Journal of Intelligent Decision and Computational Modelling

www.jidcm.reapress.com

J. Intell. Decis. Comput. Model. Vol. 1, No. 2 (2025) 107–137.

Paper Type: Original Article

A Glimpse of Input Congestion Methods in Data Envelopment Analysis

Faezeh Jokar[1], Mohammad Jalali Varnamkhasti[1,*], Abdollah Hadi Vencheh[1]

[1] Department of Science, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran; jokar-f@gmail.com; jalali.m.v@gmail.com; ahadi@khuisf.ac.ir.

Citation:

Received: 14 August 2024 Revised: 16 October 2024 Accepted: 22 December 2024	Jokar, F., & Hadi Vencheh, A. (2025). A glimpse of input congestion methods in data envelopment analysis. *Journal of intelligent decision and computational modelling, 1*(2), 107-137.

Abstract

This comprehensive review explores recent advances in measuring and analyzing congestion within Data Envelopment Analysis (DEA), a pivotal tool in efficiency assessment of Decision-Making Units (DMUs). The study categorizes multiple methodologies developed over the years, focusing on both classical and novel models to detect, quantify, and interpret congestion situations where input increases lead to output reductions or vice versa, indicating inefficiencies or overutilization. These approaches include input-oriented, output-oriented, multi-stage, weight-restriction, and models addressing undesirable outputs, as well as those accommodating integer data and production trade-offs. Several models utilize concepts such as Pareto efficiency, slack variables, and weight restrictions to formalize congestion detection, with specific attention to strong, weak, and wide congestion phenomena. The synthesis includes algorithms for projecting DMUs onto efficiency or congestion boundaries, alongside measures for the extent of congestion. Furthermore, innovative techniques addressing multiple stages, undesirable outputs, and integer data expand the applicability of DEA in complex, real-world scenarios. The findings offer researchers a robust, categorically organized toolkit for congestion analysis, advancing the understanding of resource overuse and production bottlenecks, thereby contributing to more accurate efficiency measurement and resource management strategies in diverse sectors.

Keywords: Data envelopment analysis, Decision-making units, Input-oriented, Output-oriented, Multi-stage.

1 | Introduction

Data Envelopment Analysis (DEA) is a branch of management interested in evaluating the efficiency of homogeneous Decision-Making Units (DMUs). Charnes, Cooper, and Rhodes (CCR) [1] developed DEA in 1978 in their famous article. Since 1978, there has been a spurt of broad searches on the DEA. Today, many scholars all over the world are working in this domain. The performances of DMUs are affected by the

Corresponding Author: jalali.m.v@gmail.com
https://doi.org/10.48314/jidcm.v1i2.68

number of origins that DMUs use. Usually, increases in inputs cause increases in outputs. But there are situations where an increase in one or more inputs generates a reduction in one or more outputs. For example, in an underground coal mine, too many men decrease the output of coal. In such situations, there is congestion in inputs or the production process. The definition used in this research is as follows:

Definition 1.

Input: A set of resources available to the DMU, denoted by the vector X= $(x_1, x_2, \dots\dots, x_m)$.

Output: A set of items produced from the input vector X by the DMU, denoted by the vector Y= $(y_1, y_2, \dots\dots, y_m)$.

Efficiency-extended Pareto-Koopmans

Full (100%) efficiency is attained by any DMU if and only if none of its inputs or outputs can be improved without worsening some of its other inputs or outputs. If the performances of other DMUs do not show that some of their inputs or outputs can be improved without worsening some of their other inputs or outputs [2].

Congesion 1 ([3]). Congestion is said to occur when the output that is maximally possible can be increased by reducing one or more inputs without improving any other input or output. Conversely, congestion is said to occur when some of the outputs that are maximally possible are reduced by increasing one or more inputs without improving any other input or output.

Input-oriented CCR model

$$\theta_o^* = \min \theta_o,$$

$$\text{s.t.}$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} \le \theta x_{io}, \ i = 1 \dots, m, \tag{1}$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} \le \theta y_{ro}, \ r = 1 \dots, s,$$

$$\lambda_j, s_{io}^-, s_{ro}^+) \ge 0, \ j = 1 \dots, n, \ i = 1 \dots, m, \ r = 1 \dots, s.$$

Banker, Charnes and Cooper BCC model

$$\varphi_o^* = \max \varphi_o,$$

$$\text{s.t.}$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} \le \theta x_{io}, i = 1 \dots, m,$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} \le \theta y_{ro}, r = 1 \dots, s \tag{2}$$

$$\sum_{j=1}^{n} \lambda_j = 1,$$

$$\lambda_j, s_{io}^-, s_{ro}^+) \ge 0, \ j = 1 \dots, n, i = 1 \dots, m, r = 1 \dots, s.$$

Congestion has been an under-researched topic in Western economics partly [5] because a Nobel Laureate economist questioned whether "congestion" as a subject of research should have any place in economics in his review of the "X-Efficiency" concept of [4], [6]. However, after a long period of neglect in the economics literature, Färe and Svensson [7] excogitated new research in this area by reformulating some of the concepts connected with congestion. Färe and Grosskopf [8] then gave this abstraction operational form. Later, complete the models (And methods of analysis) that they used to analyze congestion and accord them a form that would now be identified with DEA [9].

This approach was the only one accessible in the DEA literature and was therefore employed in all of the research into congestion in the numerous applications that were then undertaken, however, formulated an alternative approach which has also begun to see different extensions and applications interest in the development of alternative approaches has now started to result in additions and extensions in a various of ways [10]. This has been advantageous because new alternatives provide a perspective on shortcomings as well as advantages in the utilization of existing models. This was exhibited, for example, in the exchanges between Färe and [11–13].

Shortcomings in the [14] approach were then identified [15] in a manner that led to the exchanges between [16] and [17]. Many new applications have been reported in different fields. Congestion, as used in economics, refers to circumstances where reductions in one or more inputs generate an increase in one or more outputs without worsening any other input or output. There are two principal approaches for identifying and measuring congestion [8], [15]. To overcome these theoretical shortcomings in previous research about congestion and returns to scale, the first problem of RTS was considered, and several methods for measuring RTS [17] were developed.

2 | Conjection Model

Färe, Grosskope and Lovell approach

The Färe, Grosskope, and Lovell (FGL) method [18] proceeds in two steps. The first step utilizes an "input-oriented" *Model (1).* In this model x_{ij} is the observed amount of input i=l,...,m utilized by DMU_j and $y_{ro}$is the observed amount of output r=l,...,s produced by DMU_j. The x_{io} and y_{ro} represent the amounts of inputs i=l,...,m and outputs r=l,...,s associated with $DMU_o$where DMU_o is the $DMU_j=DMU_o$ to be evaluated dependent on all DMU_j (Including itself). The aim is to minimize all of the inputs of DMU_o in the proportion $\theta^*$where, because the $x_{io} = x_{ij}$ and $y_{ro} = y_{rj}$ for $DMU_j=DMU_o$come in on both sides of the constraints in *Eq. (1)*, the optimal $\theta^* = \theta$ does not exceed unity, and the non-negativity of the λ_o, x_{ij} and y_{rj} implies that the value of θ^* will not be negative under the optimization in *Eq. (1)*. Hence, 0 ≤ Minq =θ^* ≤1. There are the following explanations of technical efficiency and inefficiency:

Färe, Grosskope, and Lovell technical efficiency

I. Technical efficiency is gained by DMU_o if and only if θ^*=1.

II. Technical inefficiency exists in the performance of DMU_o if and only if 0 ≤ θ^* ≤1.

This definition ignores the possible presence of non-zero slacks even when the answer of *Eq. (1)* shows them to be present. This definition refers to "weak" technical efficiency. This is the term utilized in the operations research literature. In the economics literature, it is referred to as the assumption of "strong disposal." In any case, FGL then goes on to the following second step model:

$$\beta^* = \text{Min}\ \beta,$$

$$\text{s.t.}$$

$$\sum_{j=1}^{n} x_{ij}\lambda_j = \beta x_{io}, i = 1 \dots, m, \tag{3}$$

$$\sum_{j=1}^{n} y_{rj}\lambda_j \geq y_{r0},\ r = 1\dots, s,$$

$\lambda_j \geq 0,\ j = 1 \ldots, n.$

Note that the first i=l..., m inequalities in *Model (1)* are replaced by *Eq. (3)*. Therefore, slack is not possible in the inputs. The fact that only the output can yield non-zero slack is then referred to as "weak disposal" by [19]. We note that *Model (3)* is more restricted than *Model (3)* by the goodness of replacing inequalities with equations. Hence, we have $0 \leq \theta^* \leq \beta^*$. FGL utilizes this property to develop a "measure" of congestion:

$$0 \leq C(\theta^*, \beta^*) = \frac{\theta^*}{\beta^*} \leq 1.$$

Combining *Models (1)* and *(2)* in a two-step manner, FGL utilizes this measure to identify congestion in terms of the following conditions:

I. Congestion is identified as present in the performance of DMU_o if and only if

II. $C(\theta^*, \beta^*) < 1$.

III. Congestion is identified as not present in the performance of DMU_o if and only if $C(\theta^*, \beta^*) = 1$.

Cooper, Thompson, and Thrall approach

Cooper et al. [10] present another model, which was extended by Brockett et al. [11] in their study of congestion in Chinese production. See the further developments on the utilization of these results for policy guidance in [12]. This alternate method also proceeds in a two-step manner, with *Model (2)* utilized in the first step. In *Model (2)*, ξ is a non-Archimedean element smaller than any positive real number. It is used only in theory to refrain from rewriting the model. In other words, in practice, two distinct structures with their objective functions, but similar constraints, must be solved to obtain the optimal solution. Then solve *Eq. (2)* for each DMU. For an optimal solution $(\rho^*, \lambda^*, S^{+^*}, S^{-^*})$ of *Eq. (2)*, re-express the constraints in the following form:

$$\rho^* y_{ro} + s_r^{+^*} = \sum_{j=1}^{n} y_{rj} \lambda_j^*,\ r = 1 \ldots, s, \quad (4)$$

$$x_{io} - s_i^{-^*} = \sum_{j=1}^{n} x_{ij} \lambda_j^*,\ i=1\ldots,m. \quad (5)$$

In this method, the values on the Left-Hand Side (LHS) in *Eq. (4)* and *Eq. (5)* are used to explain new outputs and inputs $\hat{y_{ro}}, r = 1, \ldots, s$, $\hat{x_{io}}, i = 1, \ldots, m$ as in the following:

$$\hat{y_{ro}} = \rho^* y_{ro} + s_r^{+^*} \geq y_{ro},\ r = 1 \ldots, s, \quad (6)$$

$$\hat{x_{io}} = x_{io} - s_i^{-^*} \leq x_{io},\ i = 1 \ldots, m. \quad (7)$$

Note that $\hat{y_{ro}}, \hat{x_{io}}$ are the coordinates of points on the efficiency frontier. In the above method, inefficiency is a necessary condition for the presence of congestion. Therefore, at first, the method uses *Eq. (2)* to identify whether DMU_0 is inefficient. If it is found to be so, then the technique uses *Eq. (2)* and *Eq. (3)* to formulate *Eq. (6)*:

$$\text{Max } \sum_{i=1}^{m} \delta_i^-,$$

$$\text{s.t.}$$

$$\hat{y_{ro}} = \rho^* y_{ro} + s_r^{+^*} = \sum_{j=1}^{n} y_{rj} \lambda_j,\ r=1\ldots, s,$$

$$\hat{x_{io}} = x_{io} - s_i^{-^*} = \sum_{j=1}^{n} x_{ij} \lambda_j - \delta_i^-,\ i = 1 \ldots,$$

$$\sum_{j=1}^{n} \lambda_{j=1},\ j=1\ldots, n, \quad (8)$$

$s_i^{-*} \geq \delta_i^*$, i=1..., m,

$\delta_i^* \geq 0$, i=1..., m

$\lambda_j \geq 0, j = 1 \ldots, n.$

Finally, to determine the level of congestion, the method utilizes the input constraints at the bottom of *Model (6)* to obtain:

$$\sum_{j=1}^{n} x_{ij}\lambda_j^* - \hat{x}_{io} = \delta_i^{-*}, \quad i = 1 \ldots, m,$$
$$s_i^{-c^*} = s_i^{-*} - \delta_i^{-*}, \ i = 1 \ldots, m, \quad (9)$$

Substituting *Eq. (8)* into *Eq. (6)*, can rewrite the latter as *Eq. (9)*:

$$\text{Min} \sum_{i=1}^{m} s_i^{-c},$$
$$\text{s.t.}$$
$$\rho^* y_{ro} + s_r^{+*} = \sum_{j=1}^{n} y_{rj}\lambda_j, r = 1 \ldots, s,$$
$$x_{io} - s_i^{-c} = \sum_{j=1}^{n} x_{ij}\lambda_j, \ i = 1 \ldots, m, \quad (10)$$
$$\sum_{j=1}^{n} \lambda_{j=1}, j = 1 \ldots, n,$$
$$s_i^{-c} \geq 0, i = 1 \ldots, m,$$
$$\lambda_j \geq 0, \ j=1\ldots, n.$$

Cooper et al.'s method

Using ξ, Cooper et al. [20] combined this stage into the single *Model (10)*:

$$\text{Max } \rho + \xi \left(\sum_{r=1}^{s} s_r^{+} - \xi \sum_{i=1}^{m} s_i^{-c}\right),$$
$$\text{s.t.}$$
$$\sum_{j=1}^{n} x_{ij}\lambda_j + s_{io}^{-c} = x_{io}, \ i=1\ldots, m,$$
$$\sum_{j=1}^{n} y_{rj}\lambda_j - s_{ro}^{+} = \rho_0 y_{ro}, \ r=1\ldots, s, \quad (11)$$
$$\sum_{j=1}^{n} \lambda_{j=1}, \ j=1\ldots, n,$$
$$(\lambda_j, s_{io}^{-}, s_{ro}^{+}) \geq 0, \ j=1\ldots., n, \ i=1\ldots, m, r=1\ldots, s.$$

As referred to earlier, the use of ξ only has theoretical justification. In the method, the calculations must be performed in three stages, and three models must be solved in order to achieve the optimal solution and identify the corresponding level of congestion.

In order to detect the presence of congestion in a DMU, Cooper et al. [20] presented the following theorem for identifying and measuring congestion:

Theorem 1. Congestion is present if and only if in an optimal solution ($\rho^*, \lambda^*, s^{+^*}, s^{-^*}$) of *Eq. (2)*, at least one of the following two conditions is satisfied:

I. ρ^* >1 and there is at least one $s^{-c^*} > 0, (1 \leq i \leq m)$.

II. There exists at least one $s_r^{+^*} > 0, (1 \leq r \leq s)$ and at least one $s^{-c^*} > 0, (1 \leq i \leq m)$ For additional detail on, see [20].

Tone and Sahoo approach

In this method, there exist n DMU for j = 1..., n) and each DMU uses m inputs (i = 1..., m) to produce s outputs (r = 1..., s). The ith input and the rth output are specified by (x_{ij}, y_{rj}) for the j[th] DMU [21]. The output-oriented congestion of the kth DMU is measured by comparing the objective values of the following two (Left and right) DEA models:

Original

$$\begin{aligned}
&\text{Max } \theta,\\
&\text{s. t.}\\
&\textstyle\sum_{j=1}^{n} \lambda_j\, x_{ij} \leq x_{ik},\ \text{i=1..., m},\\
&\textstyle\sum_{j=1}^{n} y_{rj}\lambda_j - \theta y_{rk} \geq 0,\ \text{r=1..., s},\\
&\textstyle\sum_{j=1}^{n} \lambda_j = 1,\\
&\theta = \text{URS},\ \text{and } \lambda_j \geq 0,\ \text{j=1..., n}.
\end{aligned} \tag{12}$$

Congestion

$$\begin{aligned}
&\text{Max } \beta,\\
&\text{s.t.}\\
&-\textstyle\sum_{j=1}^{n} x_{ij}\lambda_j + x_{ik} \geq 0, \text{i=1..., n},\\
&\textstyle\sum_{j=1}^{n} y_{rj}\lambda_j - \beta y_{rk} \geq 0, \text{r=1..., s},\\
&\sum_{j=1}^{n} \lambda_j = 1,\\
&\beta = \text{URS and } \lambda_j \geq 0, \text{j=1..., n}.
\end{aligned} \tag{13}$$

The original (left) model is a Banker, Charnes and Cooper (BCC) model [4]. The original *Model (11)* provides a radial non-parametric measure for Technical Efficiency (TE). An important feature of the TE measure is that it avoids the assumption of constant RTS (So, variable RTS). The variable λ_j (j = 1..., n) is used to connect inputs and outputs in a data domain representing a Production Possibility Set (PPS). The variables (θ and β) in the objectives of the two models represent, respectively, a level of TE under two different production technologies. Those are Unrestricted (URS) in such a manner that each of the two variables can take any sign. As can be easily identified by comparing *Eq. (11)* with *Eq. (12)*, there is only a major difference between the two DEA models. The first set of constraints is formulated by inequality in *Eq. (11)*, while being equality in *Eq. (12)*. To extend the two DEA models further into the issue of congestion, let the PPSs of *Eq. (11)* and *Eq. (12)* be:

$$p = \left\{ (x,y) \middle| \begin{array}{c} x \geq \sum_{j=1}^{n} x_j\lambda_j, y \leq \sum_{j=1}^{n} y_j\lambda_j, \\ \sum_{j=1}^{n} \lambda_j = 1 \; \lambda_j \geq 0, j = 1, \dots, n. \end{array} \right.$$

$$p_{convex:} \left\{ (x,y) \middle| \begin{array}{c} x = \sum_{j=1}^{n} x_j\lambda_j, y \leq \sum_{j=1}^{n} y_j\lambda_j, \\ \sum_{j=1}^{n} \lambda_j = 1 \; \lambda_j \geq 0, j = 1, \dots, n. \end{array} \right.$$

The two DEA models can be expressed by max $\{\theta|(x_k,\theta y_k\}\in P$ and max max $\{\beta|(x_k,\beta y_k\}\in P_{convex}$, using the two PPSs. The two DEA *Models (11)* and *(12)* have the following dual formulations:

Banker, Charnes and Cooper

$$\begin{aligned} &\text{Min} \sum_{i=1}^{m} v_i x_{ik} - \sigma, \\ &\sum_{i=1}^{m} v_i x_{ij} + \sum_{r=1}^{s} u_r y_{rj} + \sigma \leq 0, \quad j = 1, \dots, n, \\ &\textstyle\sum_{r=1}^{s} u_r y_{rk} = 1, \\ &v_i \geq 0, u_r \geq 0, \text{and } \sigma, \text{URS}. \end{aligned} \tag{14}$$

Congestion

$$\begin{aligned} &\min \sum_{i=1}^{m} v_i x_{ik} - \sigma, \\ &\sum_{i=1}^{m} v_i x_{ij} + \sum_{r=1}^{s} u_r y_{rj} + \sigma \leq 0, j = 1, \dots, n, \\ &\sum_{r=1}^{s} u_r y_{rk} = 1, \\ &v_i, \text{URS}, u_r \geq 0 \text{ and } \sigma, \text{URS}. \end{aligned} \tag{15}$$

Here, v_i (i = 1..., m) and u_r (r = 1..., s) are dual variables (multipliers) derived from the first and second constraint sets in *Eq. (11)* and *Eq. (12)*. Similarly, r is a dual variable derived from the last constraint. Almost no difference is identified between the two dual models. However, a careful examination of the two formulations indicates that v_i is non-negative in *Eq. (3)*, but it is unrestricted URS in *Eq. (14)*.

Let the optimal value of *Eq. (11)* be θ^* and that of *Eq. (12)* be β^*, where the superscript (*) indicates optimality. Then, the output-oriented Congestion (OC) is measured by the following ratio.

$$\text{Output} - \text{oriented Congestion: OC } (\theta^*, \beta^*) = \frac{\theta^*}{\beta^*}$$

If $\text{OC } (\theta^*, \beta^*) > 1$, then the congestion occurs on the kth DMU. Meanwhile, if $\text{OC } (\theta^*, \beta^*) = 1$, then there is no congestion. The previous research extended the concept of congestion further by linking it to other economic concepts related to efficiency, proposed to measure the concept of strong efficiency, and discussed the relationship between the two economic concepts [21], [22]. The strong efficiency is defined as follows.

Definition 2. Let θ^* be the optimal value of *Eq. (11)* and let $(d^{x^*}, d^{y^*}, \lambda^*)$ be an optimal solution of

$$\begin{aligned}
&\text{Max} \sum_{i=1}^{m} d_i^x + \sum_{r=1}^{s} d_r^y,\\
&x_{ik}=\sum_{j=1}^{n} x_{ij}\lambda_j + d_j^x,\ i=1\ldots, m,\\
&\theta^* y_{rk}=\sum_{j=1}^{n} y_{rj}\lambda_j - d_r^y,\ r=1\ldots, s,\\
&\sum_{j=1}^{n} \lambda_j=1 \text{ and } \lambda_j \geq 0,\ j=1\ldots, n.
\end{aligned} \tag{16}$$

The status of "strong efficiency" is identified within P if and only if $\theta^* = 1, d^{x^*} = 0$, and $d^{y^*} = 0$. Similarly, let β^* be the optimal value of *Eq. (12)* and let (d^{y^*}, λ^*)be an optimal solution of:

$$\begin{aligned}
&\max \sum_{i=1}^{m} d_r^y,\\
&x_{ik}=\sum_{j=1}^{n} x_{ij}\lambda_j, i=1\ldots, m,\\
&\beta^* y_{rk}=\sum_{j=1}^{n} y_{rj}\lambda_j - d_r^y,\ r=1\ldots, s,\\
&\sum_{j=1}^{n} \lambda_j=1 \text{ and } \lambda_j \geq 0,\ j=1\ldots, n.
\end{aligned} \tag{17}$$

The status of "strong efficiency" is identified within p_{convex} if and only if $\beta = 1$ and $d^{y^*} = 0$.. Based upon the definition regarding strong efficiency, they redefined the concept of "weak congestion" in the following manner [22]:

Definition 3. A DMU is "weakly congested" if it is strongly efficient with respect to p_{convex} and there exists an activity in p_{convex} that uses fewer resources in one or more inputs to make more products in one or more outputs. *Definition 3* makes it possible for the following DEA model to examine whether the kth DMU suffers from an occurrence of strong congestion:

$$\begin{aligned}
&\text{Max} \sum_{i=1}^{m} v_i x_{ik} - \sum_{i=1}^{m} v_i x_{ij} + \sum_{r=1}^{s} u_r y_{rj} + \sigma \leq 0,\ j = 1, \ldots, n,\\
&\sum_{i=1}^{m} v_i x_{ij} + \sum_{r=1}^{s} u_r y_{rj} + \sigma = 0\,, \qquad j = 1, \ldots, n,\\
&\sum_{r=1}^{s} u_r y_{rk} = 1\,, v_i = \text{URS}\,, u_r \geq 0 \text{ and } \sigma, \text{URS}.
\end{aligned} \tag{18}$$

According to TS method [22], if the optimal objective value of *Eq. (17)* is negative, then the kth DMU suffers from "strong congestion". In this study, the concept of "strong congestion" implies the status of "weak congestion" under *Definition 2*. The assertion is trivial because if a DMU is under congestion, then it satisfies weak congestion. Moreover, let us consider that a DMU belongs to "congestion" if the DMU satisfies the condition related to *Eq. (15)*. The three different concepts (Congestion, weak congestion and strong congestion) have the following relationship among them.

Theorem 2. If a DMU belongs to strong congestion, then the DMU belongs to congestion. If a DMU is strongly efficient with respect to Pconvex and it belongs to congestion, then the DMU belongs to weak congestion.

2.1 | Congestion Measurement by Tone and Sahoo, under Multiple Projections

The TS technique [22] has proposed the following method to measure the Degree of Scale Elasticity (DSE) of the kth DMU on its projected point (x'_k, y'_k):

Step 1. Let θ^*be the optimal value of *Eq. (11)* and let $(d^{x^*}, d^{y^*}, \lambda^*)$ be an optimal solution of *Eq. (16)*:

If $\theta^* = 1$, $d^{x^*} = 0$, and $d^{y^*} = 0$; then (x'_k, y'_k) is efficient and not congested under variable RTS. Let

$$\bar{p} = \max\{vx'_k | \text{the same constraints as } (8), v \geq 0\}.$$

and

$$p = \min\{vx'_k | \text{the same constraints as } (2-17), v \geq 0\}.$$

Let $DSE_k = \frac{\bar{p}+p}{2}$ and stop.

If $\theta^* = 1$, $d^{x^*} \neq 0$, and $d^{y^*} = 0$; then (x'_k, y'_k) is technically inefficient and stops.

If $\theta^* = 1$, and $d^{y^*} \neq 0$; or $\theta^* > 1$ then (x'_k, y'_k) is congested, and go to *Step 2.*

Step 2. Let $\bar{p}$ be the optimal value of *Eq. (17).* If $\bar{p} < 0$, then (x'_k, y'_k) is strongly congested.

$$p = \min\{vx'_k | \text{the same constraints as } (2-17), v \geq 0\} \text{ Let } DSE_k = \frac{\bar{p}+p}{2} \text{ and stop.}$$

Otherwise $\bar{p} \geq 0$, (x'_k, y'_k) is weakly, but not strongly, congested. Solve the following problem:

$$\text{Max}\left\{\frac{1}{s}\sum_{r=1}^{s}\frac{t_r^y}{y'_{rk}} + \varepsilon_n \frac{1}{m}\sum_{i=1}^{m}\frac{t_i^x}{x'_{ik}} | x'_k - t^x, y'_k + t^y \in P, t^x \geq 0, t^y \geq 0\right\}, \tag{19}$$

where ε_n is a non-Archimedean small number.

Let (t^{x^*}, t^{y^*}) be an optimal solution of *Eq. (10).* Let $\bar{s}$ and $\overline{m}$ be the number of positive ${t_r}^{y^*}(r = 1, \ldots, s)$ and the number of positive ${t_i}^{x^*}(i = 1, \ldots, m)$. Then, the DSE is measured by

$$DSE_k = \frac{-\frac{1}{s}\sum_{r=1}^{s}\frac{t_r^{y^*}}{y'_{rk}}}{\frac{1}{m}\sum_{i=1}^{m}\frac{t_i^{x^*}}{x'_{ik}}}.$$

Toshiyuki method

Definition 4. A DMU is "widely" congested if it exists on the boundary of p_{convex} and it has an activity in p_{convex} that uses fewer resources in one or more inputs to make more products in one or more outputs [17].

This definition implies that if a DMU is widely congested, then it exists on the boundary of P_{convex}. However, the DMU doesn't need to be strongly efficient with respect to P_{convex}. If the DMU is strongly efficient with respect to p_{convex}, it exists on the boundary of P_{convex}.

Theorem 3. If a DMU is "weakly" congested, then the DMU is "widely" congested.

The change from weak congestion to wide congestion implies that strong efficiency with respect to p_{convex} is not essential in terms of congestion identification. Instead, we must examine whether a DMU exists on the boundary of p_{convex}. As a consequence of such a change, we can eliminate *Step 2* of Sueyoshi and Sekitani [17], including *Eq. (18),* from the congestion identification. Based upon *Theorem 3*, we proposed the following new approach to project all DMUs onto the boundary of p_{convex}.

Projection: Identify optimal β^* of each DMU by solving *Eq. (12)*. A projected point of the DMU is identified by:

$$x'_k \leftarrow x_k \text{ (unchanged) and } y'_k \leftarrow \beta^* y_k. \tag{20}$$

Since β^* is uniquely determined by *Eq. (12)*, the projected point (x'_k, y'_k) is also uniquely determined by *Eq. (19)*. A computational benefit of the projection is that *Eq. (19)* is not influenced by the occurrence of multiple projections. Furthermore, the projected point (x'_k, y'_k) has $\beta^* = 1$ on optimality of *Eq. (12)*. To examine whether a DMU is congested or not, we need to characterize the wide congestion in the following manner mathematically.

Theorem 4. Assume that (x_k, y_k) is on the boundary of $p_{convex}$and the optimal value of *Eq. (12)* is $\beta^* = 1$. A DMU (x_k, y_k) is widely congested if and only if any optimal solution of *Eq. (14)* (v*, u*, r*) satisfies either:

I. $v_i^* < 0$ for at least one i ∈ {1..., m}.

II. $v^* > 0$ for at least one i ∈ {1..., m}, and $u_r^* = 0$ for at least one r∈{1..., s}.

Identification of wide congestion under the occurrence of multiple projections

For the identification of wide congestion under the occurrence of multiple projections, the following approach for identifying wide congestion consists of two linear programming problems.

Step 3. Choose δ> 0 arbitrarily (Where δ is a real number) and solve the following problem

$$\begin{aligned}
&\text{Max } \varepsilon + \sum_{r=1}^{s} d_r^y,\\
&\sum_{i=1}^{m} v_i x_{ij} + \sum_{r=1}^{s} u_r y_{rj} + \sigma \leq 0, \quad j = 1, \dots, n,\\
&x_{ik} = \sum_{j=1}^{n} x_{ij}\lambda_j, \ i=1\dots, m,\\
&\sum_{r=1}^{s} u_r y_{rk} = 1,\\
&\beta y_k = \sum_{j=1}^{n} y_{rj}\lambda_j - d_r^y, \ r=1\dots, s,\\
&\sum_{i=1}^{m} v_i x_{ik} - \sigma = \beta, \sum_{j=1}^{n} \lambda_j = 1, \quad j = 1, \dots, n,\\
&v_i x_{ik} - \varepsilon \geq 0, \ i=1\dots, m,\\
&\varepsilon \leq \delta d_r^y \geq 0, v_i, \text{URS}, \geq 0 \ , \sigma, \text{URS}, \beta, \text{URS}, , \text{URS}, \lambda_j \geq 0.
\end{aligned} \tag{21}$$

Here, an arbitrary real number (σ) guarantees the existence of an optimal solution of *Eq. (14)*. Since e represents the smallest value of $v_i x_{ik}$ (i = 1..., m) in such a manner of min {min $\{v_i x_{ik} | i = 1..., m\}$,σ }, the arbitrary number (σ) functions as the upper bound. Consequently, *Eq. (20)* always has an optimal solution. All the constraints of *Eq. (20)*, except ($v_i x_{ik} - \varepsilon \geq 0$ (i=1…, m) $\sum_{i=1}^{m} v_i x_{ik} - \sigma = \beta \quad \varepsilon \leq \delta$) are obtained from *Eq. (12)* and *(14)*. Sueyoshi and Sekitani [17] provided a rationale regarding why *Eq. (20)* deals with an occurrence of multiple solutions. *Problem (13)* is a modified version of their approach for *Eq. (17)*. The proposed approach restricts the DEA dual variable in order to obtain a reduced projection range for the measurement of wide congestion.

Let $(\lambda^*, \beta^*, d^{y^*}, v^*, u^*, \sigma^*, \varepsilon^*)$be an optimal solution of *Eq. (20)*, then we can identify the wide congestion on the projected point $(x_k, \beta^* y_k)$ of the kth DMU as follows:

I. if $\varepsilon^* < 0$, then($x_k, \beta^* y_k$) is widely congested,

II. $\varepsilon^* > 0$, then $(x_k, \beta^* y_k)$ is not widely congested,

III. if $\varepsilon^* = 0$ and $\sum_{r=1}^{s} d_r^{y^*} > 0$, then $(x_k, \beta^* y_k)$ is widely congested and,

IV. if $\varepsilon^* = 0$ and $\sum_{r=1}^{s} d_r^{y^*} = 0$, then go to *Step 2*.

Step 4. Solve the following problem

$$\max \alpha$$

$$\text{s.t.}$$

$$\sum_{i=1}^{m} v_i x_{ij} + \sum_{r=1}^{s} u_r y_{rj} + \sigma \leq 0 , j = 1, \ldots, n,$$

$$\sum_{r=1}^{s} u_r y_{rk} = 1,$$

$$\sum_{i=1}^{m} v_i x_{ik} - \sigma = \beta, \; v_i \geq 0, \quad u_r \geq 0, \alpha \geq 0 \; \sigma, \text{URS}.$$

The wide congestion of the kth DMU is identified as follows:

I. if $\alpha^* > 0$, then $(x_k, \beta^* y_k)$ is not widely congested and

II. if $\alpha^* = 0$, then $(x_k, \beta^* y_k)$ is widely congested.

Jahanshahloo, Rashidi, and Parker method

Solve *Model (2)* for each DMUj, j = 1…, n, and achieve the optimal solution [23]:

$(\rho^*, \lambda^*, s^{+^*}, s^{-^*})$ denoting the ρ^* corresponding to DMUj by ρ_j^*. The set E is defined as follows:

$$E=\{j \mid \rho^* = 1 \}. \tag{22}$$

Among the DMUs in set E, there exists at least one DMU, say DMU_j, that has the highest usage in its first input component compared with the first input component of the remaining DMUs of set E. That is to say,

$$\exists (l \in E) \;\; \text{s.t. for all j} (j \in E) \Rightarrow x_{1l} \geq x_{1j}. \tag{23}$$

They denoted x_{1l} by x_1^*. Then find, again, among the DMUs in E, a DMU, say t DMU, that has the highest usage in its second input component compared to the remaining DMUs in E. In other words,

$$\exists \; (t \in E) \;\; \text{s.t} \; \forall \; j \; (j \in E) \Rightarrow x_{2t} \geq x_{2j}. \tag{24}$$

They showed x_{2t} by x_2^* Similarly, for all input components i= 1,..., m, identify a DMU in E whose ith input consumption is higher than that of all other DMUs in the set. Inputs denoted by x_j^*, i=1…, m. Note that $x_1^*, x_2^*, \ldots, x_m^*$ need not certainly be selected from a single DMU. The congestion is discussed as follows:

Definition 5. Congestion is present if and only if, in an optimal answer $(\rho^*, \lambda^*, s^{+^*}, s^{-^*})$ of *Eq. (2)* for DMUo, at least one of the following two conditions is satisfied:

I. $\varphi^* > 1$, and there is at least one $x_{io} > x_i^*$,i=1…., m.

II. There exists at least one $s_r^{+^*} > 0$ (r =1..., s), and at least one $x_{io} > x_i^*$,i=1…., m.

The amount of congestion in the ith input of DMUo is denoted by $s_i^{c'}$ where $x_{io} > x_i^*$, and define it as: $s_i^{c'} = x_{io} - x_i^*$ *Eq. (23).*

Congestion is considered not present when $x_{io} \le x_i^*$ and $s_i^{c'} = 0$. The sum of all $s_i^{c'}$ is the amount of congestion in DMUo. For purposes of explanation, consider subFig adapted from Flegg and Allen [24].

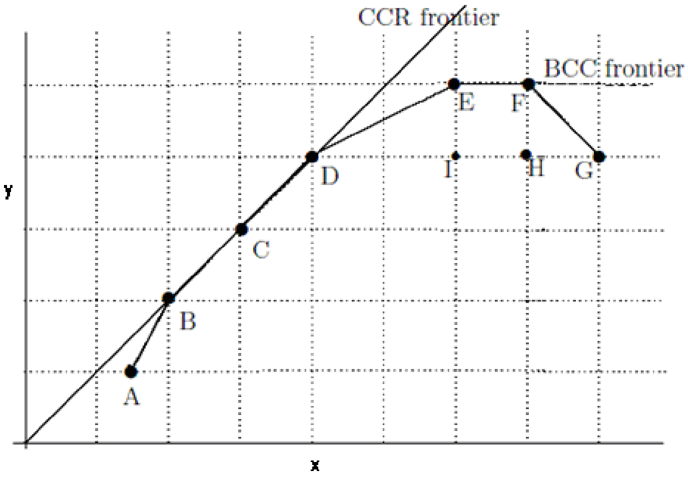

Fig. 1. Schematic adapted from Flegg & Allen (for explanatory purposes).

The set of efficient DMUs is E = {A, B, C, D, E, F}. As can be seen DMU_F has the highest input using up among the efficient DMUs, i.e.

Theorem 5. If $DMU_0^* = x_1^*, x_2^*, \ldots, x_m^*, \varphi^* y_{1o} + s_1^*, \varphi^* y_{2o} + s_2^*, \ldots, \varphi^* y_{so} + s_s^*)$, then $DMU_0^* \in PPS_{TV}$.

Theorem 6. $s_i^{c'} = s_i^{-c^*}$ where $x_{io} > x_i^*$.

Theorem 7. If for all i,i=1…, m, we have $x_{io} - x_i^* \le 0$, then there exists no congestion in DMUo.

Hosseinzadeh et al. approach (Interval congestion)

Consider n DMUs with m inputs and s outputs, with interval data, that is [9]:

$$x_{ij} \in [x_{ij}^l, x_{ij}^u],\ i=1\ldots m.$$

$$y_{rj} \in [y_{rj}^l, y_{rj}^u],\ r=1\ldots, s.$$

In other words, to measure congestion with interval data, one should first determine the efficiency interval of each DMU. To do so, they gained the efficiency of each DMU in the most pessimistic and the most optimistic cases, using the following two models introduced by Wang et al. [25] and Javanmard [26]. In the Model, which is the most pessimistic case in evaluating a DMU, they considered the unit under estimation with the highest inputs and the lowest outputs.

$$\begin{aligned} &\varphi_o^{*u} = \text{Max}\ \varphi, \\ &\text{s. t.} \\ &\textstyle\sum_{\substack{j=1 \\ j\neq o}}^{n} x_{ij}^l \lambda_j + \lambda_o x_{io}^u + s_{io}^- = x_{io}^u,\ i=1\ldots, m, \\ &\textstyle\sum_{\substack{j=1 \\ j\neq o}}^{n} y_{rj}^u \lambda_j + \lambda_o y_{ro}^l + s_{ro}^+ = \varphi_o y_{ro}^l,\ \ r=1\ldots,s, \\ &\textstyle\sum_{j=1}^{n} \lambda_j = 1,\ \ (\lambda_j, s_{io}^-, s_{ro}^+),\ j=1\ldots, n, r=1\ldots, s, i=1\ldots, m\ . \end{aligned} \tag{25}$$

As for the most optimistic case, *Model (25)* assumes the unit under evaluation with the lowest inputs and the highest outputs, and the other DMUs with the highest inputs and the lowest outputs.

$$\varphi_o^{*l} = \text{Max}\ \varphi,$$
$$\sum_{\substack{j=1\\ j\neq o}}^{n} x_{ij}^{u}\lambda_j + \lambda_o x_{io}^{l} + s_{io}^{-} = x_{io}^{l},\ i=1\ldots, m,$$
$$\sum_{\substack{j=1\\ j\neq o}}^{n} y_{rj}^{l}\lambda_j + \lambda_o y_{ro}^{u} + s_{ro}^{+} = \varphi_o y_{ro}^{u},\ r=1\ldots, s,$$
$$\sum_{j=1}^{n} \lambda_j = 1, (\lambda_j, s_{io}^{-}, s_{ro}^{+}),\quad j=1\ldots, n, r=1\ldots, s, i=1\ldots, m. \tag{26}$$

After determining the efficiency interval $[\varphi^{*l}, \varphi^{*u}]$, they suggested the following method for computing congestion. With regard to the idea in this method, the highest input value for each component is specified to compute congestion among efficient DMUs. As was stated earlier, one need not consider a single DMU for selecting all the components. To this end, they defined a set E′as follows:

$$E' = \{DMU_j | \varphi_j^{*l} = 1\}, \tag{27}$$

E′ is the largest efficient set that can maybe exist with the above data, i.e., it is the set of DMUs that are efficient in the best case. We aim to determine a congestion interval (i.e., upper and lower bounds) such that the congestion value associated with any combination of values occurring in the input and output intervals of a DMU belongs to the interval achieved. Considering the fact that inefficiency is the necessary condition for congestion, there exists no congestion in DMUs with $\varphi^{*l} = \varphi^{*u} = 1$, besides, since the DMUs in the set $E'$are efficient in their best case; they do not show congestion in this case. However, these DMUs might be inefficient in their worst case. Thus, there exists the possibility of congestion in this case for these DMUs. In computing congestion in the most optimistic case possible, the lowest input consumption of a DMU is compared with the highest input consumption of the DMUs belonging to the set E' (i.e., those efficient in the best case). To this end, they found x_i^{*u}as follows:

$$\text{For all i}\ \ i = 1, \ldots, m\ \ \exists t_i \quad s.t\ \ x_{t_i}^{u} = x_i^{*u} = \max\{x_{ij}^{u} | j \epsilon E'\}, \tag{28}$$

And in the most optimistic case possible, the highest input consumption of a DMU is contrasted with the lowest input consumption of the DMUs belonging to the set E' (i.e., those efficient in the best case). They showed x_i^{*u} as follows:

$$\text{For all i}\ \ i = 1, \ldots, m\ \ \exists k_i \quad s.t\ \ x_{k_i}^{l} = x_i^{*l} = \max\{x_{ij}^{l} | j \epsilon E'\}. \tag{29}$$

The lower bound of congestion in the ith input is denoted by $DMU_o$by s_{io}^{cl} and define it as

$$s_{io}^{cl} = x_{io}^{l} - x_{io}^{u^*},\quad i=1\ldots, m. \tag{30}$$

If $s_{io}^{cl} \geq 0$, the amount of congestion is indicated; otherwise, congestion is zero in the best case. Furthermore, they are denoted by s_{io}^{cu}. The upper bound of congestion in the ith input of the DMU is defined as:

$$s_{io}^{cu} = x_{io}^{u} - x_{io}^{l^*},\quad i=1\ldots, m. \tag{31}$$

If $s_{io}^{cu} \geq 0$, the amount of congestion is shown; otherwise, congestion is zero in the worst case

Theorem 8. The interval $[s_o^{cl}, s_o^{cu}]$ indicates an upper and a lower bound for congestion present at DMU_o.

Jahanshahloo, et al. technique

Their study started with a two-model method that was utilized to evaluate congestion [27]. Assume there are n observed DMUs, DMUj (x_j, y_j), j = 1, 2, ..., n, and all DMU_j produces the same s outputs in (Possibly) different amounts, y_{rj}, r =1, 2, ..., s, utilizing the same inputs, x_{ij}, i = 1, 2, ..., m, also in (Possibly) various amounts. All inputs and outputs are assumed to be nonnegative, but at least one input and one output are positive, i.e., $x_j = (x_{1j}, ..., x_{mj}) \geq 0$, $x_j \neq 0$ and $y_j = (y_{1j}, ..., y_{sj}) \geq 0$, $y_j \neq 0$. Then, to keep in contact with Cooper et al. [10], the method starts with the following version of a BCC model:

$$\begin{aligned}
&\text{Max } \phi + \varepsilon\left(\textstyle\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right),\\
&\text{s.t.}\\
&\textstyle\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io}, \ i=1..., m,\\
&\textstyle\sum_{j=1}^{n} \lambda_j y_{rj} + s_r^+ = y_{io}, \ r=1...., s,\\
&\textstyle\sum_{j=1}^{n} \lambda_j = 1,\\
&\lambda_j, s_i^-, s_r^+ \geq 0,\\
&j = 1, \dots, n \quad i = 1, \dots, m, \quad r = 1, \dots, s.
\end{aligned} \qquad (31)$$

Here $\varepsilon > 0$ is a non-Archimedean element, described as smaller than any positive real number. This means that ε is not a real number. The standard procedure is to prevent any need for explicitly assigning a value to ε by utilizing the following two-step process. Step one: Maximize ϕ while ignoring the slacks, s_i^-, s_r^+ in the objective. Step two: Replace ϕ with ϕ^*= max ϕ in *Model (31)* and maximize the sum of the slacks, then determine whether DMUo is efficient or inefficient in accordance with the sub definition.

Definition 6 (BCC Efficiency). DMUo is efficient if and only if the following two conditions are both satisfied:

I. $\varphi* = 1$.

II. All slack variables are zero in optimal solutions.

Definition 7 (BCC-Projection). For a BCC-inefficient DMUo,

BCC-projection, based on an optimal solution for *Model (31),* defined as follows:

$$(\widehat{x_{io}} = x_{io} - s_i^{-*}, \widehat{y_{ro}} = y_{ro} + s_r^{+*}).$$

The improved activity $(\widehat{X}_o, \widehat{Y}_o)$is BCC-efficient. As explained in Cooper et al. [13], the $(\widehat{x_{io}}, \widehat{y_{ro}})$ values together with the x_{ij} and y_{rj}, as defined in *Model (31)*, are utilized to construct the following new problem

$$\begin{aligned}
&\text{Max } \textstyle\sum_{i=1}^{m} \delta_i^-,\\
&\text{s.t.}\\
&\sum_{j=1}^{n} \lambda_j x_{ij} - \delta_i^- = \widehat{x_{io}}, i = 1, \dots, n,\\
&\sum_{j=1}^{n} \lambda_j y_{rj} = \widehat{y_{io}}, \qquad r = 1, \dots, n,\\
&\sum_{j=1}^{n} \lambda_j = 1,\\
&\delta_i^- \leq s_i^{-*}, \lambda_j \geq 0, \ j = 1, \dots, n,\\
&\delta_i^- \geq 0, i = 1, \dots, m.
\end{aligned} \qquad (32)$$

Finally, to identify the congesting inputs and to estimate their amounts, utilize i = 1, ..., m input constraints $\delta_i^- \leq s_i^{-*}$ in *Eq. (32)* to obtain:

$$s_i^{-c^*} = s_i^{-*} - \delta_i^{-*}, \text{i=1..., m}, \tag{33}$$

where δ_i^{-*} is achieved from *Eq. (32)*. $s_i^{-c^*}$ is then the congesting amount in the total slack connected with s_i^{-*} in input i = 1, ...,m, as achieved from *Model (31)* and δ_i^{-*} is the (Maximum) amount of this total slack that can be assigned to completely technical (Non-congesting) inefficiency, as achieved from *Eq. (32)*.

Computation of congestion with production trade-offs

This technique follows the signs in Podinovski's research [28], and let (P, Q), trade-offs between the inputs and/or outputs, show the possible simultaneous change to the inputs and outputs in the entire technology. Different examples of production trade-offs are discussed in [29], which assumes that k trade-offs:

$$(P_t, Q_t), \quad \text{t=1...., k}. \tag{34}$$

The use of trade-offs in *Eq. (34)* in the standard VRS technology leads to the expanded technology T_{VRS-TO}, since the abbreviation TO stands for “trade-offs“, as in the following form:

$$T_{VRS-TO}=\{(X,Y)\,|\,X\geq 0, Y \geq 0\ , X \geq \sum_{j=1}^{n} \lambda_j X_j + \sum_{t=1}^{k} \pi_t P_{t\,,}\ Y \leq \sum_{j=1}^{n} \lambda_j Y_j + \sum_{t=1}^{k} \pi_t Q_{t\,,} \sum_{j=1}^{n} \lambda_j = 1\ \lambda_j \geq 0\ , \pi_t \geq 0\ , j = 1, \dots, n\ \ , t = 1, \dots, k\}. \tag{35}$$

The output radial efficiency of DMUo in TV RS−to is defined as max $\{\phi|(\phi X_o, Y_o) \in T_{VRS-TO})$

$$\text{Max } \phi,$$

$$\text{s.t.} \sum_{j=1}^{n} \lambda_j X_j + \sum_{t=1}^{k} \pi_t P_{t\,,} + d = X_o, \tag{A1}$$

$$\sum_{j=1}^{n} \lambda_j Y_j + \sum_{t=1}^{k} \pi_t Q_{t\,,} - e = \phi\, Y_o, \tag{A2}$$

$$\sum_{j=1}^{n} \lambda_j X_j + \sum_{t=1}^{k} \pi_t P_{t\,,} + d \geq 0, \tag{A3}$$

$$\sum_{j=1}^{n} \lambda_j Y_j + \sum_{t=1}^{k} \pi_t Q_{t\,,} - e \geq 0, \tag{A4}$$

$$\sum_{j=1}^{n} \begin{matrix}\lambda_j \\ 1,\end{matrix}, \tag{A5}$$

$$\lambda, \pi, e, d \geq 0. \tag{A6}$$

The (A3) is redundant as it follows from (A1). Since $(\phi = 1, \pi_t = 0, t = 1, \dots, k, d = 0, e = 0, \lambda_j = 1, j = 1, \dots, n, j \neq o)$.

is a feasible solution for the above model and the objective function maximizing ϕ, so, $\phi^* \geq 1$. Therefore, A (4) can be deleted. This implies that the model is equal to the following LP formulation:

$$\text{Max } \phi,$$

$$\text{s.t.}$$

$$X_o \geq \sum_{j=1}^{n} \lambda_j X_j + \sum_{t=1}^{k} \pi_t P_{t\,,} \tag{35}$$

$$\phi X_o \leq \sum_{j=1}^{n} \lambda_j Y_j + \sum_{t=1}^{k} \pi_t Q_{t,} \sum_{j=1}^{n} \lambda_j,$$

$$\sum_{j=1}^{n} \lambda_j = 1, \lambda, \pi \geq 0.$$

The output radial efficiency of DMUo is equivalent to the optimal value φ∗ of the objective function in Model *Eq. (35)*. Because of the constraints of the model, the aim DMU (Xo, φ∗Yo) is a valid member of the PPS TV RS–TO and located on its boundary. Therefore, the radial efficiency φ∗ of DMUo has the meaning of a technologically feasible radial enhancement factor for the outputs of the DMUo. It is a must to be aware of the introduction of trade-offs. *Eq. (34)* in the envelopment models is equal to the incorporation of weight restrictions.

$$u^T Q_t - v^t p_t \leq 0, \ t=1\ldots, k. \tag{36}$$

In the dual multiplier forms. In VRS DEA models with production trade-offs, the second step of the optimization procedure is a test for possible non-radial improvements to the radial targets. A linear program for this reason is developed as follows:

$$\begin{aligned}
&\text{Max} \textstyle\sum_{i=1}^{m} d_i + \sum_{r=1}^{s} e_r, \\
&\text{s.t.} \\
&\sum_{j=1}^{n} \lambda_j X_j + \sum_{t=1}^{k} \pi_t P_{t,} + w + d = X_o, \\
&\sum_{j=1}^{n} \lambda_j Y_j + \sum_{t=1}^{k} \pi_t Q_{t,} - e = \phi^* Y_o, \\
&\sum_{j=1}^{n} \lambda_j X_j + \sum_{t=1}^{k} \pi_t P_{t,} + w >= 0, \\
&\sum_{j=1}^{n} \lambda_j = 1, \\
&\lambda, \pi, e, d \geq 0.
\end{aligned} \tag{37}$$

Model (37) maximizes the sum of remaining slacks subject to the clear condition that the resulting efficient target has only nonnegative inputs. This step produces a fully efficient target of DMUo: Let λ∗, π∗, e∗, w∗and∗ be any optimal solution to *Model (37)*. Define:

$$\begin{aligned}
&\hat{x}_o = \sum_{j=1}^{n} \lambda_j^* x_j + \sum_{t=1}^{k} \pi_t^* p_t + w^*, \\
&\hat{y}_o = \textstyle\sum_{j=1}^{n} \lambda_j^* x_j + \sum_{t=1}^{k} \pi_t^* Q_t + w^*.
\end{aligned} \tag{38}$$

Obviously, DMU$(\hat{x}_o, \hat{y}_o)$ is Pareto-efficient in technology $T_{V\,RS-TO}$. According to *Theorem 1*, if $\phi^* = 1$ and optimal vectors e∗ and d∗ are zero vectors, DMUo coincide with DMU $(\hat{x}_o, \hat{y}_o)$ and is accordingly efficient. Then DMUo is inefficient and $(\hat{x}_o, \hat{y}_o)$ can be regarded as its efficient target.

Since it was told in Section 2, firstly, by applying *Model (32)*, the BCC-projection of a DMUo, that is $(\hat{x}_o, \hat{y}_o)$. According to *Definition 2, which* is obtained, it is utilized to compute the congestion. Similarly, at first, there was a weight limitation.

They want to find the radius target of a DMUo utilizing *Model (35)*, which will be utilized to obtain the correspondence efficient target in *Model (37)* according to *Model (38)*. By incorporating the mentioned efficient target in *Model (10)*.

$$\max \sum_{i=1}^{m} \delta_i^-,$$

Where δ_i^{-*} is obtained from *Model (39)*. $d_i^{c^*}$ is then the congesting amount in the total slack connected with $d_i^{c^*}$ in input i = 1, ...,m, as obtained from *Model (1)*, and δ_i^{-*} is the (Maximum) amount of this total slack that can be assigned to completely technical (Non-congesting) inefficiency, as obtained from *Model (39)*.

2.2 | Multi-Stage Data Envelopment Analysis Congestion Model

$$\begin{aligned}
&\text{s.t.}\\
&\sum_{j=1}^{n} \lambda_j x_{ij} + \sum_{t=1}^{k} \pi_t p_t - \delta_i^- = \widehat{x_{io}}, \ i = 1, \dots, n,\\
&\sum_{j=1}^{n} \lambda_j y_{rj} + \sum_{t=1}^{k} \pi_t Q_t = \widehat{y_{ro}}, \ r = 1, \dots, n,\\
&\sum_{j=1}^{n} \lambda_j = 1,\\
&\delta_i^- \leq d_i^*,\\
&\lambda_j \geq 0, j = 1, \dots, n,\\
&\delta_i^- \geq 0, i = 1, \dots, m,\\
&\pi_t \geq 0.
\end{aligned} \tag{39}$$

At last, to identify the congesting inputs and to estimate their amounts, they utilized i = 1, ...,m input constraints $\delta_i^- \leq d_i^*$ in *Model (39)* to achieve [30]:

$$d_i^{c^*} = d_i^* \text{-} \delta_i^{-*}. \tag{40}$$

Assume there are n DMUs and that each DMUj (j=1; 2; . . .; n) has m inputs to the first step, and S outputs from this step $\mathcal{Z}_{sj_o}$(s=1…, s). These S outputs then become the inputs to the second cycle and $\mathcal{Z}_{op_o}$ where o=0, 1…., O is the input or entered as an input of the existing step and other subsequent steps. The outputs from the second, third, and fourth steps are denoted as yr_{jo} where r=1, 2…, R, vl_{jo} where l=1, 2…, L and w_{kjo} where g=1, 2…, G. The weights of cycle 1, cycle 2, cycle 3, and cycle 4 are η_S^A, u_r, μ_l and γ_k . The input weights of steps 1, 2, 3, and 4 are v_i, v_{op}, w_{op} and o_{0p}. They modified a one-step Färe et al. [9] method that proceeds in two steps into multi-stage DEA models. The first stage utilizes an "input-oriented" model as follows:

$$\theta^* = \min\theta, \tag{41}$$

$$\text{s.t.} \ (\textstyle\sum_{s=1}^{S} \eta_S^A z_{sjo}) + (\sum_{r=1}^{s} u_r y_{rjo}) + (\sum_{l=1}^{L} \mu_l v_{ljo}) + (\sum_{g=1}^{G} \gamma_k w_{jko}) \_y_{in} \geq 0. \tag{42}$$

$$\begin{aligned}
&(\textstyle\sum_{r=1}^{s} v_i x_{ijo}) + (\sum_{s=1}^{S} \eta_S^A z_{sjo} + \sum_{r=1}^{s} v_{op} z_{opo}) + (\sum_{r=1}^{R} u_r y_{rjo} + \sum_{q=1}^{Q} w_{op} z_{opo}) +\\
&(\textstyle\sum_{l=1}^{L} \mu_l v_{ljo} + \sum_{r=1}^{s} o_{op} z_{opo})_\theta n x_{kn} \leq 0.
\end{aligned} \tag{43}$$

$$\eta_s^A, v_i, u_r, v_{opo}, \mu_l, w_{op}, \gamma_k, o_{op} \geq 0.$$

The objective is to minimize all of the inputs of DMUo in proportion θ^* where the optimal $\theta = \theta^{**}$ does not surpass unity and the non-negativity of the $\eta_s^A, v_i, u_r, v_{opo}, \mu_l, w_{op}, \gamma_k, o_{op}$ and output implies that the value of θ^* will not be negative under the optimization in *Eq. (1)*.

$$0 \leq \min\theta = \theta^* \leq 1. \quad (44)$$

Technical efficiency is achieved by DMUo if and only if $\theta^* = 1$ Technical inefficiency is present in the performance of DMUo if and only if $0 \leq \theta^* < 1$ Second stage model:

$$\beta^* = \min \beta, \quad (45)$$

$$\text{s.t. } \left(\sum_{s=1}^{S} \eta_S^A z_{sjo}\right) + \left(\sum_{r=1}^{s} u_r y_{rjo}\right) + \left(\sum_{l=1}^{L} \mu_l v_{ljo}\right) + \left(\sum_{g=1}^{G} \gamma_k w_{jko}\right) \_y_{in} \geq 0. \quad (46)$$

$$\left(\sum_{r=1}^{s} v_i x_{ijo}\right) + \left(\sum_{s=1}^{S} \eta_S^A z_{sjo} + \sum_{r=1}^{s} v_{op} z_{opo}\right) + \left(\sum_{r=1}^{R} u_r y_{rjo} + \sum_{q=1}^{Q} w_{op} z_{opo}\right) + \left(\sum_{l=1}^{L} \mu_l v_{ljo} + \sum_{r=1}^{N} o_{op} z_{opo}\right) \_ \beta n x_{kn} \leq 0. \quad (47)$$

$$\eta_s^A, v_i, u_r, v_{opo}, \mu_l, w_{op}, \gamma_k, o_{op} \geq 0.$$

Note that the first i=1…, m qualities in *Eq. (43)* are replaced by *Eq. (47)*. Thus, slack is not possible in the inputs. The fact that only the output can yield a nonzero slack is then referred to as "weak disposal". Hence, $0 = \theta^* \leq \beta^*$ and these results can be used for developing a "measure" of congestion:

$$0 \leq C(\theta^*, \beta^*) = \frac{\theta^*}{\beta^*}. \quad (48)$$

Combining *Models (41)-(45)* in a two-stage manner, they utilized this measure to identify congestion in terms of the following conditions:

I. Congestion is identified as present in the performance of DMUo if and only if:

$$C(\theta^*, \beta^*) \leq 1. \quad (49)$$

II. Congestion is identified as not present in the performance of DMUo if and only if:

$$C(\theta^*, \beta^*) = 1.$$

Our proposed congestion model will multiply the congestion scores of each process cycle to check the presence or absence of congestion in the overall supply chain ${\theta^{o^*}}/{\beta^{o^*}}$.

Output-oriented multi-stage congestion model

$$\theta^* = \min\theta, \quad (50)$$

$$\left(\sum_{s=1}^{S} \eta_S^A z_{sjo}\right) + \left(\sum_{r=1}^{s} u_r y_{rjo}\right) + \left(\sum_{l=1}^{L} \mu_l v_{ljo}\right) + \left(\sum_{g=1}^{G} \gamma_k w_{jko}\right) \_\theta n y_{in} \geq 0. \quad (51)$$

$$\left(\sum_{r=1}^{s} v_i x_{ijo}\right) + \left(\sum_{s=1}^{S} \eta_S^A z_{sjo} + \sum_{r=1}^{s} v_{op} z_{opo}\right) + \left(\sum_{r=1}^{R} u_r y_{rjo} + \sum_{q=1}^{Q} w_{op} z_{opo}\right) + \left(\sum_{l=1}^{L} \mu_l v_{ljo} + \sum_{r=1}^{N} o_{op} z_{opo}\right) \_ x_{kn} \leq 0. \quad (52)$$

$$\eta_s^A, v_i, u_r, v_{opo}, \mu_l, w_{op}, \gamma_k, o_{op} \geq 0.$$

Technical efficiency is achieved by DMUo if and only if $\theta^{**} = 1$. Technical inefficiency is present in the performance of DMUo if and only if $0 \leq \theta^* < 1$. Second stage model:

$$\beta^* = \min \beta, \quad (53)$$

$$\text{s.t. } \left(\sum_{s=1}^{S} \eta_S^A z_{sjo}\right) + \left(\sum_{r=1}^{s} u_r y_{rjo}\right) + \left(\sum_{l=1}^{L} \mu_l v_{ljo}\right) + \left(\sum_{g=1}^{G} \gamma_k w_{jko}\right) \_\beta y_{in} \geq 0. \quad (54)$$

$$\left(\sum_{r=1}^{s} v_i x_{ijo}\right) + \left(\sum_{s=1}^{S} \eta_S^A z_{sjo} + \sum_{r=1}^{s} v_{op} z_{opo}\right) + \left(\sum_{r=1}^{R} u_r y_{rjo} + \sum_{q=1}^{Q} w_{op} z_{opo}\right) + \left(\sum_{l=1}^{L} \mu_l v_{ljo} + \sum_{r=1}^{N} o_{op} z_{opo}\right) \_ x_{kn} \leq 0. \quad (55)$$

$\eta_s^A, v_i, u_r, v_{opo}, \mu_l, w_{op}, \gamma_k, o_{op} \geq 0.$

Combining *Models (50)-(53)* in a two-stage manner, they utilized this measure to identify congestion in terms of the following conditions,

I. Congestion is identified as present in the performance of DMUo if and only if:

$$C(\theta^*, \beta^*) \leq 1. \tag{56}$$

II. Congestion is identified as not present in the performance of DMUo if and only if:

$$C(\theta^*, \beta^*) = 1. \tag{57}$$

Our proposed congestion model will multiply the congestion scores of each process cycle to check the presence or absence of congestion in the overall supply chain ${\theta^{o^*}}/{\beta^{o^*}}$. When the congestion scores of several independent stages occur together.

Wua et al. (Congestion with undesirable outputs)

Assume x, y, and u show the inputs, desirable outputs, and undesirable outputs, respectively. Proof of congestion in this scenario happens whenever reducing some inputs x can increase desirable outputs y and decrease undesirable outputs u concurrently. Before measuring this kind of congestion, the method should develop a method to solve the problem of undesirable outputs. So far, there have been several methods for addressing the undesirable output problem. In this study, the method of Seiford and Zhu [31] is chosen to deal with undesirable outputs. Since the input-oriented model may produce incorrect results in distinguishing congestion, this paper will concentrate on the output-oriented model [12]. The model is shown as follows:

$$\begin{aligned}
&\text{Max } \delta,\\
&\text{s.t. } \textstyle\sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{io},\\
&\textstyle\sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro},\\
&\textstyle\sum_{j=1}^{n} \lambda_j o_{tj} \geq \delta o_{to}.\\
&o_{tj} = -u_{tj} + \alpha_t,\\
&\textstyle\sum_{j=1}^{n} \lambda_j = 1,\\
&\lambda_j \geq 0, \quad i = 1, \dots, m\ , r = 1, \dots, s\ , t = 1, \dots, k, \qquad j = 1, \dots, n,
\end{aligned} \tag{58}$$

Where α_t is a big enough positive number that can make every o_{tj} positive. The fourth constraint is utilized to trans form the undesirable output to a new variable, whose value is the larger the better, by adding a sufficiently positive constant to the negative amount of objectionable output. The third constraint is utilized to constrain the new changeable in the Possible Production Set (PPS). We call this model the BCC SZ model. When the constraints on λ are changed, we can get some other DEA models, for example: $\sum_{j=1}^{n} \lambda_j$ is free (CCR SZ model) $\sum_{j=1}^{n} \lambda_j \leq 1$ (FG SZ model) and $\sum_{j=1}^{n} \lambda_j \geq 1$ (ST SZ model) The corresponding new model, defined as the unew model, for determining the congestion of DMU0, is shown as follows:

$$\begin{aligned}
&\text{Max } \delta,\\
&\text{s.t. } \textstyle\sum_{j=1}^{n} \lambda_j x_{ij} = x_{io},\\
&\textstyle\sum_{j=1}^{n} \lambda_j y_{rj} \geq \delta y_{ro},\\
&o_{tj} = -u_{tj} + \alpha_t,
\end{aligned}$$

$$\sum_{j=1}^{n}\lambda_j o_{tj} \geq \delta o_{to},$$ **(59)**

$$\sum_{j=1}^{n}\lambda_j = 1,$$

$$\lambda_j \geq 0, i = 1, \dots, m \;, r = 1, \dots, s \;, t = 1, \dots, k, \qquad j = 1, \dots, n$$

Comparing *Models (58)* and *(59)* showed that the difference is when they have a different first constraint. The constraint of *Model (58)* is more relaxed than that of *Model (59)*. The optimal solution of *Model (59)* is the feasible solution of *Model (58)*. Also, *Model (59)* has a different PPS, which is closed, because of the first constraint.

Definition 8. If the optimal value of the UNEW model, *Model (59)*, satisfies z^* =1, then we call DMU0 weakly efficient, corresponding to the UNEW model, or weakly efficient in short.

Definition 9. Assume DMU0 is weakly efficient in the UNEW model. If it has $(\hat{x},\hat{y},\hat{u})$ATNEW, $\hat{x} \leq x_o$ and $\hat{x} \neq x_o$, $\hat{y} \leq y_o$, $\hat{u} > u_o$, where $T_{new} = \{(x,y) | \sum_{j=1}^{n}\lambda_j x_{ij} = x_{io}, \sum_{j=1}^{n}\lambda_j y_{rj} \leq y_{ro}, \sum_{j=1}^{n}\lambda_j u_{rj} = u_{ro}, \sum_{j=1}^{n}\lambda_j = 1, \lambda_j \geq 0\}$. Then the DMU evidences congestion.

Definition 10. Assume DMU0 is weakly efficient in the UNEW model. If it has $(\hat{x},\hat{y},\hat{u})$ATNEW, $\hat{x} \leq x_o$ and $\hat{x} \neq x_o$, $\hat{y} \geq y_o$, $\hat{u} \geq u_o$, where $T_{new} = \{(x,y) | \sum_{j=1}^{n}\lambda_j x_{ij} = x_{io}, \sum_{j=1}^{n}\lambda_j y_{rj} \geq y_{ro}, \sum_{j=1}^{n}\lambda_j u_{rj} \leq u_{ro}, \sum_{j=1}^{n}\lambda_j = 1, \lambda_j \geq 0\}$. Then the DMU evidences congestion.

Theorem 9. A weakly efficient DMU0 of the UNEW model evidences congestion if and only if it is not weakly (BCC SZ) DEA efficient, if and only if it is neither weakly (FGSZ)DEA efficient nor weakly (STSZ)DEA efficient.

Abbasi et al. (Estimation of congestion in free disposal Hull models using data envelopment analysis)

In this technique, first, concisely describe some characteristic property of the FDH model. Consider n DMUs where each DMUj $(j = 1, \dots, n)$ uses m inputs xij $(i = 1, \dots, m)$ to produce s outputs yrj $(r = 1, \dots, s)$. Let $xj = (x1j, \dots, xmj)^T$ and $yj = (y1j, \dots, ysj)^T$ We will also assume that $x_j \geq 0$, $x_j \neq 0$ and $y_j \geq 0, y_j \neq 0$ The PPS T is shown as:

$$T = \{(x,y) | \in R_+^{m+s} | \text{ y can be produced from x}\}.$$ **(60)**

This set is denoted by T_{FDH}, concerning the assumptions of deterministic and free disposability of the production technology:

$$T_{FDH} = \{(x,y): \sum_{j=1}^{n}\lambda_j x_j \leq x, \sum_{j=1}^{n}\lambda_j y_j \geq y, \sum_{j=1}^{n}\lambda_j = 1, \; \lambda_j \in (0,1), j = 1, \dots, n\}.$$ **(61)**

The additive FDH model to evaluate the efficiency of an exceptional DMUp $(p \in \{1, \dots, n\})$ under the *T*FDH is as follows:

$$\text{Max} \sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+,$$

$$\text{s.t.} \sum_{j=1}^{n}\lambda_j x_{ij} + s_i^- = x_{ip}, \; i=1\dots, m,$$

$$\sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ = y_{rp}, \; r=1\dots, s,$$ **(62)**

$$\sum_{j=1}^{n}\lambda_j = 1, \; \lambda_j \in (0,1), j=1\dots, n,$$

$$s_i^- \geq 0, \; s_r^+ \geq 0, \; i=1\dots, m, \; r=1\dots, s.$$

Definition 11 (FDH efficiency). Consider *Model (62).* If the optimal objective value is zero, then DMUp is said to be FDH efficient. It is worth noting that, to the CCR and BCC models, the FDH model does not require the convexity assumption. So, this model has a discrete nature, which causes the efficient target point for an inefficient DMU to be assigned as a point among only actually observed DMUs, so the efficiency analysis is done relative to the other given DMUs instead of a hypothetical efficiency frontier. This has the advantage that the achievement goal for an inefficient DMU given by its efficient target point will be more credible than in cases of CCR and BCC models.

Definition 12 (FDH output efficiency). Consider the following model. If z_{FDH}= 0, then DMUp is said to be FDH output efficient:

$$
\begin{aligned}
& z_{FDH}=\max \textstyle\sum_{r=1}^{s} s_r^+, \\
& \text{s.t.} \textstyle\sum_{j=1}^{n} \lambda_j x_{ij} \le x_{ip},\ i=1\ldots, m, \\
& \textstyle\sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ = y_{rp},\ r=1\ldots, s, \\
& \textstyle\sum_{j=1}^{n} \lambda_j = 1\ \ \lambda_j \in (0,1), j=1\ldots, n, \\
& s_i^- \ge 0\ ,\ \ s_r^+ \ge 0,\ i=1\ldots, m,\ r=1\ldots, s.
\end{aligned}
\tag{63}
$$

Definition 13 (congestion). Evidence of congestion is here in the performance of any DMU, when a decrease in one or more inputs is connected with increases that are maximally possible in one or more outputs without worsening other inputs or outputs. Conversely, congestion is said to happen when some of the outputs that are maximally possible are reduced by increasing one or more inputs without improving any other inputs or outputs.

Definition 14 (strong congestion). If a proportionate reduction in all inputs of a DMU warrants an increase in all maximally possible outputs, then strong congestion happens.

Definition 15 (technical efficiency). Efficiency is obtained by DMU0 if and only if it is not possible to improve some of its inputs or outputs without worsening some of its other inputs or outputs.

Definition 16 (technical inefficiency). Technical inefficiency is said to exist in the performance of DMU0 when the evidence shows that it is possible to improve some input or output without worsening some other inputs or outputs.

In T_{FDH}, the efficiency covering is a staircase based on those given DMUs that are not dominated by other given DMUs. It should be noted that evaluating congestion in usual models for convex PPS has been studied on T_{NEW}, which is a PPS without an input disposability element. Let us denote T_{NEW} corresponding to T_{FDH} as T_{NFDH}, which can be defined as follows:

$$
T_{NFDH} = \{(x,y): \sum_{j=1}^{n} \lambda_j x_j = x\,, \sum_{j=1}^{n} \lambda_j y_j \ge y\,, \sum_{j=1}^{n} \lambda_j = 1,\ \ \lambda_j \in (0,1), j = 1,\ldots,n\}. \tag{64}
$$

A new set is introduced as follows:

$$
FDH^{-1} = \{(x,y), \sum_{j=1}^{n} \lambda_j x_j \ge x\,, \sum_{j=1}^{n} \lambda_j y_j \ge y\,, \sum_{j=1}^{n} \lambda_j = 1, \lambda_j \in (0,1), j = 1,\ldots,n\}. \tag{65}
$$

It seems that the set of $FDH^{-1}$is achieved by reversing the sign of the input inequalities in T_{FDH} The following model is used to deal with the congestion occurrence in the FDH model:

$$
z_{FDH^{-1}}= \max \textstyle\sum_{r=1}^{s} s_r^+
$$

$$\text{s.t.} \sum_{j=1}^{n} \lambda_j x_{ij} \geq x_{ip},\ i=1\ldots, m,$$
$$\sum_{j=1}^{n} y_{rj}\lambda_j - s_r^{+} = y_{rp},\ r=1\ldots, s, \tag{66}$$
$$\sum_{j=1}^{n} \lambda_j = 1, \lambda_j \in (0,1)$$
$$s_i^{-} \geq 0,\ s_r^{+} \geq 0,\ j=1\ldots, n, i=1\ldots, m, r=1\ldots, s.$$

They described the above model "FDH^{-1}output additive model." To see what is involved, they commented that the input (like the output) constraints take the form in this adaptation of additive models, the objective is to maximize the outputs without reducing any of the input constraints take the from $\sum_{j=1}^{n} \lambda_j x_{ij} \geq x_{ip.}$Since, in this adaptation of additive models, the objective is to maximize the outputs without reducing any of the inputs.

Definition 17 (FDH^{-1}output efficiency). Consider the *Model (66)*. If $Z_{FDH^{-1}}= 0$, then it is said to be FDH^{-1} output efficiently.

Lemma 1. DMUp is FDH^{-1}output efficient if and only if the following system has no solution:

$$\sum_{j=1}^{n} \lambda_j x_j \geq x_p,$$
$$\sum_{j=1}^{n} \lambda_j y_j \geq y_p, \sum_{j=1}^{n} \lambda_j y_j \neq y_p, \tag{67}$$
$$\sum_{j=1}^{n} \lambda_j = 1, \lambda_j \in (0,1), j = 1, \ldots, n.$$

Definition 18 (congestion in the FDH model). Le DMUp $t = (xp, yp)$ be FDH^{-1}output efficient; if there exists DMUk $= (x_k, y_k)$, such that $x_k \leq x_p, x_k \neq x_p$ and $y_p \leq y_k, y_p \neq y_k$, then DMUp has evidence of congestion.

Definition 19 (strong congestion in the FDH model). Let DMUp $= (xp, yp)$ be congested in FDH model; if there exists DMUk $= (x_k, y_k)$, such that $x_k < x_p$ and $y_k > y_p$, then DMUp has evidence of strong congestion.

Lemma 2. Let DMUp be FDH^{-1} output efficient; then DMUp has evidence of congestion if and only if the following system has a solution:

$$\sum_{j=1}^{n} \lambda_j x_j \leq x_p,$$
$$\sum_{j=1}^{n} \lambda_j y_j \geq y_p, \sum_{j=1}^{n} \lambda_j y_j \neq y_p, \tag{68}$$
$$\sum_{j=1}^{n} \lambda_j = 1\ \ \lambda_j \in (0,1), j = 1, \ldots, n.$$

Lemma 3. DMUp is not FDH output efficient if and only if the following linear system has a solution:

$$\sum_{j=1}^{n} \lambda_j y_j \geq y_p, \sum_{j=1}^{n} \lambda_j y_j \neq y_p, \tag{69}$$
$$\sum_{j=1}^{n} \lambda_j = 1\ \ \lambda_j \in (0,1),\ j = 1, \ldots, n.$$

Theorem 10. Let DMUp be FDH^{-1}output efficient; then ZFDH, DMUp has evidence of congestion if and only if DMUp is not FDH output efficient.

By utilizing *Theorem 13*, we can provide the following procedure to evaluate congestion in the FDH model.

I. Solve *Model (66)* corresponding to (xp, yp); let (λ^*, s^{+^*}) be the optimal solution of it. Let $\hat{y}_p = y_p + s^{+^*}$ It is evident that $(x_p, \hat{y}_p)$ is FDH^{-1} output efficient.

II. Solve *Model (63)* for $(x_p, \hat{y}_p)$.

III. If $z_{FDH} > 0$, then DMUp is congested.

Models (63)-(66) are mixed-integer programming, but can show that it does not need any mathematical programming problem to solve. Actually, an enumeration algorithm based on pairwise comparisons, similar to Tulken's enumeration algorithm for the case of the radial FDH model, can be utilized. Now, based upon the foregoing procedure and previous explanation, they proposed the following algorithm. The proposed algorithm includes two parts. In Part (a), the existence of congestion in the performance of DMUp, and Part (b), if DMUp is recognized to be congested in Part (a), the amount of congestion for each input, as well as the reduction amount of each output due to congestion, will be estimated.

Proposed algorithm

Part (a)

Step 5. Calculate the optimal value of *Model (7)* by the following equation:

$$FDH^{-1} = \sum_{r=1}^{s}\left(y_{rq} - y_{rp}\right) = Max_{j\epsilon D}\sum_{r=1}^{s}\left(y_{rj} - y_{rp}\right), \tag{70}$$

Where

$$Dp = \{j \in \{1, \dots, n\} \mid x_j \geq x_p \text{and}\ y_j \geq y_p\}. \tag{71}$$

Step 6. Let $\hat{y}_p = y_p + s^{+^*}$, where $s^{+^*} = y_q - y_p$ Obtain the optimal value of *Model (63)* by:

$$z_{FDH} = max_{j\epsilon\hat{D}}\sum_{r=1}^{s}\left(y_{rj} - \hat{y}_{rp}\right), \tag{72}$$

Where

$$\hat{D}_p = \{j \in \{1, \dots, n\} \mid x_j \leq x_p, y_j \geq \hat{y}_p\} \tag{73}$$

Step 7. If $Z_{FDH} > 0$, then DMUp is congested, so go to Part (b); furthermore, if there exists $j \in \hat{D}_p$ such that $x_j < x_p$ and $y_j > \hat{y}_p\}$, then, based on *Definition 4*, DMUp is strongly congested. If $Z_{FDH} = 0$, then DMUp is not congested and stops.

Part (b)

Step 8. Define Kp as follows:

$$k_p = \{j \in \hat{D}_p \mid Z_{FDH} = \textstyle\sum_{r=1}^{s}\left(y_{rj} - \hat{y}_{rp}\right)\}. \tag{74}$$

Then calculate

$$\alpha^* = min\sum_{i=1}^{m}\left(x_{ip} - x_{ij}\right). \tag{75}$$

Step 9. Define T_P as follows:

$$T_P = \{j\epsilon k_p | \alpha^* = \textstyle\sum_{i=1}^{m}\left(x_{ip} - x_{ij}\right)\}. \tag{76}$$

For $j \in T_P$ define $s_i^{c^*}$ as the amount of congestion in the i[th] input of DMUP and $\hat{s}_r^{+^*}$ as a reduction in the amount of r[th] output due to congestion, as follows:

$$s_i^{c^*} = x_{ip} - x_{ij}, \text{ i=1...., m,}$$

$$\hat{s}_r^{+^*} = y_{rj} - \hat{y}_{rp}, \text{ r=1...., s.}$$

$\alpha^* = \sum_{i=1}^{m} s_i^{c^*}$ is the total amount of congestion in all inputs of DMUp.

Corollary 1. If $\hat{D}_p = 0$, congestion has no appearance at DMU_p.

Hajihosseini et al. approach

One of the famous basic DEA models is "Multipplier form of BCC with output oriented" [13]. The efficiency for DMUo of BCC is output-oriented. The efficiency for DMUo is evaluated by this model, which states that each DMU has m inputs and s outputs:

$$\begin{aligned} &\text{Min } \frac{v^t x_0 + v_0}{u^t y_0},\\ &\text{s.t. } \frac{v^t j + v_0}{u^t j} \geq 1, \text{ j=1....,n,}\\ &u^t \geq 1_s \varepsilon,\\ &v^t \geq 1_m \varepsilon, \end{aligned} \quad (77)$$

Where the decision variables are the weight vectors $u^t = (u_1, \dots, u_s)$, $v^t = (v_1, \dots, v_m)$ and $x_j^t = (x_{1j}, \dots, x_{mj})$, $y_j^t = (y_{1j}, \dots, y_{mj})$ are the input and output vectors for DMUj (j = 1, …, n). If the optimal value of the above model is equivalent to one, DMUo is efficient; otherwise, it is inefficient. The equivalent linear programming of the above model will be substituted as follows.

$$\begin{aligned} &\text{Min } v^t x_0 + v_0,\\ &\text{s.t. } u^t y_0 = 1,\\ &v^t x_j + v_0 - u^t y_j \geq 0,\\ &u^t \geq 1_s \varepsilon,\\ &v^t \geq 1_m \varepsilon, \end{aligned} \quad (78)$$

As we can see in the feasible region of *Model (78)* $v^t x_j + v_0 - u^t y_j \geq 0$ (j = 1, …, n). In this model, the DMU is more favored when it has a smaller value. Hence, when we minimize the $v^t x_j + v_0 - u^t y_j \geq 0$(j = 1, …, n), we can reach our purpose. Therefore, rather than solving the above model, one may minimize the $v^t x_j + v_0 - u^t y_j \geq 0$ (j = 1…, n) with respect to the same region. Therefore, based on Noura and Hoseini [32] methodology, we proposed the following Multi-Objective Linear Programming (MOLP) to explain CSW:

$$\begin{aligned} &\text{Min } v^t x_j + v_0 - u^t y_j, \text{ j=1..., n,}\\ &\text{s.t. } v^t x_j + v_0 - u^t y_j \geq 0, \text{ j=1..., n,}\\ &u^t \geq 1_s \varepsilon,\\ &v^t \geq 1_m \varepsilon. \end{aligned} \quad (79)$$

The CSW with equivalent weights is applied to solve the above MOLP as follows:

$$\text{Min } \sum_{j=1}^{n} \left(v^t x_j + v_0 - u^t y_j \right),$$

$$\text{s.t. } v^t x_j + v_0 - u^t y_j \geq 0, \tag{80}$$

$$u^t \geq 1_s \varepsilon,$$

$$v^t \geq 1_m \varepsilon.$$

Suppose (u∗, v∗)t will be the optimal solution of the *Model (79)*, so it is considered as CSW, and for ranking and comparing DMUs, they utilize the efficiency score of DMU_j (j = 1, …, n) as $\varphi_j^* = \frac{v^{*t}x_j + v_0^*}{u^{*t}y_j}$ When φ_j^* is equivalent to one can conclude the DMU under evaluation (DMUj) is efficient. Now, As Noura and Hoseini [32]. E is defined as follows: $E = \{j: \varphi_j^* = 1\}$. And x_i^* is defined as the highest value of each input for all elements in the set of E. The following sub-definition is suggested to identify congestion.

Definition 20. Congestion in DMUo eventually occurs if the optimal solution of φ_j^* for DMUo, the following condition is satisfied: $\varphi_j^* > 1$ and there is at least one $x_{io} > x_i^*$, i= 1, ..., m. The amount of congestion in the ith input of DMUo is denoted by $s_i^{c'}$ as follows:

$$s_i^{c'} = x_{io} - x_i^*.$$

The sum of all $s_i^{c'}$ is the amount of congestion in DMU_o. Congestion does not present when $x_{io} \leq x_i^*$ or $s_i^{c'} = 0$ for all i =1, …, m.

Theorem 11. The amount of congestion in the proposed method is equal to the amount of congestion in the Cooper et al. method, but vice versa is not true.

Noura and Hoseini's method

Assume there are n DMUs that are evaluated in terms of m inputs and s outputs [32], [33]. Let xij and yrj be input and output values of DMUj for i=1, ..., m and r=1, …, s, spot BCC model. The efficiencies of the DMUs utilizing weight restrictions are measured by the sub-m*odel (11)*.

$$\text{Min } \sum_{i=1}^{n} v_i x_{ip} + v_o,$$

$$\text{s.t. } \sum_{r=1}^{s} u_r y_{rp} = 1, \tag{82}$$

$$\sum_{i=1}^{n} v_i x_{ip} + v_o - \sum_{r=1}^{s} u_r y_{rj} \leq 0, \text{ j=1..., m},$$

$$\sum_{i=1}^{m} v_i p_{ik} \leq 0, \ k = 1, \ldots, 2m-2,$$

$$\sum_{i=1}^{m} u_r q_{rt} \leq 0, t = 1, \ldots, 2s-2,$$

$$u_r \geq \varepsilon, \ \text{ r=1..., s},$$

$$v_i \geq \varepsilon, \ \text{i=1..., m},$$

Where p_{m*2m-2}=(pink) and q_{s*2s-2} = (qrt) are matrices that are connected with weight restrictions as described below. Such as, if the ratio of weights for the initial and ith of input and initial and rth of output is as follows:

$$l_{1i} \leq \frac{v_i}{v_1} \leq u_{1i}, \ l_{1i} v_i \leq v_i \leq u_{1i} v_1, \ i = 2, \ldots, m.$$

$$L_{1r} \leq \frac{u_r}{u_1} \leq U_{1r}, \ L_{1r} u_1 \leq u_r \leq U_{1r} u_1, \ \ r = 2, \ldots,$$

Where l_{1i} and u_{1i} are the lower and upper bounds of $\frac{v_i}{v_1}$ and L_{1r} and U_{1r} are the lower and upper bounds of $\frac{u_r}{u_1}$. In this case, the matrices P and Q are defined as follows:

$$P=\begin{bmatrix} l_{12} & -u_{12} & l_{13} & -u_{13} & \ldots\ldots \\ -1 & 1 & 0 & 0 & \vdots \\ \ldots & \ldots & \ldots & \ldots & \ldots\ldots \end{bmatrix}.$$

$$Q=\begin{bmatrix} L_{12} & -U_{12} & L_{13} & -U_{13} & \ldots\ldots \\ -1 & 1 & 0 & 0 & \vdots \\ \ldots & \ldots & \ldots & \ldots & \ldots\ldots \end{bmatrix}.$$

Congestion with weight restriction utilizing common weights: Based on Noura and Hoseini [32]. Methodology, they proposed the following MOLP with weight restriction utilizing a common set of weights, *Model (2)*.

$$\begin{aligned} & \text{Min } v^t x_j + v_o - u^t y_j,\ j=1\ldots., n, \\ & \text{s.t.} \quad v^t x_j + v_o - u^t y_j \geq 0,\ j=1\ldots., n, \\ & \sum_{i=1}^{m} v_i p_{ik} \leq 0,\ k = 1, \ldots., 2m-2, \\ & \sum_{i=1}^{m} u_r q_{rt} \leq 0, t = 1, \ldots., 2s-2, \\ & u^t \geq 1_s \varepsilon, \\ & v^t \geq 1_m \varepsilon, \end{aligned} \tag{83}$$

Where the decision variables are the weight vectors $v^t = (v_1, \ldots\ldots, v_m) u^t = (u_1, \ldots\ldots, u_s)$ and $x_j^t = (x_{1j}, \ldots\ldots, x_{mj}), y_j^t = (y_{1j}, \ldots\ldots, y_{sj})$ are the input and output vectors for $DMU_j (j = 1\ldots, n)$. The equivalent weights method is applied to solve the above MOLP. It also assumes that all weights are equivalent to one. As a result, we obtain *Model (3)*.

$$\begin{aligned} & \min \textstyle\sum_{i=1}^{n} v^t x_j + v_o - u^t y_j, \\ & v^t x_j + v_o - u^t y_j \geq 0,\ j=1\ldots.,n, \\ & \sum_{i=1}^{m} v_i p_{ik} \leq 0, k = 1, \ldots., 2m-2, \\ & \sum_{i=1}^{m} u_r q_{rt} \leq 0, t = 1, \ldots., 2s-2, \\ & u^t \geq 1_s \varepsilon, \\ & v^t \geq 1_m \varepsilon. \end{aligned} \tag{84}$$

This implies *Model (4)*.

$$\begin{aligned} & \min \textstyle\sum_{i=1}^{n} v^t x_j + v_o - u^t y_j, \\ & v^t x_j + v_o - u^t y_j - \Delta_j = 0, \\ & \sum_{i=1}^{m} v_i p_{ik} \leq 0, \\ & v^t x_j + v_o - u^t y_j - \Delta_j = 0, j = 1 \ldots., n, \\ & \sum_{i=1}^{m} v_i p_{ik} \leq 0, k = 1, \ldots., 2m-2, \\ & \sum_{i=1}^{m} u_r q_{rt} \leq 0, t = 1, \ldots., 2s-2, \end{aligned} \tag{85}$$

$$\sum_{i=1}^{m} v_i p_{ik} \leq 0,$$
$$u^t \geq 1_s \varepsilon,$$
$$v^t \geq 1_m \varepsilon,$$
$$\Delta_j \geq 0.$$

From $v^t x_j + v_o - u^t y_j - \Delta_j = 0$, we have $v^t x_j + v_o - u^t y_j = \Delta_j$. Hence, we obtain *Model (5)*.

$$\begin{aligned}
&\text{Min}, \textstyle\sum_{i=1}^{n} \Delta_j, \\
&v^t x_j + v_o - u^t y_j - \Delta_j = 0, \ j = 1 \ldots, n, \qquad \textbf{(86)}\\
&\sum_{i=1}^{m} v_i p_{ik} \leq 0, \ k = 1, \ldots, 2m-2, \\
&\sum_{i=1}^{m} u_r q_{rt} \leq 0, \ t = 1, \ldots, 2s-2, \\
&u^t \geq 1_s \varepsilon, \\
&v^t \geq 1_m \varepsilon, \\
&\Delta_j \geq 0.
\end{aligned}$$

Now assume $(u^*, v^*)^t$ will be the optimal solution of *Problem (86)*, which is named the Common Set of Weights (CSW) with weight restriction for ranking and comparing DMUs. According to the achieved CSW, the efficiency score of DMUj (j = 1, …,n) will be $\varphi_j^* = \frac{v^{*t} x_j + v_0^*}{u^{*t} y_j}$ If φ_j^* is equivalent to one, then the DMU under evaluation is efficient. Now, according to Noura and Hoseini [32], the efficient set of DMUs (E) is defined as follows:

$$E = \{j: \varphi_j^* = 1\}.$$

The highest value in each input among DMUs of E for all components is introduced with $x_i^* = \max\{x_{ij}: j \in E\}$, j=1…., n. So, the following revised definition is suggested to identify congestion.

Definition 21. Congestion in DMUo eventually happens if for the optimal solution of DMUo φ_0^*, the following condition is satisfied $\varphi_j^* > 1$ and there is at least one $x_{io} > x_i^*$, i=1…., m.

The amount of congestion in the ith input of DMUo is shown with s_{io}^c as follows: $s_{io}^c = x_{io} - x_i^*$

The sum of all s_{io}^c is the amount of congestion in DMUo.

$$s_{io}^c = \sum_{i=1}^{m} s_{io}^c.$$

Congestion does not show in DMUo when $x_{io} \leq x_i^*$ or s_{io}^c for all i=1,,m.

Karimi et al. approach

In this research, they considered n DMUs shown by $\{(x_j, y_j), \ j = 1, \ldots, n\}$, they assumed that each DMU produces the same set of outputs by consuming the same set of inputs, and the sole dissimilarity may be in the quantity of inputs and outputs. For each DMUj (Where j=1,…,n), they denoted the non-negative input and output vectors by $x_j = (x_{1j}, \ldots, x_{mj})^t$, $y_j = (y_{1j}, \ldots, y_{sj})^t$ (Where, t is the sign of transposition) [33]. For the sake of simplicity in the notations, they utilized $x_j = [x_1, \ldots, x_n]^t$ and $y_j = [y_1, \ldots, y_n]^t$ to denote, respectively, the m × n input matrix and the s × n output matrix. Ordinary DEA models assume that all data are allowed to take positive real values. Although in many practical cases, some inputs and/or outputs can only take integer values. Lozano and Villa [34], [35] were the pioneers in paying attention to this difference. They introduced integer constraints into DEA models and proposed a MILP model for evaluating the efficiency of DMUs. In the same context, Kazemi and Kuosmanen [36] did some studies in integer-valued

DEA. The works of Matin Kazemi and Kuosmanen [36] are based on some dictum. For the sake of completeness of the current paper, they presented a summarized description of their dictum. To this method, they supposed that T is the PPS of the integer-valued DEA, defined by: $T = \{(x, y)|x \in Z_+^m \text{ can produce } y \in Z_+^s\}$.

The dictum of Matin Kazemi and Kuosmanen [37] is as follows:

I. Envelopment: $(x_j, y_j) \in T$:for all $j = 1, \ldots, n$.

II. Natural disposability: (x,y)∈T and (u,v)∈ $Z_+^{(m+s)}$:y≥ $v \to (x + u, y - v) \in T$.

III. Natural convexity: If $(x_1, y_1), (x_2, y_2) \in T$ and (x,y)=$\lambda(x_1, y_1) + (1 - \lambda)(x_2, y_2)$ where $\lambda \in [0,1]$ then (x,y)∈ $Z_+^{(m+s)} \to (x, y) \in T$

IV. Natural divisibility: (x,y)∈ T and $\exists\lambda \in [0,1]$: $(\lambda x, \lambda y) \in Z_+^{(m+s)} \to (\lambda x, \lambda y) \in T$

V. Natural augment ability: (x,y)∈ T and $\exists\lambda \geq 1 (\lambda x, \lambda y) \in Z_+^{(m+s)} \to (\lambda x, \lambda y) \in T$.

These dictums are integer variants of the standard dictums in conventional DEA. Exactly, the dictums are dissimilar from standard dictums only in the type of input and output vectors that must be integer-valued. More detailed descriptions of these dictums can be found in Matin Kazemi and Kuosmanen [36], [37]. According to the presented axioms, one can construct a difference of PPSs. In this research, they used the following PPS satisfying the dictums 1, 2, and 3:

$$T_{VRS}^{IDEA} = \{(x, y) \in Z_+^{(m+s)}|x \geq \sum_{j=1}^{n} \lambda_j x_j, y \leq \sum_{j=1}^{n} \lambda_j y_j, \sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0 \text{ for all } j\}.$$

Based on this PPS, Matin Kazemi and Kuosmanen [36], [37] showed an input-oriented radial model. In their model, the input and output variables are classified into two categories. The classification is based on the type of the variables, i.e., whether they are continuous or integer. In the following, they showed an output-oriented version of their model. In this model, I stand for integer input/output, and the subsets of integer-valued and real-valued inputs, integer-valued and real-valued outputs are denoted by I^I, I^{NI}, O^IandO^{NI}respectively (where I^{NI} means non-integer). *Model (87)*:

$$\text{Max } [\varphi + \varepsilon(\textstyle\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+ + \sum_{i=1}^{m} s_r^l)], \quad \textbf{(87)}$$

$$\text{s.t.} \textstyle\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io}, \quad i \in I,$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = \varphi y_{ro}, r \in O \backslash O_I,$$

$$\textstyle\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = \hat{y}_r, \ r \in O_I,$$

$$\hat{y}_r = \varphi y_{ro} + s_r^l, \ r \in O_I,$$

$$\sum_{j=1}^{n} \lambda_j = 1,$$

$$s_i^- \geq 0, \text{for all j},$$

$$s_r^l \geq 0, \text{for all } r = 1, \ldots, p,$$

$$s_i^- \in Z_+,$$

$$s_r^+ \geq 0, \text{for all r},$$

$$\hat{y}_r \in Z_+, i \in I,$$

$$r \in O_I,$$

$\lambda_j \geq 0, j = 1, \dots, n,$

Where ε is a small non-Archimedean positive number and $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)^t$ is a column vector of unknown variables that are often referred to as structural or intensity variables. They are utilized for connecting the input and output vectors via convex combination. Variables s_r^+, s_i^-, s_r^l represent the non-radial slack variables, and in the end $\hat{y}_r \in Z_+$ is the integer-valued reference point for outputs O_I. The objective of this model amounts to finding the maximum possible extension that can be done on outputs without any need to use up more inputs.

We note that the difference between congestion and technical inefficiency relies on the fact that, in the case of technical inefficiency, it is possible to decrease the inputs without reducing output. The outputs (Or increasing the outputs without increasing the inputs). In other words, technical inefficiency can be considered as an excess that we utilize in some inputs or as a shortfall that we produce in some outputs. So, when there is congestion in the system, any decrement in the inputs leads to some increments in (At least) one or more outputs; without worsening any other inputs or outputs (Or, in a similar way, any increment of inputs leads to some decrements in, at least, one or more outputs; without improving any other input or output).

3 | Conclusion

This comprehensive review examined various methodologies developed within DEA for assessing congestion and inefficiencies in DMUs. The study categorized approaches based on their focus, such as input and output orientations, multi-stage models, weight restrictions, and techniques addressing undesirable outputs, integer data, and production trade-offs.

Notable models include the classical DEA formulations extended to detect and measure congestion phenomena, the efficiency interval and projection methods, and advanced frameworks incorporating production trade-offs and weight restrictions.The review highlighted the distinction between congestion and technical inefficiency, emphasizing that congestion occurs when increased inputs can paradoxically lead to reduced outputs without worsening other parameters. Several innovative models, like the FDH (Free disposal Hull) and its variants, were explained, providing credible, data-driven congestion measures that rely on observed data points.

The article discussed algorithms, projection techniques, and the integration of weights to identify and quantify congestion levels, including strong and weak congestion scenarios. The synthesis underscores that these varied DEA techniques offer robust tools for researchers and practitioners aiming to diagnose, measure, and improve operational efficiency amid resource overuse, ultimately enriching the analytical capabilities for performance evaluation in complex systems.

Conflict of Interest Disclosure

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial or non-financial interest in the subject matter discussed in this manuscript.

Data Availability Statement

The datasets used and/or analyzed during the current study are not publicly available due to [reason if applicable], but can be made available by the corresponding author when scientifically justified.

Funding Statement

The authors confirm that no financial support was provided for the research, authorship, or publication of this article.

References

[1] Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, *2*(6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

[2] Koopmans, T. C. (1951). Analysis of production as an efficient combination of activities. *Analysis of production and allocation*, *158*(1), 33. https://www.sciepub.com/reference/39504

[3] Färe, R., Grosskopf, S., Lovell, C. A. K., & Pasurka, C. (1989). "Multilateral productivity comparisons when some outputs are undesirable: A nonparametric approach." *Review of economics and statistics*, *71*(1), 90–98. https://doi.org/10.2307/1928055

[4] 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. https://doi.org/10.1287/mnsc.30.9.1078

[5] Stigler, G. J. (1976). The xistence of x-efficiency. *The American economic review*, *66*(1), 213–216. https://www.jstor.org/stable/1804963

[6] Leibenstein, H. (1966). Allocative efficiency vs." x-efficiency". *The American economic review*, *56*(3), 392–415. https://www.jstor.org/stable/1823775%0A

[7] Färe, R., & Svensson, L. (1980). Congestion of production factors. *Econometrica: Journal of the econometric society*, *48*(7), 1745–1753. https://doi.org/10.2307/1911932

[8] Färe, R., & Grosskopf, S. (1983). Measuring congestion in production. *Zeitschrift für nationalökonomie journal of economics*, *43*(3), 257–271. https://www.jstor.org/stable/41796223

[9] Färe, R., Grosskopf, S., & Lovell, C. A. K. (1985). *The measurement of efficiencies of production*. Boston: Kluwer-Nihoff Publishing. https://doi.org/10.1007/978-94-015-7721-2

[10] 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

[11] Brockett, P. L., Cooper, W. W., Wang, Y., & Shin, H. C. (1998). Inefficiency and congestion in Chinese production before and after the 1978 economic reforms. *Socio-economic planning sciences*, *32*(1), 1–20. https://B2n.ir/rh2344

[12] Cooper, W. W., Seiford, L. M., & Zhu, J. (2000). A unified additive model approach for evaluating inefficiency and congestion with associated measures in DEA. *Socio-economic planning sciences*, *34*(1), 1–25. https://www.deafrontier.net/papers/SEPScongestion.pdf

[13] Cooper, W. W., Seiford, L. M., & Zhu, J. (2001). Slacks and congestion: Response to a comment by R. Färe and S. Grosskopf. *Socio-economic planning sciences*, *35*(3), 205–215. https://doi.org/10.17485/ijst/2015/v8i6/61948

[14] Färe, R., Grosskopf, S., & Lovell, C. A. K. (1994). *Production frontiers*. Cambridge University Press. https://B2n.ir/nf7981

[15] Cooper, W. W., Gu, B., & Li, S. (2001). Note: Alternative treatments of congestion in DEA-a response to the Cherchye, Kuosmanen and post critique. *European journal of operational research*, *132*(1), 81–87. https://www.sciencedirect.com/science/article/pii/S0377221700001764

[16] Cherchye, L., Kuosmanen, T., & Post, T. (2001). Alternative treatments of congestion in DEA: A rejoinder to Cooper, Gu, and Li. *European journal of operational research*, *132*(1), 75–80. https://B2n.ir/qt7156

[17] 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://doi.org/10.1016/j.ejor.2007.12.022

[18] , A., Lotfi, F. H., & others. (2015). Congestion in DEA under weight restriction using common weights. *Journal of mathematical extension*, *10*, 121–133. https://ijmex.com/index.php/ijmex/article/view/367

[19] Färe, R., Grosskopf, S., & Lovell, C. K. (1985). *The measurement of efficiency of production* (Vol. 6). Springer Science & Business Media. https://B2n.ir/ff1558

[20] Cooper, W. W., Deng, H., Huang, Z. M., & Li, S. X. (2002). A one-model approach to congestion in data envelopment analysis. *Socio-economic planning sciences*, *36*(4), 231–238. https://B2n.ir/es2070

[21] 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

[22] Brockett, P., Cooper, W., Deng, H., Golden, L., & Ruefli, T. (2004). Using DEA to identify and manage congestion. *Journal of productivity analysis*, *22*, 207–226. https://doi.org/10.1007/s11123-004-7574-0

[23] Noura, A. A., Lotfi, F. H., Jahanshahloo, G. R., Rashidi, S. F., & Parker, B. R. (2010). A new method for measuring congestion in data envelopment analysis. *Socio-economic planning sciences*, *44*(4), 240-246. https://doi.org/10.1016/j.seps.2010.06.003

[24] Flegg, A. T., & Allen, D. O. (2009). Congestion in the Chinese automobile and textile industries revisited. *Socio-economic planning sciences*, *43*(3), 177–191. https://doi.org/10.1016/j.seps.2008.10.003

[25] Wang, Y. M., Greatbanks, R., & Yang, J. B. (2005). Interval efficiency assessment using data envelopment analysis. *Fuzzy sets and systems*, *153*(3), 347–370. https://doi.org/10.1016/j.fss.2004.12.011

[26] Javanmard, M., & Mishmast Nehi, H. (2019). A solving method for fuzzy linear programming problem with interval type-2 fuzzy numbers. *International journal of fuzzy systems*, *21*, 882–891. https://doi.org/10.1007/s40815-018-0591-3

[27] Davoudi, N., Hamidi, F., & Mishmast Nehi, H. (2023). A method for solving interval type-2 Triangular fuzzy Bilevel linear programming problem. *Yugoslav journal of operations research*, *33*(1), 71–90. https://doi.org/10.2298/YJOR210715027H

[28] Podinovski, V. V. (2007). Computation of efficient targets in DEA models with production trade-offs and weight restrictions. *European journal of operational research*, *181*(2), 586–591. https://doi.org/10.1016/j.ejor.2006.06.041

[29] Podinovski, V. V. (2004). Production trade-offs and weight restrictions in data envelopment analysis. *Journal of the operational research society*, *55*(12), 1311–1322. https://doi.org/10.1057/palgrave.jors.2601794

[30] Gattoufi, S., Oral, M., & Reisman, A. (2004). Data envelopment analysis literature: A bibliography update (1951--2001). *Journal of socio-economic planning sciences*, *38*(2–3), 159–229. https://B2n.ir/nx5174

[31] Seiford, L. M., & Zhu, J. (2005). A response to comments on modeling undesirable factors in efficiency evaluation. *European journal of operational research*, *161*(2), 579–581.

[32] Noura, A. A., & Hoseini, E. (2013). Measuring congestion in data envelopment analysis with common weights. *International journal of data envelopment analysis*, *3*(1), 627–632. (**In Persian**). https://www.sid.ir/FileServer/JE/5072820130301

[33] Karimi, B., Khorram, E., & Moeini, M. (2016). Identification of congestion by means of integer-valued data envelopment analysis. *Computers & industrial engineering*, *98*, 513–521. https://doi.org/10.1016/j.cie.2016.06.017

[34] Lozano, S., & Villa, G. (2006). Data envelopment analysis of integer-valued inputs and outputs. *Computers & operations research*, *33*(10), 3004–3014. https://doi.org/10.1016/j.cor.2005.02.031

[35] Lozano, S., & Villa, G. (2007). Integer DEA models: How DEA models can handle integer inputs and outputs. *Modeling data irregularities and structural complexities in data envelopment analysis*, 271–289. https://doi.org/10.1007/978-0-387-71607-7_15

[36] Matin Kazemi, R., & Kuosmanen, T. (2009). Theory of integer-valued data envelopment analysis under alternative returns to scale axioms. *Omega*, *37*(5), 988–995. https://doi.org/10.1016/j.omega.2008.11.002

[37] Kuosmanen, T., & Matin Kazemi, R. (2009). Theory of integer-valued data envelopment analysis. *European journal of operational research*, *192*(2), 658–667. https://doi.org/10.1016/j.ejor.2007.09.040