Journal of Mahani Mathematical Research
Shahid Bahonar University of Kerman
Print ISSN: 2251-7952 Online ISSN: 2645-4505

ELIMINATING CONGESTION OF DECISION-MAKING UNITS USING INVERSE DATA ENVELOPMENT ANALYSIS

T. Shahsavan, M. Sanei, Gh. Tohidi, F. Hosseinzadeh Lotfi, and S. Ghobadi ✉

Article type: Research Article
(Received: 10 September 2023, Received in revised form 25 April 2024)
(Accepted: 14 June 2024, Published Online: 15 June 2024)

Abstract. This survey proposes a new application of the inverse data envelopment analysis (InvDEA) in the problem of merging decision-making units (DMUs) to improve the performance of DMUs by removing congestion. Congestion is a factor in reducing production; therefore, removing it decreases costs and increases outputs. There are two significant subjects in the merging DMUs. Estimating the inherited inputs and outputs of a new production DMU with no congestion is the first problem while achieving a pre-specified efficiency level from the merged DMU is the second one. Both problems are examined using the ideas of inverse DEA and congestion. Using Pareto solutions to multiple-objective programming problems, sufficient conditions for inherited input/output estimates with no congestion and increasing efficiency are created. Besides, an example is perused for the reliability of the proposed approach in basic research institutes in the Chinese Academy of Science (CAS) in 2010.

Keywords: Inverse Data Envelopment Analysis (InvDEA), Congestion, Merging, Inverse DEA, Multiple-objective programming (MOP).

2020 MSC: 90C08, 90C90, 90B50, 90C29, 91B06.

1. Introduction

Data envelopment analysis is a technique to evaluate the performance of different organizations which has many applications in management and economics. Data envelopment analysis was first used by Charans et al. [10] to estimate the efficiency of the decision-making units. Identifying and eliminating inefficient DMUs and investigating the causes of DMU inefficiency are useful ways to improve the performance of production systems. Congestion is one of the reasons behind DMUs' inefficiency. Congestion occurs when raising at least one input to a DMU decreases at least one output without impairing other inputs and outputs. Färe and Svensson [21] published the first paper on congestion. Later, Färe and Grosskoph [19] obtained the congestion employing DEA. Furthermore, Färe et al. [20] identified the existence of congestion with the FGL

✉ Ghobadi@iaukhsh.ac.ir, ORCID: 0000-0002-1884-0767
https://doi.org/10.22103/jmmr.2024.22163.1506
Publisher: Shahid Bahonar University of Kerman
How to cite: T. Shahsavan, M. Sanei, Gh. Tohidi, F. Hosseinzadeh Lotfi, S. Ghobadi, *Eliminating congestion of decision-making units using inverse data envelopment analysis*, J. Mahani Math. Res. 2024; 13(2): 453 - 479.

model [20]. The FGL model has achieved radial efficiency with the presence of the strong possibility of inputs and the weak possibility of inputs, and the ratio of these two efficiencies has been used as an index to identify congestion. However, the FGL model cannot calculate the input congestion amount. Cooper et al. [1] identified the congestion using the slack variables and model CTT and obtained the input congestion amount. The CTT model is a two-step process. In the first step, the projection of a DMU is obtained by the output-oriented BCC model [6], and in the second step, maximum slacks are added to the projection inputs is calculated. The difference between the optimal slacks in steps 1. and 2 specifies the congestion value. Then, Cooper et al. [12] combined the two steps of the CTT model and proposed a single-model method to calculate congestion. Later, Brocket et al. [9] used the CTT model in China's industry. Jahanshahloo and Khodabakhshi [29] obtained the amount of input congestion with a model based on a suitable combination of input to improve outputs. Tone et al. [48] specified the weak congestion with a non-radial model and they proposed a multiplier model to identify the strong congestion. Sueyoshi and Sekitani [47] removed Tone model problems corresponding to multiple optimal solutions. They considered congestion to be the same as the negative return to return scale and found a theoretical relationship between them. Many theoretical and empirical studies in the field of congestion were done, including congestion in data envelopment analysis with uncertain data [4, 33], congestion identification in the supply chain [41, 44], congestion measurement under undesirable outputs [18], congestion detection in data envelopment analysis with negative and non-negative data [35,36,38,43]. Determining the congested border by obtaining the congestion hyperplane [15], directional congestion in specific and unspecific input or output directions [34, 53], congestion determination in dynamic DEA with inter-temporal dependence [42], textile industry of China [30], e-commerce sector [54], the forestry sector [32], china's, industries carbon congestion [56], wastewater treatment [22], shipping company [49], investment [27], and university [40] and so on. There has been no comprehensive research on how to minimize or remove congestion. Kao [32] demonstrated how to decongest certain congested DMUs using an untested method. The purpose of this study is to minimize congestion on congested units via the use of merging DMUs. In the real world, companies employ the merger phenomena to establish a new, more efficient entity. A merger occurs when at least two DMUs incorporate their activities to generate a new merged DMU with improved efficiency. Restructuring, the synergy of DMUs via consolidation, or the inverse of the synergy of DMUs through a split is considered. Gattofi et al. [23] and Amin et al. [2]), used inverse DEA to merge banks to increase productivity. Amin et al. [2] used restructuring to create the synergy of units through consolidation. Ghobadi [24] has utilized merging for DMUs with interval data. The integration of DMUs under inter-temporal dependence has been done by Zeinodin et al. [55]. Shiri et al. [45] have applied merging in systems with network structure. Amin and Qukil [3] proposed a new method for the merging problem of the unit under a flexible target setting using inverse DEA. Soltanifar et al. [46] have used merging in inverse DEAR models with negative data. The subject of merging has many practical applications, including the banking sector [5, 28, 52], the agriculture industry [8, 39], forestry [7], airline mergers [37], the effect of mergers and acquisitions in pharmaceutical industry [25], and the water sector [14]. Inverse DEA was utilized in most merging problems.

Inverse DEA identifies the critical factors affecting a DMU's efficiency, such as its input and output. In other words, inverse DEA is a method for calculating inputs and outputs in a manner that maintains or improves efficiency. Wei et al. [51] studied an inverse DEA model for estimating inputs (outputs). Jahanshahlou et al. [30] have studied the output estimation problem in the case that the efficiency is improved to a certain extent. Hadi Vinche et al. [26] have obtained the inputs of a unit so that its outputs are increased to a certain extent and its efficiency remains unchanged. Jahanshahlou et al. [31] have used inverse data envelopment analysis under internal temporal dependence. Inverse data envelopment analysis has many applications, including investment analysis for improving a production system [11], Merger of banks to improve the performance of banks [2, 23], and, other cases of merging. Also, to review the study on Inverse DEA and its applications, refer to Emrouznejad et al. [17]. In this study, merging and inverse DEA are used to reduce the congestion. In this paper, at least two congested units have combined their activities and created a new merged DMU. Thus, the new merged unit has no congestion and the efficiency of the new unit has improved compared to the units before the merger. An inverse DEA model is utilized to find the inherited data of a new unit with a predefined efficiency level. The superiority of the proposed method is that in a production system with congested units, congested units are eliminated using merging, and non-congested DMUs are generated. In other words, the proposed method shows the application of the merger phenomenon to eliminate congestion. The remainder of the study is set as follows. Section 2 prepares an introduction to multiple-objective programming and congestion. Section 3 proposes a method for eliminating at least two congested DMUs using merging. Section 4 uses a practical example in CAS institutes in 2010 to eliminate congestion. The conclusion is made in section 5.

2. Preliminaries

In this section, first multiple-objective programming is mentioned, then the BCC and congested production possibility sets and relevant output-oriented radial models are reviewed. Then, some congestion models are described.

2.1. **Multiple-objective programming (MOP).** The general form of multiple-objective programming (MOP) problems is as follows:

$$\max \{f_1(x), f_2(x), \ldots, f_k(x)\} \quad s.t.\ X = \{x \in \mathbb{R}^n |\ g_i(x) \leq 0,\ i = 1, 2, \ldots, l\}. \tag{1}$$

The optimal solution is not defined for multiple-objective problems. Pareto solution and weak Pareto solutions are defined for MOP problems as follows:

Definition 2.1. Pareto solution [16] $\bar{x} \in X$ is a Pareto solution of a MOP, if there does not exist $\hat{x} \in X$ so that:

$$f_i(\hat{x}) \geq f_i(\bar{x}),\ \forall i \in \{1, 2, \ldots, k\},$$
$$f_i(\hat{x}) > f_i(\bar{x}),\ \exists i \in \{1, 2, \ldots, k\}.$$

Definition 2.2. Weak Pareto solution [16] $\bar{x} \in X$ is a weak Pareto solution of a MOP, if there does not exist $\hat{x} \in X$ so that:

$$f_i(\hat{x}) > f_i(\bar{x}),\ \exists i \in \{1, 2, \ldots, k\}.$$

In other words, $\bar{x}$ is a weak Pareto solution of a MOP, if there is no feasible solution that improves all the objective functions simultaneously.

In multiple-objective problems, the Pareto solution is not unique. Pareto solutions are obtained using various methods. In this study, the Pareto solution is determined by the weighted sum method.

2.1.1. *Weighted sum method.* The weighted sum method for solving multi-objective planning problems is defined as a model (2) [16]:

$$\begin{aligned} &\max \sum_{j=1}^{k} w_j f_j(x), \\ &s.t.\ X = \{x \in \mathbb{R}^n |\ g_i(x) \leq 0,\ i = 1, 2, \ldots, l\}, \\ &\sum_{j=1}^{k} w_j = 1, \\ &w_j \geq 0,\ j = 1, 2, \ldots, k. \end{aligned} \tag{2}$$

Each optimal solution of model (2) is a Pareto solution of model (1). For proof, refer to Ehrgott [16]. In model (2), X is the feasible region and w_j is the weight of the objective functions, which is determined by the manager. The manager weights the objective functions based on their importance, and any more important objective is given a higher weight.

2.2. **BCC and congested production possibility sets.** In this section, firstly, the BCC and congested production possibility sets have been reviewed, and then the output-oriented radial models under the strong disposability of inputs and outputs and also, under the inputs non-disposability have been stated. Suppose there are n DMUs: $\{DMU_j \mid j = 1, \ldots, n\}$, which each

DMU generate s outputs: $\{y_{rj} \geq 0 \mid r = 1, \ldots, s\}$ by consuming m inputs: $\{x_{ij} \geq 0 \mid i = 1, \ldots, m\}$. Banker et al. [6] introduced the BCC production possibility set (PPS_{BCC}) that satisfies the principles of observation, convexity, and strong disposability of inputs and outputs. It is as follows:

$$PPS_{BCC} = \{(x, y) \mid \sum_{j=1}^{n} \lambda_j x_j \leq x,\ \sum_{j=1}^{n} \lambda_j y_j \geq y,\ \sum_{j=1}^{n} \lambda_j = 1,$$
$$\lambda_j \geq 0 \geq,\ j = 1, \ldots, n\}.$$

A DMU has congestion if an increase (decrease) in at least one input leads to a decrease (increase) in at least one output without having a worse effect on other inputs and outputs. According to the definition of congestion, there is no constant return to scale and disposability of inputs principles in a congested production possibility set. Tone and Sahoo [48] stated congested production possibility set (PPS_C) as Definition (2.3):

Definition 2.3. (PPS_C) [48] PPS_C, a production possibility set includes the principles of observations, convexity, strong disposability of outputs, and non-disposability of inputs and is defined as follows:

$$PPS_{BCC} = \{(x, y) \mid \sum_{j=1}^{n} \lambda_j x_j = x,\ \sum_{j=1}^{n} \lambda_j y_j \geq y,\ \sum_{j=1}^{n} \lambda_j = 1,$$
$$\lambda_j \geq 0 \geq,\ j = 1, \ldots, n\}.$$

According to the definitions of PPS_{BCC} and PPS_C, it is obvious that PPS_C is a subset of PPS_{BCC}. Radial efficiency in the output-oriented models of DMU_O under the PPS_{BCC} and PPS_C is obtained from the models (3) and (4) respectively.

$$\begin{aligned} \varphi^* = \quad & \max\ \varphi \\ & s.t.\ \sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{io},\ i = 1, 2, \ldots, m, \\ & \sum_{j=1}^{n} \lambda_j y_{rj} \geq \varphi y_{ro},\ r = 1, 2, \ldots, s, \\ & \sum_{j=1}^{n} \lambda_j = 1, \\ & \lambda_j \geq 0,\ j = 1, 2, \ldots, n, \end{aligned} \tag{3}$$

$$
\begin{aligned}
\psi^* = \quad & \max \ \varphi \\
& s.t. \ \sum_{j=1}^{n} \lambda_j x_{ij} = x_{io}, \ i = 1, 2, \dots, m, \\
& \sum_{j=1}^{n} \lambda_j y_{rj} \geq \varphi y_{ro}, \ r = 1, 2, \dots, s, \\
& \sum_{j=1}^{n} \lambda_j = 1, \\
& \lambda_j \geq 0, \ j = 1, 2, \dots, n,
\end{aligned} \tag{4}
$$

In models (3) and (4), $(\lambda,\ \varphi)$ is the vector of variables. If the optimal solutions of model (3) and (4) are 1 i.e. $\varphi^*,\ \psi^* = 1$ then DMU_O is radial efficiency in the BCC and congestion production possibility sets. It is obvious that the feasible region of model (3) is a subset of the feasible region of model (4) and $\varphi^* \geq \psi^*$. Models (3) and (4) are called BCC and congestion models, respectively.

2.3. **Review of congestion.** In this section, FGL , WY-TS , and Cooper models are mentioned for determining congestion.

2.3.1. *FGL model.* This method was invented by Färe et al. [20]. In this method, first, using model (3), the maximum radial increase of DMU_O outputs i.e. φ^* has been obtained, then with the model (5), the maximum radial increase of DMU_O outputs i.e. β^* under weak disposability of inputs has been determined.

$$
\begin{aligned}
& \max \ \beta \\
& s.t. \ \sum_{j=1}^{n} \lambda_j x_{ij} = \tau x_{io}, \ i = 1, 2, \dots, m, \\
& \sum_{j=1}^{n} \lambda_j y_{rj} \geq \beta y_{ro}, \ r = 1, 2, \dots, s, \\
& \sum_{j=1}^{n} \lambda_j = 1, \\
& 0 < \tau \leq 1, \\
& \lambda_j \geq 0, \ j = 1, 2, \dots, n,
\end{aligned} \tag{5}
$$

If $\frac{\beta^*}{\varphi^*} < 1$, then DMU_O is congestion. The FGL model identifies congested units, but does not calculate the amount of congestion.

2.3.2. *WY-TS model.* Wei and Yan [50] and Tone and Shaoo [48] have introduced the WY-TS method for the existence of congestion. To identify the congestion, they obtained the ratio of $C = \frac{\psi^*}{\varphi^*} < 1$. φ^* and ψ^* are the optimal solutions of models (3) and (4), respectively. They specified that the DMU_O is congested if $C < 1$. The WY-TS model does not determine the amount of congestion.

2.3.3. *Cooper model.* Cooper et al. [12, 13] obtained the congestion value by using a model based on slack variables, which is known as the CTT model. The CTT model consists of two stages. In the first step, using the model (6), the projection of the unit under evaluation, i.e. DMU_O, is determined.

$$
\begin{aligned}
&\max\ \eta + \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+) \\
&s.t.\ \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io},\ i = 1, 2, \ldots, m, \\
&\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = \eta y_{ro},\ r = 1, 2, \ldots, s, \\
&\sum_{j=1}^{n} \lambda_j = 1, \\
&s_i^- \geq 0,\ s_r^+ \geq 0,\ \lambda_j \geq 0,\ \forall i,\ r,\ j.
\end{aligned} \tag{6}
$$

The projection of DMU_O is as follows:

$$
\begin{aligned}
\hat{x}_{io} &= x_{io} - s_{io}^{-*},\ i = 1, 2, \ldots, m, \\
\hat{y}_{ro} &= \eta^* y_{ro} + s_{ro}^{+*},\ r = 1, 2, \ldots, s.
\end{aligned}
$$

That s_{io}^{-*}, s_{ro}^{+*} and η^* are obtained from the optimal solution of the model (6).

Then, in the second step of the CTT method using the model (7), the maximum increase has been found in the projection inputs i.e. $\hat{x}_{io}$ so that the projection outputs remain at the same level as before i.e. $\hat{y}_{ro}$.

$$
\begin{aligned}
&\max \sum_{j=1}^{n} \sigma_{io}^{-} \\
&s.t. \sum_{j=1}^{n} \lambda_j x_{ij} = \hat{x}_{io} + \sigma_{io}^{-}, \; i = 1, 2, \ldots, m, \\
&\sum_{j=1}^{n} \lambda_j y_{rj} = \hat{y}_{ro}, \; r = 1, 2, \ldots, s, \\
&\sum_{j=1}^{n} \lambda_j = 1, \\
&\sigma_{io}^{-} \leq s_{io}^{-*}, \; i = 1, 2, \ldots, m, \\
&\lambda_j \geq 0, \; j = 1, 2, \ldots, n.
\end{aligned} \tag{7}
$$

The congestion amount of i^{th} input of DMU_O is obtained from equation (8).

$$
s_{io}^{c} = s_{io}^{-*} - \sigma_{io}^{-*}, \; i = 1, 2, \ldots, m, \tag{8}
$$

Where σ_{io}^{-*} is the optimal solution of model (7).

Also, Cooper et al. [12] presented a single-model method (model (9)) by combining two models (6) and (7).

$$
\begin{aligned}
&\max \; \eta + \varepsilon \sum_{r=1}^{s} s_{ro}^{+} - \varepsilon^2 \sum_{i=1}^{m} s_{io}^{c} \\
&s.t. \sum_{j=1}^{n} \lambda_j x_{ij} + s_{io}^{c} = x_{io}, \; i = 1, 2, \ldots, m, \\
&\sum_{j=1}^{n} \lambda_j y_{rj} - s_{ro}^{+} = \eta y_{ro}, \; r = 1, 2, \ldots, s, \\
&\sum_{j=1}^{n} \lambda_j = 1, \\
&s_{io}^{c} \geq 0, \; s_{ro}^{+} \geq 0, \; \lambda_j \geq 0, \; \forall i, \; r, \; j.
\end{aligned} \tag{9}
$$

In model (9) η expresses the output radial increase rate. s_{ro}^{+} and s_{io}^{c} demonstrate the r^{th} output slack variable of DMU_O and the i^{th} congestion variable of DMU_O respectively. Also, ε is a non-Archimedean small positive number. Cooper et al. [12] used Theorem (2.4) to obtain the amount of congestion.

Theorem 2.4. *[12] DMU_O is congested if and only if in an optimal solution of model (5) i.e. (η^*, s_{io}^{c*}, s_{ro}^{+*}, λ_j^*) at least one of the following two statuses is defined:*

1) $\eta^* > 1$ and there is at least one i $(i = 1, 2, \ldots, m)$ so that $s_{io}^{c*} > 0$.

2) There is at least one r $(r = 1, 2, \ldots, s)$ so that $s_{ro}^{+*} > 0$, and there is at least one i $(i = 1, 2, \ldots, m)$ so that $s_{io}^{c*} > 0$.

Conditions 1 or 2 show that the decrease of at least one input, i.e. s_{io}^{c*}, causes the increase of at least one output, i.e. s_{ro}^{+*}. In other words, congestion has occurred in i^{th} input from DMU_O. The amount of congestion is s_{io}^{c*}.

3. A proposed method for eliminating input congestion using merging DMUs

Suppose that in a production system whose production technology is under inputs non-disposability, there are n DMUs. In addition, at least two of these n DMUs have congestion. Congested DMUs have been merged and a new DMU has been produced without congestion and a predefined efficiency target. Moreover, the merged units are removed after the merger. Using the inverse DEA of the new unit the inherited inputs and outputs are obtained from the merged DMUs. Also, the output-oriented radial efficiency of the new DMU in a congested production possibility set is predetermined by the manager. In the following, first, the phenomenon of congestion removal by merging is described with an example, and then a general model for congestion removal by merging is proposed.

3.1. **hypothetical example.** This section explains the proposed method with a hypothetical example. By merging two congested DMUs, a non-congested DMU with predefined efficiency is created. Table (1) shows 8 DMUs with one input and one output. The data set is taken from Cooper et al. [12].

Table 1. The data set of the hypothetical example.

DMUs	A	B	C	D	E	F	G	H
input	1	2	3	5	4	4	4.5	3
output	0.5	2	2	1	1	1.2	1.2	1

Using models (4) and (9) the efficiency amount: φ, and the congestion amount of 8 DMUs shown in Table (2).

Table 2. The efficiency and congestion amount of DMUs.

DMUs	A	B	C	D	E	F	G	H
φ^*	1	1	1	1	1.5	1.25	1.04	2
s_{io}^{c*}	0	0	0	2	1	1	1.5	0

According to Table (2), we find that the two DMUs E and F have congestion. Using the proposed model, two DMUs E and F have been merged a new merged DMU (M) has been produced as follows.

$$x_M = \alpha_E + \alpha_F,$$
$$y_M = \beta_E + \beta_F.$$

α_E and α_F are the inherited inputs of the merged new unit, from DMUs E and F, respectively. Also, β_E and β_F are the inherited outputs of the merged new unit, from DMUs E and F, respectively. To find the inherited inputs and outputs of the new DMU from units E and F, model (10) is proposed.

(10)

$$\begin{aligned}
&\min\ ((\alpha_E + \alpha_F) - (\beta_E + \beta_F))\\
&s.t.\ \lambda_A + 2\lambda_B + 3\lambda_C + 5\lambda_D + 4.5\lambda_G + 3\lambda_H + (\alpha_E + \alpha_F)\lambda_M = (\alpha_E + \alpha_F),\\
&0.5\lambda_A + 2\lambda_B + 2\lambda_C + \lambda_D + 1.2\lambda_G + \lambda_H + (\beta_E + \beta_F)\lambda_M = (\beta_E + \beta_F),\\
&\alpha_E + \alpha_F \le x_E - s_E^{C*} = 4 - 1,\\
&\alpha_E + \alpha_F \le x_F - s_F^{C*} = 4 - 1,\\
&\beta_E + \beta_F \le \max_{j \ne E,F} \{y_j\} = 2,\\
&\lambda_M \ge 0, \lambda_j \ge 0,\ j \ne E, F\\
&\alpha_E \ge 0,\ \alpha_F \ge 0,\ \beta_E \ge 0,\ \beta_F \ge 0,
\end{aligned}$$

The objective function of Model (10) determines the minimum input value of the merged unit, which is at least equal to inputs of the congested units before the merging, while their congestion amount is removed. It also obtains the maximum output of the integrated unit, which is at least equal to the maximum amount of the observed outputs. In Model (10), φ_M is pre-defined as follows.

$$1 \le \varphi_M = 1.1 \le \min\{\varphi_E^* = 1.5,\ \varphi_F^* = 1.25\}.$$

Because the merged new DMU is inefficient, then $\lambda_M = 0$, so the nonlinear model (10) becomes a linear model and the optimal solution is as follows:

$$\begin{aligned}
&\alpha_E^* = 2,\ \alpha_F^* = 0,\ \beta_E^* = 0,\ \beta_F^* = 1.82,\\
&x_M = \alpha_E^* + \alpha_F^* = 2,\\
&y_M = \beta_E^* + \beta_F^* = 1.82,
\end{aligned}$$

To use the resources of both units E and F in the production of a merged new DMU, (DMU_M) the weight vector in the objective function is used as follows:

$$\begin{aligned}
&\min\ ((v_E\alpha_E + v_F\alpha_F) - (w_E\beta_E + w_F\beta_F))\\
&v_E = 0.3,\ v_F = 0.7,\ w_E = 0.4,\ w_F = 06,
\end{aligned}$$

According to Figure (1), DMU_M is without congestion. In addition, units E and F must be removed from the congestion production possibility set. Figure (1) shows the production possibility set of the post-merger DMUs.

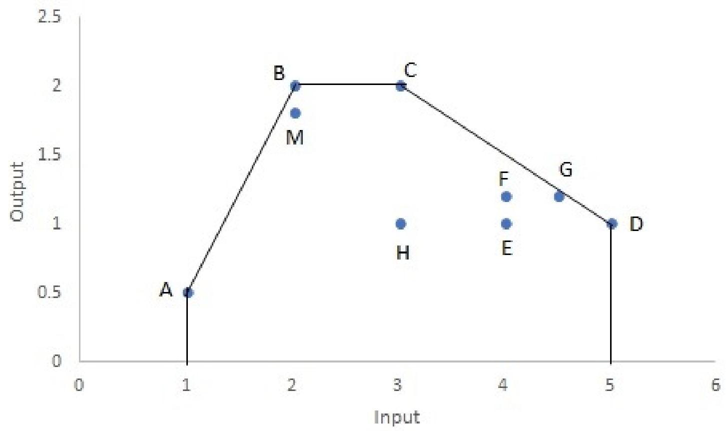

FIGURE 1. The production possibility set of the post-merger DMUs.

3.2. **A general model for eliminating inputs congestion.** In this section, a general method for eliminating input congestion using merging congested DMUs is described. Suppose there are n DMUs so that at least two DMUs such as E and F have congestion. The following multiple-objective model is proposed to integrate DMU_K and DMU_L and to specify the inputs and outputs of the merged new unit with predefined efficiency.

(11)

$$
\begin{aligned}
&\min\ (\alpha_{1K}+\alpha_{1L},\ldots,\alpha_{mK}+\alpha_{mL})\\
&\max\ (\beta_{1K}+\beta_{1L},\ldots,\beta_{sK}+\beta_{sL})\\
&s.t.\ \sum_{j\neq K,L}^{n}\lambda_j x_{ij}+(\alpha_{iK}+\alpha_{iL})\lambda_M=\alpha_{iK}+\alpha_{iL},\ i=1,2,\ldots,m,\\
&\sum_{j\neq K,L}^{n}\lambda_j y_{rj}+(\beta_{rK}+\beta_{rL})\lambda_M=\beta_{rK}+\beta_{rL},\ r=1,2,\ldots,s,\\
&\sum_{j\neq K,L}^{n}\lambda_j+\lambda_M=1,\\
&\alpha_{iK}+\alpha_{iL}\leq x_{iK}-s_{iK}^{C*},\ i=1,2,\ldots,m,\\
&\alpha_{iK}+\alpha_{iL}\leq x_{iL}-s_{iL}^{C*},\ i=1,2,\ldots,m,\\
&\beta_{rK}+\beta_{rL}\leq \max_{j\neq K,L}\{y_{rj}\},\ r=1,2,\ldots,s,\\
&\lambda_M\geq 0,\lambda_j\geq 0,\ j\neq K,\ L,\\
&\alpha_{iK}\geq 0,\ \alpha_{iL}\geq 0,\ \beta_{rK}\geq 0,\ \beta_{rL}\geq 0,\ \forall i,\ r.
\end{aligned}
$$

In model (11), $(\alpha_{iK}+\alpha_{iL}, \beta_{rK}+\beta_{rL}, \lambda_j, \lambda_M)$ is the vector of variables. $\alpha_{iK}+\alpha_{iL}$ is the i^{th} input of the inherited new unit from merging DMUs. $\beta_{rK}+\beta_{rL}$ is the r^{th} output of the inherited new unit from merging DMUs. λ_j is the j^{th} amount of intensity variable. φ_M is the predefined value as:

$$1 \leq \varphi_M \leq \min\{\varphi_K,\ \varphi_L\},$$

So φ_K and φ_L are the optimal values of model (4) in the assessment of units K and L, respectively. s_{iK}^{C*} and s_{iL}^{C*} are the amounts of i^{th} input congestion of DMU_K and DMU_L respectively, obtained by model (9). Using the weight-sum method, model (11) can be converted to the following nonlinear single-objective problem:

$$\min \sum_{i=1}^{m}(v_{iK}\alpha_{iK}+v_{iL}\alpha_{iL}) - \sum_{r=1}^{s}(w_{rK}\beta_{rK}+w_{rL}\beta_{rL}) *$$

Remark 3.1. For the new DMU not to be the same as other units, we can use the following constraints in the model (11).

$$\begin{array}{l} \lambda_j + a_j = 1,\ j=1,2,...,n, \\ a_j \geq \epsilon,\ j=1,2,...,n, \end{array}$$

where ϵ is a non- Archimedean small positive number.

Theorem 3.2. *If the new unit (DMU_M) is within the congestion production possibility set of the pre-merging DMUs, then the nonlinear multiple-objective model (11) converts the following linear multiple-objective model:*

$$\begin{array}{ll} & \min\ (\alpha_{1K}+\alpha_{1L},\ldots,\alpha_{mK}+\alpha_{mL}) \\ & \max\ (\beta_{1K}+\beta_{1L},\ldots,\beta_{sK}+\beta_{sL}) \\ & s.t.\ \sum_{j\neq K,L}^{n}\lambda_j x_{ij} = \alpha_{iK}+\alpha_{iL},\ i=1,2,\ldots,m, \\ & \sum_{j\neq K,L}^{n}\lambda_j y_{rj} = \beta_{rK}+\beta_{rL},\ r=1,2,\ldots,s, \\ (12) & \sum_{j\neq K,L}^{n}\lambda_j = 1, \\ & \alpha_{iK}+\alpha_{iL} \leq x_{iK} - s_{iK}^{C*},\ i=1,2,\ldots,m,\ (a) \\ & \alpha_{iK}+\alpha_{iL} \leq x_{iL} - s_{iL}^{C*},\ i=1,2,\ldots,m,\ (b) \\ & \beta_{rK}+\beta_{rL} \leq \max_{j\neq K,L}\{y_{rj}\},\ r=1,2,\ldots,s,\ (c) \\ & \lambda_j \geq 0,\ j\neq K,\ L, \\ & \alpha_{iK}\geq 0,\ \alpha_{iL}\geq 0,\ \beta_{rK}\geq 0,\ \beta_{rL}\geq 0,\ \forall i,\ r. \end{array}$$

Proof. If DMU_M is within the PPS_C of the pre-merging DMUs, then DMU_M is the internal or boundary unit of the PPS_C of the pre-merging DMUs.

If DMU_M is an internal unit of the PPS_C of the pre-merging DMUs, then DMU_M is produced by a convex combination of DMUs other than units K and L. Therefore, in each Pareto solution of model (11) $\lambda_M = 0$, and, model (11) becomes a linear model.

If DMU_M is a boundary unit of the PPS_C of the pre-merging DMUs, then the new unit is displayed as a convex combination of efficient DMUs. In this case, too, we have $\lambda_M = 0$, so model (11) converts to a linear model. □

Remark 3.3. In this study, according to theorem (2.4), the merged new unit must be within the PPS_C of pre-merging DMUs. Also, for the corresponding pre and post-merging congested production possibility sets not to change, the merged units must not be extreme.

Definition 3.4. An efficient congested unit is an efficient extreme unit if it is not obtained from a convex combination of other units. In other words, the optimal solution of the following model is zero.

$$
\begin{aligned}
& \max \sum_{j \neq o} \lambda_j \\
& s.t. \sum_{j=1}^{n} \lambda_j x_{ij} = x_{io}, \; i = 1, 2, \ldots, m, \\
& \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro}, \; r = 1, 2, \ldots, s, \\
& \sum_{j=1}^{n} \lambda_j = 1, \\
& \lambda_j \geq 0, \; j = 1, 2, \ldots, n,
\end{aligned} \tag{13}
$$

In model (13) DMU_O is an efficient congested unit of model (4).

Theorem 3.5. *Suppose that DMU_M is within the PPS_C of the pre-merging DMUs, also $(\alpha^*_{iK} + \alpha^*_{iL}, \beta^*_{rK} + \beta^*_{rL}, \lambda^*_j)$ is a Pareto solution of model (12) if*

$$
\begin{aligned}
x_{iM} &= \alpha^*_{iK} + \alpha^*_{iL}, \; i = 1, 2, \ldots, m, \\
y_{rM} &= \beta^*_{rK} + \beta^*_{rL}, \; r = 1, 2, \ldots, s.
\end{aligned}
$$

Then $(x_{iM}, y_{rM}) = DMU_M$ has no congestion. Moreover, the radial efficiency under the input non-possibility of DMU_M, in other words, the optimal amount of model (4) is φ_M.

Proof. To prove the no congestion of DMU_M the following model is considered:

$$
\begin{aligned}
&\max\ \eta + \varepsilon \sum_{r=1}^{s} s_{rM}^{+} - \varepsilon^{2} \sum_{i=1}^{m} s_{iM}^{c} \\
&s.t.\ \sum_{j\neq K,L}^{n} \lambda_j x_{ij} + \lambda_M x_{iM} + s_{iM}^{c} = x_{iM},\ i = 1,2,\dots,m, \\
&\sum_{j\neq K,L}^{n} \lambda_j y_{rj} + \lambda_M y_{rM} - s_{rM}^{+} = \eta y_{rM},\ r = 1,2,\dots,s, \\
&\sum_{j\neq K,L}^{n} \lambda_j + \lambda_M = 1, \\
&s_{iM}^{c} \geq 0,\ s_{rM}^{+} \geq 0,\ \lambda_M \geq 0,\ \lambda_j \geq 0,\ \forall i,\ r,\ j.
\end{aligned} \tag{14}
$$

Assume DMU_M has congestion, i.e. the optimal solution $(\tilde{\eta}, \tilde{\lambda}_j : j \neq K, L, \tilde{\lambda}_M, \tilde{s}_i^{C}, \tilde{s}_r^{+})$ f model (14) satisfies to at least one of the following two cases:

$$
\begin{aligned}
&\tilde{\eta} > 1,\ \exists i\ s.t.\ \tilde{s}_i^{C} > 0, \\
&\exists r\ \tilde{s}_r^{+} > 0,\ \exists i\ s.t.\ \tilde{s}_i^{C} > 0.
\end{aligned}
$$

Since the optimal solution is also feasible, we conclude from the constraints of model (14):

$$
\begin{aligned}
&\sum_{j\neq K,L}^{n} \tilde{\lambda}_j x_{ij} = x_{iM} - \tilde{s}_{iM}^{c},\ i = 1,2,\dots,m,\ (A1) \\
&\sum_{j\neq K,L}^{n} \tilde{\lambda}_j y_{rj} = \tilde{\eta} y_{rM} + \tilde{s}_{rM}^{+},\ r = 1,2,\dots,s,\ (A2) \\
&\sum_{j\neq K,L}^{n} \tilde{\lambda}_j = 1,\ (A3)
\end{aligned}
$$

Because DMU_M is congested, it is inefficient, as a result, in equations (A1), (A2), and (A3) $\tilde{\lambda}_M = 0$.

Because the feasible region of model (14) is a subset of the feasible region of model (4), it is that the radial efficiency in model (14) i.e. $(\tilde{\eta})$ is greater than or equal to the radial efficiency in model (4) i.e. (φ_M). In other words, we have $\tilde{\eta} \geq \varphi_M$.

According to the equation (A2):

$$\sum_{j\neq K,L}^{n} \tilde{\lambda}_j y_{rj} = \tilde{\eta} y_{rM} + \tilde{s}^+_{rM} \geq \varphi_M y_{rM} + \tilde{s}^+_{rM} = \varphi_M (y_{rM} + \frac{\tilde{s}^+_{rM}}{\varphi_M}) \ \Rightarrow$$
$$\sum_{j\neq K,L}^{n} \tilde{\lambda}_j y_{rj} \geq \varphi_M (y_{rM} + \frac{\tilde{s}^+_{rM}}{\varphi_M}). \ (A4)$$

On the other hand, according to the constraints of model (12) we have:

$$x_{iM} - \tilde{s}^{C*}_{iM} \leq x_{iK} - s^{C*}_{iK}, \ i = 1, 2, \ldots, m, \ (A5)$$
$$x_{iM} - \tilde{s}^{C*}_{iM} \leq x_{iL} - s^{C*}_{iL}, \ i = 1, 2, \ldots, m, \ (A6)$$

Since, the DMU_M projection i.e. $(x_{iM} - \tilde{s}^{C*}_{iM}, \ \tilde{\eta} y_{rM} + \tilde{s}^+_{rM})$ of model (14) is located in BCC and congestion production possibility sets, hence:

$$\tilde{\eta} y_{rM} + \tilde{s}^+_{rM} \leq \max_{j\neq K,L} \{y_{rj}\}, \ (A7)$$

$$\tilde{\eta} y_{rM} + \tilde{s}^+_{rM} \geq \varphi_M y_{rM} + \tilde{s}^+_{rM} = \varphi_M (y_{rM} + \frac{\tilde{s}^+_{rM}}{\varphi_M}) \geq y_{rM} + \frac{\tilde{s}^+_{rM}}{\varphi_M}, \ (A8)$$

By equation (A7) and (A8):

$$(y_{rM} + \frac{\tilde{s}^+_{rM}}{\varphi_M}) \leq \max_{j\neq K,L} \{y_{rj}\}, \ (A9)$$

According to equations (A1), (A2), (A3), (A4), (A5), (A6), and (A9) $(\bar{\alpha}_{iK} = \alpha^*_{iK} - \tilde{s}^c_{iM}; \forall i, \bar{\alpha}_{iL} = \alpha^*_{iL}; \forall i, \bar{\beta}_{rK} = \beta^*_{rK} + \frac{\tilde{s}^+_{rM}}{\varphi_M}; \forall r, \bar{\beta}_{rL} = \beta^*_{rL}; \forall r, \bar{\lambda}_j = \lambda^*_j; j \neq K, L)$ is a feasible solution of model (12), in which:

$$(\bar{\alpha}_{1K} + \bar{\alpha}_{1L}, \ldots, \bar{\alpha}_{mK} + \bar{\alpha}_{mL}) \lneqq (\alpha^*_{1K} + \alpha^*_{1L}, \ldots, \alpha^*_{mK} + \alpha^*_{mL})$$
$$(\bar{\beta}_{1K} + \bar{\beta}_{1L}, \ldots, \bar{\beta}_{sK} + \bar{\beta}_{sL}) \gneqq (\beta^*_{1K} + \beta^*_{1L}, \ldots, \beta^*_{sK} + \beta^*_{sL}).$$

This contradicts the assumption that $(\alpha^*_{iK} + \alpha^*_{iL}, \beta^*_{rK} + \beta^*_{rL}, \lambda^*_j)$ is a Pareto solution of model (12). Therefore, DMU_M has no congestion.

Now we prove that the optimal solution of model (4) in the evaluation of DMU_M is equal to φ_M.

Since $(\alpha^*_{iK} + \alpha^*_{iL}, \beta^*_{rK} + \beta^*_{rL}, \lambda^*_j)$ is the Pareto solution of model (12), hence:

$$\sum_{j\neq K,L}^{n} \lambda^*_j x_{ij} = \alpha^*_{iK} + \alpha^*_{iL} = x_{iM}, \ (A10),$$
$$\sum_{j\neq K,L}^{n} \lambda^*_j y_{rj} = \varphi_M (\beta^*_{rK} + \beta^*_{rL}) = \varphi_M y_{rM} \geq y_{rM}, \ (A11),$$
$$\sum_{j\neq K,L}^{n} \lambda^*_j = 1, \ (A12)$$

Therefore, $(\lambda_j^*; j \neq K, L, \lambda_M^* = 0, \varphi_M)$ is a feasible solution of model (4), and we prove that this is also the optimal solution.

Suppose the optimal solution of model (4) in the DMU_M assessment is: $(\bar{\lambda}_j; j \neq K, L, \bar{\lambda}_M, \bar{\varphi})$ where $\bar{\varphi} > \varphi_M$.

Therefore:

$$\sum_{j \neq K,L}^{n} \bar{\lambda}_j x_{ij} + \bar{\lambda}_M x_{iM} = x_{iM}, \ (A13),$$
$$\sum_{j \neq K,L}^{n} \bar{\lambda}_j y_{rj} + \bar{\lambda}_M y_{rM} = \bar{\varphi} y_{rM}, \ (A14),$$
$$\sum_{j \neq K,L}^{n} \bar{\lambda}_j + \bar{\lambda}_M = 1, \ (A15)$$

By placing equations (A10) and (A11) in equations (A13) and (A14), respectively, the following relations are obtained:

$$\sum_{j \neq K,L}^{n} \bar{\lambda}_j x_{ij} + \bar{\lambda}_M (\sum_{j \neq K,L}^{n} \lambda_j^* x_{ij}) = x_{iM} \Rightarrow \sum_{j \neq K,L}^{n} (\bar{\lambda}_j + \bar{\lambda}_M \lambda_j^*) x_{ij} = x_{iM},$$

We set:

$$\hat{\lambda}_j = \bar{\lambda}_j + \bar{\lambda}_M \lambda_j^*. \ (A16)$$

Therefore:

$$\sum_{j \neq K,L}^{n} \hat{\lambda}_j x_{ij} = x_{iM}. \ (A17)$$

$$\sum_{j \neq K,L}^{n} \bar{\lambda}_j y_{rj} + \bar{\lambda}_M (\sum_{j \neq K,L}^{n} \lambda_j^* y_{rj}) \geq \bar{\varphi} y_{rM} > \varphi_M y_{rM} \Rightarrow$$
$$\sum_{j \neq K,L}^{n} (\bar{\lambda}_j + \bar{\lambda}_M \lambda_j^*) y_{rj} > \varphi_M y_{rM},$$

By equation (A16) we have:

$$\sum_{j \neq K,L}^{n} \hat{\lambda}_j y_{rj} > \varphi_M y_{rM} \Rightarrow \sum_{j \neq K,L}^{n} \hat{\lambda}_j y_{rj} \geq \varphi_M y_{rM} + a_r = \varphi_M (y_{rM} + \frac{a_r}{\varphi_M}) \Rightarrow$$
$$\sum_{j \neq K,L}^{n} \hat{\lambda}_j y_{rj} \geq \varphi_M (y_{rM} + \frac{a_r}{\varphi_M}). \ (A18)$$

By equation (A12) and (A15) we have:

$$\sum_{j\neq K,L}^{n} \hat{\lambda}_j = \sum_{j\neq K,L}^{n} (\bar{\lambda}_j + \bar{\lambda}_M \lambda_j^*) = \sum_{j\neq K,L}^{n} \bar{\lambda}_j + \bar{\lambda}_M \sum_{j\neq K,L}^{n} \lambda_j^* = 1 \Rightarrow$$
$$\sum_{j\neq K,L}^{n} \hat{\lambda}_j = 1. \ (A19)$$

By equation (A14) we have that φy_{rM} belongs to the congested production possibility set i.e.

$$\varphi y_{rM} \leq \max_{j\neq K,L} \{y_{rj}\}, \ (A20),$$

Also, we have:

$$\bar{\varphi} y_{rM} > \varphi_M y_{rM} \Rightarrow \bar{\varphi} y_{rM} \geq \varphi_M y_{rM} + a_r = \varphi_M (y_{rM} + \frac{a_r}{\varphi_M}) \geq (y_{rM} + \frac{a_r}{\varphi_M}). \ (A21)$$

According to (A20) and (A21):

$$y_{rM} + \frac{a_r}{\varphi_M} \leq \max_{j\neq K,L} \{y_{rj}\}, \ (A22),$$

By equations (A17), (A18), (A19), (A22), and constraints (a) and (b) of model (12) it is obvious that

$$(\bar{\alpha}_{iK} = \alpha_{iK}^*; \forall i, \bar{\alpha}_{iL} = \alpha_{iL}^*; \forall i, \bar{\beta}_{rk} = \beta_{rk}^* + \frac{a_r}{\varphi_M}; \forall r, \bar{\beta}_{rL} = \beta_{rL}^*; \forall r, \bar{\lambda}_j = \hat{\lambda}_j; j \neq K, L, \bar{\lambda}_M = 0),$$

is a feasible solution of model (12), in which:

$$(\bar{\beta}_{1K} + \bar{\beta}_{1L}, \ldots, \bar{\beta}_{sK} + \bar{\beta}_{sL}) \gneqq (\beta_{1K}^* + \beta_{1L}^*, \ldots, \beta_{sK}^* + \beta_{sL}^*).$$

This contradicts the assumption that $(\alpha_{iK}^* + \alpha_{iL}^*, \beta_{rK}^* + \beta_{rL}^*, \lambda_j^*)$ is a Pareto solution of model (12). Therefore, efficiency of DMU_M is equal φ_M. □

Remark 3.6. The new unit produced with the model (12) is such that the current PPS does not change, in other words, the efficiency and congestion of other units are unaffected by the new unit. For this purpose, the constraints of Model (12) must be adjusted in such a way that model (12) is feasible. The input constraints ((a) and (b)) cause the new unit produced in the non-congested region, and the output constraint (c) finds the maximum possible output of the new unit in the non-congested region. If the produced new unit changes the efficiency frontier, in this case, the model (15) is used.

Remark 3.7. If $\beta_{rK} + \beta_{rL} \geq y_{rK} + y_{rL}$ constraint is added to the model (12), model (12) becomes infeasible because the units before merging are congested. Our goal in merging units is to eliminate congestion and improve efficiency. Therefore, to reach these targets, the inputs are reduced and the outputs are increased. However, the increase in the outputs of the new unit is not necessarily equal to the outputs of the pre-merging DMUs. Rather, the increase in the outputs of the new unit is the amount that the new unit is located in the PPS of pre-merging DMUs.

When model (12) is infeasible, i.e., the new unit produced is not in the PPS of per-merging DMUs, in this case, according to Amin et al. [12], the following variable change is used in model (11) to solve this problem.

$$\alpha_{iq}\lambda_M = \hat{\alpha}_{iq},\ i = 1, 2, \ldots, m,\ q = K, L,$$
$$\beta_{rq}\lambda_M = \hat{\beta}_{rq},\ r = 1, 2, \ldots, s,\ q = K, L,$$

where

$$\begin{aligned}
&\lambda_M \in \{0, 1\},\\
&\hat{\alpha}_{iq} \le \alpha_{iq} x_{iq},\ i = 1, 2, \ldots, m,\ q = K, L,\\
&\alpha_{iq} - (1 - \lambda_M)x_{iq} \le \hat{\alpha}_{iq} \le \alpha_{iq},\ i = 1, 2, \ldots, m,\ q = K, L,\\
&\hat{\beta}_{rq} \le \beta_{rq} y_{rq},\ r = 1, 2, \ldots, s,\ q = K, L,\\
&\beta_{rq} - (1 - \lambda_M)y_{rq} \le \hat{\beta}_{rq} \le \beta_{rq},\ r = 1, 2, \ldots, s,\ q = K, L.
\end{aligned}$$

Therefore, the nonlinear model (11) is linearized as follows:
(15)

$$\begin{aligned}
&\min\ (\alpha_{1K} + \alpha_{1L}, \ldots, \alpha_{mK} + \alpha_{mL})\\
&\max\ (\beta_{1K} + \beta_{1L}, \ldots, \beta_{sK} + \beta_{sL})\\
&s.t.\ \sum_{j \ne K,L}^{n} \lambda_j x_{ij} + (\hat{\alpha}_{iK} + \hat{\alpha}_{iL})\lambda_M = \alpha_{iK} + \alpha_{iL},\ i = 1, 2, \ldots, m,\\
&\sum_{j \ne K,L}^{n} \lambda_j y_{rj} + (\hat{\beta}_{rK} + \hat{\beta}_{rL})\lambda_M = \beta_{rK} + \beta_{rL},\ r = 1, 2, \ldots, s,\\
&\sum_{j \ne K,L}^{n} \lambda_j + \lambda_M = 1,\\
&\alpha_{iK} + \alpha_{iL} \le x_{iK} - s_{iK}^{C*},\ i = 1, 2, \ldots, m,\\
&\alpha_{iK} + \alpha_{iL} \le x_{iL} - s_{iL}^{C*},\ i = 1, 2, \ldots, m,\\
&\beta_{rK} + \beta_{rL} \le \max_{j \ne K,L}\{y_{rj}\},\ r = 1, 2, \ldots, s,\\
&\hat{\alpha}_{iq} \le \alpha_{iq} x_{iq},\ i = 1, 2, \ldots, m,\ q = K, L,\\
&\alpha_{iq} - (1 - \lambda_M)x_{iq} \le \hat{\alpha}_{iq} \le \alpha_{iq},\ i = 1, 2, \ldots, m,\ q = K, L,\\
&\hat{\beta}_{rq} \le \beta_{rq} y_{rq},\ r = 1, 2, \ldots, s,\ q = K, L,\\
&\beta_{rq} - (1 - \lambda_M)y_{rq} \le \hat{\beta}_{rq} \le \beta_{rq},\ r = 1, 2, \ldots, s,\ q = K, L.\\
&\lambda_M \in \{0, 1\}, \lambda_j \ge 0,\ j \ne K,\ L,\\
&\alpha_{iK} \ge 0,\ \alpha_{iL} \ge 0,\ \beta_{rK} \ge 0,\ \beta_{rL} \ge 0,\ \forall i,\ r.
\end{aligned}$$

Model (15) is a linear integer programming model. When the PPSs pre and post-merging of DMUs are not the same, model (15) can be used. Also, when the constraint $\beta_{rK} + \beta_{rL} \geq y_{rK} + y_{rL}$ is added to model (12), model (15) can be employed.

Remark 3.8. We can extend the proposed model in this study to merge more than two congested DMUs as follows:

$$
\begin{aligned}
&\min\ (\alpha_{ij},\ i = 1,2,\ldots,m,\ \forall j \in C)\\
&\max\ (\beta_{rj},\ r = 1,2,\ldots,s,\ \forall j \in C)\\
&s.t.\ \sum_{j\in Q}\lambda_j x_{ij} + \lambda_M \sum_{j\in C}\alpha_{ij} = \sum_{j\in C}\alpha_{ij},\ i = 1,2,\ldots,m,\\
&\sum_{j\in Q}\lambda_j y_{rj} + \lambda_M \sum_{j\in C}\beta_{rj} \geq \varphi_M \sum_{j\in C}\beta_{rj},\ r = 1,2,\ldots,s,\\
&\sum_{j\in C}\lambda_j + \lambda_M = 1,\\
&\sum_{j\in C}\alpha_{ij} \leq x_{ij} - s_{ij}^{C*},\ i = 1,2,\ldots,m, \forall j \in C,\\
&\sum_{j\in C}\beta_{rj} + \beta_{rL} \leq \max_{j\in Q}\{y_{rj}\},\ r = 1,2,\ldots,s,\\
&\lambda_M \geq 0, \lambda_j \geq 0,\ j = 1,2,\ldots,n,\\
&\alpha_{ij} \geq 0,\ \beta_{rj} \geq 0,\ \forall i,\ r, j \in C.
\end{aligned}
\tag{16}
$$

In model (16) we define:

$$
\begin{aligned}
C &= \{DMU_j \mid j \in congestedDMUs\},\\
Q &= \{DMU_j \mid j = 1,2,\ldots,n\} - C.
\end{aligned}
$$

Also, in the model (16) s_{ij}^{c*} is the amount of i^{th} input congestion of j^{th} congested unit. α_{ij} and β_{rj} are the new unit the inherited i^{th} input and r^{th} output from j^{th} congested DMU respectively. φ_M is the predefined value as:

$$1 \leq \varphi_M \leq \min_{j\in C}\{\varphi_j\},$$

If new DMU i.e. $DMU_M = (\sum_{j\in C}\alpha_{ij}; \forall i,\ \sum_{j\in C}\beta_{rj}; \forall r)$ is within PPS_C of pre-merging DMUs, then the nonlinear multiple-objective model (16) converts the linear multiple-objective model (17):

$$
\begin{aligned}
&\min\ (\alpha_{ij},\ i=1,2,\ldots,m,\ \forall j\in C)\\
&\max\ (\beta_{rj},\ r=1,2,\ldots,s,\ \forall j\in C)\\
&s.t.\ \sum_{j\in Q}\lambda_j x_{ij}=\sum_{j\in C}\alpha_{ij},\ i=1,2,\ldots,m,\\
&\sum_{j\in Q}\lambda_j y_{rj}\geq \varphi_M\sum_{j\in C}\beta_{rj},\ r=1,2,\ldots,s,\\
&\sum_{j\in C}\lambda_j=1,\\
&\sum_{j\in C}\alpha_{ij}\leq x_{ij}-s_{ij}^{C*},\ i=1,2,\ldots,m,\forall j\in C,\\
&\sum_{j\in C}\beta_{rj}+\beta_{rL}\leq \max_{j\in Q}\{y_{rj}\},\ r=1,2,\ldots,s,\\
&\lambda_j\geq 0,\ j=1,2,\ldots,n,\\
&\alpha_{ij}\geq 0,\ \beta_{rj}\geq 0,\ \forall i,\ r,j\in C.
\end{aligned}
\tag{17}
$$

4. **Empirical example**

This section presents an application of the introduced method in CAS research institutes 2010. The dataset is taken from Yang et al. [53]. The dataset includes 16 DMUs, and each DMU has two inputs to generate four outputs. The inputs are: the full-time equivalent of full-time research staff (x_1) and the amount of total income of each institute (x_2), as well as outputs are: the number of international papers indexed by the Web of Science from Thompson Reuters (y_1), the number of high-quality papers published in top research (y_2), the number of graduate student enrolment in 2009 (y_3), and the amount of external research funding from research contracts (y_4), which are displayed in Table 3.

Using models (9) and (13) we find DMU_3, DMU_8, DMU_9, DMU_{10}, DMU_{11}, DMU_{12}, DMU_{15}, and DMU_{16} have congestion. Moreover, DMU_8, DMU_9, DMU_{10}, DMU_{11}, and DMU_{15} are extremely congested units. Radial efficiency in the output-oriented of congested DMUs is determined by the model (4). The congestion and efficiency values of congested DMUs are shown in Table 4.

In Table 4, s_i^c is the i^{th} input surplus caused by congestion. In other words, s_i^c is the amount of inefficiency at the i^{th} input due to congestion. By merging two DMU12 and DMU16 using model (12), a new DMU without congestion and with the pre-defined efficiency was obtained. To use the inputs and outputs of both DMUs 12 and 16 in the new DMU, the weighted objective function (*)

Table 3. The data set of 16 CAS research institutes in 2010.

DMUs	x_1	x_2	y_1	y_2	y_3	y_4
1	252	117.945	436	133	184	31.558
2	37	29.431	243	127	43	15.3041
3	240	101.425	164	70	89	33.8365
4	356	368.483	810	276	247	183.8434
5	310	195.862	200	55	111	12.9342
6	201	188.829	104	49	33	60.7366
7	157	131.301	113	49	45	72.5368
8	236	77.439	8	1	44	23.7015
9	805	396.905	371	118	89	216.9885
10	886	411.539	607	216	168	88.5561
11	623	221.428	314	49	89	45.3597
12	560	264.341	261	79	131	41.1156
13	1344	900.509	627	168	346	645.4150
14	508	344.312	971	518	335	205.4528
15	380	161.331	395	180	117	90.0373
16	132	83.972	229	138	62	32.6111

Table 4. The congestion and efficiency values of congested DMUs.

Congested DMUs	The amount of congestion (s_i^C)		The amount of efficiency(φ)
3	$s_1^C = 71.68$	$s_2^C = 0$	1.18
8	$s_1^C = 127.16$	$s_2^C = 0$	1 (extreme efficient)
9	$s_1^C = 216.80$	$s_2^C = 0$	1 (extreme efficient)
10	$s_1^C = 378.00$	$s_2^C = 67.23$	1 (extreme efficient)
11	$s_1^C = 298.81$	$s_2^C = 0$	1 (extreme efficient)
12	$s_1^C = 142.44$	$s_2^C = 0$	1.45
15	$s_1^C = 145.45$	$s_2^C = 0$	1 (extreme efficient)
16	$s_1^C = 13.42$	$s_2^C = 0$	1.34

with the following weights was utilized.

$$v_{i12} = (0.4, 0.6), w_{r12} = (0.2, 0.1, 0.4, 0.3),$$
$$v_{i12} = (0.3, 0.7), w_{r12} = (0.3, 0.4, 0.2, 0.1).$$

The inputs and outputs amount of the merged DMU with pre-defined efficiency in two cases ($\varphi_M = 1.2$ and $\varphi_M = 1$) are displayed in Table 5.

In the case that $\varphi_M = 1$, the results of Table 5 show that the amounts of the inherited inputs for the merged new DMU are $0 + 118.58$ and $83.67 + 0$ and the inherited outputs are $0 + 369.09$, $0 + 194.72$, $93.58 + 0$, and $48.24 + 0$. In other words:

Table 5. The inputs and outputs amount of the merged DMU with pre-defined efficiency.

pre-defined efficiency	Inherited inputs of the merged DMU				Inherited outputs of the merged DMU							
φ_M	$\alpha_{1,12}$	$\alpha_{2,12}$	$\alpha_{1,16}$	$\alpha_{2,16}$	$\beta_{1,12}$	$\beta_{2,12}$	$\beta_{3,12}$	$\beta_{4,12}$	$\beta_{1,16}$	$\beta_{2,16}$	$\beta_{3,16}$	$\beta_{4,16}$
$\varphi_M = 1$	0	83.97	118.58	0	0	0	93.58	48.24	369.09	194.72	0	0
$\varphi_M = 1.2$	0	83.97	118.58	0	0	0	77.98	40.20	307.58	162.27	0	0

$$x_M = (118.58, 83.97),\ y_{r1M} = (369.09, 194.72, 93.58, 48.24).$$

In addition, when $\varphi_M = 1.2$, the inputs and outputs of DMU_M are obtained as follows:

$$x_M = (118.58, 83.97),\ y_{r1M} = (307.58, 162.27, 77.98, 40.20).$$

In both cases, with less staff and less revenue spent on research, a new unit has been created without congestion and with improved efficiency. In other words, two research institutes with input surplus have been merged in such a way that the merged research institute does not have the input surplus and its efficiency has increased compared to the two research institutions before the merger.

5. Conclusion

There are many studies on identifying and determining the congestion amount of the inputs of DMUs, but there are fewer studies on reducing or eliminating congestion. To the best of the authors' knowledge, the only study on congestion removal has been done by Kao [32]. Kao [32], using an example and without proof, has studied the elimination of congestion in some congested units. In this paper, using inverse DEA and multiple-objective planning, the activities at least of two congestion DMUs are combined, and a new merged DMU is created. The feature of the new unit is that, firstly, it has no congestion. Secondly, the efficiency of the new unit has been improved in the congested production possibility set. The advantage of the suggested strategy is that, since congestion is a sort of inefficiency, removing congested DMUs benefits production systems. Also, by removing congestion, improper allocation of resources is prevented. Another advantage of the proposed method is that it has used related theorems to prove the claim. Moreover, a real case in CAS research institutes has been utilized to show the suggested method's capacity. According to this paper, the future research topics are as follows:

1- Examining the congestion and efficiency of other DMUs when PPS changes post-merging of DMUs.

2- Providing methods to reduce or eliminate the congestion of DMUs in dynamic DEA.

Acknowledgment

We would like to thank Editor-in-Chief and anonymous reviewers for their thoughtful comments and efforts towards improving our manuscript.

Conflict of interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] Amin, G. R., Emrouznejad, A., Gattoufi, S., (2017). Minor and major consolidations in inverse dea: Definition and determination. Computers Industrial Engineering 103, 193-200. https://doi.org/10.1016/j.cie.2016.11.029

[2] Amin, G. R., Emrouznejad, A., Gattoufi, S., (2017). Modelling generalized firms' restructuring using inverse dea. Journal of Productivity Analysis 48, 51-61. https://doi.org/10.1007/s11123-017-0501-y

[3] Amin, G. R., Qukil, A., (2019). Flexible target setting in mergers using inverse data envelopment analysis. International Journal of Operational Research 35, 301-317. https://doi.org/10.1504/ijor.2019.10022710

[4] Asgharian, M., Khodabakhshi, M., Neralic, L., (2010). Congestion in stochastic data envelopment analysis: an input relaxation approach. International Journal of Statistics and Management System 5 (1), 84-106.

[5] Avkiran, N., (1999). The evidence on efficiency gains: the role of mergers and the benefits to the public. Journal of Banking & Finance 23 (7), 991-1013. https://doi.org/10.1016/S0378-4266(98)00129-0

[6] Banker, R. D., Charnes, A., Cooper, W. W., (1984). Some models for estimating technical and scale efficiencies in data envelopment analysis. Management Science 30, 1078-1092. http://www.jstor.org/stable/2631725

[7] Bogetoft, P., Strange, N., Thorsen, B., (2003). Efficiency and merger gains in the danish forestry extension service. Forest Science 49 (4), 585-595. https://doi.org/10.1093/forestscience/49.4.585

[8] Bogetoft, P., Wang, D., (2005). Estimating the potential gains from mergers. Journal of Productivity Analysis 23, 145-171. https://doi.10.22004/ag.econ.24213

[9] Brockett, P. L., Cooper, W. W., Shin, H. C.,Wang, Y., (1998). Inefficiency and congestion in chinese production before and after the 1978 economic reforms. Socio- Economic planning sciences 32, 1-20. https://doi.org/10.1016/S0038-0121(97)00020-7

[10] Charnes, A., Cooper, W. W., Rhodes, E., (1978). Measuring the efficiency of decision making units. European Journal of Operational Research 2, 429-444. https://doi.org/10.1016/0377-2217(78)90138-8

[11] Chen, L., Wang, Y., Lai, F., Feng, F., (2017). An investment analysis for china's sustainable development based on inverse data envelopment analysis. Journal of Cleaner Production 142 (4), 1638-1649. https://doi.org/10.1016/j.jclepro.2016.11.129

[12] Cooper, W. W., Deng, H., Huang, Z. M., Li, S. X., (2002). One model approach to congestion in data envelopment analysis. Socio-Economic planning sciences 36, 231-238. https://doi.org/10.1016/S0038-0121(02)00008-3

[13] Cooper, W. W., Thompson, R. G., Thrall, R. M., (1996). Introduction: extensions and new developments in dea. Annals of operations research 66, 3-45. https://doi.org/10.1007/BF02125451

[14] De Witte, K., Dijkgraaf, E., (2010). Mean and bold? on separating merger economies from structural efficiency gains in the drinking water sector. Journal of the Operational Research Society 61 (2), 222-234. https://doi.org/10.1057/jors.2008.129

[15] Ebrahimzadeh Adimi, M., Rostamy-Malkhalifeh, M., Hossienzadeh lotfi, F., Mehrjoo, R., (2019). A new linear method to find the congestion hyperplane in dea. Mathematical Sciences 13, 43-52. https://doi.org/10.1007/s40096-019-0277-5

[16] Ehrgott, M., (2005). Multicriteria optimization. Springer Berlin.

[17] Emrouznejad, A., Amin, G. R., Ghiyasi, M., Michali, M., (2023). A review of inverse data envelopment analysis: origins, development and future directions. IMA Journal of Management Mathematics 34 (3), 421-440. https://doi.org/10.1093/imaman/dpad006

[18] Fang, L., (2015). Congestion measurement in nonparametric analysis under the weakly disposable technology. European Journal of Operational Research 245, 203-208. https://doi.org/10.1016/j.ejor.2015.03.001

[19] Färe, R., Grosskopf, S., (1983). Measuring congestion in production. Zeitschr. f. Nationalökonomie 43 (3), 257-271. https://doi.org/10.1007/BF01283574

[20] Färe, R., Grosskopf, S., Lovel, C. A. K., (1985). The measurement of efficiency of production. Kluwer-Nijhoff Publishing, Boston.

[21] Färe, R., Svensson, L., (1980). Congestion of production factors. The Econometric Society 48, 1745-1752.

[22] Fuentes, R., Bellver-Domingo, Á., Hernánndez-Chover, V., Hernánndez- Sancho, F., (2020). Identification and correction of congestion in wastewater treatment plants in the community of valencia, spain. Zeitschr. f. Nationalökonomie 27, 15729-15742. https://doi.org/10.1007/s11356-020-08063-1

[23] Gattoufi, S., Amin, G. R., Emrouznejad, A., (2014). A new inverse dea method for merging banks. IMA Journal of Management Mathematics 25, 73-87. https://doi.org/10.1093/imaman/dps027

[24] Ghobadi, S., (2021). Merging decision-making units with interval data. RAIRO-Operations Research 55, 1605-1630. https://doi.org/10.1051/ro/2020029

[25] Gholoom, S., Zervopoulos, P., (2023). The effect of mergers and acquisitions on efficiency: evidence from the pharmaceutical industry. Advances in Science and Technology 129, 77-95. https://doi.org/10.4028/p-qLBAB6

[26] Hadi-Vencheh, A., Foroughi, A. A., (2006). A generalized dea model for inputs/ outputs estimation. Mathematical and Computer Modelling 43 (5-6), 447-457. https://doi.org/10.1016/j.mcm.2005.08.005

[27] Hajaji, H., Yousefi, S., Sean, R. F., Hassanzadeh, A., (2021). Recommending investment opportunities given congestion by adaptive network data envelopment analysis model: assessing sustainability of supply chains. RAIRO - Operations Research 55, 21-49. https://doi.org/10.1051/ro/2019059

[28] Halkos, G., Tzeremes, N., (2013). Estimating the degree of operating efficiency gains from a potential bank merger and acquisition: a dea bootstrapped approach. Journal of Banking & Finance 37 (5), 1658-1668. https://doi.org/10.1016/j.jbankfin.2012.12.009

[29] Jahanshahloo, G. R., khodabakhshi, M., (2004). Suitable combination of input for improving outputs in dea with determining input congestion: considering textile industry of china. Applied Mathematics and Computation 151 (1), 263-273. https://doi.org/10.1016/S0096-3003(03)00337-0

[30] Jahanshahloo, G. R., Lotfi, F. H., Shoja, N., Tohidi, G., Razavyan, S., (2004). Input estimation and identification of extra inputs in inverse dea models. Applied Mathematics and Computation 156 (2), 427-437. https://doi.org/10.1016/j.amc.2003.08.001

[31] Jahanshahloo, G. R., Soleimani-damaneh, M., Ghobadi, S., (2015). Inverse dea under inter-temporal dependence using multiple-objective programming. European Journal of Operational Research 240, 447-456. https://doi.org/10.1016/j.ejor.2014.07.002

[32] Kao, C., (2010). Congestion measurement and elimination under the framework of data envelopment analysis. International Journal of Production Economics 123 (2), 257-265. https://doi.org/10.1016/j.ijpe.2009.06.044

[33] Kheirollahi, H., Hessari, P., (2017). An input relaxation model for evaluating congestion in fuzzy dea. Croatian Operational Research Review 8 (2), 391-408. DOI: 10.17535/crorr.2017.0025

[34] Khezri, S., Dehnokhalaji, A., Hosseinzadeh Lotfi, F., (2021). A full investigation of the directional congestion in data envelopment analysis. RAIRO-Operations Research 55, 571-591. https://doi.org/10.1051/ro/2019092

[35] Khoveyni, M., Eslami, R., Yang, G., (2017). Negative data in dea: Recognizing congestion and specifying the least and the most congested decision making units. Computers & Operations Research 79, 39-48. https://doi.org/10.1016/j.cor.2016.09.002

[36] Khoveyni, M., Eslami, R., khodabakhshi, M., Jahanshahloo, G. R., Hosseinzadeh Lotfi, F., (2013). Recognizing strong and weak congestion slack based in data envelopment analysis. Computers & Industrial Engineering 64 (2), 731-738. DOI: 10.1016/j.cie.2012.11.014

[37] Kong, C., Chow, W., Ka, M., Fung, Y., (2012). Measuring the effects of china's airline mergers on the productivity of state-owned carriers. Journal of Air Transport Management 25, 1-4. DOI: 10.1016/j.jairtraman.2011.08.006

[38] Mehdiloozad, M., Zhu, J., Sahoo, B., (2018). Identification of congestion in data envelopment analysis under the occurrence of multiple projections: A reliable method capable of dealing with negative data. European Journal of Operational Research 265 (2), 644-654. DOI: 10.1016/j.ejor.2017.07.065

[39] Oukil, A., (2022). Investigating prospective gains from mergers in the agricultural sector through inverse dea. IMA Journal of Management Mathematics 34 (3), 465-490. DOI: 10.1093/imaman/dpac004

[40] Ren, X. T., Fukuyama, H., Yang, G., (2022). Eliminating congestion by increasing inputs in r&d activities of chinese universities. Omega 110, 102618. https://doi.org/10.1016/j.omega.2022.102618

[41] Shadab, M., Saati, S., Farzipoor Saen, R., Khoveyni, M., Mostafaee, A., (2021). Measuring congestion in sustainable supply chain based on data envelopment analysis. Neural Computing and Applications 33, 12477-12491. https://doi.org/10.1007/s00521-021-05889-9

[42] Shahsavan, T., Sanei, M., Tohidi, G., Lotfi, F. H., Ghobadi, S., (2022). A new method of determining decision-making unit congestion under inter-temporal dependence. Soft Comput 26, 2063-2073. https://doi.org/10.1007/s00500-021-06566-8

[43] Shahsavan, T., Sanei, M., Tohidi, G., Lotfi, F. H., Ghobadi, S., (2023). Determining the amount of the excess input and the output shortage of the congested decision-making units with negative data. Mathematical Sciences 17, https://doi.org/10.1007/s40096-023-00511-6.

[44] Sharma, M. J., Song Jin, Y., (2013). Multi-stage data envelopment analysis congestion model. Operational Research 13, 399-413. https://doi.org/10.1007/s12351-012-0128-8

[45] Shiri Daryani, Z., Tohidi, G., Daneshian, B., Razavyan, S., Hosseinzadeh Lotfi, F., (2023). Inverse data envelopment analysis for merging two-stage network systems. International Journal of Applied Operational Research 11 (4), 39-49. http://ijorlu.liau.ac.ir/article-1-647-en

[46] 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. https://doi.org/10.1093/imaman/dpac001

[47] Sueyoshi, T., Sekitani, K., (2009). Dea congestion and returns to scale under an occurrence of multiple optimal projections. European Journal of Operational Research 194 (2), 592-607. https://ideas.repec.org/a/eee/ejores/v194y2009i2p592-607

[48] Tone, K., Sahoo, B. K., (2004). Degree of scale economies and congestion: A unified dea approach. European Journal of Operational Research 158 (3), 755-772. https://doi.org/10.1016/S0377-2217(03)00370-9

[49] Wang, Z., Wu, X., Lo, K., Mi, J., (2021). Assessing the management efficiency of shipping company from a congestion perspective: A case study of hapag-lloyd. Ocean Coast Manag 209, 4-23. DOI: 10.1016/j.ocecoaman.2021.105617

[50] Wei, Q. L., Yan, H., (2004). Congestion and returns to scale in data envelopment analysis. European Journal of Operational Research 153 (3), 641-660. DOI: 10.1016/S0377-2217(02)00799-3

[51] Wei, Q. L., Zhang, J. Z., Zhang, X. S., (2000). An inverse dea model for inputs/outputs estimate. European Journal of Operational Research 121 (1), 151-163. https://doi.org/10.1016/S0377-2217(99)00007-7

[52] Wu, D., Birge, J.,(2011). Serial chain merger evaluation model and application to mortgage banking. Decision Sciences 43 (1), 5-360. https://doi.org/10.1111/j.1540-5915.2011.00340.x

[53] Yang, G., Ren, X., Khoveyni, M., Eslami, R., (2020). Directional congestion in the framework of data envelopment analysis. Journal of Management Science and Engineering 5 (1), 57-75. https://doi.org/10.1016/j.jmse.2020.02.001

[54] Yang, Z., Shi, Y., Yan, H., (2017). Analysis on pure e-commerce congestion effect, productivity effect and profitability in China. Socio-Economic Planning Sciences 57, 35-49. DOI: 10.1016/j.seps.2016.08.002

[55] Zenodin, E., Ghobadi, S., (2020). Merging decision-making units under intertemporal dependence. IMA Journal of Management Mathematics 31 (2), 139-16. https://doi.org/10.1093/imaman/dpz005

[56] Zhang, Y. J., Liu, J. Y., Su, B., (2020). Carbon congestion effects in china's industry: Evidence from provincial and sectoral levels. Energy Economics 86, 104635. DOI: 10.1016/j.eneco.2019.104635

Tahereh Shahsavan
Orcid number: 0009-0007-3639-3842
Department of Mathematics
Central Tehran Branch, Islamic Azad University,
Tehran, Iran.
Email address: t.shahsavan@yahoo.com

Masoud Sanei
Orcid number: 0000-0002-2097-9354
Department of Mathematics
Central Tehran Branch, Islamic Azad University,
Tehran, Iran.
Email address: masoudsanei49@yahoo.com

Ghasem Tohidi
Orcid number: 0000-0003-0199-3465
Department of Mathematics
Central Tehran Branch, Islamic Azad University,
Tehran, Iran.
Email address: gh.tohidi@yahoo.com

Farhad Hosseinzadeh Lotfi
Orcid number: 0000-0001-5022-553X
Department of Mathematics
Science and Research Branch, Islamic Azad University,
Tehran, Iran.
Email address: farhad@hosseinzadeh.ir

Saeid Ghobadi
Orcid number: 0000-0002-1884-0767
Department of Mathematics
Khomeinishahr Branch, Islamic Azad University,
Isfahan, Iran.
Email address: Email: Ghobadi@iaukhsh.ac.ir