 2018). Second, DEA does not involve any assumption about the functional form, hence not relating the independent and dependent variables (i.e., inputs and outputs) in any specific way. In turn, it evaluates each DMU relative to other DMUs, while not requiring any prior specification of weights. Finally, a great advantage of DEA lies in its simplicity and generality. It captures the efficiency in utilizing the inputs to produce the outputs, all expressed in various units, with a single, easily interpretable performance measure.

For the last forty years, many extensions of DEA have been proposed (see Cook and Seiford 2009; Emrouznejad and Yang 2018). The traditional DEA models assumed that the consumed inputs and produced outputs could be precisely expressed with numerical values on a ratio scale. However, in many real-world problems, this is not possible for a few reasons (see Aparicio et al. 2019; Cooper et al. 1999; Shokouhi et al. 2010). These reasons include inexact specification of inputs and outputs, the uncertainty of data used to compute the consumed inputs or desired outputs, subjectivity involved in this process, and high costs in terms of time or financial resources needed for conducting the accurate measurements (Corrente et al. 2017). As a result, the measurements of inputs and outputs often remain imperfect. This, in turn, requires methodological developments that could handle such uncertain or inaccurate evaluations.

In the context of DEA, two types of imperfect inputs and outputs received particular attention (Liu et al. 2013). On the one hand, the basic idea to capture the uncertainty is using an imprecise evaluation in terms of the interval of possible values. On the other hand, ordinal assessments can be considered. The latter is helpful if only qualitative information is available, some binary features are involved in the analysis, or it is possible to obtain the ranking of units in terms of some input or output instead of precise quantitative measurements.

To handle the imprecision of inputs and outputs, Cooper et al. (1999) proposed Imprecise DEA (IDEA), where precise performance values were replaced with

intervals. This methodology has been further revised and enriched in different ways. For example, Kim et al. (1999) accounted for strong and weak ordinal relations as well as ratio interval data. Furthermore, Despotis and Smirlis (2002) dealt with transforming the interval performances into precise ones, incorporating them into a standard DEA model to optimize the computational performance of the problem. Moreover, Zhu (2003) developed a linear programming model handling strong ordinal inputs and outputs. Also, Ebrahimi and Khalili (2018) proposed the models—incorporating preference information—that find the most preferred DMU and rank other efficient DMUs. The DEA models handling imprecise data have been successfully used in the telecommunication sector (Cooper et al. 2001), machinery industry (Kao and Liu 2005), wheat farming (Hadi-Vencheh and Matin 2011), port efficiency assessment (Zahran et al. 2020), and healthcare (see Azadi and Saen 2013; Karsak and Karadayi 2017).

In the traditional DEA and IDEA methods, only the most favorable input/output weight vector is considered when evaluating each DMU's performance. This may be criticized for a few reasons. First, choosing the individual weight vector for each DMU makes the comparison of efficiencies questionable due to the non-uniqueness of the most advantageous weight vectors and lack of a common basis to analyze the attained scores (Lahdelma and Salminen 2006). Second, such an analysis is focused on a minimal set of scenarios while ignoring other feasible weight vectors that could provide helpful information on the variety of efficiency scores (Salo and Punkka 2011). Third, the efficient frontier, which forms the basis for evaluating the DMUs, requires prior assumptions of the return-to-scale. Besides, it strongly depends on the set of considered DMUs (see Zhu 1996; Seiford and Zhu 1998). Fourth, using a single efficiency measure that divides the DMUs into efficient and inefficient ones offers too limited capabilities for discriminating between the units (see Adler et al. 2002; Hosseinzadeh Lotfi et al. 2013). All these drawbacks motivated the development of robustness analysis methods, which quantify the stability of efficiency results for different feasible weight vectors. Given imprecise inputs and outputs, the need to include uncertainties when working out the results is even more evident. The robust conclusions should be valid in all or most scenarios (see Kadziński and Tervonen 2013; Liang et al. 2020), with a scenario being equivalent to a set of possible values for data of the problem and the efficiency model parameters.

Some essential methodological advancements oriented toward robustness analysis for IDEA have been proposed over the last two decades. In particular, Despotis and Smirlis (2002) derived the optimistic and pessimistic efficiency scores for each DMU. Both are computed with the most favorable weight vectors for a given unit while assuming the most and the least advantageous scenarios for the inputs and outputs. Based on these results, the units can be divided into three groups: efficient in the most pessimistic scenario, inefficient even in the most optimistic scenario, and an intermediate class including DMUs with unitary optimistic efficiency and pessimistic efficiency lesser than one. This classification was further analyzed in Jahanshahloo et al. (2004) to consider the “radius of stability”. For each DMU, it is defined with a pair of values,  $\alpha$  and  $\beta$ , indicating, respectively, a decrease of the upper bounds of input and output intervals and an increase of the respective lower bounds for which the efficiency class remains unchanged. Furthermore, Kao

(2006) proposed mathematical models for computing the optimistic and pessimistic efficiency scores in the presence of both interval and ordinal inputs and outputs. A similar aim of deriving an efficiency interval for each DMU—though in different settings—was considered in Ebrahimi and Toloo (2020) and Park (2007). Also, in this context, Ebrahimi et al. (2021) and Toloo et al. (2021) accounted for the dual-role factors, which can be interpreted as input and output at the same time. In turn, Haghighat and Khorram (2005) proposed non-linear models for deriving the maximal and minimal numbers of efficient units when the input and output performances are given as intervals. The Monte Carlo simulation was incorporated into the stochastic DEA to derive the distribution of efficiency scores in the setting where inputs and outputs were expressed as intervals formed by values gathered in different years (Kao and Liu 2009). Dehnokhalaji et al. (2022) proposed a robust optimization framework for performance measurement and cross-efficiency inspired ranking of DMUs. An additive value DEA model was considered in Gouveia et al. (2013) to construct the efficiency intervals and find the maximal percentage tolerance by which one could deteriorate the inputs or outputs of a given DMU so that it remains efficient. Finally, Azizi et al. (2015) proposed a slack-based method to find the optimistic and pessimistic efficiency intervals for DMUs for DEA involving imprecise data. Specifically, two classifications of DMUs into efficient and inefficient units were proposed considering the optimistic and pessimistic settings. In addition, the procedures for obtaining an overall interval score as well as constructing a complete ranking of DMUs were introduced.

The most important contribution of this paper consists of proposing a rich framework for robustness analysis in the context of imprecise inputs and outputs. As opposed to the existing approaches that extend IDEA, our methodology considers uncertainty related to the interval or ordinal data and all feasible weight vectors simultaneously. In particular, we propose tools for analyzing the robustness of three types of outcomes: efficiency scores, efficiency preference relations, and efficiency ranks.

On the one hand, we derive extreme, robust results using dedicated mathematical programming models exploiting all scenarios involving imprecise input/output data and feasible weight vectors. We show how to compute the extreme efficiency scores and ranks and verify the truth or falsity of the necessary and possible efficiency preference relations (Kadziński et al. 2017). The efficiency bounds and ranking intervals reveal the pessimistic and optimistic performance of each unit (Salo and Punkka 2011). In turn, the two relations focus on the pairwise comparisons that need to be validated for all or at least one feasible scenario (Kadziński et al. 2017).

On the other hand, we implement the stochastic analysis to derive the distribution of different measures and results (Lahdelma and Salminen 2006). We employ the Monte Carlo simulations to analyze a sufficiently large and representative set of feasible weight vectors and input/output performances consistent with the imprecise information. For this purpose, we apply a suitably adjusted Hit-And-Run algorithm (see Ciomek and Kadziński 2021; Tervonen et al. 2013). The outcomes are quantified through Efficiency Acceptability Interval Indices, Efficiency Rank Acceptability Indices, and Pairwise Efficiency Outranking Indices (see Lahdelma and Salminen 2006; Kadziński et al. 2017). The stochastic indices capture the shares of feasible

scenarios that guarantee a given score or rank to a particular DMU or confirm that one DMU is at least as good as the other. Also, we estimate the expected efficiency scores and ranks for all DMUs. These measures can be the basis for constructing a complete ranking of DMUs based on the robust outcomes derived from analyzing feasible weights, inputs, and outputs. From the methodological perspective, the proposed methodology can be seen as an extension and adjustment of an integrated framework for robustness analysis proposed in Kadziński et al. (2017) to the case of imprecise (interval or ordinal) evaluations.

We also present open-source software that implements the proposed framework for robustness analysis. The software consists of modules available on the *diviz* platform (Meyer and Bigaret 2012). These modules accept the specification of linear constraints concerning the weights related to inputs and outputs. Moreover, they have been designed to admit their combination into complex algorithmic workflows. The latter can be employed to share the methodological developments and results of case studies among users.

Finally, we illustrate the use of both the framework for robustness analysis and software in real-world studies concerning efficiency analysis of Chinese ports (Jiang et al. 2021) and industrial robots (Saen 2006). The units are described in terms of precise, interval, and ordinal factors. These examples demonstrate the practical usefulness of robust results concerning scores, ranks, and pairwise preference relations. Also, we emphasize the complementarity of exact and stochastic results. Moreover, we demonstrate that both the space of feasible weight vectors as well as imprecise input and output performances influence the robustness of attained efficiency results.

The remainder of the paper is organized in the following way. In Sect. 2, we discuss the proposed methods for robustness analysis within the scope of Imprecise Data Envelopment Analysis. In Sect. 3, we present the algorithmic modules implementing the proposed methodological framework on the *diviz* platform. Section 4 is devoted to an illustrative case study concerning the efficiency analysis of Chinese ports. The results of the study on industrial robots are reported in the e-Appendix (supplementary material available online). Section 5 concludes the paper and outlines avenues for future work.

## 2 Robustness analysis for Imprecise Data Envelopment Analysis

### 2.1 Notation and basic concepts

The following notation is used in the paper:

- $\mathcal{D} = \{DMU_1, \dots, DMU_K\}$ —a set of considered DMUs, where  $K$  is the number of DMUs ( $K = |\mathcal{D}|$ );
- $x_m$ — $m$ -th input,  $m \in \{1, \dots, M\}$ ;
- $y_n$ — $n$ -th output,  $n \in \{1, \dots, N\}$ ;
- $PI$ ,  $II$  and  $OI$ —subsets of precise, interval, and ordinal inputs, respectively;
- $PO$ ,  $IO$  and  $OO$ —subsets of precise, interval, and ordinal outputs, respectively;
- $x_{mo}$ —the value of  $m$ -th input consumed by  $DMU_o \in \mathcal{D}$ ,  $m \in PI \cup OI$ ;

- $y_{no}$ —the value of  $n$ -th output produced by  $DMU_o \in \mathcal{D}$ ,  $n \in PO \cup OO$ ;
- $[x_{mo*}, x_{mo}^*]$ —an interval value of  $m$ -th input of  $DMU_o$ ,  $m \in II$ ;
- $[y_{no*}, y_{no}^*]$ —an interval value of  $n$ -th output of  $DMU_o$ ,  $n \in IO$ ;
- $X_{mo}$ —the value of  $v_m \cdot x_{mo}$  for ordinal inputs,  $m \in OI$ ;
- $Y_{no}$ —the value of  $u_n \cdot y_{no}$  for ordinal outputs,  $n \in OO$ ;
- $v = \{v_1, \dots, v_M\}$ —a vector of input weights;
- $u = \{u_1, \dots, u_N\}$ —a vector of output weights;
- $\eta, \chi$ —values representing the minimal ratios between the successive values of ordinal inputs and ordinal outputs,  $\eta, \chi > 1$  (in this paper, we set  $\eta = \chi = 1.1$ );
- $S_v = \{v = (v_1, \dots, v_M)^T \neq 0 \mid v \geq 0, A_v v \leq 0\}$  and  $S_u = \{u = (u_1, \dots, u_N)^T \neq 0 \mid u \geq 0, A_u u \leq 0\}$ —spaces of feasible input and output weights, respectively;  $A_v$  and  $A_u$  are matrices of coefficients involved in the linear constraints on weights derived from the user’s preferences.

To illustrate the notation, let us refer to an example presented in Table 1, which is derived from Despotis and Smirlis (2002). The set of DMUs is composed of five units,  $\mathcal{D} = \{D_1, D_2, D_3, D_4, D_5\}$ . They consume two inputs—one precise ( $PI = \{i_1\}$ ) and the other interval ( $II = \{i_2\}$ ), and produce two outputs—one precise ( $PO = \{o_1\}$ ) and the other ordinal ( $OO = \{o_2\}$ ). The weights associated with the inputs are denoted by  $v_1$  and  $v_2$ , and the respective weights for the outputs are  $u_1$  and  $u_2$ . When it comes to unit  $D_1$ , its precise input is  $x_{11} = 100$  and the interval input is  $[x_{21*}, x_{21}^*] = [0.6, 0.7]$ . The respective outputs are  $y_{11} = 2000$  and  $y_{21} = 4$ . The latter will be represented in the following mathematical models as  $Y_{21} = y_{21} \cdot u_2$ , and the following order  $Y_{24} < Y_{22} < Y_{25} < Y_{21} < Y_{23}$  will be maintained.

In what follows, we discuss the methods for robustness analysis in the context of Imprecise DEA. They can be divided into two subgroups. One of them is devoted to the exact analysis using linear programming techniques. In contrast, the other aims to estimate some stochastic acceptability indices through the Monte Carlo simulations. The analysis is conducted given all feasible efficiency scenarios, where each scenario corresponds to a specific, admissible realization of both weights and performances on inputs and outputs.

**Table 1** Example set of Decision Making Units involving imprecise data

$DMU_o$	$i_1$ (precise)	$i_2$ (interval)	$o_1$ (precise)	$o_2$ (ordinal)
$D_1$	100	[0.6, 0.7]	2000	4
$D_2$	150	[0.8, 0.9]	1000	2
$D_3$	150	[1.0, 1.0]	1200	5
$D_4$	200	[0.7, 0.8]	900	1
$D_5$	200	[1.0, 1.0]	600	3

### 2.2 Exact robustness analysis with linear programming

In this section, we discuss how to derive exact robust outcomes using mathematical programming. These results capture the extreme cases observed for all feasible efficiency scenarios  $(\mathbf{u}, \mathbf{v}, \mathbf{x}, \mathbf{y})$  defined by the sets of admissible weights as well as values of inputs and outputs. They concern the following three perspectives: scores, ranks, and pairwise preference relations. We refer to the concepts of extreme scores and ranks and the necessary and possible preference relations that have been introduced in the literature. However, the models for their computation that are presented in this section are original and specifically adjusted to the context of IDEA.

To conduct a robustness analysis given interval inputs and outputs, we need to consider the most and the least advantageous (i.e., optimistic and pessimistic) scenarios for each DMU. On the one hand, the optimistic scenario for  $DMU_o$  is realized by assuming that its inputs are the least possible and its outputs are the greatest admissible by the specified intervals. In contrast, for the remaining units, both the inputs and outputs are the least advantageous, i.e.:

$$x_{mk} = \begin{cases} x_{mk**}, & \text{if } m = o, \\ x_{mk}^*, & \text{otherwise,} \end{cases} \tag{1}$$

$$y_{nk} = \begin{cases} y_{nk}^*, & \text{if } n = o, \\ y_{nk**}, & \text{otherwise.} \end{cases} \tag{2}$$

On the other hand, the pessimistic scenario for  $DMU_o$  is realized by assuming that its imprecise inputs and outputs are replaced with the least favorable values. For the remaining DMUs, the minimal inputs and the maximal outputs are considered, i.e.:

$$x_{mk} = \begin{cases} x_{mk}^*, & \text{if } m = o, \\ x_{mk**}, & \text{otherwise,} \end{cases} \tag{3}$$

$$y_{nk} = \begin{cases} y_{nk**}, & \text{if } n = o, \\ y_{nk}^*, & \text{otherwise.} \end{cases} \tag{4}$$

When the dataset involves the ordinal factors, the products  $v_m \cdot x_{mk}$  or  $u_n \cdot y_{nk}$  are replaced by one variable, respectively,  $X_{mk}$  or  $Y_{nk}$ . Additionally, the constraints respecting the character of ordinal evaluations need to be included in the model. In particular, the constraints imposing a strong ordinal relation should not take an additive form, e.g.,  $X_2 \geq X_1 + \epsilon$ , where  $\epsilon$  is a small positive constant. In turn, as the original ordinal evaluations  $x_{mk}$  and  $y_{nk}$  are transformed into variables  $X_{mk}$  or  $Y_{nk}$  involving multiplication by a common weight ( $v_m$  and  $u_n$ ), the ratios of subsequent  $X_{mk}$  or  $Y_{nk}$  values needs to be greater than one, i.e.:

$$\begin{cases} \chi \cdot Y_{ni} \leq Y_{nj}, & (i,j) \in \{(i,j) : y_{ni} \leq y_{nj}\}, n \in OO, \\ \eta \cdot X_{mi} \leq X_{mj}, & (i,j) \in \{(i,j) : x_{mi} \leq x_{mj}\}, m \in OI, \end{cases} \tag{5}$$

where  $\chi, \eta > 1$  (Zhu 2003).

Following (Zhu 2003), we consider the efficiency of  $DMU_o$  defined as a ratio of a single virtual output to a single virtual input:

$$E_o = \frac{\sum_{n \in PO} u_n y_{no} + \sum_{n \in IO} u_n y_{no} + \sum_{n \in OO} Y_{no}}{\sum_{m \in PI} v_m x_{mo} + \sum_{m \in II} v_m x_{mo} + \sum_{m \in OI} X_{mo}} \tag{6}$$

where  $y_{no} \in [y_{no*}, y_{no}^*]$  for  $n \in IO$  and  $x_{mo} \in [x_{mo*}, x_{mo}^*]$  for  $m \in II$ . The virtual output and input aggregate multiple outputs or inputs while ensuring that each relevant factor contributes to an overall measure of efficiency. Please note that the contributions from these factors are dimensionless. This is due to multiplying the precise and imprecise performances by the weights and using dedicated components for the ordinal factors. Still, their major role is to maintain the desired relationships between the efficiencies of various units implied by their input and output values. In fact, the above expression ensures that  $E_o$  does not deteriorate if one (i) increases the output values or decreases the input values in  $DMU_o$  or (ii) decreases the output values or increases the input values in other DMUs. At the same time, this representation eliminates the scale transformations (Zhu 2003), reducing the computational burden in applications.

### 2.2.1 Extreme efficiency scores

When it comes to the efficiencies, for each  $DMU_o$ , we determine the maximal  $E_o^*$  and minimal  $E_{o*}$  scores that it can attain for at least one feasible scenario (see Despotis and Smirlis 2002; Kadziński et al. 2017; Kao 2006). To find the greatest (optimistic) efficiency score for  $DMU_o$ , the following Linear Programming (LP) model needs to be solved:

$$\begin{aligned} & \text{Maximize: } E_o^* = \sum_{n \in PO} u_n y_{no} + \sum_{n \in IO} u_n y_{no}^* + \sum_{n \in OO} Y_{no} \\ \text{s.t. } & [E^* - C1] \quad \sum_{m \in PI} v_m x_{mo} + \sum_{m \in II} v_m x_{mo}^* + \sum_{m \in OI} X_{mo} = 1, \\ & [E^* - C2] \quad \sum_{n \in PO} u_n y_{no} + \sum_{n \in IO} u_n y_{no}^* + \sum_{n \in OO} Y_{no} \leq 1, \\ & [E^* - C3] \quad \sum_{n \in PO} u_n y_{nk} + \sum_{n \in IO} u_n y_{nk}^* + \sum_{n \in OO} Y_{nk} \leq \sum_{m \in PI} v_m x_{mk} + \sum_{m \in II} v_m x_{mk}^* + \sum_{m \in OI} X_{mk}, \quad k = 1, \dots, K; k \neq o, \\ & [E^* - C4] \quad \chi Y_{ni} \leq Y_{nj}, \quad (i, j) \in \{(i, j) : y_{ni} \leq y_{nj}\}, n \in OO, \\ & [E^* - C5] \quad \eta X_{mi} \leq X_{mj}, \quad (i, j) \in \{(i, j) : x_{mi} \leq x_{mj}\}, m \in OI, \\ & [E^* - C6] \quad (v, u) \in (S_{\rightarrow}, S_{\leftarrow}). \end{aligned} \tag{7}$$

Model (7) is equivalent to the classical CCR model for DEA with imprecise data. It finds the most favorable weight vector for  $DMU_o$  in its best input/output scenario and the worst possible scenarios for the remaining DMUs. The space of variables is composed of the following weights:  $v_m$  for  $n \in PI \cup II$ ,  $X_{mk}$  for  $m \in OI$  and  $DMU_k \in \mathcal{D}$ ,  $u_n$  for  $n \in PO \cup IO$  and  $Y_{nk}$  for  $n \in OO$  and  $DMU_k \in \mathcal{D}$ . It is constrained so that the virtual input of  $DMU_o$  is equal to one ( $[E^* - C1]$ ), the efficiency scores for all DMUs are not greater than one ( $[E^* - C2]$  and  $[E^* - C3]$ ), the monotonicity relations derived from the analysis of ordinal inputs and outputs are preserved ( $[E^* - C4]$  and  $[E^* - C5]$ ), and the constraints on the admissible values of input and output weights are satisfied ( $[E^* - C6]$ ). The last three constraints are present in all



following LP models. The optimal value of  $E_o^*$  is between zero and one. The DMUs with optimal  $E_o^* = 1$  are considered as efficient.

The minimal (pessimistic) efficiency score for  $DMU_o$  using the CCR model with imprecise information can be derived by solving the following Mixed-Integer Linear Programming (MILP) model:

$$\begin{aligned}
 &\text{Minimize: } E_{o*} = \sum_{n \in PO} u_n y_{no} + \sum_{n \in IO} u_n y_{no*} + \sum_{n \in OO} Y_{no} \\
 \text{s.t. } [E_* - C1] &\quad \sum_{m \in PI} v_m x_{mo} + \sum_{m \in II} v_m x_{mo}^* + \sum_{m \in OI} X_{mo} = 1, \\
 [E_* - C2] &\quad \sum_{n \in PO} u_n y_{no} + \sum_{n \in IO} u_n y_{no*} + \sum_{n \in OO} Y_{no} \geq 1 - C(1 - b_o), \\
 [E_* - C3] &\quad \sum_{n \in PO} u_n y_{nk} + \sum_{n \in IO} u_n y_{nk}^* + \sum_{n \in OO} Y_{nk} \geq \sum_{m \in PI} v_m x_{mk} + \sum_{m \in II} v_m x_{mk}^* + \sum_{m \in OI} X_{mk} - C(1 - b_k), \quad k = 1, \dots, K; k \neq o, \\
 [E_* - C4] &\quad \sum_{k=1}^K b_k \geq 1, \\
 [E_* - C5] &\quad b_k \in \{0, 1\}, \quad k = 1, \dots, K, \\
 [E_* - C6] &\quad \chi y_{ni} \leq y_{nj}, \quad (i, j) \in \{(i, j) : y_{ni} \leq y_{nj}\}, n \in OO, \\
 [E_* - C7] &\quad \eta x_{mi} \leq x_{mj}, \quad (i, j) \in \{(i, j) : x_{mi} \leq x_{mj}\}, m \in OI, \\
 [E_* - C8] &\quad (v, u) \in (S_v, S_u).
 \end{aligned} \tag{8}$$

The above model allows for finding the least favorable weight vector for  $DMU_o$  in terms of its efficiency while considering the worst possible scenario for  $DMU_o$  and the best admissible scenarios for the remaining units  $DMU_k, k = 1, \dots, K$  and  $k \neq o$ . Under these conditions, we constrain the space of feasible solutions by imposing—without loss of generality—that the virtual input of  $DMU_o$  equals one ( $[E_* - C1]$ ), assuming that at least one unit is efficient (its efficiency score must be greater than or equal to one;  $[E_* - C2]$ – $[E_* - C5]$ ), preserving the ordinal factors’ monotonicity ( $[E_* - C6]$ – $[E_* - C7]$ ), and satisfying the pre-defined constraints on the admissible values of input and output weights ( $[E_* - C8]$ ). Apart from the weights already considered in model (7), we include the binary variables  $b_k \in \{0, 1\}, k = 1, \dots, K$ . The optimal value of  $E_{o*}$  is between zero and one. Overall,  $[E_{o*}, E_o^*]$  can be deemed as an efficiency interval (Salo and Punkka 2011).

Note that  $C$  is a large positive constant. Irrespective of which  $DMU_o$  is considered, it is sufficient that  $C > \max_{DMU_i, DMU_k \in \mathcal{D}} \{\max\{\max_{m \in PI} \{x_{mk}/x_{mi}\}, \max_{m \in II} \{x_{mk*}/x_{mi}^*\}\}, \text{ if } OI \neq \emptyset : \eta^K\}$ . It is so because to minimize  $E_{o*}$ , the solver also minimizes  $E_k$  for  $k = 1, \dots, K$ . Since constraint  $[E_* - C1]$  imposes  $\sum_{m \in PI} v_m x_{mo} + \sum_{m \in II} v_m x_{mo}^* + \sum_{m \in OI} X_{mo} = 1$ , then for  $k = 1, \dots, K$ ,  $C$  is greater than  $\sum_{m \in PI} v_m x_{mk} + \sum_{m \in II} v_m x_{mk}^* + \sum_{m \in OI} X_{mk}$ . Consequently, when binary variable  $b_k$  equals 0 for  $k = 1, \dots, K$ , constraint  $[E_* - C2]$  (when  $k = o$ ) or constraint  $[E_* - C3]$  (when  $k = 1, \dots, K, k \neq o$ ) is satisfied for all values of the variables. However, constraint  $[E_* - C4]$  imposes that at least one  $b_k$  for  $k = 1, \dots, K$  is equal to one. Then, the respective efficiency  $E_k$  is greater or equal to one since the virtual output of  $DMU_k$  is greater or equal to its virtual input.

*Illustrative example* In Table 2, we present the extreme efficiencies derived for five units contained in the illustrative example introduced in Sect. 2.1. They reveal that two units ( $D_1$  and  $D_3$ ) are efficient, attaining the maximal efficiency score equal to one. The

**Table 2** Exact robust results derived with mathematical programming for the illustrative example

$DMU_o$	Extreme scores		Extreme ranks		Robust relations				
	$E_o^*$	$E_{o*}$	$R_o^*$	$R_{o*}$	$D_1$	$D_2$	$D_3$	$D_4$	$D_5$
$D_1$	1.000	0.013	1	2	N	N	P	N	N
$D_2$	0.723	0.011	2	5		N	P	P	P
$D_3$	1.000	0.367	1	4	P	P	N	P	N
$D_4$	0.751	0.010	2	5		P	P	N	P
$D_5$	0.637	0.012	3	5		P		P	N

efficiency intervals are relatively wide and span over the range of over 0.6 for all units. For example, the minimal efficiencies of  $D_1$  and  $D_3$  are, respectively, 0.013 and 0.367.

### 2.2.2 Extreme efficiency ranks

As far as efficiency ranks are concerned, we determine the best  $R_o^*$  and the worst  $R_{o*}$  ranks that are attained by  $DMU_o$  for at least one feasible scenario (see Kadziński et al. 2012, 2017; Salo and Punkka 2011). Given a fixed input/output weight vector and precise feasible performances for  $DMU_o$ , it attains  $k$ -th rank if exactly  $k - 1$  other units attain higher efficiency scores than  $DMU_o$ . To find the minimal (i.e., the best) efficiency rank for  $DMU_o$ , the following MILP model needs to be solved:

$$\begin{aligned}
 &\text{Minimize:} && R_o^* = 1 + \sum_{k=1, k \neq o}^K b_k \\
 \text{s.t. } &[R^* - C1] && \sum_{n \in PO} u_n y_{no} + \sum_{n \in IO} u_n y_{no}^* + \sum_{n \in OO} Y_{no} = 1, \\
 &[R^* - C2] && \sum_{m \in PI} v_m x_{mo} + \sum_{m \in MO} v_m x_{mo} + \sum_{m \in OI} X_{mo} = 1, \\
 &[R^* - C3] && \sum_{n \in PO} u_n y_{nk} + \sum_{n \in IO} u_n y_{nk}^* + \sum_{n \in OO} Y_{nk} \leq \sum_{m \in PI} v_m x_{mk} + \sum_{m \in IO} v_m x_{mk}^* + \sum_{m \in OI} X_{mk} + Cb_k, && k = 1, \dots, K; k \neq o, \\
 &[R^* - C4] && b_k \in \{0, 1\}, && k = 1, \dots, K; k \neq o, \\
 &[R^* - C5] && \chi Y_{ni} \leq Y_{nj}, && (i, j) \in \{(i, j) : y_{ni} \leq y_{nj}\}, n \in OO, \\
 &[R^* - C6] && \eta X_{mi} \leq X_{mj}, && (i, j) \in \{(i, j) : x_{mi} \leq x_{mj}\}, m \in OI, \\
 &[R^* - C7] && (v, u) \in (S_v, S_u).
 \end{aligned} \tag{9}$$

The above model sets the efficiency score of  $DMU_o$  in its most optimistic realization equal to one ( $[R^* - C1]$ – $[R^* - C2]$ ). For the remaining units, we assume their pessimistic realizations ( $[R^* - C3]$ ) and minimize the number of DMUs with efficiency scores greater than for  $DMU_o$ . This is attained by introducing the binary variables  $b_k$  for each  $DMU_k$ ,  $k = 1, \dots, K$ , and  $k \neq o$  ( $[R^* - C5]$ ). When the efficiency score of  $DMU_k$  cannot be lower than or equal to one, then  $b_k$  is set to one, and the respective constraint  $[R^* - C4]$  is always satisfied for all possible variable values. This is implied by the use of a large positive constant  $C$ . Analogously to the reasoning for model (8), irrespective of which  $DMU_o$  is considered, it is sufficient that  $C > \max_{DMU_l, DMU_k \in D} \{ \max\{ \max_{n \in PO} \{y_{nk}/y_{nl}\}, \max_{n \in IO} \{y_{nk}^*/y_{nl}^*\} \}, \text{ if } OO \neq \emptyset : \chi^K \}$ . In turn, if the efficiency score of  $DMU_o$  is greater or equal to the efficiency of  $DMU_k$ ,  $b_k$  is set to zero. Thus, by minimizing the

sum of  $b_k, k = 1, \dots, K$ , and  $k \neq o$ , we can obtain the best possible rank of  $DMU_o$ . The optimal value of  $R_o^*$  is between one and  $K$ .

The worst (i.e., the maximal) possible rank of  $DMU_o$  can be computed with the following MILP model:

$$\begin{aligned}
 &\text{Maximize:} && R_{o^*} = 1 + \sum_{k=1, k \neq o}^K b_k \\
 \text{s.t. } &[R_* - C1] && \sum_{n \in PO} u_n y_{no} + \sum_{n \in IO} u_n y_{no^*} + \sum_{n \in OO} Y_{no} = 1, \\
 &[R_* - C2] && \sum_{m \in PI} v_m x_{mo} + \sum_{m \in II} v_m x_{mo}^* + \sum_{m \in OI} X_{mo} = 1, \\
 &[R_* - C3] && \sum_{m \in PI} v_m x_{mk} + \sum_{m \in II} v_m x_{mk}^* + \sum_{m \in OI} X_{mk} \leq \sum_{n \in PO} u_n y_{nk} + \sum_{n \in IO} u_n y_{nk}^* + \sum_{n \in OO} Y_{nk} + C(1 - b_k), && k = 1, \dots, K; k \neq o, \\
 &[R_* - C4] && b_k \in \{0, 1\}, && k = 1, \dots, K; k \neq o, \\
 &[R_* - C5] && \chi_{ni}^Y \leq y_{nj}, && (i, j) \in \{(i, j) : y_{ni} \leq y_{nj}\}, n \in OO, \\
 &[R_* - C6] && \eta_{mi}^X \leq x_{mj}, && (i, j) \in \{(i, j) : x_{mi} \leq x_{mj}\}, m \in OI, \\
 &[R_* - C7] && (v, u) \in (S_v, S_u).
 \end{aligned} \tag{10}$$

The above model maximizes the number of DMUs with efficiency scores greater or equal to  $DMU_o$ . Again, we assume that the efficiency of  $DMU_o$  is equal to one ( $[R_* - C1] - [R_* - C2]$ ). However, at this time, we consider the pessimistic realization of  $DMU_o$ . Then, we introduce the constraints imposing that the efficiencies of the remaining DMUs in their optimistic realizations are not lower than one ( $[R_* - C3]$ ). The component  $C \cdot (1 - b_k)$  included in the respective constraint implies that the latter can be violated. If binary variable  $b_k$  ( $[R_* - C4]$ ) is equal to one, constraint  $[R_* - C3]$  holds, whereas for  $b_k = 0$ —it is satisfied for any variables' values. Note that  $C$  should be set similarly as for model (8). When maximizing the sum of  $b_k, k = 1, \dots, K$ , and  $k \neq o$ , we minimize the number of DMUs for which constraint  $[R_* - C3]$  is violated. Thus, the sum of  $b_k$  increased by one corresponds to the worst possible rank of  $DMU_o$ . The optimal value of  $R_{o^*}$  is between one and  $K$ .

*Illustrative example* The extreme ranks for the illustrative example introduced in Sect. 2.1 are presented in Table 2. The efficient units attain the first rank in the best case. Although the minimal efficiency of  $D_1$  is worse than for  $D_3$ , in the worst case its rank can drop only to the second position ( $R_{1^*} = 2$ ), whereas  $D_3$  can be ranked even fourth ( $R_{3^*} = 4$ ) in the most pessimistic scenario. The inefficient units can be ranked second ( $D_2$  and  $D_4$ ) or third ( $D_5$ ) in the best case, while all are ranked at the bottom in the least advantageous scenario.

### 2.2.3 Necessary and possible efficiency preference relations

When it comes to the stability of comparisons observed for pairs of DMUs given all feasible scenarios, we consider the necessary ( $\succsim_E^N$ ) and possible ( $\succsim_E^P$ ) efficiency preference relations (see Greco et al. 2008; Kadziński et al. 2017). They are defined in the following way:

- $DMU_o$  is necessarily preferred to  $DMU_l$  ( $DMU_o \succsim_E^N DMU_l$ ) if  $DMU_o$  attains at least as good efficiency as  $DMU_l$  for all feasible scenarios defined by the sets of admissible weights, as well as values of inputs and outputs, or, equivalently, if

for all feasible weight vectors the efficiency of  $DMU_o$  in its pessimistic realization is not worse than the efficiency of  $DMU_l$  in its optimistic realization;

- $DMU_o$  is possibly preferred to  $DMU_l$  ( $DMU_o \succsim_E^P DMU_l$ ) if  $DMU_o$  attains at least as good efficiency as  $DMU_l$  for at least one feasible scenario defined by the sets of admissible weights, as well as values of inputs and outputs, or, equivalently, if for at least one feasible weight vector the efficiency of  $DMU_o$  in its optimistic realization is not worse than the efficiency of  $DMU_l$  in its pessimistic relations.

To verify the truth of the necessary efficiency preference relation  $DMU_o \succsim_E^N DMU_l$  for pair  $(DMU_o, DMU_l)$ , we need to solve the following LP model:

$$\begin{aligned}
 \text{Minimize: } E_{o*} &= \sum_{n \in PO} u_n y_{no} + \sum_{n \in IO} u_n y_{nos} + \sum_{n \in OO} Y_{no} \\
 \text{s.t. } \sum_{m \in PI} v_m x_{mo} + \sum_{m \in II} v_m x_{mo}^* + \sum_{m \in OI} X_{mo} &= 1, \\
 \sum_{n \in PO} u_n y_{nl} + \sum_{n \in IO} u_n y_{nl}^* + \sum_{n \in OO} Y_{nl} &= \sum_{m \in PI} v_m x_{ml} + \sum_{m \in II} v_m x_{ml}^* + \sum_{m \in OI} X_{ml}, \\
 \chi Y_{ni} &\leq Y_{nj}, & (i, j) \in \{(i, j) : y_{ni} \leq y_{nj}\}, n \in OO, \\
 \eta X_{mi} &\leq X_{mj}, & (i, j) \in \{(i, j) : x_{mi} \leq x_{mj}\}, m \in OI, \\
 (v, u) &\in (S_v, S_u).
 \end{aligned} \tag{11}$$

The above model finds the minimal efficiency score of  $DMU_o$  in its pessimistic realization while assuming that the efficiency of  $DMU_l$  in its optimistic realization is equal to one. If the obtained optimal value of  $E_{o*}$  is greater than or equal to one, then for all weight vectors  $(\mathbf{u}, \mathbf{v})$ , the efficiency of  $DMU_o$  is not worse than efficiency of  $DMU_l$ , i.e.,  $DMU_o \succsim_E^N DMU_l$ . Otherwise,  $\text{not}(DMU_o \succsim_E^N DMU_l)$ .

The following LP model allows verifying the truth of the possible efficiency preference relation  $DMU_o \succsim_E^P DMU_l$  for pair  $(DMU_o, DMU_l)$ :

$$\begin{aligned}
 \text{Maximize: } E_o^* &= \sum_{n \in PO} u_n y_{no} + \sum_{n \in IO} u_n y_{no}^* + \sum_{n \in OO} Y_{no} \\
 \text{s.t. } \sum_{m \in PI} v_m x_{mo} + \sum_{m \in II} v_m x_{mo}^* + \sum_{m \in OI} X_{mo} &= 1, \\
 \sum_{n \in PO} u_n y_{nl} + \sum_{n \in IO} u_n y_{nl}^* + \sum_{n \in OO} Y_{nl} &= \sum_{m \in PI} v_m x_{ml} + \sum_{m \in II} v_m x_{ml}^* + \sum_{m \in OI} X_{ml}, \\
 \chi Y_{ni} &\leq Y_{nj}, & (i, j) \in \{(i, j) : y_{ni} \leq y_{nj}\}, n \in OO, \\
 \eta X_{mi} &\leq X_{mj}, & (i, j) \in \{(i, j) : x_{mi} \leq x_{mj}\}, m \in OI, \\
 (v, u) &\in (S_v, S_u).
 \end{aligned} \tag{12}$$

The above model computes the maximal efficiency of  $DMU_o$  in its optimistic realization while assuming that the efficiency of  $DMU_l$  in its pessimistic realization is equal to one. If the optimal value of  $E_o^*$  is greater than or equal to one, then there exists at least one weight vector  $(\mathbf{u}, \mathbf{v})$  for which the efficiency of  $DMU_o$  is not worse than efficiency of  $DMU_l$ , i.e.,  $DMU_o \succsim_E^P DMU_l$ . Otherwise,  $\text{not}(DMU_o \succsim_E^P DMU_l)$ .

*Illustrative example* The necessary and possible relations for the illustrative example are presented in Table 2. Note that the necessary relation is transitive and implies the truth of the possible relation. Let us observe that unit  $D_1$  is necessarily preferred to the three inefficient units ( $D_2, D_4$ , and  $D_5$ ), whereas  $D_3$  is robustly at least as good only when compared to  $D_5$ . The efficient units are incomparable in terms of the necessary relation while being possibly preferred over each other. The inefficient units are not preferred over any other unit for all feasible scenarios. However, they

are possibly preferred over each other (e.g.,  $D_2 \succ_E^P D_4$  and  $D_4 \succ_E^P D_2$ ). Moreover,  $D_2$  and  $D_4$  are at least as good as  $D_3$  for at least one feasible scenario, whereas none of the inefficient units attains the score of  $D_1$  for any feasible setting.

### 2.3 Stochastic analysis with the Monte Carlo simulation

In this section, we discuss how to derive stochastic outcomes using the Monte Carlo simulation. These results capture the share or distribution of feasible efficiency scenarios  $(\mathbf{u}, \mathbf{v}, \mathbf{x}, \mathbf{y})$  that confirm a given outcome referring to the attained scores or ranks, or the truth of pairwise preference relation.

To conduct such a stochastic analysis, we need to sample a representative subset of all feasible efficiency scenarios. This requires the assumption about the probability distributions of the joint density function of the feasible input/output weight vectors and the precise performances within the specified interval values on various inputs and outputs (Lahdelma and Salminen 2006). In general, the proposed approach can be used with any arbitrarily selected distribution. However, when the expert does not impose the parameter distribution, we assume the uniform distribution of weights and performances (see Kadziński et al. 2017; Lahdelma and Salminen 2001).

To simulate the feasible efficiency scenarios, we need to derive the weights and performances from the feasible space. For sampling weights from the uniform distribution, we use the Hit-And-Run (HAR) algorithm (Tervonen et al. 2013). Since it requires the space of sampling to be bounded, we perform normalization of possible input/output weights:

$$\sum_{n=1}^N u_n = \sum_{m=1}^M v_m = 1. \quad (13)$$

When it comes to sampling the performances, a dedicated treatment has been designed to deal with the interval and ordinal factors. For the interval inputs and outputs, for each  $DMU_o$ , we randomly select the exact values from the intervals  $[x_{mo}^*, x_{mo}^*]$  or  $[y_{no}^*, y_{no}^*]$  using HAR. Regarding dealing with the ordinal factors, we adopt the SMAA-O approach (Lahdelma et al. 2003). Specifically, we assume that a function simulating some ordinal inputs or outputs is increasing. We assume that the exact values corresponding to the ordinal performances are drawn from the  $[0, 1]$  interval without losing generality. Hence we randomly choose a set of  $K$  numbers from this range. The obtained values are sorted and considered as a single sample of precise performances of DMUs consistent with the order imposed by the original ordinal performances (e.g., a unit with the worst ordinal output or the best ordinal input is assigned the least precise value).

The samples concerning the weights and the input and output values are put together to simulate the feasible efficiency scenarios. For each of them, we compute the efficiencies for all DMUs. The results obtained for all sampled scenarios are summarized in stochastic acceptability indices concerning scores, ranks, and pairwise relations. Since their values are approximated using the Monte Carlo simulation

rather than computed exactly through analytical methods, we consider the estimations of the true indices in practice. However, with a sufficiently large number of samples, such values can be estimated up to a pre-defined accuracy (Tervonen and Lahdelma 2007).

### 2.3.1 Distribution of efficiency scores

The Efficiency Acceptability Interval Index  $EAI(DMU_o, b_i)$  is defined as the share of feasible scenarios for which the efficiency score of  $DMU_o$  is contained in the sub-interval  $b_i \subset [0, 1]$ , where  $i = 1, \dots, B$ , and  $B$  is the number of efficiency sub-intervals considered in the analysis. By default, the sub-intervals are assumed to be disjoint and to span over the same widths. Note that for each  $DMU_o \in \mathcal{D}$ ,  $\sum_{i=1}^B EAI(DMU_o, b_i) = 1$ . Moreover, by analyzing the scores obtained by  $DMU_o$ , we may compute an expected efficiency (denoted by  $EE_o$ ) as an average of efficiencies derived for all sampled scenarios. Such efficiency may be used to impose a complete ranking on the set of DMUs (Labijak-Kowalska and Kadziński 2021).

*Illustrative example* In Table 3, we present the estimates of  $EAI$ s computed based on 10,000 samples derived with the Monte Carlo simulation for the illustrative example introduced in Sect. 2.1. We have selected five buckets ( $B = 5$ ), and hence the considered sub-intervals are  $[0, 0.2]$ ,  $(0.2, 0.4]$ ,  $(0.4, 0.6]$ ,  $(0.6, 0.8]$ ,  $(0.8, 1.0]$ . The most probable efficiency ranges for  $D_1$  and  $D_2$  are, respectively,  $(0.8, 1.0]$  ( $EAI(D_1, (0.8, 1.0]) = 0.958$ ) and  $(0.2, 0.4]$  ( $EAI(D_2, (0.2, 0.4]) = 0.716$ ). On the other extreme, the estimated probability of  $D_1$  attaining an efficiency score lower than 0.2 or  $D_2$  attaining a score greater than 0.6 is zero. However, the analysis of extreme efficiency scores presented in Sect. 2.2.1 reveals that it is possible. Nonetheless, when combining this information with the analysis of  $EAI$ s, we know that such a scenario is improbable. As far as expected efficiencies are concerned, they impose the following ranking on the set of DMUs:  $D_1 > D_3 > D_2 > D_5 > D_4$ , hence allowing discrimination between both efficient and inefficient units.

In the e-Appendix, we present a detailed step-by-step description of calculating the  $EAI$ s and other stochastic measures for the considered example. To make the description self-contained and its size reasonable, we use only ten samples as opposed 10,000 samples considered in the main paper.

### 2.3.2 Efficiency rank acceptability indices

Efficiency Rank Acceptability Index  $ERAI(DMU_o, r)$  for  $DMU_o \in \mathcal{D}$  and a specific rank  $r \in \{1, 2, \dots, K\}$  is defined as the share of feasible scenarios for which  $DMU_o$  is placed at the  $r$ -th position in the ranking imposed by the efficiency scores of all DMUs in  $\mathcal{D}$ . Note that for each  $DMU_o \in \mathcal{D}$ ,  $\sum_{r=1}^K ERAI(DMU_o, r) = 1$ . These stochastic indices can be used to approximate an expected efficiency rank (denoted by  $ER_o$ ) for  $DMU_o$  in the following way:  $ER_o = \sum_{r=1}^K r \cdot ERAI(DMU_o, r)$  (Ang et al. 2021). Similar to the expected efficiencies, the expected efficiency ranks can be used to order the units from the best to the worst.

**Table 3** Stochastic results derived with the Monte Carlo simulation for the illustrative example

$DMU_o$	$EAIIs$					$EERs$					$PEOIs$				
	$[0,0,0.2]$	$(0.2,0.4]$	$(0.4,0.6]$	$(0.6,0.8]$	$(0.8,1.0]$	$EE_o$					$ER_o$				
						1	2	3	4	5	$D_1$	$D_2$	$D_3$	$D_4$	$D_5$
$D_1$	0.000	0.002	0.008	0.032	0.958	0.980	0.83	0.17	0.00	0.00	1.17	1.00	1.00	0.83	1.00
$D_2$	0.182	0.716	0.102	0.000	0.000	0.289	0.00	0.00	0.66	0.34	3.34	0.00	1.00	1.00	0.66
$D_3$	0.000	0.081	0.101	0.440	0.378	0.749	0.17	0.83	0.00	0.00	1.83	0.17	1.00	1.00	1.00
$D_4$	0.790	0.210	0.000	0.000	0.000	0.134	0.00	0.00	0.00	0.16	4.84	0.00	0.00	1.00	0.16
$D_5$	0.220	0.694	0.086	0.000	0.000	0.272	0.00	0.00	0.34	0.50	3.82	0.00	0.34	0.00	0.84

*Illustrative example* The rank acceptabilities for the illustrative examples are presented in Table 3. Based on the derived samples' analysis, only  $D_1$  and  $D_3$  can be ranked in the first two positions. However, the probability of  $D_1$  being ranked at the top is higher than for  $D_3$  ( $ERAI(D_1, 1) = 0.83 > ERAI(D_3, 1) = 0.17$ ). Even though the minimal rank for  $D_3$  indicated that it could be ranked fourth in the most pessimistic case, the analysis of  $ERAI$ s suggests that the scenarios for which it drops out of the top two are very unlikely. The distribution of ranks for  $D_5$  confirms that it is ranked fourth for half of the scenarios. Further, the probabilities of attaining the third and fifth positions by  $D_5$  are equal to, respectively, 34% and 16%. The expected efficiency ranks impose the following order on the set of DMUs:  $D_1 > D_3 > D_2 > D_5 > D_4$ . Even though it exploits the ordinal results (i.e., ranks) rather than cardinal ones (i.e., efficiencies), this ranking is the same as when considering the expected efficiencies.

### 2.3.3 Pairwise efficiency outranking indices

The Pairwise Efficiency Outranking Index  $PEOI(DMU_o, DMU_l)$  is defined as the share of feasible scenarios for which  $DMU_o$  is at least as efficient as  $DMU_l$ . Note that for  $(DMU_o, DMU_l) \in \mathcal{D} \times \mathcal{D}$ ,  $0 \leq PEOI(DMU_o, DMU_l) \leq 1$  and  $0 \leq PEOI(DMU_o, DMU_l) + PEOI(DMU_l, DMU_o) \leq 2$ .

*Illustrative example* The  $PEOI$ s derived for the illustrative example are presented in Table 3. Note that when for the pairs for which the necessary relation holds (e.g.,  $(D_1, D_2)$  and  $(D_3, D_5)$ ),  $PEOI$  is equal to one, whereas for the pairs for which the possible relation is false (e.g.,  $(D_2, D_1)$  and  $(D_5, D_1)$ ),  $PEOI$  is zero. The analysis of  $PEOI$ s is the most informative for pairs that are not related by the necessary relation. For example, the share of scenarios for which  $D_1$  attains higher efficiency than  $D_3$  is five times greater than the share for which the inverse relation holds. In the same spirit,  $D_2$  is more efficient than  $D_5$  for twice as many scenarios as  $D_5$  being more favorable than  $D_2$ . Having compared  $D_3$  with  $D_2$  or  $D_4$  using the exact robust analysis methods, we know that these pairs are not related by  $\succ_{\mathcal{E}}^N$ . However,  $PEOI$ s indicate that the scenarios for which  $D_2$  and  $D_4$  are strictly better than  $D_3$  are extremely limited ( $PEOI(D_2, D_3) = 0$  and  $PEOI(D_4, D_3) = 0$ ).

To demonstrate the impact that joint consideration of variable weights and imprecise inputs and outputs has on the obtained robust results, in the e-Appendix, we reconsider the illustrative example. Specifically, we analyze five scenarios while replacing performances on a single or two imprecise factors with the respective precise data. For each scenario, we discuss the six types of results. In this way, we demonstrate that imprecision of inputs and outputs contributes to the uncertainty of efficiency outcomes in the same way as the multiplicity of weights associated with these factors.



### 3 Implementation on the *diviz* platform

Diviz is an open-source platform that allows designing and executing algorithmic workflows implementing operational research methods (Meyer and Bigaret 2012). The software consists of two major components: i) a Java client, which allows users to design workflows using existing computational and graphical modules, and ii) servers, where the computations are performed and the results are generated. The greatest number of contributions on *diviz* concern Multiple Criteria Decision Analysis (MCDA) (see Cinelli et al. 2022; Greco et al. 2016). All *diviz* modules take input data and produce outputs in XMCD, a dedicated XML-based format.

#### 3.1 Implemented modules

All methods for robustness analysis in Imprecise DEA have been implemented in R and made available on the *diviz* platform as independent modules (web services). Their source code is available at [https://github.com/alabijak/diviz\\_DEA/tree/master/ImpreciseDEACCR](https://github.com/alabijak/diviz_DEA/tree/master/ImpreciseDEACCR). They can be used individually or combined into complex workflows. Each module accepts five input files:

- *units* containing information about the considered DMUs;
- *inputs/outputs* listing information on the inputs and outputs and their scales (quantitative or qualitative (ordinal));
- *performance* providing information on the DMUs' precise performances or, if the problem involves interval inputs and outputs, the minimal performances of DMUs;
- *max performance* is an optional file used in case the interval inputs/outputs are considered; it defines the DMUs' maximal performances;
- *weights constraints* is an optional file containing linear constraints on the weights of inputs and outputs, defining the space of feasible weight vectors.

The modules admit the specification of some additional parameters. The most important ones are *samplesNo* indicating the number of samples derived with the Monte Carlo simulation realized with the HAR algorithm and *tolerance* (in %) used to convert the precise performances into interval ones. For example, a precise value  $x$  is transformed into the interval  $[(1 - tolerance) \cdot x; (1 + tolerance) \cdot x]$ .

The following modules for robustness analysis in IDEA have been implemented on *diviz*:

- *ImpreciseDEA-CCR\_efficiencies* computes the minimal and maximal efficiencies ( $E^*$  and  $E_*$ ) for each DMU using linear programming techniques;
- *ImpreciseDEA-CCR\_extremeRanks* computes the best and the worst efficiency ranks ( $R^*$  and  $R_*$ ) for each DMU using MILP;
- *ImpreciseDEA-CCR\_preferenceRelations* verifies the truth of the necessary and possible efficiency preference relations for all pairs of DMUs using linear programming;

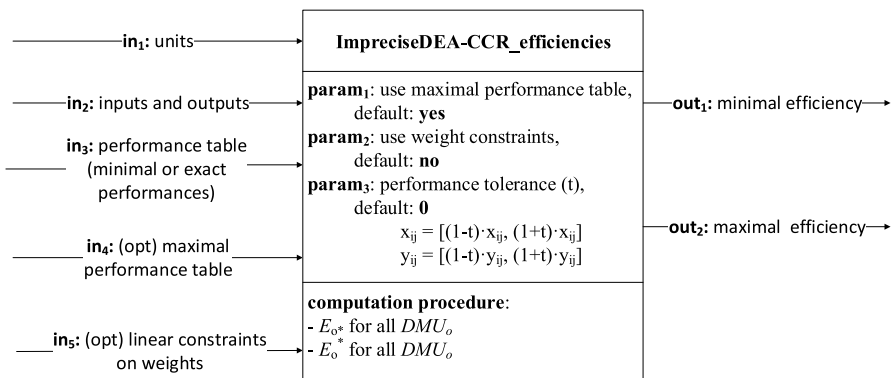
- *ImpreciseDEA-CCR-SMAA\_efficiencies* computes the efficiency distribution, the extreme efficiency scores observed in the analyzed sample of feasible scenarios, and an expected efficiency for each DMU, using the HAR algorithm; it additionally requires specification of the number of buckets as a method parameter;
- *ImpreciseDEA-CCR-SMAA\_preferenceRelations* computes *PEOIs* for all pairs of DMUs using HAR;
- *ImpreciseDEA-CCR-SMAA\_ranks* computes *ERAI*s for all DMUs and ranks, extreme efficiency ranks observed in the analyzed sample of feasible scenarios, and an expected rank for each DMU using HAR.

The structures of two exemplary modules, *ImpreciseDEA-CCR\_efficiencies* and *ImpreciseDEA-CCR-SMAA\_efficiencies*, are presented in Figs. 1 and 2, respectively. They perform computations according to the methods presented in Sects. 2.2.1 and 2.3.1, respectively.

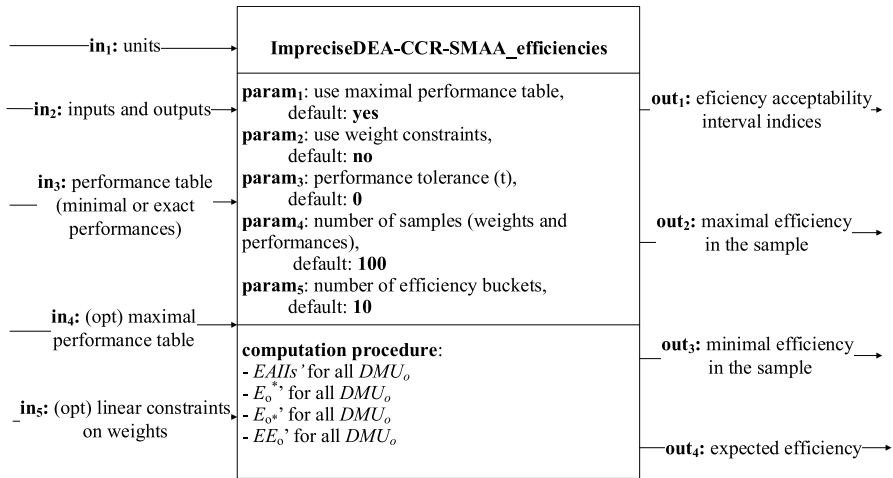
The implemented modules can be combined into an algorithmic workflow with other available computational or visualization modules. Such a workflow can be easily exported and shared with other users. Moreover, the infrastructure of *diviz* allows storing the history of past executions, which is very useful when comparing the results for different settings (e.g., with and without preference information specified by the user). The workflow designed to obtain the results for the case study discussed in Sect. 4 is graphically presented in Fig. 3.

### 4 Illustrative case study

To illustrate the practical usefulness of the proposed framework, we performed the robustness analysis for two studies concerning 27 industrial robots and 17 Chinese ports. The former is based on data derived from Saen (2006), and the detailed results are given in the e-Appendix. The latter builds on data from Jiang et al. (2021),



**Fig. 1** The structure of the *diviz* module which computes the extreme efficiency scores for each DMU using MILP for the Imprecise DEA model



**Fig. 2** The structure of the *diviz* module which computes the Efficiency Acceptability Interval Indices, observed extreme efficiency scores, and expected efficiency for each DMU using the Imprecise DEA model and the Monte Carlo simulation

and the outcomes are reported in this section. The workflows and input data in the XMCD format (ver. 2) for both studies are available at [https://github.com/alabijak/diviz\\_DEA/tree/master/ImpreciseDEACCR](https://github.com/alabijak/diviz_DEA/tree/master/ImpreciseDEACCR).

The ports are described in terms of two precise inputs (labor population and energy consumption), two desirable outputs (cargo throughput—precise and employee satisfaction—ordinal), and one undesirable output (water pollutants—interval). Following (Jiang et al. 2021), the latter factor is treated as an input during the analysis. To obtain the same magnitude for all precise and interval factors, we used the mean normalization before running the methods (see Sarkis 2007; Tomažević et al. 2016; Widiarto and Emrouznejad 2015). The performances of ports on all inputs and outputs are presented in Table 4.

In what follows, we discuss the results of robustness analysis considering the three perspectives: efficiency scores, efficiency ranks, and preference relations. The values of stochastic acceptability indices are estimated based on the 10,000 sampled feasible scenarios. To illustrate that the methods can handle linear weight constraints, we assess water pollutants as less important factor than the other two inputs, i.e.,  $u_{wp} \leq u_{lp}$  and  $u_{wp} \leq u_{ec}$ , where  $u_{wp}, u_{lp}, u_{ec}$  are, respectively, weights assigned to water pollutants, labor population, and energy consumption. Moreover, we introduce two other constraints preventing the overwhelming role of any input, i.e.,  $w_{lp} \leq w_{ec} + w_{wp}$  and  $w_{ec} \leq w_{lp} + w_{wp}$ .

### 4.1 Efficiency scores

Figure 4 presents the extreme efficiency scores ( $E^*$  and  $E_*$ ) for all DMUs. Regarding the maximal (optimistic) efficiencies, they indicate six efficient ports (Yingkou, Tianjin, Yantai, Ningbo-zhoushan, Fuzhou, and Shantou) with  $E^* = 1$ . Among the

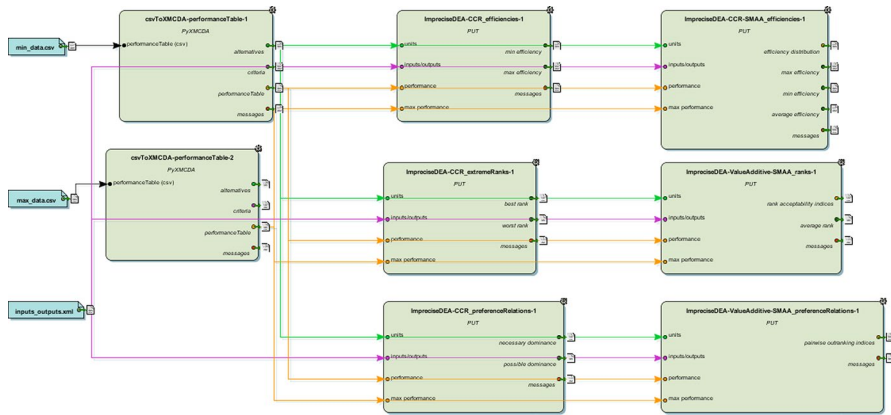


Fig. 3 The *diviz* workflow used to perform the efficiency analysis for a case study (see Sect. 4)

inefficient ports, the best efficiency is attained by Fangcheng (0.909) and Zhanjiang (0.887). On the other extreme, there are two ports with maximal efficiency scores lower than 0.6 (Shanghai (0.539) and Qinhuangdao (0.598)). When analyzing the minimal (pessimistic) efficiencies, the most advantageous port is Ningbo-zhoushan (0.158). The minimal efficiency scores of all other ports are significantly lower (close to zero).

In general, the efficiency intervals are relatively wide. For this reason, analyzing the distribution of efficiency scores is desirable. In Table 5, we present the Efficiency Acceptability Interval Indices, while assuming  $B = 10$  sub-intervals from  $[0, 0.1]$  to  $(0.9, 1.0]$ . When it comes to the efficient ports, the greatest *EAI* for the best interval is attained by Tianjin ( $EAI(Tianjin, (0.9, 1.0]) = 57.3%$ ) and Yantai ( $EAI(Yantai, (0.9, 1.0]) = 68.3%$ ). Only three other ports attained an efficiency greater than 0.9 for at least one sample, but the respective *EAI*s are significantly lower (16.8% for Shantou and less than 9% for others). Interestingly, Fuzhou—deemed efficient—has not achieved an efficiency score in the best interval for any weight sample. Obviously, such scores are feasible (as confirmed with the analysis of exact extreme scores), but *EAI*s indicate that they are improbable.

For some ports, the analysis of *EAI*s allows indicating the most probable ranges of efficiencies even if the efficiency intervals are relatively wide. For example, the efficiency score for Yantai is in the best three buckets for 98.6% of feasible scenarios, with the vast majority (68.3%) in the last bucket. In the same spirit, the efficiency score of Qinhuangdao is between 0.2 and 0.4 for 89.8% of feasible scenarios, and there is no sample for which its efficiency is greater than 0.5. However, there is also a group of ports with efficiency scores strongly dependent on the selected weight and performance vectors. For example, for Shantou, *EAI*s greater than 16% are attained for the two very different intervals,  $(0.9, 1]$  and  $(0.3, 0.4]$ . Also, for this port and eight buckets representing efficiency scores between 0.2 and 1.0, *EAI*s are greater than zero. Similarly, Fuzhou has a positive share of feasible scenarios for nine sub-intervals.

**Table 4** Input and output values for considered Chinese ports (for employee satisfaction, 1 and 17 indicate, respectively, the worst and the best ordinal performances)

Port	Inputs		Outputs		Undesirable output
	Labor population	Energy consumption	Cargo throughput	Employee satisfaction	Water pollutants
Dalian	0.961	0.997	1.255	15	[1.137, 1.421]
Yingkou	0.641	1.093	1.012	16	[0.796, 1.023]
Qinhuangdao	1.827	0.406	0.537	9	[0.455, 0.682]
Tianjin	1.022	0.533	1.583	15	[1.251, 1.706]
Yantai	0.208	0.798	0.763	15	[0.569, 0.853]
Qingdao	1.686	1.701	1.438	15	[1.137, 1.706]
Rizhao	0.763	0.830	1.006	9	[0.910, 1.137]
Shanghai	2.524	2.811	1.854	11	[1.706, 2.161]
Lianyungang	0.574	0.721	0.577	9	[0.455, 0.682]
Ningbo-zhoushan	1.895	2.533	2.651	17	[1.990, 2.843]
Fuzhou	0.434	0.465	0.417	6	[0.341, 0.455]
Xiamen	0.980	0.600	0.601	6	[0.398, 0.569]
Shantou	0.601	0.141	0.143	11	[0.284, 0.398]
Shenzhen	0.989	0.444	0.615	1	[0.512, 0.625]
Guangzhou	1.096	1.499	1.502	6	[1.706, 1.990]
Zhanjiang	0.353	0.971	0.736	6	[0.569, 0.682]
Fangcheng	0.447	0.456	0.307	6	[0.341, 0.512]

The distribution of efficiency scores can be translated into a single, easily understandable measure, i.e., expected efficiency (see Fig. 4). These scores impose a complete order on the set of ports. Yantai is the best, with an expected efficiency of 0.930. This means it is either efficient or very close to being efficient for most feasible scenarios. The other two ports in the top three are Tianjin (0.877) and Yingkou (0.768). Dalian attains the next highest expected efficiency. Even though it is inefficient, it is ranked better in terms of *EE* than the remaining three efficient ports (Ningbo-zhoushan, Fuzhou, and Shantou). The three ports with the least expected efficiencies are Shenzhen (0.376), Shanghai (0.334), and Qinhuangdao (0.307).

### 4.2 Efficiency ranks

The extreme efficiency ranks ( $R^*$  and  $R_*$ ) for all ports are presented in Fig. 5. Only the six ports deemed as efficient have the best rank equal to one. Furthermore, the inefficient units with relatively high maximal efficiency scores attain the best rank equal to two (see Dalian, Rizhao, Zhanjiang, and Fangcheng). Only one additional inefficient port (Lianyungang) is ranked at the podium in the best case. Four ports are always ranked outside the top five (see Shanghai, Xiamen,

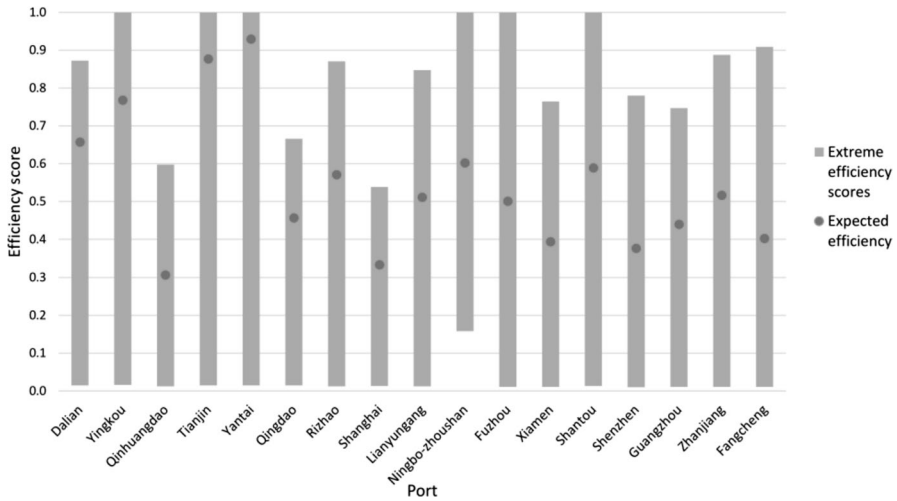


Fig. 4 Graphical representation of extreme efficiency scores and expected efficiencies for seventeen ports

Shenzhen, and Guangzhou). The least advantageous among them is Shanghai because even in the optimistic scenario, as many as 7 other ports attain higher efficiency ( $R^* = 8$ ).

The most preferred ports in terms of the worst efficiency ranks are Yantai ( $R_* = 7$ ) and Yingkou ( $R_* = 11$ ). When it comes to other efficient ports, their ranks drop to 14-th (for Tianjin), 16-th (for Ningbo-zhoushan and Fuzhou), or 17-th (for Shantou) positions in the most pessimistic case. Hence the ranks of Shantou range between the most extreme possible ones. In general, for 13 out of 17 ports, the pessimistic rank is lower than 15-th. Among them, six inefficient ports (Qinhuangdao, Shanghai, Xiamen, Shenzhen, Guangzhou, and Fangcheng) are ranked at the bottom in the least advantageous scenario.

Since the rank intervals for all ports are broad, we have estimated Efficiency Rank Acceptability Indices (see Table 6). They reveal the distribution of ranks attained by each DMU across the feasible weight and performance vectors. Four of six efficient ports attained the top position for at least one feasible scenario. For Yantai, this occurs for 58.9% of samples, Tianjin and Shantou are ranked at the top for similar shares of scenarios (19.6% and 15.3%, respectively), whereas for Yingkou, the ERAI for the first position is significantly lesser (6.2%). Regarding the remaining efficient ports, Ningbo-zhoushan attained at most second rank, and Fuzhou was at most third for relatively negligible shares of feasible scenarios.

For many ports, it is possible to indicate a single rank or a relatively narrow range of ranks attained for the vast majority of feasible scenarios. For example, Yingkou is ranked third for more than 60% scenarios, Yantai is ranked in the top three for all sampled scenarios, and Dalian is placed between fourth and sixth for more than 85% scenarios. Some other ports attain a more extensive range of ranks for a significant

**Table 5** Efficiency distribution for considered ports obtained with the Monte Carlo simulation

Port	[0.0, 0.1]	(0.1, 0.2]	(0.2, 0.3]	(0.3, 0.4]	(0.4, 0.5]	(0.5, 0.6]	(0.6, 0.7]	(0.7, 0.8]	(0.8, 0.9]	(0.9, 1.0]
Dalian	0.000	0.000	0.002	0.007	0.052	0.154	0.443	0.335	0.007	0.000
Yingkou	0.000	0.000	0.000	0.000	0.001	0.021	0.221	0.447	0.221	0.089
Qinhuangdao	0.002	0.047	0.390	0.508	0.053	0.000	0.000	0.000	0.000	0.000
Tianjin	0.000	0.000	0.002	0.000	0.015	0.056	0.090	0.113	0.151	0.573
Yantai	0.000	0.000	0.000	0.000	0.001	0.001	0.012	0.093	0.210	0.683
Qingdao	0.000	0.002	0.039	0.168	0.475	0.309	0.007	0.000	0.000	0.000
Rizhao	0.000	0.012	0.040	0.084	0.140	0.169	0.426	0.127	0.002	0.000
Shanghai	0.004	0.112	0.203	0.411	0.269	0.001	0.000	0.000	0.000	0.000
Lianyungang	0.000	0.002	0.026	0.089	0.262	0.483	0.135	0.003	0.000	0.000
Ningbo-zhoushan	0.000	0.000	0.010	0.064	0.115	0.196	0.444	0.163	0.007	0.001
Fuzhou	0.003	0.019	0.040	0.114	0.221	0.446	0.144	0.012	0.001	0.000
Xiamen	0.011	0.050	0.135	0.240	0.436	0.126	0.002	0.000	0.000	0.000
Shantou	0.000	0.000	0.139	0.165	0.132	0.118	0.111	0.085	0.082	0.168
Shenzhen	0.043	0.091	0.144	0.185	0.359	0.175	0.003	0.000	0.000	0.000
Guangzhou	0.022	0.086	0.107	0.141	0.166	0.351	0.127	0.000	0.000	0.000
Zhanjiang	0.004	0.033	0.068	0.112	0.155	0.303	0.255	0.070	0.000	0.000
Fangcheng	0.004	0.026	0.081	0.350	0.429	0.092	0.016	0.002	0.000	0.000

share of feasible scenarios. For example, Shantou has non-zero *ERAI* for all ranks, with the acceptability indices ranging from 1.2% to 18.6%. Furthermore, Rizhao and Lianyungang have *ERAI*s greater than 10% for five consecutive ranks. Also, the analysis of *ERAI*s leads to identifying some ranks which are feasible according to the exact robustness analysis while being less probable, as confirmed by the stochastic analysis conducted with the Monte Carlo simulation. For example, Qinhuangdao could be ranked fifth in the best case, but the probabilities of ranks above 8 are already close to zero. In the same spirit, Tianjin could be ranked 14-th in the worst case, but the shares of scenarios ranking it worse than 9 are negligible. In general, the analysis of *ERAI*s points out the ports for which the attained ranks are rather stable, or the ranks' variability is great. This means that their position strongly depends on the particular scenario (weights and performances).

*ERAI*s can be transformed into expected ranks that allow ordering the ports from the best to the worst (see Fig. 5). In particular, for Yantai, *ER* is equal to 1.438, which confirms its superiority over the remaining ports. The following two positions are attained by Tianjin (2.827) and Yingkou (2.918). Even though Dalian is inefficient, its expected rank is better than those attained by the remaining three efficient ports. Qinhuangdao, Shanghai, and Shenzhen attain the worst expected ranks. They confirm that these three ports are placed at the bottom for most feasible scenarios.

### 4.3 Preference relations

The third analyzed perspective concerns the stability of efficiency-based pairwise preference relations. In Table 7, we present the matrix summarizing the truth of the necessary (N) and possible (P) preference relations. First, let us note that the necessary relation is reflexive, which is confirmed with “N” on the main diagonal of Table 7. Furthermore, the necessary relation holds for 20 pairs involving different

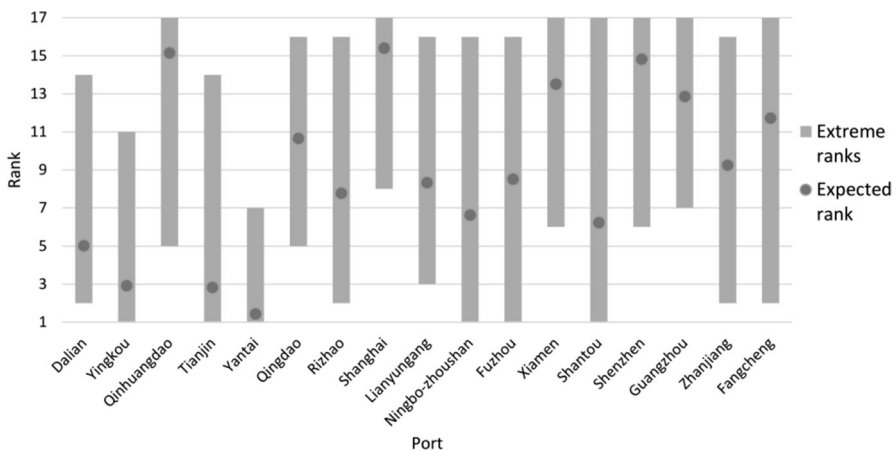


Fig. 5 Graphical representation of extreme efficiency ranks and estimated expected ranks (note that the closer to the bottom of the figure, the better the ranks)



**Table 6** Efficiency Rank Acceptability Indices for analyzed ports obtained with the Monte Carlo simulation

Port	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Dalian	0.000	0.000	0.046	0.290	0.431	0.133	0.045	0.036	0.017	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000
Yingkou	0.062	0.169	0.613	0.119	0.023	0.010	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Qinhuangdao	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.005	0.008	0.051	0.022	0.039	0.071	0.123	0.110	0.189	0.382
Tianjin	0.196	0.260	0.174	0.303	0.040	0.017	0.006	0.003	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Yantai	0.589	0.384	0.027	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Qingdao	0.000	0.000	0.000	0.000	0.000	0.014	0.092	0.062	0.114	0.095	0.260	0.198	0.102	0.055	0.008	0.000	0.000
Rizhao	0.000	0.000	0.005	0.027	0.086	0.182	0.211	0.142	0.125	0.103	0.076	0.037	0.006	0.000	0.000	0.000	0.000
Shanghai	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.002	0.005	0.013	0.018	0.039	0.141	0.214	0.336	0.228
Lianyungang	0.000	0.000	0.000	0.015	0.039	0.113	0.150	0.215	0.176	0.203	0.067	0.021	0.001	0.000	0.000	0.000	0.000
Ningbo-zhoushan	0.000	0.001	0.025	0.078	0.224	0.300	0.111	0.078	0.064	0.039	0.046	0.021	0.005	0.008	0.000	0.000	0.000
Fuzhou	0.000	0.000	0.004	0.024	0.061	0.090	0.107	0.163	0.212	0.169	0.132	0.034	0.004	0.000	0.000	0.000	0.000
Xiamen	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.032	0.103	0.373	0.329	0.133	0.022	0.006
Shantou	0.153	0.186	0.100	0.122	0.057	0.051	0.023	0.025	0.024	0.020	0.019	0.039	0.012	0.016	0.022	0.070	0.061
Shenzhen	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.007	0.024	0.128	0.105	0.099	0.218	0.168	0.248
Guangzhou	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.128	0.059	0.058	0.061	0.122	0.083	0.108	0.145	0.170	0.064
Zhanjiang	0.000	0.000	0.004	0.015	0.023	0.056	0.201	0.085	0.126	0.187	0.122	0.094	0.063	0.023	0.001	0.000	0.000
Fangcheng	0.000	0.000	0.002	0.007	0.016	0.034	0.048	0.054	0.069	0.060	0.125	0.146	0.136	0.098	0.149	0.045	0.011

ports. This means that for all feasible scenarios, one port attains efficiency at least as good as the other port, confirming the robustness of its advantage. For example, Tianjin is necessarily preferred to Qingdao, Guangzhou, and Shanghai.

The necessary relation can be presented graphically using the Hasse diagram (see Fig. 6). Such a diagram does not represent the truth of relations that can be derived from the transitivity. For example, since Dalian is necessarily preferred to Qingdao, and Qingdao is necessarily preferred to Shanghai, Dalian is also preferred to Shanghai. Note that no other port is necessarily preferred over six efficient ports. However, there are also four inefficient (Dalian, Rizhao, Shenzhen, and Fangcheng) for which there does not exist any other port confirming its necessary advantage over them.

The port which proves its robust superiority over the greatest number of other ports is Yantai. It is necessarily preferred over seven other ports. An interesting situation can be observed for Fuzhou and Shantou. Although they are efficient, they are not necessarily preferred over any other port. The latter (i.e., no outgoing arc) also holds for seven other ports. Among them, Shanghai and Guangzhou are necessarily worse than the most significant number of other ports (6 and 5, respectively). The necessary relation graph suggests the improvement paths that inefficient ports can follow to improve their efficiency gradually. For example, for Shanghai, we can construct the following example paths: Shanghai–Qingdao–Dalian or Shanghai–Qingdao–Yingkou. Alternatively, it can directly learn from Ningbo-zhoushan.

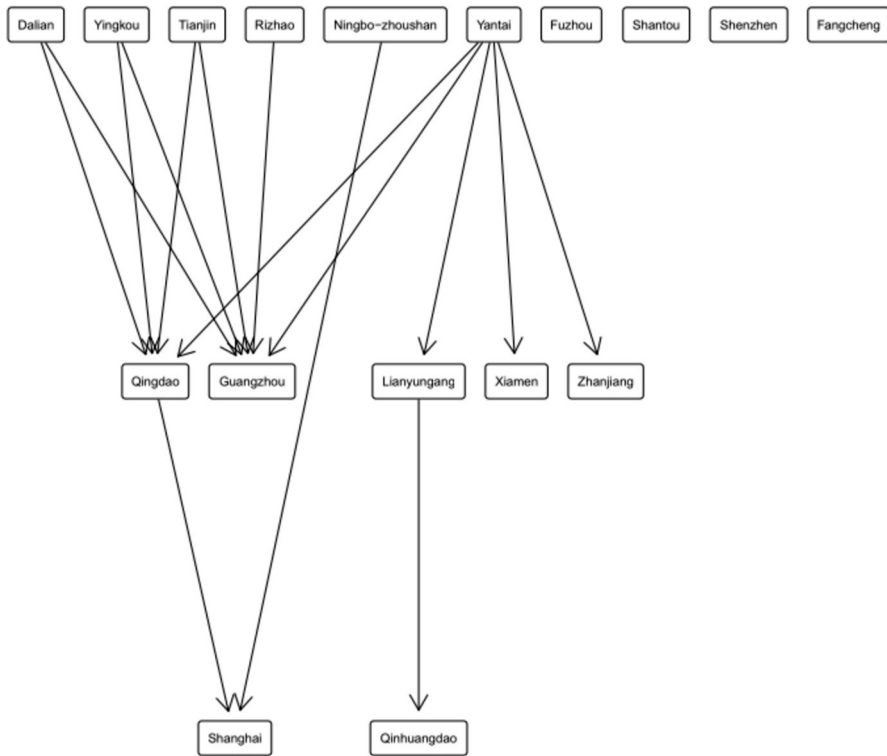
Regarding the possible preference relation (see Table 7), let us emphasize that the necessary relation implies the truth of the possible one. However, the latter one holds also for pairs that are not related by the necessary preference. For example, all efficient ports are incomparable in terms of the necessary relation. This means there is at least one feasible scenario for which one port is preferred to the other and at least one feasible scenario for which this relation is inverse (e.g., Yingkou and Tianjin). Such a situation also occurs for pairs of inefficient ports (e.g., Dalian and Rizhao or Qingdao and Guangzhou). When the possible relation for a given pair of ports is false (e.g., (Qingdao, Dalian) or (Shanghai, Qingdao)), then one port is less efficient than the other for all feasible scenarios.

When the necessary relation is true or the possible relation is false, a pair of ports is compared in the same way for all feasible scenarios. However, when the ports are incomparable in terms of the necessary relation, it is interesting to analyze the shares of feasible scenarios that rank one of the ports at least as good as the other. Pairwise Efficiency Outranking Indices capture such shares (see Table 8).

For some pairs, these indices confirm the superiority of one port over the other. For example,  $PEOI(\text{Yingkou}, \text{Zhanjiang}) = 0.995$  and  $PEOI(\text{Rizhao}, \text{Qinhuangdao}) = 0.999$  indicate that one port is at least as efficient as the other for over 99% scenarios, and the inverse relation is extremely unlikely with  $PEOIs$  close to zero. For some other pairs, the  $PEOIs$  indicate that the shares of scenarios confirming the advantage of either port over another are very similar (e.g.,  $PEOI(\text{Shanghai}, \text{Qinhuangdao}) = 0.506$  and  $PEOI(\text{Qinhuangdao}, \text{Shanghai}) = 0.494$  or  $PEOI(\text{Shenzhen}, \text{Qinhuangdao}) = 0.540$  and  $PEOI(\text{Qinhuangdao}, \text{Shenzhen}) = 0.460$ ). For these pairs, the indication of a more preferred port strongly depends on the selected weights and performance vectors.

**Table 7** The necessary and possible efficiency preference relations for all pairs of ports (N—necessary relation; P—possible relation; empty cell—the relation is not even possible)

Port	Dalian	Yingkou	Qinhuangdao	Tianjin	Yantai	Qingdao	Rizhao	Shanghai	Lianyungang	Ningbo	Fuzhou	Xiamen	Shantou	Shenzhen	Guangzhou	Zhanjiang	Fangcheng
Dalian	N	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P
Yingkou	P	N	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P
Qinhuangdao	P	P	N	P	P	P	P	P	P	P	P	P	P	P	P	P	P
Tianjin	P	P	P	N	P	P	P	P	P	P	P	P	P	P	P	P	P
Yantai	P	P	P	P	N	P	P	P	P	P	P	P	P	P	P	P	P
Qingdao	P	P	P	P	P	N	P	P	P	P	P	P	P	P	P	P	P
Rizhao	P	P	P	P	P	P	N	P	P	P	P	P	P	P	P	P	P
Shanghai	P	P	P	P	P	P	P	N	P	P	P	P	P	P	P	P	P
Lianyungang	P	P	P	P	P	P	P	P	N	P	P	P	P	P	P	P	P
Ningbo	P	P	P	P	P	P	P	P	P	N	P	P	P	P	P	P	P
Zhoushan	P	P	P	P	P	P	P	P	P	P	N	P	P	P	P	P	P
Fuzhou	P	P	P	P	P	P	P	P	P	P	P	N	P	P	P	P	P
Xiamen	P	P	P	P	P	P	P	P	P	P	P	P	N	P	P	P	P
Shantou	P	P	P	P	P	P	P	P	P	P	P	P	P	N	P	P	P
Shenzhen	P	P	P	P	P	P	P	P	P	P	P	P	P	P	N	P	P
Guangzhou	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	N	P
Zhanjiang	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	N
Fangcheng	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P



**Fig. 6** The Hasse diagram representing the necessary efficiency preference relation

## 5 Conclusions and future work

We have introduced a rich framework for robustness in the context of Imprecise Data Envelopment Analysis. The proposed methods are applicable in the context of three major types of uncertainty that occur in real-world decision problems (Pelissari et al. 2021). First, we consider ambiguity in the input and output performances that could be interpreted differently due to their ordinal or interval character. Second, we account for stochasticity by considering discrete and continuous probability distributions. Third, we deal with partial information on the input and output weights by exploiting the space of feasible weights delimited with a set of linear constraints. In this way, the proposed approaches for robustness analysis can be applied to real-world problems for which it is difficult to express the knowledge or collect precise data, the variables are unquantifiable, some errors in measurements occur, or the users are not able or willing to express their complete preferences (see Dehnokhalaji et al. 2022; Pelissari et al. 2021).

When considering the stability of results that can be attained for the possible performances and weights, we focus on three types of efficiency-based outcomes: scores, preference relations, and ranks. On the one hand, the mathematical

programming models compute the extreme efficiency scores and ranks and verify the truth of the necessary and possible preference relations. These outcomes reveal each DMU's performance for the most and the least advantageous scenarios and collate the efficiencies of all pairs of DMUs for all or at least one feasible scenario. Thus, they offer an exact perspective on the DMUs' performance. On the other hand, we incorporate the stochastic analysis driven by the Monte Carlo simulations to derive the probability distribution of different outcomes and expected results. These stochastic outcomes complement the exact results derived using mathematical programming. They also provide means for analyzing trends or some prevailing scenarios and imposing the ranking on the set of DMUs in the line of expected scores or ranks. Above all, such various outcomes have not been offered by any previous IDEA method.

The practical usefulness of the proposed framework has been illustrated in real-world case studies concerning the evaluation of Chinese ports (Jiang et al. 2021) and industrial robots (Saen 2006). The data sets involved precise, interval, and ordinal factors. The results were computed with modules implemented in R and available on the *diviz* platform (Meyer and Bigaret 2012). They incorporate MILP solver and advanced sampling techniques.

The main limitations of the proposed framework are three-fold. First, when the number of units runs over a few hundred, the linear programs are too big and too many, posing a significant problem for contemporary solvers. This is particularly true for results such as extreme ranks that are established using binary variables. Moreover, for such big data problems, the robust results, such as the ranking induced by the necessary preference relation, cannot be presented to the user because of their high complexity. Then, it is more beneficial to refer to complete orders of units based on expected efficiencies or ranks. Still, let us emphasize that large scale-applications are less common in DEA. Second, the results of the stochastic analysis depend on the assumed distribution of weights and performances within the intervals as well as the hypotheses made when representing the ordinal performances. Clearly, the developed framework is applicable with other types of distributions than uniform and arbitrary performance ranges from which the ordinal performances could be sampled. These choices may affect the values of stochastic acceptabilities. Also, whichever assumptions are made, even if the indices can be estimated within the acceptable error bound, they are not accurate. Third, we accounted for the standard imprecision types considered in IDEA, including interval and ordinal performances. As proved by a comprehensive review by Pelissari et al. (2021), uncertain performances can also be modeled differently. The most popular approaches for this purpose include fuzzy numbers, non-uniform probabilistic distributions, evidential reasoning, and grey numbers. Their combinations with DEA have gained in popularity in recent years. Hence adopting the framework for robustness analysis to their context remains an appealing direction for future research.

We develop the methods introduced in this paper in the following directions. First, we adapt them to hierarchical structures of inputs and outputs. In this way, the robustness of efficiency outcomes given imprecise performances can be investigated at the comprehensive and local levels (Shen et al. 2013). The latter corresponds to a more elementary perspective or particular sub-area of DMUs' functioning. Indeed,

**Table 8** Pairwise Efficiency Outranking Indices for all pairs of ports obtained with the Monte Carlo simulation

Port	Dalian	Yingkou	Qinhuangdao	Qinhuangdao	Tianjin	Tianjin	Yantai	Qingdao	Rizhao	Shanghai	Lianyungang	Ningbo	Fuzhou	Xiamen	Shantou	Shenzhen	Guangzhou	Zhanjiang	Fangcheng
Dalian	1.000	0.069	1.000	1.000	0.002	1.000	0.979	1.000	0.910	0.834	0.893	1.000	0.410	1.000	1.000	0.967	0.937		
Yingkou	0.931	1.000	1.000	0.968	0.499	1.000	1.000	1.000	0.970	0.990	1.000	0.598	1.000	1.000	0.995	0.995			
Qinhuangdao	0.000	0.000	1.000	0.052	0.000	0.494	0.001	0.000	0.020	0.058	0.218	0.062	0.460	0.385	0.118	0.092			
Tianjin	0.998	0.501	1.000	1.000	1.000	0.207	1.000	1.000	0.981	0.973	0.968	1.000	0.552	1.000	1.000	0.999	0.973		
Yantai	1.000	0.931	1.000	1.000	0.793	1.000	1.000	1.000	0.996	1.000	1.000	0.809	1.000	1.000	1.000	1.000	1.000		
Qingdao	0.000	0.000	0.948	0.000	0.000	0.194	1.000	1.000	0.204	0.018	0.300	0.881	0.236	0.944	0.721	0.349	0.643		
Rizhao	0.021	0.032	0.999	0.000	0.000	0.806	1.000	0.995	0.511	0.307	0.645	0.990	0.318	1.000	1.000	0.781	0.759		
Shanghai	0.000	0.000	0.506	0.000	0.000	0.000	0.005	1.000	0.006	0.000	0.029	0.157	0.146	0.412	0.210	0.048	0.218		
Lianyungang	0.090	0.000	1.000	0.796	0.019	0.000	0.489	0.994	1.000	0.280	0.484	0.999	0.272	0.993	0.815	0.602	0.840		
Ningbo	0.166	0.030	0.980	0.982	0.027	0.004	0.693	1.000	0.720	1.000	0.763	0.966	0.376	0.997	1.000	0.856	0.839		
Fuzhou	0.107	0.010	0.942	0.700	0.032	0.000	0.355	0.971	0.516	0.237	1.000	1.000	0.274	0.998	0.808	0.558	0.981		
Xiamen	0.000	0.000	0.782	0.119	0.000	0.000	0.010	0.843	0.001	0.034	0.000	1.000	0.184	0.717	0.487	0.004	0.284		
Shantou	0.590	0.402	0.938	0.764	0.448	0.191	0.682	0.854	0.728	0.624	0.726	0.816	1.000	0.812	0.758	0.711	0.830		
Shenzhen	0.000	0.000	0.540	0.056	0.000	0.000	0.000	0.588	0.007	0.003	0.002	0.283	0.188	1.000	0.174	0.000	0.266		
Guangzhou	0.000	0.000	0.615	0.279	0.000	0.000	0.000	0.790	0.185	0.000	0.192	0.513	0.242	0.826	1.000	0.004	0.418		
Zhanjiang	0.033	0.005	0.882	0.651	0.001	0.000	0.219	0.952	0.398	0.144	0.442	0.996	0.289	1.000	0.996	1.000	0.618		
Fangcheng	0.063	0.005	0.908	0.357	0.027	0.000	0.241	0.782	0.160	0.161	0.019	0.716	0.170	0.734	0.582	0.382	1.000		

the hierarchical structure is helpful for decomposing complex decision problems into smaller, manageable sub-problems. This is particularly useful in scenarios with high numbers of inputs and outputs, which often happens in medicine, energy, banking, and finances. Moreover, we design the Multiple Objective Optimization algorithms to determine different scenarios of improvements (i.e., reductions of consumed resources or increases of produced results) required for attaining or maintaining a particular target. These targets refer to many robust results, e.g., being necessarily ranked in the top three or attaining an efficiency score of at least 0.7 for all feasible scenarios (Ciomek et al. 2018). Moreover, the proposed framework offers flexibility to the Decision Maker, who can indicate which factors should be modified and to which extent. The obtained solutions reflect the trade-offs between modifications needed on various factors. Their analysis may lead to selecting the most preferred solution to be implemented in practice. Finally, we develop the methods for robustness analysis in the context of value-based DEA (see Gouveia et al. 2008; Labijak-Kowalska et al. 2023). This model is based on concepts from Multiple Criteria Decision Analysis, allowing the incorporation of managerial preferences on different levels. In this regard, the uncertainty is related to performances, weights, and the shape of value functions for inputs and outputs. However, the output types produced by these methods are similar to those discussed in this paper.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12351-023-00755-z>.

**Author contributions** AL-K: Conceptualization, Methodology, Software, Investigation, Data curation, Writing- Original draft, Writing-review and editing, Visualization. MK: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing- Original draft, Writing-review and editing, Supervision, Project administration, Funding acquisition.

**Funding** The research of Anna Labijak-Kowalska was supported by the Polish Ministry of Education and Science, grant no. 0311/SBAD/0726. Miłosz Kadziński acknowledges support from the Polish National Science Center under the SONATA BIS project (Grant No. DEC-2019/34/E/HS4/00045).

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Adler N, Friedman L, Sinuany-Stern Z (2002) Review of ranking methods in the Data Envelopment Analysis context. *Eur J Oper Res* 140(2):249–265
- Ang S, Zhu Y, Yang F (2021) Efficiency evaluation and ranking of supply chains based on stochastic multicriteria acceptability analysis and Data Envelopment Analysis. *Int Trans Oper Res* 28:3190–3219
- Aparicio J, Ruiz JL, Sirvent I (2007) Closest targets and minimum distance to the Pareto-efficient frontier in DEA. *J Prod Anal* 28(3):209–218
- Aparicio J, Cordero JM, Ortiz L (2019) Measuring efficiency in education: the influence of imprecision and variability in data on DEA estimates. *Socioecon Plan Sci* 68:100698
- Azadi M, Saen RF (2013) A combination of QFD and imprecise DEA with enhanced Russell graph measure: a case study in healthcare. *Socioecon Plan Sci* 47(4):281–291
- Azizi H, Kordrostami S, Amirteimoori A (2015) Slacks-based measures of efficiency in imprecise Data Envelopment Analysis: an approach based on Data Envelopment Analysis with double frontiers. *Comput Ind Eng* 79:42–51
- Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. *Eur J Oper Res* 2(6):429–444
- Charnes A, Cooper W, Lewin A, Seiford L (1994) *Data Envelopment Analysis: theory, methodology, and applications*. Springer, Netherlands
- Chen L, Wang Y-M (2020) DEA target setting approach within the cross efficiency framework. *Omega* 96:102072
- Cinelli M, Kadziński M, Miebs G, Gonzalez M, Słowiński R (2022) Recommending multiple criteria decision analysis methods with a new taxonomy-based decision support system. *Eur J Oper Res* 302(2):633–65
- Ciomek K, Kadziński M (2021) Polyrun: a Java library for sampling from the bounded convex polytopes. *SoftwareX* 13:100659
- Ciomek K, Ferretti V, Kadziński M (2018) Predictive analytics and disused railways requalification: insights from a Post Factum Analysis perspective. *Decis Support Syst* 105:34–51
- Cook WD, Seiford LM (2009) Data Envelopment Analysis (DEA)—thirty years on. *Eur J Oper Res* 192(1):1–17
- Cooper WW, Park KS, Yu G (1999) IDEA and AR-IDEA: models for dealing with imprecise data in DEA. *Manag Sci* 45(4):597–607
- Cooper WW, Park KS, Yu G (2001) An illustrative application of IDEA (Imprecise Data Envelopment Analysis) to a Korean mobile telecommunication company. *Oper Res* 49(6):807–820
- Cooper W, Seiford L, Zhu J (2014) *Handbook on Data Envelopment Analysis*. International series in operations research & management science. Springer, New York
- Corrente S, Greco S, Słowiński R (2017) Handling imprecise evaluations in multiple criteria decision aiding and robust ordinal regression by n-point intervals. *Fuzzy Optim Decis Mak* 16(2):127–157
- Dehnohalaji A, Khezri S, Emrouznejad S (2022) A box-uncertainty in DEA: a robust performance measurement framework. *Expert Syst Appl* 187:115855
- Despotis DK, Smirlis YG (2002) Data Envelopment Analysis with imprecise data. *Eur J Oper Res* 140(1):24–36
- Ebrahimi B, Khalili M (2018) A new integrated AR-IDEA model to find the best DMU in the presence of both weight restrictions and imprecise data. *Comput Ind Eng* 125:357–363
- Ebrahimi B, Toloo M (2020) Efficiency bounds and efficiency classifications in imprecise DEA: an extension. *J Oper Res Soc* 71(3):491–504
- Ebrahimi B, Tavana M, Kleine A, Dellnitz A (2021) An epsilon-based Data Envelopment Analysis approach for solving performance measurement problems with interval and ordinal dual-role factors. *OR Spectr* 43:1103–1124
- Emrouznejad A, Yang G-L (2018) A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socioecon Plan Sci* 61:4–8
- Farrell MJ (1957) The measurement of productive efficiency. *J R Stat Soc Ser A (Gen)* 120(3):253–290
- Gouveia M, Dias L, Antunes C (2008) Additive DEA based on MCDA with imprecise information. *J Oper Res Soc* 59:54–63
- Gouveia M, Dias LC, Antunes CH (2013) Super-efficiency and stability intervals in additive DEA. *J Oper Res Soc* 64(1):86–96



- Greco S, Mousseau V, Słowiński R (2008) Ordinal regression revisited: multiple criteria ranking using a set of additive value functions. *Eur J Oper Res* 191(2):416–436
- Greco S, Ehrgott M, Figueira J (2016) Multiple criteria decision analysis—state of the art surveys. *International series in operations research & management science*. Springer, New York
- Hadi-Vencheh A, Matin RK (2011) An application of IDEA to wheat farming efficiency. *Agric Econ* 42(4):487–493
- Haghighat MS, Khorram E (2005) The maximum and minimum number of efficient units in DEA with interval data. *Appl Math Comput* 163(2):919–930
- Hosseinzadeh Lotfi F, Jahanshahloo GR, Khodabakhshi M, Rostamy-Malkhelifeh M, Moghaddas Z, Vaez-Ghasemi M (2013) A review of ranking models in Data Envelopment Analysis. *J Appl Math* 2013:492421
- Jahanshahloo GR, Lofti FH, Moradi M (2004) Sensitivity and stability analysis in DEA with interval data. *Appl Math Comput* 156(2):463–477
- Jiang B, Yang C, Dong Q, Li J (2021) Ecological efficiency evaluation of China's port industries with imprecise data. *Evol Intell*. <https://doi.org/10.1007/s12065-021-00638-2>
- Kadziński M, Tervonen T (2013) Robust multi-criteria ranking with additive value models and holistic pair-wise preference statements. *Eur J Oper Res* 228(1):169–180
- Kadziński M, Greco S, Słowiński R (2012) Extreme ranking analysis in robust ordinal regression. *Omega* 40(4):488–501
- Kadziński M, Labijak A, Napieraj M (2017) Integrated framework for robustness analysis using ratio-based efficiency model with application to evaluation of Polish airports. *Omega* 67:1–18
- Kao C (2006) Interval efficiency measures in Data Envelopment Analysis with imprecise data. *Eur J Oper Res* 174(2):1087–1099
- Kao C, Liu S-T (2005) Data Envelopment Analysis with imprecise data: an application of Taiwan machinery firms. *Int J Uncertain Fuzziness Knowl Based Syst* 13(2):225–240
- Kao C, Liu S-T (2009) Stochastic Data Envelopment Analysis in measuring the efficiency of Taiwan commercial banks. *Eur J Oper Res* 196(1):312–322
- Karsak E, Karadayi M (2017) Imprecise DEA framework for evaluating health-care performance of districts. *Kybernetes* 46(4):706–727
- Kim S-H, Park C-G, Park K-S (1999) An application of Data Envelopment Analysis in telephone offices evaluation with partial data. *Comput Oper Res* 26(1):59–72
- Labijak-Kowalska A, Kadziński M (2021) Experimental comparison of results provided by ranking methods in Data Envelopment Analysis. *Expert Syst Appl* 173:114739
- Labijak-Kowalska A, Kadziński M, Spychała I, Dias LC, Fiallos J, Patrick J, Michalowski W, Farion K (2023) Performance evaluation of emergency department physicians using robust value-based additive efficiency model. *Int Trans Oper Res* 30(1):503–544
- Lahdelma R, Salminen P (2001) SMAA-2: stochastic multicriteria acceptability analysis for group decision making. *Oper Res* 49(3):444–454
- Lahdelma R, Salminen P (2006) Stochastic multicriteria acceptability analysis using the data envelopment model. *Eur J Oper Res* 170(1):241–252
- Lahdelma R, Miettinen K, Salminen P (2003) Ordinal criteria in stochastic multicriteria acceptability analysis (SMAA). *Eur J Oper Res* 147(1):117–127
- Liang Q, Liao X, Shang J (2020) A multiple criteria approach integrating social ties to support purchase decision. *Comput Ind Eng* 147:106655
- Liu JS, Lu LY, Lu W-M, Lin BJ (2013) A survey of DEA applications. *Omega* 41(5):893–902
- Meyer P, Bigaret S (2012) Diviz: a software for modeling, processing and sharing algorithmic workflows in MCDA. *Intell Decis Technol* 6(4):283–296
- Park K (2007) Efficiency bounds and efficiency classifications in DEA with imprecise data. *J Oper Res Soc* 58(4):533–540
- Pelissari R, Oliveira MC, Abackerli AJ, Ben-Amor S, Assumpcao MRP (2021) Techniques to model uncertain input data of multi-criteria decision-making problems: a literature review. *Int Trans Oper Res* 28(2):523–559
- Saen RF (2006) Technologies ranking in the presence of both cardinal and ordinal data. *Appl Math Comput* 176(2):476–487
- Salo A, Punkka A (2011) Ranking intervals and dominance relations for ratio-based efficiency analysis. *Manag Sci* 57(1):200–214
- Sarkis J (2007) Preparing your data for DEA. In: Zhu J, Cook WD (eds) *Modeling data irregularities and structural complexities in Data Envelopment Analysis*. Springer, Berlin, pp 305–320

- Seiford L, Zhu J (1998) Stability regions for maintaining efficiency in Data Envelopment Analysis. *Eur J Oper Res* 108:127–139
- Shen Y, Hermans E, Brijs T, Wets G (2013) Data Envelopment Analysis for composite indicators: a multiple layer model. *Soc Indic Res* 114(2):739–756
- Shokouhi AH, Hatami-Marbini A, Tavana M, Saati S (2010) A robust optimization approach for imprecise Data Envelopment Analysis. *Comput Ind Eng* 59(3):387–397
- Tervonen T, Lahdelma R (2007) Implementing stochastic multicriteria acceptability analysis. *Eur J Oper Res* 178(2):500–513
- Tervonen T, van Valkenhoef G, Baştürk N, Postmus D (2013) Hit-And-Run enables efficient weight generation for simulation-based multiple criteria decision analysis. *Eur J Oper Res* 224(3):552–559
- Toloo M, Keshavarz E, Hatami-Marbini A (2021) An interval efficiency analysis with dual-role factors. *OR Spectr* 43:255–287
- Tomažević N, Seljak J, Aristovnik A (2016) TQM in public administration organisations: an application of Data Envelopment Analysis in the police service. *Total Qual Manag Bus Excell* 27(11–12):1396–1412
- Widiarto I, Emrouznejad A (2015) Social and financial efficiency of Islamic microfinance institutions: a Data Envelopment Analysis application. *Socioecon Plan Sci* 50:1–17
- Wu J, Yu Y, Zhu Q, An Q, Liang L (2018) Closest target for the orientation-free context-dependent DEA under variable returns to scale. *J Oper Res Soc* 69(11):1819–1833
- Zahran SZ, Alam JB, Al-Zahrani AH, Smirlis Y, Papadimitriou S, Tsioumas V (2020) Analysis of port efficiency using imprecise and incomplete data. *Oper Res Int J* 20(1):219–246
- Zhu J (1996) Robustness of the efficient DMUs in Data Envelopment Analysis. *Eur J Oper Res* 90:451–460
- Zhu J (2003) Imprecise Data Envelopment Analysis (IDEA): a review and improvement with an application. *Eur J Oper Res* 144(3):513–529

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.