Available online at http://ijdea.srbiau.ac.ir

Int. J. Data Envelopment Analysis (ISSN 2345-458X)

Vol. 12, No. 3, Year 2024 Article ID IJDEA-00422, Pages 58-66
Research Article

International Journal of Data Envelopment Analysis

Science and Research Branch (IAU)

Estimating and improving the efficiency of decision-making units with interval data

F. Hosseinzadehlotfi [1], T. Allahviranloo [2*], B. Rahmaniperchkolaei[2]

[1] Department of Industrial Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran,

[2] Department of Mathematics, Science and Research Branch, Islamic Azad University, Tehran, Iran.

Received 4 February 2024, Accepted 21 June 2024

Abstract

Performance measurement is always considered one of the most important tasks of managers. Hence, management knowledge is measurement knowledge and if we cannot measure something, we certainly cannot control it and consequently we cannot manage it. In this paper, we examine data envelopment analysis models for improving inefficient units. In this study, 20 bank branches in Tehran were selected and mathematical models were presented for estimating inputs with interval data.

The findings of this research highlight the importance of integrating advanced analytical tools like DEA into management practices. By quantifying inefficiencies and offering clear pathways for improvement, DEA empowers managers to make data-driven decisions that enhance overall performance. This approach is particularly valuable in competitive environments, such as the banking sector, where efficiency and service quality directly impact customer satisfaction and profitability.

Keywords: Data envelopment analysis, interval data, Estimate, Bank Branch.

* Corresponding author: Email: allahviranloo@gmail.com

1. Introduction

Since ancient times, especially after the end of World War II, senior decision-making managers have realized that any decision-making process lacking the application of scientific methods inevitably involves personal biases. Therefore, the use of scientific methods for evaluating units is essential and indispensable.

Understanding the performance of subordinate units is the most critical responsibility of a manager for making appropriate decisions to guide them effectively. The complexity of information, the vast volume of performance data, the impacts of external factors, the influence of competing units, the limitations of units in making appropriate decisions (for instance, due to their governmental nature), sudden policy changes due to reactive approaches to severe problems (such as unemployment and inflation), are among the challenges. Without a scientific approach, managers cannot adequately understand the performance of their units nor make effective decisions to enhance efficiency and productivity [1-3].

The aforementioned issue is often relevant in both public and private economic systems.

2. Literature Review:

1.1. Research on DEA with Uncertain Data (Interval Data):

The foundational discussion of DEA with uncertain data was conducted by Cooper and colleagues. They introduced a method to address situations where cost, price, or other information is not precisely known, but only upper and lower bounds of these values are available. Using the concept of confidence regions in DEA models, they developed a general framework for handling interval data in DEA [4].

Lee and colleagues extended the idea of DEA with uncertain data to the collective DEA model. Antani et al. proposed a DEA model that calculates interval efficiencies for each DMU (Decision-Making Unit) from optimistic and pessimistic perspectives. Initially, their model was applicable to precise data but was later expanded to include interval and fuzzy data. Wang and colleagues introduced paired interval DEA models, addressing some of the limitations in previous models. Notably, Antani's model could integrate a combination of interval, precise, and fuzzy data but was limited to a single input and a single output [5,6].

2.1. Research on Inverse DEA:

One of the key issues in DEA is estimation, first introduced by Wei and colleagues. They explored scenarios where, if the inputs of a DMU are increased by a certain amount while maintaining constant efficiency, what level of outputs must be produced. The model developed to answer this question is referred to as the inverse DEA model [7].

In their approach, they assumed that efficiency remains constant with increased input levels and focused on estimating the required outputs. They converted the inverse DEA problem into a multi-objective linear programming (MOLP) problem, demonstrating that in some specific cases, the problem could be simplified to a single-objective linear programming model [8].

Wei and Yan's inverse DEA model estimated output levels while maintaining constant efficiency with increasing inputs. Jahanshahloo and colleagues extended Yan's model by discussing output estimation when input levels are increased under the assumption of improved DMU efficiency [9].

Foroughi and Hadi-Vanche proposed a comprehensive model for input/output estimation problems. Their work generalized the prior models by considering not only increases but also decreases in input/output levels. This broader approach reflects scenarios where units may seek adjustments in both directions. Foroughi and Vanche incorporated desired changes (increases or decreases) in input/output levels into a multi-objective linear programming (MOLP) model to estimate outputs (or inputs) [10].

Suppose there are n decision-making units as $\{DMU_j ; j = 1, \dots, n\}$, each of which uses m inputs to produce S outputs (the decision-making units are homogeneous). The input and output data values of the decision-making units are in the form of bounded bases, that is, the exact amount of data is not available and instead the upper and lower bounds of each data are determined. So that for $(j=1,\cdots,n) DMU_j$ we have $(i=1,\cdots,m) x_{ij}^1 \le x_{ij} \le x_{ij}^U$ or in other words $x_{ij} \in [x_{ij}^1, {}_{ij}^U]$; $i=1,\cdots,m$; $j=1,\cdots,n$ and also $y_{rj}^L \le y_{rj} \le y_{rj}^U$ or in other words $y_{rj} \in [y_{rj}^L, y_{rj}^U]$; $r=1,\cdots,s$; $j=1,\cdots n$ where y_{rj}^L, x_{rj}^L is the lower bound of the intervals and y_{rj}^U, x_{ij}^U is the upper bound of the intervals. Also, their values are assumed to be constant and strictly positive and always $y_{rj}^L \le y_{ij}^U, x_{ij}^1 \le x_{ij}^U$ and if we have $x_{ij}^L = x_{ij}^U$, then the i-th input of the j-th decision-making unit has an exact and certain value [11].

Now the CCR model with interval data is expressed as follows.

$$
\begin{aligned}
Max \quad & \sum_{r=1}^{s} u_r \left[y_{rp}^l, y_{rp}^u \right] \\
& \sum_{r=1}^{m} v_i \left[x_{tp}^l, x_{tp}^u \right] = l \\
& \sum_{r=1}^{s} u_r \left[y_{rp}^l, y_{rp}^u \right] - \sum_{r=1}^{m} v_i \left[x_{tp}^l, x_{tp}^u \right] \le 0 \\
& u_r, v_i \ge 0; r = l, \cdots, s; l, \cdots, m
\end{aligned} \tag{1}
$$

The above model is a nonlinear programming model. Despotis and Smirlis transformed the above nonlinear model into a linear model by making some changes.

To find the upper and lower bounds of the unit under evaluation, models (2) and (3) must be solved.

$$
\begin{aligned}
\theta^L = Max & \sum_{r=1}^{s} u_r y_{rp}^L \\
s.t. \quad & \sum_{r=1}^{s} u_r y_{rj}^U - \sum_{i=1}^{m} v_i x_{ij}^L \le 0, \quad j = 1, \dots, n \; (j \ne p) \\
& \sum_{r=1}^{s} u_r y_{rp}^L - \sum_{i=1}^{m} v_i x_{ip}^U \le 0, \\
& \sum_{i=1}^{m} v_i x_{ip}^U = 1 \\
& u_r, v_i \ge 0; r = l, \dots, s; i = l, \dots, m
\end{aligned} \tag{2}
$$

$$
\begin{aligned}
\theta^L = Max & \sum_{r=1}^{s} u_r y_{rp}^U \\
s.t. \quad & \sum_{r=1}^{s} u_r y_{rj}^L - \sum_{i=1}^{m} v_i x_{ij}^U \le 0, \quad j = 1, \dots, n \; (j \ne p) \\
& \sum_{r=1}^{s} u_r y_{rp}^U - \sum_{i=1}^{m} v_i x_{ip}^L \le 0, \\
& \sum_{i=1}^{m} v_i x_{ip}^L = 1 \\
& u_r, v_i \ge 0; r = l, \dots, s; i = l, \dots, m
\end{aligned} \tag{3}
$$

2.2 Estimating input levels in DEA with interval data:

We intend to increase the output levels of the unit under evaluation and look for inputs that will produce this amount of output (new outputs) in such a way that the efficiency level of the unit under evaluation remains constant or has the desired improvement of the decision maker or decision makers. The importance of this issue is that sometimes managers intend to make policies and plans for the future of the organization or the unit under their management, and therefore, according to the plan and decisions they have, they predict the level of production or output of the organization for a period of time and they intend to estimate the resources or inputs necessary for their desired production so that they can examine the practicality of their decisions. Sometimes the manager is satisfied with keeping the efficiency level constant and his only goal is to increase the level of output or production [13].

Suppose there are n decision-making units in the form of $\{DMU_j; j=1,\ldots,n\}$, each of which uses m inputs to produce s outputs (the decision-making units are homogeneous). The input and output data values of the decision-making units are in bounded intervals, that is, the exact amount of data is not available and instead the upper and lower bounds of each data are determined. So that for $(j=l,\cdots,n)DMU_j$ we have $x_{ij}^{l} \le x_{ij} \le x_{ij}^{U}, (i=l,\cdots,m)$ or in other words $x_{rj} \in [x_{rj}^{L}, x_{rj}^{U}]; \quad r=l,\cdots,m;$ $j=1,\cdots n$ and also $y_{rj}^{L} \le y_{rj} \le y_{rj}^{U}$ or in other words $y_{rj} \in [y_{rj}^{L}, y_{rj}^{U}];\ r=1,\cdots,s;$ $j=1,\cdots n$ where y_{rj}^{L}, x_{rj}^{L} is the lower bound of the intervals and y_{rj}^{U}, x_{ij}^{U} is the upper bound of the intervals. And also, their values are assumed to be constant and strictly positive. $y_{rj}^{L} \le y_{ij}^{U}, x_{ij}^{1} \le x_{ij}^{U}$ always holds. As stated in the DEA section with base data, a specific efficiency interval such as $\left[\theta_*^{L}, \theta_*^{U}\right]$ can be calculated for each of the $(j=1,\cdots,n)DMU_j$ decision-making units. Where θ_*^{L} is the lower bound of DMU efficiency, which is calculated by model (2), and θ_*^{U} is the upper bound of DMU efficiency. θ_*^{U} and θ_*^{L} are calculated using input-driven CCR envelope models with interval data. Since the goal is to estimate inputs, input-driven models are used.

Consider the p-th unit under evaluation, i.e., DMU_p. It is assumed that the outputs of this unit, or in other words, the lower and upper bounds of the outputs of DMU_p, i.e., $(r=l,\cdots,s)\,y_{rp}^{U}, y_{rp}^{L}$, have increased by $(r=l,\cdots,s)\,\Delta y_{rp}^{U}, \Delta y_{rp}^{L}$ respectively, and it is also assumed that $(r=l,\cdots,s)\,\Delta y_{rp}^{U}, \Delta y_{rp}^{L}$ (note that $\Delta y_{rp}^{U}, \Delta y_{rp}^{L}$ are the vectors of the amount of change in the lower and upper bounds of the input data DMU_p, respectively), so the new output values of DMU_p will be $\beta_{rp}^{L} = y_{rp}^{L} + \Delta y_{rp}^{L}; r=1,\cdots,s$ and $\beta_{rp}^{U} = y_{rp}^{U} + \Delta y_{rp}^{U}; r=1,\cdots,s$, or in other words, the new output values will be in the bounded interval $i=l,\cdots,m; \left[\beta_{rp}^{L}, \beta_{rp}^{U}\right]$.
In this case, we intend to estimate the new input values of DMU_p assuming that the efficiency remains constant and the efficiency level improves to DMU_p. We

assume that $\alpha_p = \left(\alpha_{1p}, \cdots, \alpha_{ip}, \cdots, \alpha_{mp}\right)$ is the new input value of DMU_p. Now, to estimate the input levels of DMU_p, we write Model (4) (Model for estimating inputs when output levels increase assuming efficiency remains constant) in Chapter 3 as follows.

$$Min\alpha_p = \left(\alpha_{1p}, \cdots, \alpha_{ip}, \cdots, \alpha_{mp}\right) \tag{4}$$

$$\sum_{j=1}^{n} \lambda_j \left[x_{ij}^L, x_{ij}^U\right] \le \left[\theta_*^L, \theta_*^U\right] \alpha_{ip}; i = l, \cdots, m$$

$$\sum_{j=1}^{n} \lambda_j \left[y_{ij}^L, y_{ij}^U\right] \ge \left[\beta_{rp}^L, \beta_{rp}^U\right]; r = l, \cdots, s$$

$$\lambda_j \ge 0; j = 1, \cdots n$$

Now suppose

$$\begin{cases} x_{ij}^L \le x_{ij} \le x_{ij}^U \\ \theta^{*L} \le \theta^* \le \theta^{*U} \\ y_{rj}^L \le y_{rj} \le y_{rj}^U \\ \beta_{rp}^L \le \beta_{rp} \le \beta_{rp}^U \end{cases}$$

$$\Rightarrow \begin{cases} x_{ij} = x_{ij}^L + p'_{ij}\left(x_{ij}^U, x_{ij}^L\right) \ \forall i \ \forall j & 0 \le p'_{ij} \le 1 \\ \theta^* = \theta^{*L} + \alpha\left(\theta^{*U} - \theta^{*L}\right) & 0 \le \alpha \le 1 \\ y_{rj} = y_{rj}^L + q'_{rj}\left(y_{rj}^U - y_{rj}^L\right) \ \forall r \ \forall j & 0 \le q'_{rj} \le 1 \\ \beta_{rp} = \beta_{rp}^L + \delta'_{rj}\left(\beta_{rp}^U - \beta_{rp}^L\right) \ \forall r & 0 \le \delta'_{rp} \le 1 \end{cases}$$

By inserting this variable change into model (4) and applying $p'_{ij}, \lambda_j = p_{ij}$, $q'_{rj}\lambda_j = q_{rj}$, $\alpha.\alpha_{ip} = \alpha'_{ip}$ changes, model (5) is obtained.

$$Min\alpha_p = \left(\alpha_{1p}, \cdots, \alpha_{ip}, \cdots, \alpha_{mp}\right) \tag{5}$$

$$\begin{cases} \sum_{j=1}^{n} \lambda_j x_{ij}^L + \sum_{j=1}^{n} p_{ij}\left(x_{ij}^L - x_{ij}^L\right) \\ \le \theta_0^L \alpha_{ip} + \alpha_{ip}^l\left(\theta_0^U, \theta_0^L\right); i = l, \cdots, m \end{cases}$$

$$\begin{cases} \sum_{j=1}^{n} \lambda_j y_{rj}^L + \sum_{j=1}^{n} q_{rj}\left(y_{rj}^L - y_{rj}^U\right) \\ \ge \beta_{rp}^L + \delta_{rp}\left(\beta_{rp}^U - \beta_{rp}^L\right); r = l, \cdots, s \end{cases}$$

$$0 \le p_{ij} \le \lambda_j; i = l, \cdots, m; j = l, \cdots, n$$

$$0 \le q_{rj} \le \lambda_j; r = l, \cdots, s; j = l, \cdots, n$$

$$0 \le \alpha_{ip}^l \le \alpha_{ip}; i = l, \cdots, m$$

$$0 \le \delta_{rp} \le l; r = l, \cdots, s$$

$$0 \le \gamma_{ip} \le l; i = l, \cdots, m$$

$$\lambda_j \ge 0; j = l, \cdots, n$$

The above model is a linear programming model. As you know, the values $y_{ij}^U, y_{ij}^L, \beta_{rp}^U, \beta_{rp}^L, x_{ij}^L, x_{ij}^U$ and θ_*^L, θ_*^U are given. Solving this model in practice is difficult. Therefore, another solution is used in practice to estimate the output levels. This solution is that we estimate the lower and upper bounds of the new outputs, or in other words, we estimate as $\left[\alpha_{ip}^L, \alpha_{ip}^U\right], i = 1, ..., m$, that the corresponding models are given in models (6) and (7).

$$Min \sum_{i=1}^{m} w_i \alpha_{ip}^U \tag{6}$$

$$s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij}^L + \lambda_p x_{ip}^U \le \theta_*^L \alpha_{ip}^U, \quad i = 1, ..., m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj}^U + \lambda_p y_{rp}^L \le \beta_{rp}^L, \quad r = 1, ..., s$$

$$\lambda_j \ge 0, \quad j = 1, ..., n$$

$$
\begin{aligned}
& Min \sum_{i=1}^{m} w_i \alpha_{ip}^{L} \qquad (7)\\
& s.t. \ \sum_{j=1}^{n} \lambda_j x_{ij}^{U} + \lambda_p x_{ip}^{L} \le \theta_{*}^{U} \alpha_{ip}^{L}, \ i=1,\dots,m \\
& \quad \sum_{j=1}^{n} \lambda_j y_{rj}^{L} + \lambda_p y_{rp}^{U} \le \beta_{rp}^{U}, \quad r=1,\dots,s \\
& \quad \lambda_j \ge 0, \qquad j=1,\dots,n
\end{aligned}
$$

3. Model Implementation

In this study, the data in question were collected using the indicators introduced in the previous section from 20 branch bank branches in Tehran in 2015 and analyzed using GAMS software. The results of this analysis are presented below.

The input and output indicators considered for branches are shown in Figure 1.

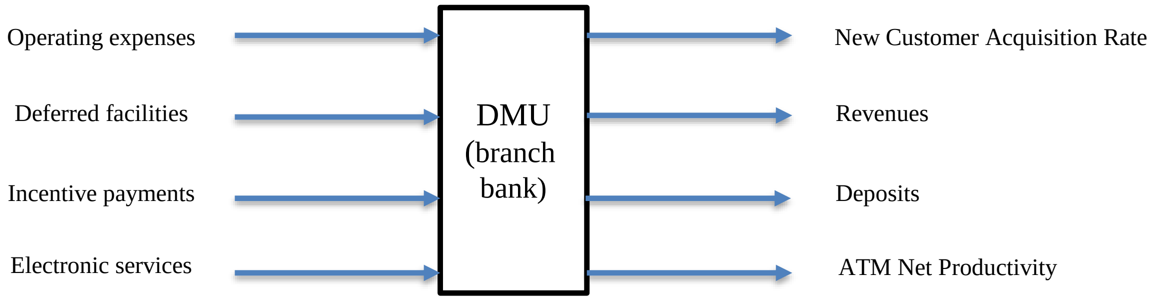

Fig 1. Branch entry and exit indicators

Table 1. Branch efficiency values

$[\theta_{*}^{L}, \theta_{*}^{U}]$	Name Branch	$[\theta_{*}^{L}, \theta_{*}^{U}]$	Name Branch
[1,1]	11	[1,1]	1
[0.68958,0.74113]	12	[0.47904,0.52987]	2
[0.67796,1]	13	[1,1]	3
[1,1]	14	[0.99629,1]	4
[0.841194,0.90133]	15	[0.93980,0.96094]	5
[1,1]	16	[0.90702,1]	6
[0.88034,0.89934]	17	[1,1]	7
[0.81120,0.82897]	18	[1,1]	8
[1,1]	19	[1,1]	9
[0.93997,1]	20	[0.79352,0.90709]	10

We consider the second unit under evaluation, which is an idle unit. The estimated output of the new inputs is expected to increase compared to the previous inputs because it is assumed that the new outputs have increased and the efficiency level is also constant. Therefore, with increasing outputs, in order for the efficiency level to remain constant, the output of the new inputs must increase. In the graphs below, the trend of changes in the upper bounds of each of the inputs can be observed in exchange for an increase in the first output level.

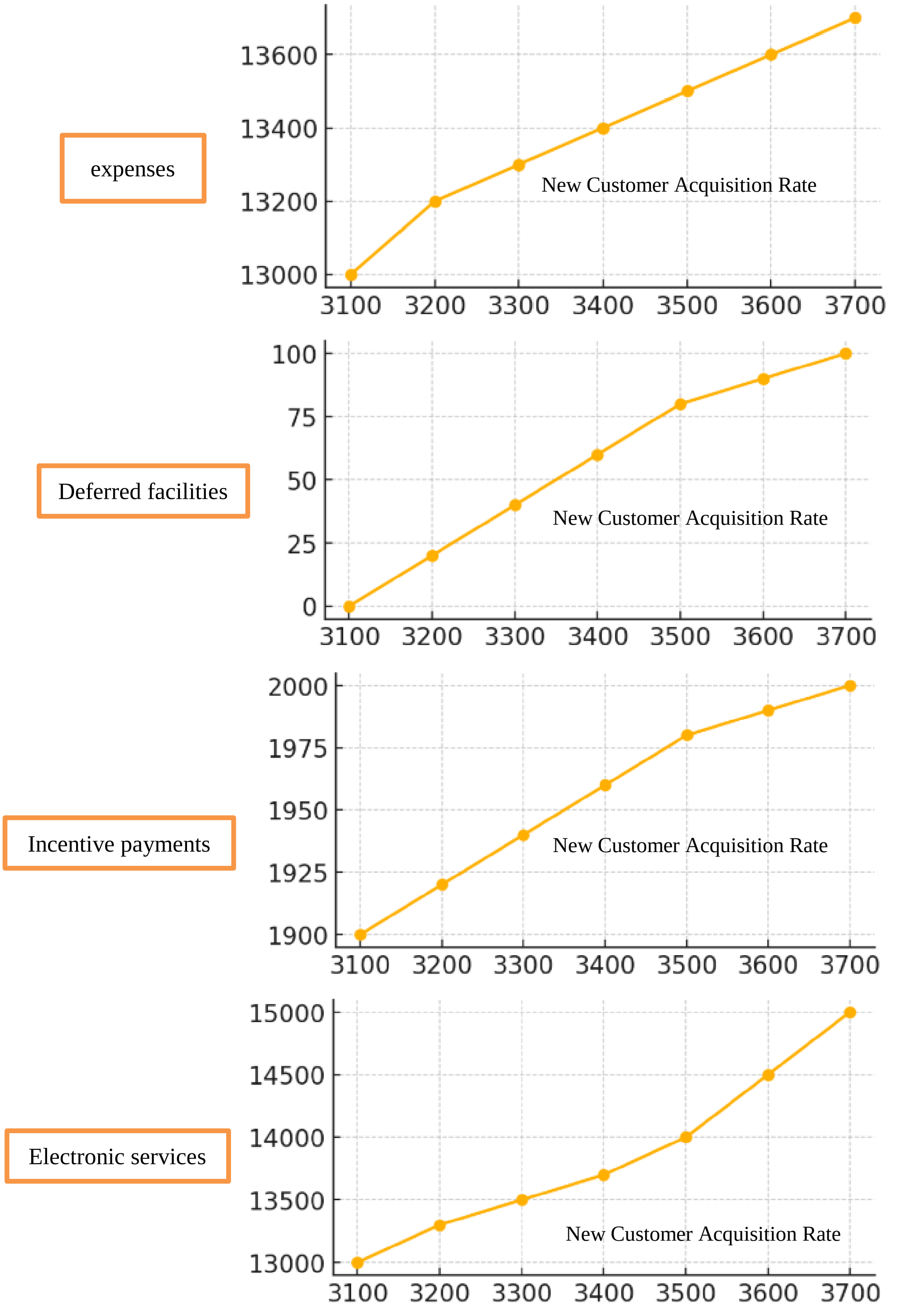

Fig 2. The trend of changes in the upper bounds of inputs in exchange for increasing outputs

Similar to the above chart, we can analyze the charts for the other indicators as well.

5. Conclusion:

The findings of this research highlight the importance of integrating advanced analytical tools like DEA into management practices. By quantifying inefficiencies and offering clear pathways for improvement, DEA empowers managers to make data-driven decisions that enhance overall performance. This approach is particularly valuable in competitive environments, such as the banking sector, where efficiency and service quality directly impact customer satisfaction and profitability.

Definite data values are not always available, so assuming that the data is interval-based, we calculated the efficiency of each branch compared to other branches. Considering the interval-based nature of the available data, an efficiency range was determined for each branch as an upper and lower bound of efficiency.

References

[1] Farrell, M. "The Measurement of productive Efficiency". Journal of Research Statistics society 20 253-290. (1957).

[2] Charnes, W. Cooper, W. Rhodes, E." Measuring the efficiency of decision-making units". European Journal of Operational Research 120.429-444. (1978).

[3] O'Neal. P, Ozcan. Y, Yangiang. M, "Benchmarking mechanical Ventilation Services in teaching hospitals". Journal of medical Systems 26 .227-240. (2002).

[4] Simar. L, Wilson. P, "Statical inference in non-parametric frontier models. The state of art Journal of productivity Analysis 13. 49-78. (2000).

[5] Cooper. W, Seiford. L, Tone. K, "Data Envelopment Analysis. A Comprehensive text with models. application. references and DEA-solver software". Journal of kluwer academic publishers.p.253. (1999).

[6] Kuosmanen. T, "Modeling blank data entries in Data Envelopment Analysis". Journal of Econometrics working paper Archive at wustl. (2002).

[7] Kao. C, Liu. S, " Data Envelopment Analysis with missing data; An application to university Libraries in Taiwan". Journal of the operational Research society 51.897-905.2000

[8] Smirlis, Y. G., Maragos, E. K., & Despotis, D. K. (2006). Data envelopment analysis with missing values: An interval DEA approach. Applied Mathematics and Computation, 177(1), 1-10.

[9] Soltanifar, M., Ghiyasi, M., Emrouznejad, A., & Sharafi, H. (2024). A novel model for merger analysis and target setting: A CSW-Inverse DEA approach. Expert Systems with Applications, 249, 123326.

[10] Razipour-GhalehJough, S., Lotfi, F. H., Rostamy-Malkhalifeh, M., & Sharafi, H. (2021). Benchmarking bank branches: A dynamic DEA approach. Journal of Information and Optimization Sciences, 42(6), 1203-1236.

[11] Soltanifar, M., Ghiyasi, M., & Sharafi, H. (2023). Inverse DEA-R models for merger analysis with negative data. IMA Journal of Management Mathematics, 34(3), 491-510.

[12] Ghiyasi, M., Soltanifar, M., & Sharafi, H. (2022). A novel inverse DEA-R model with application in hospital efficiency. Socio-economic planning sciences, 84, 101427.