RAIRO-Oper. Res. 58 (2024) 3093–3106
https://doi.org/10.1051/ro/2024113

RAIRO Operations Research
www.rairo-ro.org

A SIMULATION-BASED DEA APPROACH FOR MULTIPLE CRITERIA DECISION-MAKING PROBLEMS WITH UNCERTAIN MIXED-CRITERIA VALUES

Shiling Song[1], Ye Zhang[1], Jianhui Xie[2], Sheng Ang[3,4,*] and Feng Yang[5]

Abstract. In ex-ante decision scenarios, predicting criterion values accurately is difficult for decision makers (DMs). Inconsiderable work is normally required for measuring criteria by uncertain random values or ordinal values. However, in the classical data envelopment analysis (DEA) model, criterion values are the constants that limit the application of the classical DEA model in ex-ante decision scenarios. This paper presents a simulation-based DEA approach, which captures random and ordinal criterion values by a simple and direct simulation-based approach. The approach includes three steps. In the first step, Monte Carlo simulation methods are used to convert uncertain random values or ordinal values into cardinal data. In the second step, we use traditional DEA methods to compute the efficiency score of decision-making units (DMUs). In the third step, we ranked all DMUs by calculating the DEA-efficient acceptability of each DMU in multiple simulations and then selected the optimal DMU. The proposed approach is illustrated by experimental examples and a case study of a municipal wastewater treatment system.

Mathematics Subject Classification. 91B06.

Received June 21, 2022. Accepted May 14, 2024.

1. Introduction

Data envelopment analysis (DEA) was originally proposed by Charnes *et al.* [4] as a method for evaluating the relative efficiency of decision-making units (DMUs) performing essentially the same task. The classical DEA model assumes that the input and output of observation samples are deterministic data. Thus, the classical DEA is mainly used as an ex-post evaluation tool for classifying DMUs into inefficient and efficient ones. As in the case of post-evaluation, the criterion values are mainly deterministic. The classical DEA model may have difficulty on supporting multiple criteria decision making (MCDM) problems because typical multiple criteria decision making problems are often ex-ante, and obtaining accurate criterion values is very time-consuming

Keywords. Multiple criteria decision-making, data envelopment analysis, uncertainty, Monte Carlo simulation.

[1] College of Mathematics and Informatics, South China Agricultural University, 483 Wushan Road, Guangzhou 510000, P.R. China.

[2] International School of Business Finance, Sun Yat-Sen University, Zhuhai Campus, Tangjiawan, Zhuhai 519082, P.R. China.

[3] School of Management, Hefei University of Technology, 193 Tunxi Road, Hefei 230009, P.R. China.

[4] Intelligent Interconnection System Anhui Provincial Laboratory, Hefei 230009, P.R. China.

[5] School of Management, University of Science and Technology of China, 96 Jinzhai Road, Hefei 230026, P.R. China.

*Corresponding author: shengang@hfut.edu.cn

and expensive. Inconsiderable work is normally required for measuring criteria by uncertain random values or ordinal values, which rank the alternatives criterion-wise. In this study, we introduce a simulation-based DEA approach to manage the mixed random cardinal and ordinal criteria data.

The simulation-based DEA approach consists of three steps. In the first step, ordinal and random values are mapped into the cardinal criterion values. In the second step, we use traditional DEA methods to compute the relative efficiency of DMUs. In the third step, we calculate the relative efficiency acceptability of each DMU by counting the relative effective times of each DMU in all simulations.

Compared to other multi-criterion decision-making methods, which use weights to represent criteria's relative importance, the simulation-based DEA approach proposed in this study has three advantages. First, the approach does not need the decision makers to express their preferences information. Meanwhile, other methods, such as analytic hierarchy process (AHP [21]), and technique for order of preference by similarity to ideal solution (TOPSIS), tend to infer the weights of the criterion based on the decision-maker's preference information. If little preference information is available, these methods cannot solve the problem effectively. Second, the proposed approach is able to deal with simultaneously the uncertain ordinal criterion values, random criterion value and uncertain weight information. Third, the approach is more intuitive and accurate than other approximate methods, such as rough set theory [12] and fuzzy set theory [13].

The contributions of this study are displayed in two aspects. First, we propose a simulation-based DEA approach for the multi-criterion decision-making problems with mixed-criteria values and unknown weight information. Second, we extend classical DEA methods to solve multi-criterion decision-making problems with random and ordinal data.

The proceeding parts of this paper are structured as follows. Section 2 reviews related literature. Section 3 introduces the simulation-based DEA approach. Section 4 presents experimental examples and a case study for illustration. Discussion and conclusion of the paper is given in Sections 5 and 6, respectively.

2. Related literature

Multi-criterion decision-making problems refers to problems in which the alternatives are evaluated based on a set of criteria by considering the priority weights of the criteria and alternatives' performance [27]. The most commonly used multi-criterion decision-making methods include the AHP [21], ANP [22], and TOPSIS.

Many researchers have considered uncertainty in multi-criterion decision-making problems, and several useful research methods have been proposed. The two main uncertainties are concerned with criterion values and weights [29]. Several uncertainty factors (*e.g.*, forecasting errors, lack of time, limitation of cost, and environmental uncertainty) lead to the uncertainties of criterion values. When considering uncertainties, criterion values are no longer constants but random variables with certain probability distributions or ordinal values. For uncertain criterion values, the mainly existing research methods are fuzzy set theory and stochastic theory [10]. Fuzzy theory uses fuzzy logic to represent specific criterion values characterized by vagueness or inaccuracy, and stochastic theories employ random values with random distribution to represent uncertain criteria values.

Chen [5] implemented a fuzzy approach to select the location of a distribution center. In this research, triangular fuzzy numbers were used to model uncertain criteria. Fuzzy logic was also used by Yeo *et al.* [28], who proposed a port choice method with uncertain and incomplete information about port evaluation. Jana and Mohanty [15] also employed fuzzy set theory to capture the linguistic expression in MCDMs. Stochastic multiple criteria methods were also used to deal with MCDM problems with uncertain criterion values. Hokkanen *et al.* [14] employed stochastic multi-criteria acceptability analysis to evaluate nine municipal solid waste management projects. Proposed a unique stochastic AHP that handles uncertain criterion values in MCDM problems. By utilizing the beta distribution and approximating its median, their approach converts various types of expert evaluations into crisp values. Although fuzzy set theory and stochastic multiple attribute decision-making model solve the uncertainty of criterion values to some extent, few existing methods have paid attention to mixed random and ordinal criterion values simultaneously.

Meanwhile, criteria weights used to represent criteria's relative importance have been found to be a potential source of uncertainty in MCDMs. Many useful methods have been proposed to capture uncertain criterion weight. Most of the current methods, such as the AHP [11], ANP [16], and TOPSIS [20], tend to elicit criterion weight according to the DMs' explicit or implicit preference information. However, these methods would not work well when the preference information is completely unknown.

Lahdelma *et al.* [17] proposed a method called stochastic multiple-criterion acceptability analysis (SMAA) to deal with uncertain criterion values and unknown weight information simultaneously. When no weight information is available, the stochastic multiple-criterion acceptability analysis methods assume that weights are evenly distributed in the feasible weight space. Weights that make the alternative the most optimal one form the preferred weight space. An alternative with a large preferred weight space implies a higher probability of being the optimal alternative

In this study, we propose a simulation-based DEA approach, which takes uncertain mixed-criteria values and unknown criterion weights into account simultaneously. Data envelopment analysis is a series of methods originally proposed by Charnes *et al.* [4]. Banker *et al.* [2] extended DEA methods by proposing variable returns to scale models. Doyle and Green [9] developed a cross-efficiency model to make the DEA method discriminative in ranking DMUs. Andersen and Peterson [1] developed a super efficiency model to rank efficient DMUs. Simar and Wilson [23] applied bootstrap method to analyze the sensitivity of efficiency scores that measured by deterministic DEA method. Moreover, Cooper *et al.* [7] extended the DEA method to rank DMUs with imprecise data. Wang *et al.* [26] proposed a DEA-based method to rank DMUs with interval inputs and outputs. Cook and Zhu [6] provided a general DEA framework for dealing rank order data. However, to the best of our knowledge, few studies are available in the DEA literatures that deals with the bias raised by mix random and interval data. The simulation-based DEA approach proposed in the current study will address the problem.

3. A simulation-based DEA approach

Suppose that the decision makers select one optimal DMU from m candidate DMUs that are measured in terms of the multiple criteria classified as inputs to be minimized and outputs to be maximized. Let $A = \{a_1, \ldots, a_m\}$ denote the set of m candidate DMUs. Without losing generality, we assume that there are n_{input} input criteria and n_{output} output criteria.

3.1. Relative efficiency of DMUs

Let x_i^j denote the criterion values of DMU a_i based on the jth criterion. Following Charnes *et al.* [4] and Banker *et al.* [2], the relative efficiency of each DMU is computed by DEA Charnes–Cooper–Rhodes (CCR) model or Banker–Charnes–Cooper (BCC) model. The CCR model is based on constant return to scale (CRS) while the BCC model is based on variable return to scale (VRS).

The relative CRS efficiency of DMU a_i is given by the solution to the fractional programming model [4]:

$$\begin{aligned} &\max \frac{\sum_{j\in\text{output}} w_j x_{ij}}{\sum_{j\in\text{input}} w_j x_{ij}} \\ &\text{s.t. } \frac{\sum_{j\in\text{output}} w_j x_{kj}}{\sum_{j\in\text{input}} w_j x_{kj}} \le 1, \quad k = 1,\ldots,m, \quad w \ge 0. \end{aligned} \tag{1}$$

Applying the Charnes–Cooper transformation proposed by Charnes and Cooper [3], and let $\mu_j = tw_j, j \in$ output and $v_j = tw_j, j \in$ input, where $t = \frac{1}{\sum_{j\in\text{input}} w_j x_{ij}}$, then the fractional programming problem (1) can be equivalently converted to the following linear programming model:

$$e_i = \max \sum_{j\in\text{output}} \mu_j x_{ij}$$

$$\begin{aligned} \text{s.t.} \; & \sum_{j\in\text{input}} v_j x_{ij} = 1 \\ & \sum_{j\in\text{output}} \mu_j x_{kj} - \sum_{j\in\text{input}} v_j x_{kj} \le 0, \quad k = 1,\dots,m \\ & \mu, v \ge 0. \end{aligned} \tag{2}$$

By duality, this problem is also equivalent to the following linear programming problem:

$$\begin{aligned} & \min \theta_i \\ \text{s.t.} \; & \sum_{k=1}^{m} \lambda_k x_{kj} \le \theta_i x_{ij}, \qquad j \in \text{input} \\ & \sum_{k=1}^{m} \lambda_k x_{kj} \ge x_{ij}, \qquad j \in \text{output} \\ & \lambda, s_j^+, s_j^- \ge 0. \end{aligned} \tag{3}$$

The optimal solution, θ_i^*, is the efficiency score of DMU a_i. The constraints of model (3) imply that $\theta_i^* \le 1$. $\theta_i^* < 1$ means the DMU a_i is inefficient while $\theta_i^* = 1$ means the DMU a_i is on the efficient frontier. It is known that one DMU on the efficient frontier could still reduce inputs (increase outputs) without change outputs (inputs). Thus, slack variables s_j^+ and s_j^- for inputs and outputs are introduced, and model (3) becomes the following model (4).

$$\begin{aligned} & \min \theta_i - \varepsilon \left(\sum_{j\in\text{output}} s_j^+ + \sum_{j\in\text{input}} s_j^- \right) \\ \text{s.t.} \; & \sum_{k=1}^{m} \lambda_k x_{kj} + s_j^- = \theta_i x_{ij}, \qquad j \in \text{input} \\ & \sum_{k=1}^{m} \lambda_k x_{kj} - s_j^+ = x_{ij}, \qquad j \in \text{output} \\ & \lambda, s_j^+, s_j^- \ge 0 \end{aligned} \tag{4}$$

where ε is a non-Archimedean element defined to be smaller than any positive real number. Let θ_i^*, s_j^+ and s_j^- be optimal solutions from model (4), we have following two definitions on DEA strongly efficient [8].

Definition 1. DMU a_i is CRS strongly efficient if and only if $\theta_i^* = 1$ and $s_j^+ = s_j^- = 0, \forall j$.

According to Charnes *et al.* [4], the relative VRS efficiency of DMU a_i can be estimated by the following fractional programming:

$$\begin{aligned} \max \; & \frac{\sum_{j\in\text{output}} w_j x_{ij} + u_o}{\sum_{j\in\text{input}} w_j x_{ij}} \\ \text{s.t.} \; & \frac{\sum_{j\in\text{output}} w_j x_{kj} + u_o}{\sum_{j\in\text{input}} w_j x_{kj}} \le 1, \quad k = 1,\dots,m, \quad w \ge 0. \end{aligned} \tag{5}$$

Similarly, model (5) can be converted to a linear programming *via* Charnes–Cooper transformation and is equivalent to the following dual model:

$$
\begin{aligned}
&\min\ \theta_i \\
&\text{s.t.}\ \sum_{k=1}^{m} \lambda_k x_{kj} \leq \theta_i x_{ij}, \qquad j \in \text{input} \\
&\qquad \sum_{k=1}^{m} \lambda_k x_{kj} \geq x_{ij}, \qquad j \in \text{output} \\
&\qquad \sum_{k=1}^{m} \lambda_k = 1 \\
&\qquad \lambda, s_j^+, s_j^- \geq 0.
\end{aligned} \tag{6}
$$

By solving model (6), the optimal efficiency score θ_i^* of DMU a_i can be obtained and it is used to the following mode for solving optimal slack variables s_j^+ and s_j^-.

$$
\begin{aligned}
&\min\ \theta_i - \varepsilon \left(\sum_{j \in \text{output}} s_j^+ + \sum_{j \in \text{input}} s_j^- \right) \\
&\text{s.t.}\ \sum_{k=1}^{m} \lambda_k x_{kj} + s_j^- = \theta_i x_{ij}, \qquad j \in \text{input} \\
&\qquad \sum_{k=1}^{m} \lambda_k x_{kj} - s_j^+ = x_{ij}, \qquad j \in \text{output} \\
&\qquad \sum_{k=1}^{m} \lambda_k = 1 \\
&\qquad \lambda, s_j^+, s_j^- \geq 0.
\end{aligned} \tag{7}
$$

From model (7), the optimal efficiency score θ_i^* and slack variables s_j^+ and s_j^- can be obtained. Thus, similar with Definition 1, we can verify if the evaluated DMU i is VRS strongly efficient. That is,

Definition 2. DMU a_i is VRS strongly efficient if and only if $\theta_i^* = 1$, and the slacks $s_j^+ = s_j^- = 0$ for each criterion j.

As shown above, the CCR and BCC models can discriminate efficient and inefficient DMUs well, where inputs and outputs data are assumed continuous and deterministic. Moreover, those models provide efficiency scores for evaluating the relative performance of the DMUs. In many real-world settings, however, the decision makers are facing random data or ordinal criteria. For example, to select the optimal municipal sewage treatment plan, the decision makers are facing criteria of cost, technical difficulty, regional health effects, number of employees and sewage treatment capacity. Obviously, the data on technical difficulty may be random and ordinal. As discussed by Cook and Zhu [6], there may be bias in using classical DEA models to deal with random data.

To solve this problem, we propose a novel conception of efficient acceptability to evaluate the relative performance for DMUs with random cardinal values and ordinal values and develop a simulation-based DEA approach to compute them. The proposed simulation-based approach is able to deal with data set with more complex data structure.

3.2. Definition of CRS and VRS efficient acceptability

Uncertain criterion values of DMUs are measured by random cardinal values and ordinal values. First, the ordinal values are processed to random cardinal values.

For criteria values measured by ordinal values, let criterion values r_{ij} be represented by rank number $r = 1, \ldots, j_{\max}$ where $j_{\max}$ is the worst rank according criterion j. Following Lahdelma *et al.* [19], we convert those ordinal criterion values to random cardinal values. Define the unknown cardinal criterion values corresponding to ordinal criterion value r_{ij} be x_{ij}^r and the ordinal-to-cardinal mapping be $x_{ij}^r = v_j(r_{ij})$. Without the loss of generality, the scope of x_{ij}^r is limited to the interval $[0, 1]$, where 1 is the best value and is the worst value. Let $\Delta x_{ij}^r = x_{ij}^r - x_{ij}^{r+1}$ be the scale interval between x_{ij}^{r+1} and x_{ij}^r. Obviously Δx_{ij}^r has the following properties:

$$\sum_{r=1}^{j_{\max}-1} \Delta x_{ij}^r = 1. \tag{8}$$

Thus, the unknown scale intervals are represented by the following scale interval space.

$$x_j = \left\{ \Delta x_{ij}^r \in R^{j_{\max}-1} | \Delta x_{ij}^r > 0, \sum_r x_{ij}^r = 1 \right\}. \tag{9}$$

If no additional information about Δx_{ij}^r is given, then there is a natural assumption that the scale interval space distributes uniformly.

All uncertain criterion values of DMUs are now measured only by random cardinal values, represented by random variables x with joint density function $f(x)$ in space $X \in R^{n_{\text{input}}+n_{\text{output}}}$. In the case where each criterion is independent of each other, their joint probability density function $f(x)$ is calculated by multiplying their respective probability density functions. Let $x \in X$ be an arbitrary set of input and output sample criterion values in X, and we can calculate the efficiency score of DMU a_i through linear programming (4). Let the efficiency score of DMU a_i be $\theta_i^*(x)$, and the slack variables for inputs and outputs criteria be $s_j^+(x)$ and $s_j^-(x)$, respectively. Then the favorable set for DMU a_i for being CRS efficient can then be defined as

$$X_i^* = \left\{ x \in X : \theta_i^*(x) = 1 \wedge s_j^+(x) = s_j^+(x) = 0 \text{ for all } j \right\}. \tag{10}$$

The CRS efficient acceptability of DMU i is defined as

$$P_i^* = \int_{x \in X_i^*} f(x) \, \mathrm{d}x. \tag{11}$$

Similarly, we calculate the VRS efficient score $\theta_i^*(x)_{\text{VRS}}$ of DMU a_i through linear programming (7). Let the slack variables for inputs and outputs criteria be $s_j^+(x)_{\text{VRS}}$ and $s_j^-(x)_{\text{VRS}}$, respectively. The favorable set for DMU a_i for being strongly efficient can then be defined as

$$X_{i_{\text{VRS}}}^* = \left\{ x \in X : \theta_i^*(x)_{\text{VRS}} = 1 \wedge s_j^+(x)_{\text{VRS}} = s_j^+(x)_{\text{VRS}} = 0 \text{ for all } j \right\}. \tag{12}$$

The VRS efficient acceptability of DMU i is defined as

$$P_{i_{\text{VRS}}}^* = \int_{x \in X_{i_{\text{VRS}}}^*} f(x) \, \mathrm{d}x. \tag{13}$$

The CRS and VRS efficient acceptability defined in formula (11) and (13) are not easy to calculate since the integrals have high dimensions [24]. Next we propose a novel Monte-Carlo simulation procedure to calculate the DMUs' CRS and VRS efficient acceptability.

3.3. The simulation-based procedures for calculating efficient acceptability

In this subsection, we introduce a novel Monte Carlo simulation procedure for calculating the CRS and VRS efficient acceptability of each DMU. Both the CRS efficient acceptability and VRS efficient acceptability

TABLE 1. Experimental parameters.

Parameter	Values
Number of DMUs	$m \in \{10, 15, 20\}$
Number of input and output	$(n_{\text{input}}, n_{\text{output}}) \in \{(2,2), (3,3)\}$
Distribution of criteria values	1: Uniform distribution;
	2: Normal distribution;
	3: Ordinal values;

are calculated as the proportion of efficient times in a number of simulations. Let M as the total number of Monte Carlo simulation trials. Let $\text{SE}_i(\text{CRS})$ be the number of times that DMU i is CRS strongly efficient. Let $\text{SE}_i(\text{VRS})$ be the number of times that DMU i is VRS strongly efficient. Denote $\text{SEA}_i(\text{CRS})$ as the CRS efficient acceptability of DMU i. Denote $\text{SEA}_i(\text{VRS})$ as the VRS efficient acceptability of DMU i. The CRS efficient acceptability of DMU i is computed as the ratio of $\text{SE}_i(\text{VRS})$ to M, that is $\text{SEA}_i(\text{CRS}) = \frac{\text{SE}_i(\text{CRS})}{M}$. The VRS efficient acceptability of DMU i is computed as the ratio of $\text{SE}_i(\text{VRS})$ to M, that is $\text{SEA}_i(\text{VRS}) = \frac{\text{SE}_i(\text{VRS})}{M}$. For example, suppose that the CRS strongly efficient number of times that DMU i obtains in 10 000 simulation trials is 2000, the CRS efficient acceptability is calculated by $2000/10\,000 = 0.2$. A DMU with a larger value of the CRS efficient acceptability implies the DMU performs better.

The following algorithm presents a step-by-step procedure of the proposed simulation-based DEA approach for calculating the DMUs' CRS efficient acceptability and VRS efficient acceptability.

Step 1. Determine the trials number of Monte Carlo simulation. The trials number M is typically set as 10 000–1 000 000. Let $\text{SE}_i(\text{CRS}) = 0$ and $\text{SE}_i(\text{VRS}) = 0$.
Step 2. Generate DMUs' criterion values in each trial from their corresponding distributions.
Step 3. Compute the relative efficiency θ_i^* for DMU $i(i = 1, \ldots, m)$ through linear programming (4) and (7). Compute the slack variables s_j^+ and s_j^- through linear programming (4). If $\theta_i^* = 1$ and $s_j^+ = s_j^- = 0, \forall j$, $\text{SE}_i(\text{CRS}) = \text{SE}_i(\text{CRS}) + 1$. Compute the slack variables s_j^+ and s_j^- through linear programming (7). If $\theta_i^* = 1$ and $s_j^+ = s_j^- = 0, \forall j$, $\text{SE}_i(\text{VRS}) = \text{SE}_i(\text{VRS}) + 1$.
Step 4. Repeat steps 2 and 3 until the number of simulation iterations ends. Compute the CRS efficient acceptability by $\text{SEA}_i(\text{CRS}) = \frac{\text{SE}_i(\text{CRS})}{M}$. Compute the VRS efficient acceptability by $\text{SEA}_i(\text{VRS}) = \frac{\text{SE}_i(\text{VRS})}{M}$.
Step 5. The DUMs are evaluated and ranked according to DMU's CRS efficient acceptability or VRS efficient acceptability.

4. EXPERIMENTAL EXAMPLES AND A CASE STUDY

In this section, the simulation-based DEA approach is demonstrated through experimental examples and a case study. In the experimental examples the rankings of DMUs based on CRS efficient acceptability and VRS efficient acceptability are compared. The case study is about selecting a municipal sewage treatment system in a city in south China.

4.1. Experimental examples

Table 1 shows an overview of the experimental parameters. For each parameter value, 50 problems are randomly generated using the method described below.

In uniform distribution cases, $m \times (n_{\text{input}} + n_{\text{output}})$ uniformly distributed integer data from 100 to 200 are generated. Then, the integer data is set as the lower bound of a uniform distribution. The upper bound of the uniform distribution is defined as 1.2 times of the integer data.

In the normal distribution case, $m \times (n_{\text{input}} + n_{\text{output}})$ uniformly distributed integer data from 100 to 200 are generated. These integer data are set as the means of normal distributions. The standard deviation of the normal distributions is set as 0.1 times of the data.

In the ordinal values case, a sequence of integers from 1 to m are generated. Then the order of each data is randomly shuffled. Then the scrambled integer data are used to represent the ordinal value for each criterion.

After the criteria values are generated, the simulation-based procedures are used to calculate the CRS efficient acceptability and VRS efficient acceptability. To compare the similarity between the rankings of alternatives based on CRS efficient acceptability and VRS efficient acceptability in different problems, Kendall correlation coefficients [25] is used. Figure 1 shows the distribution of Kendall correlation coefficients between the rankings of DMUs based on CRS efficient acceptability and VRS efficient acceptability in different problems with varying parameter values. As shown in Figure 1, the Kendall correlation coefficients of the two rankings is close to one under each parameter value, indicating that rankings of DMUs based on CRS efficient acceptability and VRS efficient acceptability are highly consistent.

4.2. A case study

In the bustling metropolis of Guangzhou, China, rapid economic growth has led to an increase in sewage production. In response, the local government has set its sights on implementing a municipal sewage treatment system to address this issue. The objective is to identify the most suitable option for managing wastewater effectively for the system. With the aid of experts, ten alternatives are evaluated on the basis of five criteria:

(1) Net cost per ton (NCPT).
(2) Technical difficulty (TD).
(3) Regional health effects (RHE).
(4) Number of employees (NOE).
(5) Sewage treatment capacity (STC).

Among the five criteria, the first three are input criteria, and the last two are output criteria. Moreover, criteria 1, 4, and 5 are measured by random data. The uniform distribution is used to measure these three criteria. Criteria 2 and 3 are measured by ordinal data because these two criteria are difficult to measure by cardinal data. The ordinal values of criteria 2 and 3 are formed by experts based on a qualitative analysis of the alternatives without actual measurements. When the difference between criteria measurements is smaller than the originally defined indifference threshold, the alternatives are given the same rank. Table 2 shows the criterion values of the 10 alternatives, where NCPT is measured by Yuan per ton, the NOE is measured by the number of persons, and the STC is measured by thousand tons per day. The number of simulations is set to be 10 000. Table 3 presents the CRS efficient acceptability and VRS efficient acceptability of each alternative.

The following conclusions are formulated on the basis of Table 3. According to the CRS efficient acceptability, the order of the 10 alternatives is $G \succ A \succ B \succ F \succ H \succ I \succ J \succ D \succ E \succ C$. According to the VRS efficient acceptability, the order of the 10 alternatives is $G \succ A \succ B \succ F \succ E \succ H \succ I \succ J \succ D \succ C$.

Next, we examine the influence of the variance of the random cardinal criterion value on the results. Table 4 reports the criterion values of the 10 alternatives, where mean value of each cardinal criterion remains the same while the range is doubled.

Table 5 reports the efficient acceptability results of each alternative with data set in Table 4. Results show that the order of the 10 alternatives is $G \succ B \succ F \succ H \succ A \succ I \succ J \succ C \succ E \succ D$ according to the CRS efficient acceptability, while the order of the 10 alternatives is $G \succ B \succ F \succ A \succ H \succ I \succ J \succ C \succ E \succ D$ according to the VRS efficient acceptability.

The optimal alternatives are the same from results in Tables 3 and 5. It finds that their acceptability will decrease with the increase of variance, from changes of the alternatives' CRS efficient acceptability and VRS efficient acceptability in two different data sets.

Subsequently, we examine the influence of ordinal criterion values on the results. Suppose that all the criterion values are measured in ordinal values. The random cardinal criterion values are converted into corresponding

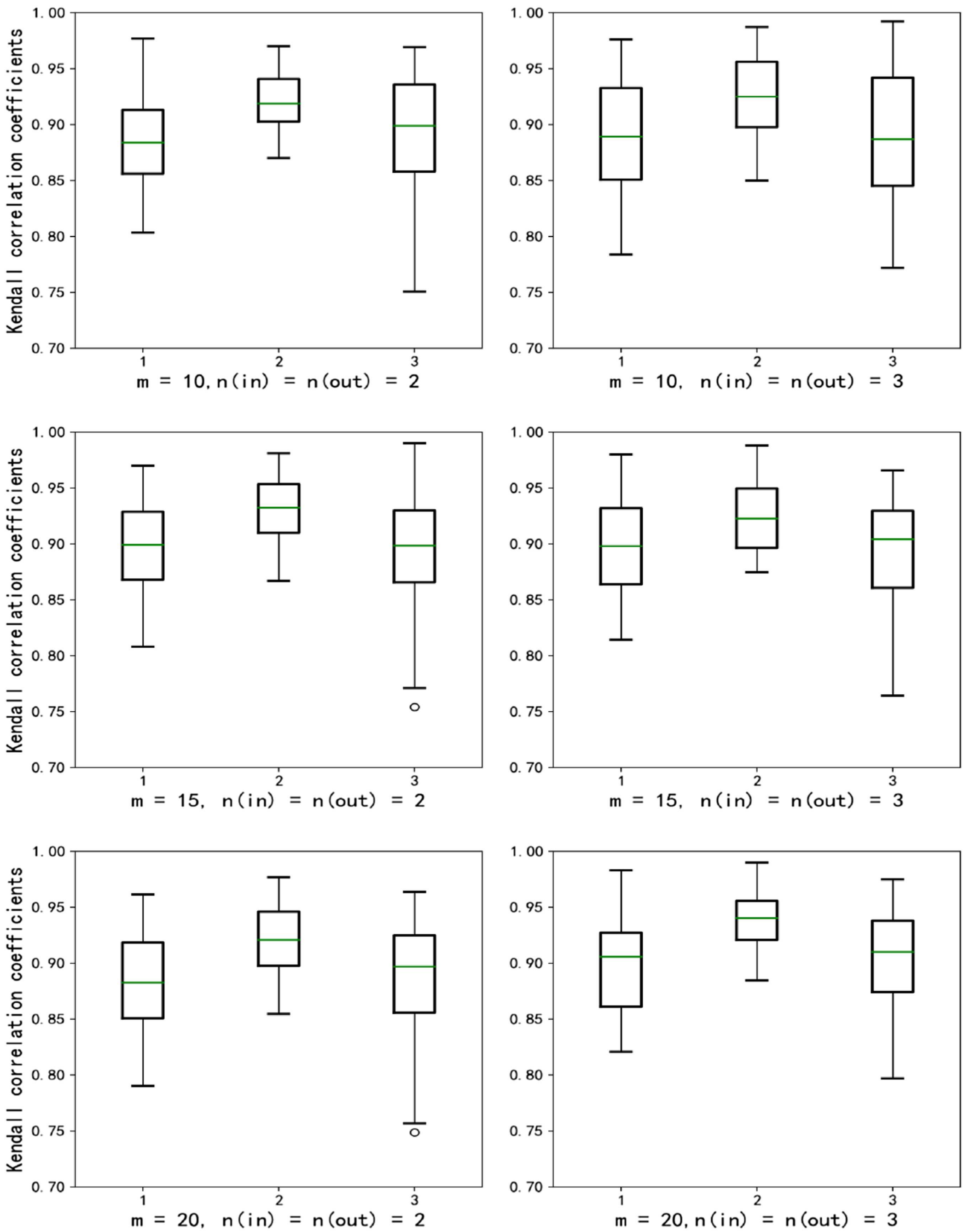

Figure 1. Distribution of Kendall correlation coefficients.

TABLE 2. Data set of the municipal sewage treatment system.

Alternatives	NCPT	TD	RHE	NOE	STC
A	$[0.4, 0.6]$	2	2	$[500, 600]$	$[30, 40]$
B	$[0.3, 0.5]$	2	1	$[400, 500]$	$[40, 60]$
C	$[0.6, 1.0]$	4	5	$[200, 300]$	$[15, 20]$
D	$[0.6, 0.8]$	3	6	$[100, 150]$	$[12, 18]$
E	$[0.5, 0.7]$	3	4	$[300, 400]$	$[25, 35]$
F	$[0.7, 1.0]$	4	5	$[600, 700]$	$[20, 30]$
G	$[0.15, 0.25]$	1	4	$[200, 300]$	$[70, 100]$
H	$[1, 1.5]$	5	3	$[500, 650]$	$[30, 40]$
I	$[0.8, 1.4]$	5	4	$[400, 600]$	$[20, 30]$
J	$[1, 1.3]$	5	5	$[300, 500]$	$[15, 25]$

TABLE 3. Efficient acceptability results of ten alternatives.

DMUs	CRS efficient acceptability	VRS efficient acceptability
A	0.548	0.630
B	0.477	0.567
C	0.000	0.110
D	0.040	0.114
E	0.006	0.258
F	0.446	0.449
G	0.990	0.990
H	0.302	0.245
I	0.215	0.238
J	0.083	0.127

TABLE 4. Data set with a doubled range for random cardinal criterion values.

Alternatives	NCPT	TD	RHE	NOE	STC
A	$[0.3, 0.7]$	2	2	$[450, 650]$	$[25, 45]$
B	$[0.2, 0.6]$	2	1	$[400, 500]$	$[30, 70]$
C	$[0.4, 1.2]$	4	5	$[150, 350]$	$[12.5, 22.5]$
D	$[0.5, 0.9]$	3	6	$[75, 175]$	$[9, 21]$
E	$[0.4, 0.8]$	3	4	$[25, 450]$	$[20, 40]$
F	$[0.55, 1.15]$	4	5	$[550, 750]$	$[15, 35]$
G	$[0.1, 0.3]$	1	4	$[150, 350]$	$[55, 115]$
H	$[0.75, 1.75]$	5	3	$[425, 725]$	$[25, 45]$
I	$[0.5, 1.7]$	5	4	$[300, 700]$	$[1535]$
J	$[0.85, 1.45]$	5	5	$[200, 600]$	$[12.5, 32.5]$

TABLE 5. Efficient acceptability results of alternatives with a doubled range in cardinal criterion values.

Alternatives	CRS efficient acceptability	VRS efficient acceptability
A	0.327	0.358
B	0.484	0.490
C	0.050	0.142
D	0.007	0.120
E	0.020	0.126
F	0.465	0.440
G	0.941	0.940
H	0.359	0.345
I	0.274	0.277
J	0.157	0.196

TABLE 6. Data set with ordinal criterion values.

Alternatives	NCPT	TD	RHE	NOE	STC
A	8	2	2	3	3
B	9	2	1	5	2
C	5	4	5	8	7
D	6	3	6	9	8
E	7	3	4	7	4
F	4	4	5	1	5
G	10	1	4	8	1
H	1	5	3	2	3
I	3	5	4	4	5
J	2	5	5	6	6

ordinal values based on their mean value. Table 6 shows the ordinal criterion values of the 10 alternatives and Table 7 reports the results.

The results are obtained from Table 7. (1) According to the CRS efficient acceptability, the order of the 10 alternatives is $G \succ A \succ B \succ F \succ H \succ I \succ E \succ C \succ D \succ J$. (2) According to the VRS efficient acceptability, the order of the 10 alternatives $G \succ B \succ A \succ F \succ E \succ I \succ H \succ C \succ J \succ D$.

Above experiments shows that the variance of CRS efficient acceptability for alternatives is much larger than that of VRS efficient acceptability and thus the former one has more power for discrimination. As a result, the simulation-based approach with CRS efficient acceptability is more suggested to search the optimal alternative.

It is known that the SMAA method often is applied to MCDM problems with random cardinal and ordinal criterion values and does not need DMs' subjective preference information. Here a comparison between the SMAA method and the proposed simulation-based DEA approach is made.

Table 8 reports the result of SMAA and the simulation-based DEA approach for data set in Table 2. The table shows that alternative G obtains the highest holistic acceptability of 0.638 and that alternative C receives a lowest holistic acceptability of 0.123. The ranking of alternatives based on their holistic acceptability is: $G \succ H \succ F \succ I \succ J \succ D \succ B \succ A \succ E \succ C$. Thus, alternative G is the optimal alternative, whereas alternative C is the worst alternative. This is consistent with the simulation-based DEA approach.

TABLE 7. Efficient acceptability results when all criterion values are ordinal.

Alternatives	CRS efficient acceptability	VRS efficient acceptability
A	0.643	0.514
B	0.561	0.556
C	0.016	0.061
D	0.006	0.041
E	0.065	0.202
F	0.309	0.267
G	0.806	0.799
H	0.215	0.140
I	0.092	0.161
J	0.001	0.055

TABLE 8. SMAA's holistic acceptability and DEA efficient acceptability of each alternative.

Alternatives	SMAA's holistic acceptability	CRS efficient acceptability	VRS efficient acceptability
A	0.167	0.548	0.630
B	0.205	0.477	0.567
C	0.123	0.000	0.110
D	0.207	0.040	0.114
E	0.124	0.006	0.258
F	0.355	0.446	0.449
G	0.638	0.990	0.990
H	0.479	0.302	0.245
I	0.325	0.215	0.238
J	0.306	0.083	0.127

Comparison results shows that the proposed simulation-based DEA approach has two advantages over the SMAA method in this case study. First, the simulation-based DEA approach ranks alternatives on the basis of resource utilization efficiency, whereas SMAA does not pay enough attention to resource utilization efficiency. Second, the simulation-based DEA approach provides a better distinction between different alternatives.

5. Discussion

One should be noted that the analysis quality of the simulation-based DEA approach is dependent on the data. If all the criteria are measured by ordinal data, then the quality of the analysis results will be slightly decreased, as shown in the case data study in Section 4.2. If criteria are measured by random data with relatively small variance, then the quality of analysis results will be significantly improved.

The simulation-based DEA approach can be used in many stages of decision-making. In the early stage, the criterion values may be imprecise. The simulation-based DEA approach helps decision-makers identify possibly inferior and superior DMUs quickly. In the later stage, with the improvement of data quality, the analysis quality of the simulation-based DEA approach will be improved. DMs obtain more useful information from analysis results of the approach.

6. Conclusion

In the study, we present a simulation-based DEA approach for aiding MCDM problems. The approach intends to help the DMs in selecting the optimal alternative when little weight information is available and parts of or all the criteria values are measured as ordinal data or stochastic values. The computations of the approach are performed using the Monte Carlo simulation.

The simulation-based DEA approach has certain limitations which provide ways for future research. First, the criteria are assumed independent of each other. Relaxing such assumption will extend the approach. Second, the efficient acceptability computation needs specifying the number of simulations. A tradeoff between the accuracy of the calculation results and the calculation time of the approach should be carefully analyzed. Combination of the simulation-based DEA approach and other efficient simulation algorithm is of interest for future research.

Acknowledgements

The authors thank the editors and reviewers for their helpful comments. The study was supported by National Natural Science Foundation of China (72101091, 71991464, 71921001, 72188101, 32372002), Scientific and Technological Planning Project of Guangzhou City (202201010184) and Fundamental Research Funds for the Central Universities of China (PA2024GDGP0033).

References

[1] P. Andersen and N.C. Petersen, A procedure for ranking efficient units in data envelopment analysis. *Manage. Sci.* **39** (1993) 1261–1264.

[2] R.D. Banker, A. Charnes and W.W. Cooper, Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manage. Sci.* **30** (1984) 1078–1092.

[3] A. Charnes and W.W. Cooper, Programming with linear fractional functionals. *Nav. Res. Logistics Q.* **9** (1962) 181–186.

[4] A. Charnes, W.W. Cooper and E. Rhodes, Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* **2** (1978) 429–444.

[5] C.T. Chen, A fuzzy approach to select the location of the distribution center. *Fuzzy Sets Syst.* **118** (2001) 65–73.

[6] W.D. Cook and J. Zhu, Rank order data in DEA: a general framework. *Eur. J. Oper. Res.* **174** (2006) 1021–1038

[7] W.W. Cooper, K.S. Park and G. Yu, IDEA and AR-IDEA: models for dealing with imprecise data in DEA. *Manage. Sci.* **45** (1999) 597–607.

[8] W.W. Cooper, L.M. Seiford and J. Zhu, editiors. Handbook on Data Envelopment Analysis. Vol. 164. Springer Science & Business Media (2011).

[9] J. Doyle and R. Green, Efficiency and cross-efficiency in DEA: derivations, meanings and uses. *J. Oper. Res. Soc.* **45** (1994) 567–578.

[10] R.M. García-Morales, A. Baquerizo and M.A. Losada, Port management and multiple-criteria decision making under uncertainty. *Ocean Eng.* **104** (2015) 31–39.

[11] S.I. Gass, The analytic hierarchy process – an exposition. *Oper. Res.* **49** (2001) 469-486.

[12] S. Greco, B. Matarazzo and R. Slowinski, Rough set approach to multi-attribute choice and ranking problems, in Multiple Criteria Decision Making. Lecture Notes in Economics and Mathematical Systems, edited by G. Fandel and T. Gal. Vol. 448. Springer, Berlin, Heidelberg (1997).

[13] R.W. Hipel, Fuzzy set techniques in decision making. *IFAC Proc. Vol.* **15** (1982) 423–431.

[14] J. Hokkanen, R. Lahdelma and P. Salminen, Multicriteria decision support in a technology competition for cleaning polluted soil in Helsinki. *J. Environ. Manage.* **60** (2000) 339–348.

[15] B. Jana and S.N. Mohanty, An intuitionistic fuzzy logic models for multicriteria decision making under uncertainty. *J. Inst. Eng. (India) Ser. C* **98** (2017) 197–201.

[16] S. Jharkharia and R. Shankar, Selection of logistics service provider: an analytic network process (ANP) approach. *Omega* **35** (2007) 274–289.

[17] R. Lahdelma, J. Hokkanen and P. Salminen, SMAA-stochastic multiobjective acceptability analysis. *Eur. J. Oper. Res.* **106** (1998) 137-143.

[18] R. Lahdelma and P. Salminen, SMAA-2: stochastic multicriteria acceptability analysis for group decision making. *Oper. Res.* **49** (2001) 444–454.
[19] R. Lahdelma, K. Miettinen and P. Salminen, Ordinal criteria in stochastic multicriteria acceptability analysis (SMAA). *Eur. J. Oper. Res.* **147** (2003) 117–127.
[20] A. Rebai, Bbtopsis: a bag based technique for order preference by similarity to ideal solution. *Fuzzy Sets Syst.* **60** (1993) 143–162.
[21] T.L. Saaty, How to make a decision: the analytic hierarchy process. *Eur. J. Oper. Res.* **48** (1990) 9–26.
[22] T.L. Saaty, Decision making – the analytic hierarchy and network processes (AHP/ANP). *J. Syst. Sci. Syst. Eng.* **13** (2004) 1–35.
[23] L. Simar and P.W. Wilson, Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models. *Manage. Sci.* **44** (1998) 49–61.
[24] T. Tervonen and R. Lahdelma, Implementing stochastic multicriteria acceptability analysis. *Eur. J. Oper. Res.* **178** (2007) 500–513.
[25] D. Valencia, R.E. Lillo and J. Romo, A Kendall correlation coefficient between functional data. *Adv. Data Anal. Classification* **13** (2019) 1083–1103.
[26] Y.M. Wang, R. Greatbanks and J.B. Yang, Interval efficiency assessment using data envelopment analysis. *Fuzzy Sets Syst.* **153** (2005) 347–370.
[27] R.R. Yager, On ordered weighted averaging aggregation operations in multicriteria decision making. *IEEE Trans. Syst. Man Cybern.* **18** (1988) 80–87.
[28] G.T. Yeo, A.K.Y. Ng, T.W. Lee and Z. Yang, Modelling port choice in an uncertain environment. *Maritime Policy Manage.* **41** (2014) 251–267.
[29] F. Zhu, P.A. Zhong and Y. Sun, Multi-criteria group decision making under uncertainty: application in reservoir flood control operation. *Environ. Modell. Softw.* **100** (2018) 236–251.