Islamic Azad Univer Sity Firuzkuh Branch
Journal Of Industrial Strategic Management
https://sanad.iau.ir/journal/mgmt Volume 13, Issue 2, Summer 2025, 1-15

Research Article

Robust Optimization FDH Model for Data Envelopment Analysis in Education

Nasrollah Keshiri[1]

1. Ph.D. Student of Industrial Management, Production and Operations Management ,Fi.C., Islamic Azad University, Firoozkooh, Iran.
Email: n_kashiri@yahoo.com

ABSTRACT

DOR: https://doi.org/10.71584/MGMT.2025.54139

Article Info:

Received Date:
2025/07/29
Accepted Date:
2025/12/05

Published Online:
2025/12/14

Abstract:

Traditional Data Envelopment Analysis (DEA) models generally rely on the assumption that all input and output data are deterministic and known with certainty. However, in many real-world scenarios, data uncertainty and measurement errors are unavoidable, which can significantly affect the reliability of efficiency evaluations. This study addresses this limitation by developing a robust version of the Free Disposal Hull (FDH) model, a non-convex DEA variant that does not require the convexity assumption. The proposed approach integrates robust optimization—a powerful methodology in operations research for managing uncertainty—into the FDH framework. Initially, an equivalent linear programming formulation of the FDH model is constructed to facilitate the application of robust optimization techniques. To improve the model's discrimination ability and to mitigate the occurrence of zero weights in the input and output variables, a method for determining appropriate lower bounds for weights is introduced. Subsequently, the robust counterpart of the FDH model is derived, incorporating uncertainty directly into the efficiency assessment. A practical case study is conducted to demonstrate the applicability and effectiveness of the proposed robust FDH model in real-world performance evaluation scenarios. The results confirm that the model maintains computational tractability while offering enhanced reliability in environments characterized by data ambiguity.

Keywords:
Performance Evaluation, Robust Optimization, Free Disposal Hull (FDH), Data Envelopment Analysis (DEA), Uncertainty

How to Cite: Keshiri, N., (2025). Robust Optimization FDH Model for Data Envelopment Analysis in Education, *Journal of Industrial Strategic Management*, 13(2): 1-15.

1. Introduction

Data Envelopment Analysis (DEA) is a non-parametric programming method that was initially introduced by Charnes et al. (1978) and later extended by Banker et al. (1984). It is used to evaluate the relative efficiency of decision-making units (DMUs) based on deterministic input and output data. Essentially, the DEA model involves a fractional linear programming problem, which can be transformed into a linear programming problem under the assumption of convexity and constant returns to scale (CRS).

Another non-parametric and deterministic model useful for evaluating the efficiency of DMUs, which does not impose the convexity assumption and is based on the principle of free disposability, was introduced by Deprins et al. (1984) and is known as the Free Disposal Hull (FDH) model. Subsequently, Tulkens (1993) proposed an integer linear programming formulation for solving the FDH model. Additionally, Kerstens and Eeckaut (1999) introduced different types of returns to scale for the FDH model, including non-increasing, non-decreasing, and constant returns to scale. In another study, Agrell and Tind (2001) proposed a linear programming formulation for the FDH model, which was later extended by Leleu (2006) to account for variable returns to scale using a cost function.

Both DEA and FDH models require input and output data to be precisely known. However, in many cases, the observed data are imprecise and subject to uncertainty. In this context, Hougaard (1999) outlined several reasons why traditional DEA models may not be suitable for modeling real-world problems and assessing their performance. One of the key reasons mentioned is that both the efficiency scores and the efficient frontier are highly sensitive to fluctuations and noise in the data, indicating that caution must be exercised when applying such models to uncertain environments. Similar limitations also apply to FDH models.

2. Literature Review and Research Background

Various approaches have been proposed to address DEA models under imprecise data conditions. Among these are interval-based methods (Cooper, Park, & Yu, 1999), fuzzy methods (Hatami-Marbini et al., 2011), stochastic approaches (Olesen & Petersen, 2016), and robust optimization techniques. In this study, we adopt the robust optimization approach to extend the DEA models, and thus, we review several key studies related to this method.

Before delving into robust optimization, it is useful to briefly highlight some of the limitations of the other approaches, namely interval, stochastic, and fuzzy methods. The primary shortcoming of interval-based approaches is that they only provide efficiency scores for the best-case and worst-case scenarios of inputs and outputs. Stochastic DEA methods allow deterministic data to be replaced by statistical and probabilistic values. In such approaches, the uncertainty in inputs and outputs is modeled through probability distributions, resulting in quadratic or nonlinear programming models. Fuzzy DEA methods, on the other hand, utilize membership functions to handle uncertainty and, similar to interval methods, provide upper and lower bounds for relative efficiency scores.

As previously stated, the focus of this study is on the robust optimization approach. This method first considers a level of uncertainty in problem parameters and then estimates a robust solution accordingly. The robust optimization method was first introduced by Soyster (1973), but the solutions it produced were overly conservative. The approach was later developed by Ben-Tal and Nemirovski (1998, 1999, 2000)—referred to hereafter as BN—who proposed replacing the original linear programming problem with its robust counterpart. Their method guarantees feasibility of the robust solution when the problem parameters follow a normal distribution, and is more flexible than Soyster's method, although the robust counterpart becomes a second-order cone program (SOCP).

More recently, Bertsimas and Sim (2003, 2004)—hereafter referred to as BS—and Bertsimas et al. (2004) introduced a new approach that allows the decision-maker to control the level of constraint protection against uncertainty and adjust the conservatism of the solution. This method is highly flexible in handling uncertainty while preserving linearity in the robust counterpart, which is particularly valuable in real-world applications.

Since its introduction, the robust optimization approach has been applied in various operations research domains, including facility location, production planning, supply chain management, and data envelopment analysis. In what follows, we review a number of significant studies that applied robust optimization to handle uncertainty in DEA models.

Sadjadi and Omrani (2008) were the first to use the robust optimization approach to evaluate the performance of electric power distribution companies in Iran, applying the multiplier form of DEA and incorporating output uncertainty. They employed both BN and BS approaches in their model development. Shokouhi et al. (2010) used the BS robustification approach under interval data uncertainty to propose a unified model that captures both optimistic and pessimistic DEA efficiency bounds using a tuning parameter. Sadjadi et al. (2011), using the BN method and the envelopment form of DEA, proposed a robust super-efficiency model to assess the performance of provincial gas companies in Iran. Omrani (2013) applied the BS robust optimization approach to the model proposed by Zohrehbandian et al. (2010), which aimed at obtaining common weights, and developed a robust model that simultaneously considered uncertainty in both inputs and outputs.

Arabmaldar et al. (2017), using the equivalent multiplier form introduced by Toloo (2014)—which replaces the equality normalization constraint with an inequality—proposed a robust counterpart of the multiplier form assuming constant returns to scale and uncertain inputs and outputs. They also presented a robust super-efficiency model using the envelopment form and employed the BS approach in their model. Toloo and Mensah (2019), assuming variable returns to scale and accounting for data uncertainty, developed a robust DEA model for evaluating the performance of 50 European banks, with emphasis on the non-negativity of decision variables. More recently, Salahi et al. (2020) proposed a robust model for deriving common weights in the presence of data uncertainty. Peykani et al. (2020) conducted a comprehensive review of 73 studies involving robust optimization in DEA models.

To the best of the authors' knowledge, only a limited number of studies have addressed the FDH model under uncertainty. Jahanshahloo et al. (2004) presented upper and lower bounds for performance evaluation of DMUs using the FDH model with interval uncertainty in input and output data. In another study, Khanjani Shiraz et al. (2014) proposed a model using chance-constrained programming, assuming fuzzy input and output data in the FDH framework. Later, Khanjani et al. (2016) proposed a model for cost efficiency evaluation by considering hybrid fuzzy-rough data uncertainty.

As the above review indicates, research on FDH models under uncertainty is still quite limited. This gap in the literature has motivated the present study. The key contributions of this paper can be summarized as follows:

1. Proposing a linear equivalent formulation of the FDH model by converting the normalization equality constraint into an inequality;
2. Introducing a method to determine a lower bound for the weights associated with inputs and outputs;
   And
3. Developing a robust FDH model using the robust optimization approach to evaluate the performance of decision-making units under data uncertainty.

The structure of this paper is as follows: Section 2 presents the foundational methods and the proposed models. In Section 3, a real-world case study is introduced and the performance of the given problem is evaluated using the developed model. Section 4 provides detailed analysis and discussion of the results.

3. Research Methodology

In this section, the main objective of the present study is to propose a robust counterpart of the Free Disposal Hull (FDH) model, which is capable of analyzing and accounting for the impact of uncertainty in input and output data. Since the FDH model is an integer programming model, we begin by reviewing its equivalent linear programming formulation and its dual counterpart, as presented by Agrell and Tind (2001) and Jahanshahloo et al. (2004), respectively.

Next, to incorporate uncertainty conditions into the FDH model, we introduce an equivalent formulation that facilitates this extension. Moreover, to ensure that both input and output weights are adequately considered in evaluating the efficiency of decision-making units (DMUs), we propose a method to determine appropriate lower bounds for these weights. Finally, by applying the robust optimization approach developed by Bertsimas and Sim (2004), we develop the robust counterpart of the FDH model.

3-1. The FDH Model and Its Linear Formulation

Suppose there are nn decision-making units (DMUs), where the jthj^{th} unit, denoted by DMUj\text{DMU}_j, has mm inputs and ss outputs represented as Xj=(x1j,…,xmj)TX_j = (x_{1j}, \ldots,

$$
\begin{aligned}
&\min \theta - \varepsilon\,(\mathbf{1}s_i^+ + \mathbf{1}s_r^-) \\
&\text{s.t.} \\
&\textstyle\sum_{j=1}^{n} x_{ij}\lambda_j + s_i^+ = \theta x_{io} && i = 1 \ldots m \\
&\textstyle\sum_{j=1}^{n} y_{rj}\lambda_j - s_r^- = y_{ro} && r = 1 \ldots s \\
&\textstyle\sum_{j=1}^{n} \lambda_j = 1 \\
&\lambda_j \in \{01\} && j = 1 \ldots . n \\
&s_i^+ \;\; s_r^- \geq 0 && \forall i, \forall r, \\
&\theta \;\; \text{free}
\end{aligned} \tag{1}
$$

where ε>0 is a non-Archimedean constant and 1 is a row vector with all its components equal to one. It should be noted that model (1) is a mixed integer programming problem whose linear equivalent form is given by Agrell & Tind, (2001) in the input nature as follows:

$$
\begin{aligned}
&\min \textstyle\sum_{j=1}^{n} \theta_j - \varepsilon \sum_{j=1}^{n}\left(\mathbf{1}s_j^+ + \mathbf{1}s_j^-\right) \\
&\text{s.t.} \\
&\textstyle\sum_{j=1}^{n} x_{ij}\lambda_j + s_j^- = \theta_j x_{io} && j = 1 \ldots n \\
&\textstyle\sum_{j=1}^{n} (y_{rj} - y_{ro})\lambda_j + s_j^+ = 0 && j = 1 \ldots n \\
&\textstyle\sum_{j=1}^{n} \lambda_j = 1 \\
&\lambda_j \geq 0 && j = 1 \ldots . n \\
&s_j^- \;\; s_j^+ \geq 0 && j = 1 \ldots n \\
&\theta_j \;\; \text{free} && j = 1 \ldots n
\end{aligned} \tag{2}
$$

e dual of model (2) is obtained as follows:

$$\begin{aligned}
&\max \quad z\,,\\
&\text{s.t.}\\
&\textstyle\sum_{i=1}^{m} x_{io} v_{ij} = 1\,, && j = 1,\dots,n,\\
&\textstyle\sum_{r=1}^{s} (y_{rj} - y_{ro}) u_{rj} - \sum_{i=1}^{m} x_{ij} v_{ij} + z \le 0\,, && j = 1,\dots,n,\\
&u_{rj}, v_{ij} \ge \varepsilon && \forall i, r, j.
\end{aligned} \tag{3}$$

As discussed in Jahanshahloo et al. (2004) $z^{*} = \min_j\{\sum_{i=1}^{m} x_{ij} v_{ij} - \sum_{r=1}^{s} (y_{rj} - y_{ro}) u_{rj}\} \le 1$

Moreover, model (4) is linear and, in large-scale problems, variable relaxation techniques can bemployed to solve it (Bazaraa, Jarvis, & Sherali, 2009).

Certainly! Here's the precise English translation of your paragraph, tailored for academic writing in operations research and robust optimization:

FDH model for considering uncertain data

Let x_ij and y_rj be nominal values for uncertain inputs and outputs of each DMU_j and let x ̂_ij and y ̂_rj be the deviations from the nominal inputs and outputs, respectively. In robust optimization, often uncertain data are defined as symmetric and unequal intervals. Therefore, equality constraints that have uncertainty in their data need to be unequal for proper analysis because equality constraints can restrict the feasible region and sometimes make it unfeasible (Ben-Tal A. et al., 2009; Gorissen, et al., 2015). Therefore, since in model (3) the first constraint, namely $\sum_{i=1}^{m} x_{io} v_{ij} = 1$, is equality, we present a method to convert this equality constraint to inequality. For this purpose, suppose we consider the normalizer constraint in model (3) with any positive parameter t as follows:

$$\begin{aligned}
&\max \quad z\,,\\
&\text{s.t.}\\
&\textstyle\sum_{i=1}^{m} x_{io} v_{ij} = t\,, && j = 1,\dots,n,\\
&\textstyle\sum_{r=1}^{s} (y_{rj} - y_{ro}) u_{rj} - \sum_{i=1}^{m} x_{ij} v_{ij} + z \le 0\,, && j = 1,\dots,n,\\
&u_{rj}, v_{ij} \ge \varepsilon && \forall i, r, j.
\end{aligned} \tag{4}$$

It is easy to check that (u^*,v^*,z^*) is the optimal solution of model (3) if and only if (tu^*,tv^*,tz^*) is the optimal solution of model (4).

Theorem 1. The following model is equivalent to the FDH and linear model (3).

Theorem 1.

The following model is equivalent to the original FDH model and its linear formulation (Model 3).

$$\begin{aligned}
&\max \quad z\,,\\
&\text{s.t.}\\
&\textstyle\sum_{i=1}^{m} x_{io} v_{ij} \le t\,, && j = 1,\dots,n,\\
&\textstyle\sum_{r=1}^{s} (y_{rj} - y_{ro}) u_{rj} - \sum_{i=1}^{m} x_{ij} v_{ij} + z \le 0\,, && j = 1,\dots,n,\\
&u_{rj}, v_{ij} \ge \varepsilon && \forall i, r, j.
\end{aligned} \tag{5}$$

Proof. According to the primal-dual relationship, to demonstrate that Models (4) and (5) are equivalent, it is sufficient to compare their corresponding dual formulations, which are derived as follows

$$
\begin{aligned}
&\min \textstyle\sum_{j=1}^{n} t\theta_j - \varepsilon\ \sum_{j=1}^{n}(\mathbf{1}s_j^{+} + \mathbf{1}s_j^{-}),\\
&\text{s.t.}\\
&\textstyle\sum_{j=1}^{n} x_{ij}\lambda_j + s_j^{-} = \theta_j x_{io}, && j = 1, \ldots, n,\\
&\textstyle\sum_{j=1}^{n}(y_{rj} - y_{ro})\lambda_j + s_j^{+} = 0, && j = 1, \ldots, n,\\
&\textstyle\sum_{j=1}^{n} \lambda_j = 1,\\
&\lambda_j \geq 0, && j = 1, \ldots., n,\\
&s_j^{-},\ s_j^{+} \geq 0, && j = 1, \ldots, n,\\
&\theta_j\ \text{free}, && j = 1, \ldots, n,
\end{aligned}
\tag{6}
$$

and

$$
\begin{aligned}
&\min \textstyle\sum_{j=1}^{n} t\theta_j - \varepsilon\ \sum_{j=1}^{n}(\mathbf{1}s_j^{+} + \mathbf{1}s_j^{-}),\\
&\text{s.t.}\\
&\textstyle\sum_{j=1}^{n} x_{ij}\lambda_j + s_j^{-} = \theta_j x_{io}, && j = 1, \ldots, n,\\
&\textstyle\sum_{j=1}^{n}(y_{rj} - y_{ro})\lambda_j + s_j^{+} = 0, && j = 1, \ldots, n,\\
&\textstyle\sum_{j=1}^{n} \lambda_j = 1,\\
&\lambda_j \geq 0, && j = 1, \ldots., n,\\
&s_j^{-},\ s_j^{+} \geq 0, && j = 1, \ldots, n,\\
&\theta_j \geq 0, && j = 1, \ldots, n,
\end{aligned}
\tag{7}
$$

Assume that θ_j^* is the optimal solution for each j in the dual models. By carefully examining the constraints of model (6), it can be easily shown that θ_j^* ≥ 0 for all j. This is because if θ_j^* < 0 for some j, it would imply that λ_j^* < 0 for that j, violating the convexity constraint ∑_(j=1)^n λ_j = 1, which is impossible. Therefore, models (6) and (7) are equivalent, and their corresponding primal models are also equivalent.

It is worth noting that in this paper, without loss of generality, we set the parameter t equal to 1,and., t = 1. This approach provides a method to analyze the uncertainty in the inputs and outputs of the FDH model..

3-2. A Model for Determining the Maximum Value of the Non-Archimedean ε

One of the main challenges in DEA models is enhancing the discrimination power among efficient units, in which the non-Archimedean ε plays a significant role. The maximum value of the non-Archimedean ε, denoted by ε*, is a positive number such that the model remains feasible for any ε ∈ [0, ε*], and becomes infeasible for any ε > ε*. The interval (0, ε*] and the set ε ∈ [0, ε*] are referred to as the *confidence interval* and the *confidence value* for the non-Archimedean ε, respectively.

On one hand, selecting a small value for ε guarantees the feasibility of DEA models (Amin & Toloo, 2004). On the other hand, choosing a larger value can enhance the discrimination power of DEA models in identifying efficient units (Cook, Kress, & Seiford, 1996). Therefore, the optimal choice of ε in equality-based DEA models is the maximum confidence value, or equivalently, the maximum non-Archimedean ε.

The following model determines the maximum non-Archimedean ε for model (5), assuming t =1:

$$
\begin{aligned}
&\varepsilon^* = \max\ \varepsilon\,,\\
&\text{s.t.}\\
&\textstyle\sum_{i=1}^{m} x_{ij}v_{ij} \le 1\,,\quad j = 1,\dots,n,\\
&\textstyle\sum_{r=1}^{s}(y_{rj} - y_{ro})u_{rj} - \sum_{i=1}^{m} x_{ij}v_{ij} + z \le 0\,,\quad j = 1,\dots,n,\\
&\varepsilon - u_{rj} \le 0,\ \ \forall r,j,\\
&\varepsilon - v_{ij} \le 0,\ \ \forall i,j,\\
&\varepsilon \ge 0
\end{aligned}
\tag{8}
$$

where ε is a nonnegative decision variable.:

Robust FDH Model Presentation

Model (3) represents a case where inputs and outputs are deterministic, meaning it uses nominal and precise data. Therefore, when the inputs and outputs are uncertain, model (3) is not suitable for analyzing the impact of uncertainty in FDH models. In this section, we apply robust optimization methods to extend the FDH model to its robust counterpart for handling imprecise data.

A common technique for presenting the robust counterpart of DEA models is to consider the worst-case scenario for uncertain inputs and outputs, allowing simultaneous investigation of both the level of uncertainty in the problem parameters and the efficiency scores. Suppose that the uncertain input and output variables are expressed as follows:

$$
\begin{aligned}
&\max\quad z\,,\\
&\text{s.t.}\\
&\textstyle\sum_{i=1}^{m} \tilde{x}_{io}v_{ij} \le 1\,, && j = 1,\dots,n,\\
&\textstyle\sum_{r=1}^{s}(\tilde{y}_{rj} - \tilde{y}_{ro})u_{rj} - \sum_{i=1}^{m} \tilde{x}_{ij}v_{ij} + z \le 0\,, && j = 1,\dots,n, j \ne o\\
&z \le 1,\\
&u_{rj}, v_{ij} \ge \varepsilon && \forall i,r,j,\\
&\forall(\tilde{x},\tilde{y}) \in \mathcal{U}
\end{aligned}
\tag{9}
$$

Theorem 3. The robust counterpart of the FDH model (9), under the uncertainty defined above, can be obtained as the following linear model:

$$
\begin{aligned}
&\max\quad z\,,\\
&\text{s.t.}\\
&\textstyle\sum_{i=1}^{m} x_{io}v_{ij} + p_j^x\Gamma_j^x + \sum_{i=1}^{m} q_{ij}^x \le 1\,, && \forall j\\
&\textstyle\sum_{r=1}^{s}(y_{rj} - y_{ro})u_{rj} - \sum_{i=1}^{m} x_{ij}v_{ij} + z + p_j^y\Gamma_j^y + p_j^x\Gamma_j^x + \sum_{r=1}^{s} q_{rj}^y + \sum_{i=1}^{m} q_{ij}^x \le 0\,, && \forall j \ne o\\
&z \le 1\\
&p_j^y + q_{rj}^y \ge (\hat{y}_{rj} - \hat{y}_{ro})u_{rj} && \forall j, \forall r\\
&p_j^x + q_{ij}^x \ge \hat{x}_{ij}v_{ij} && \forall j, \forall i\\
&p_j^y, p_j^x \ge 0 && \forall j\\
&q_{rj}^y, q_{ij}^x \ge 0 && \forall i,r,j,\\
&u_{rj}, v_{ij} \ge \varepsilon && \forall i,r,j.
\end{aligned}
\tag{10}
$$

Proof. By applying the uncertainty defined at the beginning of this section for$\tilde{x}_{ij}, \tilde{y}_{rj}, \forall i, r$ model (9) can be reformulated as follows

$$\max\ z,$$
$$\text{s.t.}$$
$$\sum_{i=1}^{m}(x_{io}+\eta_{io}^{x}\hat{x}_{io})v_{ij}\leq 1,\ \ \forall j,$$
$$\sum_{r=1}^{s}(y_{rj}+\eta_{rj}^{y}\hat{y}_{rj})u_{rj}-\sum_{r=1}^{s}(y_{ro}+\eta_{ro}^{y}\hat{y}_{ro})u_{rj}-\sum_{i=1}^{m}(x_{ij}+\eta_{ij}^{x}\hat{x}_{ij})v_{ij}+z\leq 0,\ \ \forall j\neq o, \tag{11}$$
$$z\leq 1,$$
$$u_{rj},v_{ij}\geq\varepsilon,\ \forall i,r,j.$$

By performing the necessary algebraic manipulations on the constraints of model (11), the following model can be readily obtained:

$$\max\ z,$$
$$\text{s.t.}$$
$$\sum_{i=1}^{m}x_{io}v_{ij}+\sum_{i=1}^{m}\eta_{io}^{x}\hat{x}_{io}v_{ij}\leq 1,\ \ \forall j,$$
$$\sum_{r=1}^{s}y_{rj}u_{rj}+\sum_{r=r}^{s}\eta_{rj}^{y}\hat{y}_{rj}u_{rj}-\sum_{r=1}^{s}y_{ro}u_{rj}-\sum_{r=1}^{s}\eta_{ro}^{y}\hat{y}_{ro}u_{rj}-\sum_{i=1}^{m}x_{ij}v_{ij}-\sum_{i=1}^{m}\eta_{ij}^{x}\hat{x}_{ij}v_{ij} \tag{12}$$
$$z\leq 1,$$
$$u_{rj},v_{ij}\geq\varepsilon,\ \forall i,r,j.$$

Considering that in the robust optimization approach proposed by Bertsimas and Sim (2004), the constraints of the model under study must hold under the worst-case realizations of uncertain input and output parameters, and taking into account the properties of maximization and minimization functions, the following model can be derived:

$$\max\ z,$$
$$\text{s.t.}$$
$$\sum_{i=1}^{m}x_{io}v_{ij}+\max_{\eta_{io}^{x}}\sum_{i=1}^{m}\eta_{io}^{x}\hat{x}_{io}v_{ij}\leq 1,\ \ \forall j,$$
$$\sum_{r=1}^{s}(y_{rj}-y_{ro})u_{rj}-\sum_{i=1}^{m}x_{ij}v_{ij}+z+$$
$$\max_{\eta_{rj}^{y}}\sum_{r=r}^{s}\eta_{rj}^{y}\hat{y}_{rj}u_{rj}+\max_{\eta_{ro}^{y}}\sum_{r=1}^{s}\eta_{ro}^{y}\hat{y}_{ro}u_{rj}+\max_{\eta_{ij}^{x}}\sum_{i=1}^{m}\eta_{ij}^{x}\hat{x}_{ij}v_{ij}\leq 0,\ \forall j,\ j\neq o \tag{13}$$
$$z\leq 1,$$
$$u_{rj},v_{ij}\geq\varepsilon,\ \forall i,r,j.$$

$\sum_{i\in I_j}|\eta_{ij}^{x}|\leq\Gamma_j^{x}, -1\leq\eta_{ij}^{x}\leq 1$ و $\sum_{r\in R_j}|\eta_{rj}^{y}|\leq\Gamma_j^{y}, -1\leq\eta_{rj}^{y}\leq 1$.where

It should be noted that the constraints in model (13) involve a maximization problem, and similar to standard optimization problems, they contain their own objective function, constraints, and decision variables. However, some decision variables of the outer problem are treated as fixed parameters within the inner problems.

In other words, the optimal value of the objective function in the inner problem constitutes a part of the corresponding constraint in the overall (outer) problem. Therefore, in what follows, we aim to derive an explicit form of these inner problems to complete the proof of the theorem.

$$\max\ \ \sum_{r=1}^{s}\eta_{rj}^{y}(\hat{y}_{rj}-\hat{y}_{ro})u_{rj}+\sum_{i=1}^{m}\eta_{ij}^{x}\hat{x}_{ij}v_{ij},$$
$$\text{s.t.}$$
$$\sum_{r=1}^{s}\eta_{rj}^{y}\leq\Gamma_j^{y},\ \ \forall j$$
$$\sum_{i=1}^{m}\eta_{ij}^{x}\leq\Gamma_j^{x},\ \ \forall j \tag{14}$$
$$0\leq\eta_{rj}^{y}\leq 1,\ \ \ \forall r,\forall j$$
$$0\leq\eta_{ij}^{x}\leq 1,\ \ \forall i,\forall j$$

In which u_{rj} and v_{ij} are parameters, and η^{x}_{ij} و η^{y}_{rj} are decision variables. Consider the dual of problem (14) as follows

$$\begin{aligned}
&\min\ p^{y}_{j}\Gamma^{y}_{j} + p^{x}_{j}\Gamma^{x}_{j} + \textstyle\sum_{r=1}^{s} q^{y}_{rj} + \sum_{i=1}^{m} q^{x}_{ij} \\
&\text{s.t.} \\
&p^{y}_{j} + q^{y}_{rj} \geq (\hat{y}_{rj} - \hat{y}_{ro})\, u_{rj} && \forall j, \forall r \\
&p^{x}_{j} + q^{x}_{ij} \geq \hat{x}_{ij} v_{ij} && \forall j, \forall i \\
&p^{y}_{j}, p^{x}_{j} \geq 0 && j = 1, \dots, n \\
&q^{y}_{rj}, q^{x}_{ij} \geq 0 && \forall j, \forall i, \forall r \\
&u_{rj}, v_{ij} \geq \varepsilon && \forall j, \forall i, \forall r
\end{aligned} \tag{15}$$

this context $p^{x}_{j}, p^{y}_{j} \in \mathbb{R}$ and $q^{x}_{ij}, q^{y}_{rj} \in \mathbb{R}^{m+s}$ denote the dual variables corresponding to the first to fourth constraints of model (14).

Similarly, it is possible to derive an equivalent constraint corresponding to the first constraint in model (13).

By substituting the equivalent constraints obtained in (15) into the constraint set of model (13), model (10) is derived. Hence, the proof is complete. □

Numerical Example

To evaluate and analyze the proposed model in this paper, the performance of 33 universities in Iran during the year 2004 (1383 in the Iranian calendar) has been considered (Khodabakhshi & Kheirollahi, 2013). Table 1 presents the inputs and outputs of the problem under study, which include 7 inputs and 5 outputs.

The proposed robust model has been applied to rank the 33 universities to demonstrate its practical applicability. The results of the deterministic model (3) and its robust counterpart, model (10), are shown in Table 2, which includes efficiency scores and rankings. A university is considered efficient and lies on the FDH efficiency frontier if its efficiency score equals one. The efficient universities are indicated in the second column of Table 2. Using the deterministic FDH model, 25 universities are identified as efficient. It should be noted that the efficiency scores decrease when applying the robust models.

Also, note that in solving models (3) and (10), the value of ε\varepsilonε is first obtained using the proposed model, which is 0.0154

Table 1. Inputs and Outputs of One of the Universities in Iran

oput					**input**							**DMU**
O5	**O4**	**O3**	**O2**	**O1**	**I7**	**I6**	**I5**	**I4**	**I3**	**I2**	**I1**	
0.88	0.18	0.02	0.09	0.28	0.1	0.23	0.6	0.67	0.46	0.27	0.16	1
0.89	0.39	0.23	0.2	0.55	0.31	0.35	0.86	0.48	0.53	0.38	0.29	2
0.99	0.14	0.04	0.09	0.24	0.13	0.18	0.35	0.66	0.34	0.3	0.17	3
0.9	0.2	0.09	0.09	0.29	0.09	0.19	0.39	0.66	0.45	0.3	0.16	4
0.9	0.47	0.05	0.15	0.47	0.33	0.42	0.81	0.63	0.49	0.37	0.27	5
0.91	1	1	1	1	1	1	0.59	0.9	0.99	1	1	6
0.88	0.2	0.07	0.1	0.31	0.09	0.19	0.46	0.57	0.47	0.27	0.13	7
0.91	0.17	0.07	0.05	0.34	0.15	0.34	0.84	0.42	0.72	0.32	0.12	8
0.9	0.29	0.05	0.13	0.47	0.23	0.3	0.53	0.56	0.68	0.53	0.22	9

oput					input							DMU
O5	**O4**	**O3**	**O2**	**O1**	**I7**	**I6**	**I5**	**I4**	**I3**	**I2**	**I1**	
0.92	0.36	0.3	0.19	0.43	0.4	0.39	0.32	0.36	0.39	0.34	0.33	10
0.92	0.3	0.14	0.16	0.4	0.29	0.43	0.6	0.24	0.62	0.38	0.25	11
0.91	0.49	0.11	0.54	0.56	0.51	0.58	0.68	0.77	0.5	0.41	0.34	12
0.93	0.51	0.26	0.21	0.64	0.4	0.49	0.52	0.71	0.75	0.53	0.36	13
0.9	0.24	0.07	0.11	0.39	0.19	0.24	0.53	0.46	0.36	0.24	0.17	14
0.9	0.26	0.05	0.08	0.42	0.17	0.22	0.73	0.4	0.56	0.36	0.17	15
0.92	0.21	0.06	0.05	0.36	0.14	0.17	0.48	0.9	0.37	0.23	0.15	16
0.86	0.09	0.02	0.03	0.13	0.16	0.07	0.38	0.33	0.23	0.15	0.06	17
0.91	0.04	0.02	0.01	0.09	0.04	0.09	0.51	0.32	0.2	0.11	0.06	18
0.92	0.09	0.02	0.02	0.24	0.04	0.12	0.59	0.57	0.24	0.19	0.09	19
0.9	0.11	0.08	0.03	0.12	0.06	0.11	0.5	0.45	0.23	0.17	0.05	20
0.92	0.06	0.01	0.02	0.11	0.04	0.09	0.66	0.75	0.2	0.09	0.05	21
0.93	0.04	0.02	0.01	0.2	0.09	0.14	0.19	0.16	0.69	0.34	0.1	22
0.9	0.1	0.02	0.05	0.12	0.07	0.12	0.37	0.57	0.27	0.2	0.09	23
0.9	0.07	0.04	0.02	0.11	0.04	0.1	0.51	0.5	0.21	0.14	0.06	24
0.92	0.05	0.01	0.05	0.09	0.04	0.1	0.89	0.89	0.17	0.13	0.07	25
0.96	0.05	0.04	0.02	0.09	0.03	0.07	0.54	0.44	0.16	0.14	0.06	26
0.92	0.12	0.03	0.05	0.16	0.06	0.11	0.35	0.36	0.22	0.12	0.07	27
0.88	0.08	0.03	0.03	0.17	0.04	0.1	0.62	0.43	0.23	0.14	0.07	28
0.92	0.08	0.03	0.04	0.12	0.07	0.1	0.29	0.38	0.42	0.2	0.8	29
0.9	0.07	0.07	0.03	0.18	0.04	0.08	0.64	0.7	0.2	0.15	0.6	30
0.92	0.04	0.02	0.01	0.06	0.03	0.06	0.36	0.24	0.18	0.07	0.4	31
0.93	0.03	0.03	0.01	0.08	0.05	0.08	0.63	0.59	0.13	0.07	0.03	32
0.9	0.04	0.01	0.02	0.09	0.04	0.06	0.39	0.32	0.17	0.11	0.04	33

Using the robust FDH model with data subject to uncertainty, it should be noted that even a small amount of disturbance can significantly affect the feasibility and complexity of solving the problem. The optimal solution tends to decrease as the level of uncertainty in the inputs and outputs increases. Therefore, the efficiency scores obtained from the robust model are generally

lower than those from the conventional model (see Figure 1).

Table 2. Efficiency Scores and Sensitivity Analysis of the Models

Model (10)				Model (3)		
$\Gamma_x = \Gamma_y = 5, e = 0.1$		$\Gamma_x = \Gamma_y = 5, e = 0.05$				
Grade	Efficiency	Grade	Efficiency	Grade	Efficiency	NO
32	0/8170	32	0/8600	30	0/9094	1
9	0/9152	9	0/9634	1	1	2
12	0/9035	12	0/9511	1	1	3

Model (10)				Model (3)		
22	0/8750	22	0/9211	1	1	4
20	0/8819	20	0/9283	1	1	5
1	0/9363	1	0/9856	1	1	6
13	0/9031	13	0/9506	1	1	7
26	0/8560	26	0/9010	27	0/9871	8
31	0/8294	31	0/8731	32	0/9031	9
8	0/9200	8	0/9684	1	1	10
4	0/9276	4	0/9764	1	1	11
5	0/9258	5	0/9745	1	1	12
14	0/9013	14	0/9487	1	1	13
25	0/8572	25	0/9023	26	0/9983	14
19	0/8848	19	0/9314	1	1	15
17	0/8927	17	0/9397	1	1	16
18	0/8898	18	0/9366	1	1	17
29	0/8382	29	0/8823	31	0/9054	18
15	0/8992	15	0/9456	1	1	19
2	0/9344	2	0/9836	1	1	20
24	0/8697	24	0/9155	1	1	21
3	0/9341	3	0/9833	1	1	22
33	0/7635	33	0/8037	33	0/8782	23
28	0/8466	28	0/8912	29	0/9253	24
27	0/8517	27	0/8965	28	0/9495	25
21	0/8798	21	0/9261	1	1	26
10	0/9114	10	0/9594	1	1	27
23	0/8725	23	0/9184	1	1	28
30	0/8322	30	0/8760	1	1	29
11	0/9047	11	0/9523	1	1	30
6	0/9240	6	0/9726	1	1	31
7	0/9202	7	0/9686	1	1	32
16	0/8952	16	0/9423	1	1	33

In this study, the robust efficiency results are calculated for an equal level of protection against input and output uncertainty, with $\Gamma x = \Gamma y = 5$ and a disturbance level of $\varepsilon = 0.05$. As previously mentioned, and as shown in Table 2 and Figure 1, the efficiency scores decrease as data disturbance increases. Consequently, the number of efficient units also declines, and only a few universities maintain efficiency scores close to one when the disturbance level increases from 0.05 to 0.10. The average efficiency derived from model (10) for disturbance levels of 0.05 and 0.10 is 0.9312 and 0.8845, respectively. This indicates to educational managers that as the level of uncertainty in the data increases, a significant cost must be incurred to manage and mitigate its impact.

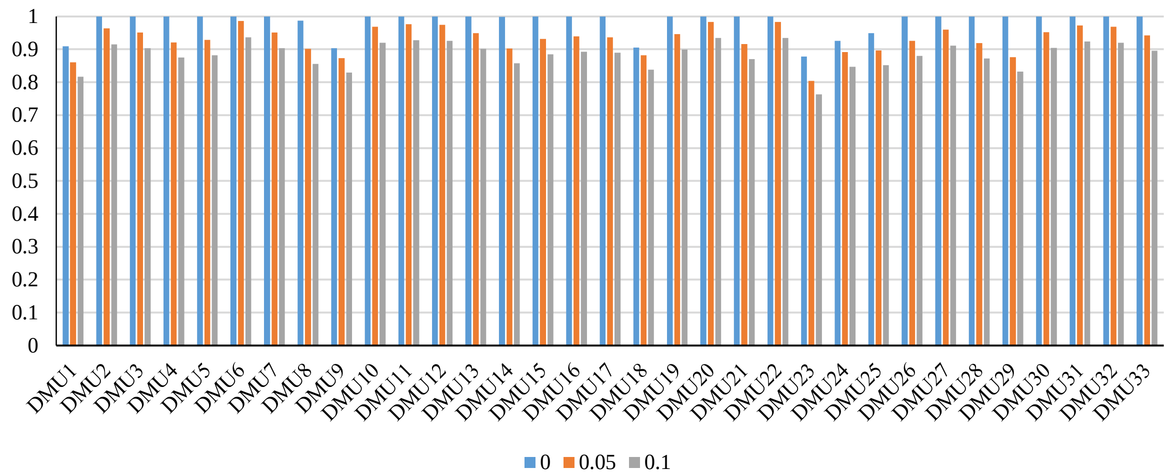

Figure 1. Comparison of robust performance under different levels of disturbance"

Results

Robust optimization is a powerful approach in operations research that is widely used to obtain optimal solutions for problems under uncertainty. In the context of Data Envelopment Analysis (DEA), investigating the impact of uncertainty on model efficiency and performing sensitivity analysis are of critical importance. This study focuses on developing a robust version of the Free Disposal Hull (FDH) model, which is known for its non-convex and non-linear structure and is often used in efficiency analysis when the convexity assumption is not justified.

To achieve this goal, we first formulated the classical FDH model as a linear programming problem. We demonstrated that the normalization constraint in the FDH model, traditionally expressed as an equality, can be relaxed into an inequality. This relaxation allows for the incorporation of uncertainty into the model formulation, which is a prerequisite for applying robust optimization techniques.

Furthermore, to prevent the degeneracy problem caused by zero-valued input or output weights, we proposed a method to establish lower bounds on these weights. This ensures a more meaningful and realistic representation of decision-making units (DMUs), especially under uncertainty.

In this paper, we utilized the robust optimization framework introduced by Bertsimas and Sim (2004), which allows for a controlled level of conservatism in handling uncertainty through the concept of a “budget of uncertainty.” This approach enables decision-makers to adjust the level of protection against uncertainty without significantly compromising solution quality.

We applied the proposed robust FDH model to evaluate the performance of 33 universities in Iran. The results indicate a clear trend: as the level of uncertainty (i.e., the disturbance level) increases, the efficiency scores of the DMUs generally decrease. This outcome highlights the sensitivity of performance evaluations to data perturbations and emphasizes the necessity for higher education managers and policy-makers to incorporate uncertainty considerations into their decision-making processes. In practice, this means that reliable, robust strategies should be adopted to mitigate the adverse effects of data inaccuracy and to ensure more stable and trustworthy performance assessments.

The proposed model contributes to the existing literature by offering a robust analytical framework capable of capturing uncertainty in non-convex efficiency models. It opens new avenues for future research in robust DEA models and provides practical insights for organizations operating in environments with incomplete or imprecise data.

Funding

There is no funding support.

Declaration of Competing Interest

The author has no conflicts of interest to declare that are relevant to the content of this article.

ACKNOWLEDGEMENTS

We would like to express our gratitude to the anonymous reviewers for their valuable comments, which greatly contributed to improving our work.

References

Agrell, P. J., & Tind, J. (2001). A Dual Approach to Nonconvex Frontier Models. *Journal of Productivity Analysis*, *16*(2), 129–147.

Amin, G. R., & Toloo, M. (2004). A polynomial-time algorithm for finding ε in DEA models. *Computers and Operations Research*, *31*(5), 803–805.

Arabmaldar, A., Jablonsky, J., & Saljooghi, F. H. (2017). A new robust DEA model and super-efficiency measure. *Optimization*, *66*(5), 723–736.

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

Bazaraa, M. S., Jarvis, J. J., & Sherali, H. D. (2009). *Linear Programming and Network Flows* (4th ed.). Wiley.

Ben-Tal A., El Ghaoui, L. and Nemirovski, A. (2009). *Robust Optimization*. Princeton University Press.

Ben-Tal, A., & Nemirovski, A. (1998). Robust convex optimization. *Mathematics of Operations Research*, *23*(4), 769–805.

Ben-Tal, A., & Nemirovski, A. (1999). Robust solutions of uncertain linear programs. *Operations Research Letters*, *25*(1), 1–13.

Ben-Tal, Aharon, & Nemirovski, A. (2000). Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming*, *88*(3), 411–424.

Bertsimas, D., Pachamanova, D., & Sim, M. (2004). Robust linear optimization under general norms. *Operations Research Letters*, *32*(6), 510–516.

Bertsimas, D., & Sim, M. (2003). Robust discrete optimization and network flows. *Mathematical Programming*, *98*(1), 49–71.

Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research*, *52*(1), 35–53.

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

Cook, W. D., Kress, M., & Seiford, L. M. (1996). Data Envelopment Analysis in the Presence of Both Quantitative and Qualitative Factors. *The Journal of the Operational Research Society*, *47*(7), 945.

Cooper, W. W., Park, K. S., & Yu, G. (1999). IDEA and AR-IDEA: models for dealing with imprecise data in DEA. *Management Science*, *45*(4), 597–607.

Deprins, D., Simar, L., & Tulkens, H. (1984). Measuring labor-efficiency in post offices. In *The performance of public enterprises: concepts and measurements* (pp. 243–268). Amsterdam: North-Holland.

Gorissen, B. L., Yanikoğlu, I., & den Hertog, D. (2015). A practical guide to robust optimization. *Omega*, *53*, 124–137.

Hatami-Marbini, A., Emrouznejad, A., & Tavana, M. (2011). A taxonomy and review of the fuzzy data envelopment analysis literature: Two decades in the making. *European Journal of Operational Research*, *214*(3), 457–472.

Hougaard, J. L. (1999). Fuzzy scores of technical efficiency. *European Journal of Operational Research*, *115*(3), 529–541.

Jahanshahloo, G. R., Matin, R. K., & Vencheh, A. H. (2004). On FDH efficiency analysis with interval data. *Applied Mathematics and Computation*, *159*(1), 47–55.

Kerstens, K., & Eeckaut, P. Vanden. (1999). Estimating returns to scale using non-parametric deterministic technologies: A new method based on goodness-of-fit. *European Journal of Operational Research*, *113*(1), 206–214.

Khanjani, R., Fukuyama, H., Tavana, M., & Di, D. (2016). An integrated data envelopment analysis and free disposal hull framework for cost-efficiency measurement using rough sets. *Applied Soft Computing Journal*, *46*, 204–219.

Khanjani Shiraz, R., Tavana, M., & Paryab, K. (2014). Fuzzy free disposal hull models under possibility and credibility measures. *International Journal of Data Analysis Techniques and Strategies*, *6*(3), 286–306.

Khodabakhshi, M., & Kheirollahi, H. (2013). Performance evaluation of Iran universities with Stochastic Data Envelopment Analysis (SDEA). *International Journal of Data Envelopment Analysis*, *1*(1), 7–13.

Leleu, H. (2006). A linear programming framework for free disposal hull technologies and cost functions : Primal and dual models. *European Journal of Operational Research*, *168*, 340–344.

Olesen, O. B., & Petersen, N. C. (2016). Stochastic data envelopment analysis - A review. *European Journal of Operational Research*, *251*(1), 2–21.

Omrani, H. (2013). Common weights data envelopment analysis with uncertain data: A robust optimization approach. *Computers & Industrial Engineering*, *66*(4), 1163–1170.

Peykani, P., Mohammadi, E., Saen, R. F., Sadjadi, S. J., & Rostamy-Malkhalifeh, M. (2020). Data envelopment analysis and robust optimization: A review. *Expert Systems*, (September 2019), e12534.

Sadjadi, S. J., & Omrani, H. (2008). Data envelopment analysis with uncertain data: An application for Iranian electricity distribution companies. *Energy Policy*, *36*(11), 4247–4254.

Sadjadi, S. J., Omrani, H., Abdollahzadeh, S., Alinaghian, M., & Mohammadi, H. (2011). A robust super-efficiency data envelopment analysis model for ranking of provincial gas companies in Iran. *Expert Systems with Applications*, *38*(9), 10875–10881.

Salahi, M., Toloo, M., & Torabi, N. (2020). A new robust optimization approach to common weights formulation in DEA A new robust optimization approach to common weights formulation. *Journal of the Operational Research Society*, 1–13.

Shokouhi, A. H., Hatami-Marbini, A., Tavana, M., & Saati, S. (2010). A robust optimization approach for imprecise data envelopment analysis. *Computers & Industrial Engineering*, *59*(3), 387–397.

Soyster, A. L. (1973). Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, *21*(5), 1154–1157.

Toloo, M. (2014). An epsilon-free approach for finding the most efficient unit in DEA. *Applied Mathematical Modelling*, *38*(13), 3182–3192.

Toloo, M., & Mensah, E. K. (2019). Robust optimization with nonnegative decision variables: A DEA approach. *Computers & Industrial Engineering*, *127*, 313–325.

Tulkens, H. (1993). On FDH Efficiency Analysis : Some Methodological Issues and Applications to Retail Banking , Courts , and Urban Transit, *210*, 183–210.

Zohrehbandian, M., Makui, A., & Alinezhad, A. (2010). A compromise solution approach for finding common weights in DEA: an improvement to Kao and Hung's approach. *Journal of the Operational Research Society*, *61*(4), 604–610.