Operational Research (2024) 24:4
https://doi.org/10.1007/s12351-023-00811-8

ORIGINAL PAPER

Embracing fairness within a cross-efficiency hierarchical network DEA system

Siwei Xiao[1] · Marios Dominikos Kremantzis[1] · Leonidas Sotirios Kyrgiakos[2] · George Vlontzos[2] · Panos M. Pardalos[3]

Received: 4 August 2023 / Revised: 26 November 2023 / Accepted: 4 December 2023 /
Published online: 27 December 2023

Abstract
Several scholars have utilized hierarchical network Data Envelopment Analysis modeling techniques to assess the performance of complex structures. However, there has been limited consideration given to the integration of a peer-appraisal setting within a self-evaluation hierarchical context. This aims to enhance discriminatory power and mitigate the issue of unrealistic weighting scheme. To this end, our study extends the single-stage hierarchical additive self-evaluation model of Kao (Omega 51:121–127, 2015. https://doi.org/10.1016/j.omega.2014.09.008), by integrating the well-established cross-efficiency method. An original combination of a maxmin secondary goal model and the Criteria Importance Through Inter-criteria Correlation (CRITIC) method is proposed, to expand the basic hierarchical self-evaluation model. The maxmin model addresses the issue of the non-unique optimal multipliers obtained from the self-evaluation model, ensuring a more realistic weight scheme. The CRITIC method, that tackles the aggregation problem by objectively determining weights of criteria, rewards the minority and is conducive to a fairer evaluation. Results indicate that the proposed approach is more likely to obtain a unique efficiency and ranking score for the units under consideration. This study entails a numerical experimentation aimed at evaluating the efficiency of a set of 20 universities while validating the applicability of our proposed approach. To conclude, the practical applications of this methodological framework could encompass assessing services within the higher education sector or fostering sustainable development across various operations within a hierarchical structure.

Keywords Data envelopment analysis · Hierarchical network · Cross-efficiency · Fairness · CRITIC · Higher education

Extended author information available on the last page of the article

1 Introduction

1.1 Traditional DEA and the emergence of network DEA

Data envelopment analysis (DEA) is a mathematical programming-based approach for the comparison of the relative efficiency of typically homogeneous decision-making units (DMUs) that make use of inputs to yield outputs (Sinuany-Stern 2023). The potency of the DEA approach has become increasingly compelling as its application expands across a diverse array of fields such as agriculture (Kyrgiakos et al. 2023), banking (Omrani et al. 2023), healthcare (Azadi et al. 2023), information technology (Zhang et al. 2023), supply chain management and transportation (Saen et al. 2022; Miškić et al. 2022; Wong and Wong 2007), sustainable development (Goto and Sueyoshi 2023; Vaez-Ghasemi et al. 2022), and higher education (Kounetas et al. 2023; Kremantzis et al. 2022a, b; Ghasemi et al. 2020; Kao 2015).

The DEA model was originally proposed by Charnes et al. (1978) and later expanded by Banker et al. (1984). A traditional DEA structure essentially treats each DMU as a comprehensive system, without considering its internal operational mechanism. This traditional black-box model can potentially cause measurement bias and weakness in detecting inefficient components (Zhong et al. 2021; Liu and Wang 2015; Lewis and Sexton 2004). Additionally, since a black-box structure evaluates a DMU at its best possible light, the non-unique input and output multipliers of DMUs pose a challenge to how realistic and acceptable the weighting scheme is. To cope with those issues, the concept of network DEA has been proposed (Kao 2017; Färe and Grosskopf 2000). Network DEA is capable of providing managerial insights for organizational resource control, particularly in finding inefficient components and devising targeted remedial measures (Ratner et al. 2023; Lozano and Khezri 2021; Chu and Zhu 2021; Avkiran 2009).

When measuring the efficiency of a network structure, two fundamental approaches are commonly employed: efficiency decomposition and efficiency aggregation (Lu et al. 2022; Kao 2016). These methods diverge in their approach to defining the efficiencies of both the overall system and its constituent sub-systems. Prior research has shown that network system efficiency can be formulated either through additive aggregation of divisional efficiencies adjusted by a factor, or via a multiplicative aggregation form adjusted by a factor greater than one (Kao 2017, 2016).

Numerous network DEA models have been studied, and this paper delves into two fundamental categories: parallel and hierarchical structures. A parallel system treats each production unit as a sub-system under a whole mechanism, which highlights that each unit acts autonomously. Parallel-based models have been applied to multiple sectors, such as banking, retail sales, railway, city transportation, universities, and energy (Sun et al. 2023; Ding et al. 2023; Bian et al. 2015; Lozano 2015; Chao et al. 2010; Vaz et al. 2010; Yu 2008). Kao (2009) proposed a parallel system model in which the system efficiency is equal to the weighted average of the constituent efficiencies.

1.2 Challenges and advances in hierarchical DEA models

Compared with the previous structure, the hierarchical structure is more complicated to analyse due to its features of multiple levels of functions (Kao 2014). Since almost all real-world entities are operating under a hierarchical structure, some studies have been done using hierarchical-structured DEA models. Yu et al. (2022) utilized a hierarchical DEA model to measure the capital index of global airlines. Hendrawan et al. (2021) proposed a hierarchical DEA model to compare the undergraduate departments of a University in Indonesia. Zhang and Chen (2019) integrated a series two-stage model into a single stage hierarchical network DEA system to assess the high-tech sector in China. Gan et al. (2020) considered the integration of a multi-layer hierarchy into a generalised two-stage series network to evaluate the international shipping industry. Kashim et al. (2018) introduced an improved hierarchical network DEA approach for evaluating the relative performance of a set of 14 university faculties. Kao (2015) discussed the shared-input problem among departments of teaching, research, and service for divisions under a single stage hierarchical system. Through an additive decomposition model, Kao demonstrated that the overall efficiency is equivalent to the weighted arithmetic average of the bottom-level units. A number of studies have been reported in this direction such as Chou et al. (2023), Zhu et al. (2023), and Kashim et al. (2017).

While Kao's (2015) findings illustrate strong discriminatory capabilities in ranking system and division efficiencies, they still exhibit certain limitations within the context of this study. To begin with, the self-evaluation hierarchical network model offers alternative optimal solutions that could potentially result in unrealistic input and output multipliers. Consequently, an inefficient DMU may be erroneously assessed as efficient, leading to potentially misleading insights. Moreover, recent research employing one-stage hierarchical DEA models faces the challenge of inadequately distinguishing inefficiencies at the bottom-level components, particularly when integrating a greater number of sub-functions, as observed in studies by Kashim et al. (2018) and Gan et al. (2020). Lastly, the inherent nature of self-efficiency scores precludes the incorporation of peer opinions, which is a noteworthy limitation.

1.3 Addressing fairness in cross-efficiency approaches

The cross-efficiency evaluation method of Sexton et al. (1986) has been proposed to overcome such shortcomings by enabling each DMU to not only be evaluated based on their own optimal multipliers, but by additionally considering the optimal multipliers of each of the remaining (n-1) DMUs. Nevertheless, given that the evaluation system ignores the relationship among peers, less attention is paid to fairness in the evaluation outcomes. Beullens et al. (2012) defines fairness as a principle corresponding to a system in which the evaluation and ranking results are acceptable for all DMUs. Fairness, as highlighted by Cui et al. (2007), refers to a concept in which a player shows concern not only for their own benefits but also for how profits are attained by other players. Such an understanding is anticipated to motivate

DMUs to accept their measurement results during DEA evaluations. In the pursuit of enhancing fairness, an alternative secondary goal can be incorporated into the cross-efficiency framework, wherein the optimal self-evaluated score for the target DMU is considered. For instance, Zhu et al. (2021) proposed a maxmin secondary goal model aimed at maximising the efficiency of the worst-performing DMUs.

The aggregation of peer-efficiency scores is a critical concern in cross-efficiency applications. Conventional methods, such as the arithmetic mean aggregation strategy, treat opinions under each DMU equally (Wang and Chin 2010; Liang et al. 2008), which overlooks the significance of the assigned weights to scores (Wang and Wang 2013). Additionally, the average cross-efficiency approach might disregard contradictions arising from different principles, including irrational risky attitudes and competitive relationships among them (Zhang et al. 2022; Zuo and Guan 2017).

To cope with these drawbacks, several methods have been introduced in a DEA context to derive the weights of the criteria. The first method of aggregating cross-efficiency scores considers players in a cooperative mode, see studies (Wang et al. 2021; Dong et al. 2020). The second approach, as proposed by Angize et al. (2013), converts the original cross-efficiency matrix into a ranking table and applies a first-order model to obtain aggregation weights. Compared to the two previous methods, the third approach takes advantage of the information value, which is also known as entropy (Rezvani and Khazaei 2013). Sharafi et al. (2020) introduced a fixed-cost allocation method within cross efficiency, ensuring a Pareto cross-efficient outcome. Song and Liu (2018) developed the cross-efficiency DEA model based on Shannon entropy to derive an objective weight distribution among a set of evaluators. Maddahi et al. (2014) suggested a proportional weight assignment approach within the cross-efficiency. Wang and Chin (2011) proposed the use of ordered weighted averaging (OWA) operator weights for cross efficiency aggregation in DEA to consider the decision maker's optimism level in the final overall efficiency assessment. Contreras et al. (2021) introduced a new method for cross-efficiency evaluation in DEA by using bargaining problems and the Kalai-Smorodinsky solution; this enabled the discrimination between optimal weighting profiles for DMUs by agreeing upon input and output multipliers among peer DMUs. Kremantzis et al. (2022a, b) proposed a Criteria Importance Through Inter-criteria Correlation (CRITIC) peer-evaluation approach applied to a generalised two-stage network system, allowing each stage to achieve a higher efficiency score, and encouraging units to accept the evaluation process. Numerous studies have explored this aspect, as exemplified by Borrás et al. (2023), Ning et al. (2023), and Xu et al. (2023).

1.4 Research contributions

Hierarchical network DEA models have proven valuable for assessing efficiency and ranking DMUs within hierarchical structures. More precisely, multi-level decision-making can act as a support tool to identify inefficient units within an organization or entity and recommend resource allocation strategies to enhance overall system efficiency. From an operational research standpoint, this aids in strategic planning, ultimately resulting in a more practical and timely decision-making process. Nonetheless, existing hierarchical DEA models suffer from issues such as unrealistic weights, non-unique efficiency results, and lack of peer opinion that may lead to a less acceptable ranking. Notably, there has been limited research addressing fairness concerns within a single-stage hierarchical structure when using a DEA cross-efficiency approach. This paper contributes to two different aspects. Firstly, we propose a novel approach that combines a maxmin secondary objective with a CRITIC cross-efficiency approach to extend Kao's (2015) additive self-evaluation model. This combination is advantageous as it better captures the internal structures of DMUs, leading to more realistic weightings and enhanced discriminatory power, thereby improving efficiency outcomes. Secondly, our proposed cross-efficiency DEA approach allows the system to highlight minority viewpoints alongside mainstream ones. This inclusion promotes a more acceptable ranking outcome, contributing to fairness improvement within the hierarchical network DEA framework. The proposed model is compared with two of the most prevalent hierarchical DEA approaches (Liu et al. 2022; Kao 2015) within the operational research field, providing insights on the effectiveness of the CRITIC cross-efficiency approach.

The remainder of this paper is organized as follows. Section 2 describes the methodological background while Sect. 3 proposes a new methodological approach to ensure fairer evaluation and ranking results in DMUs under a single stage hierarchical structure. A numerical example of 20 universities will be used to examine the empirical results and discuss important implications in Sect. 4. Section 5 will summarize the key findings, provide managerial implications, and suggest future guidelines.

2 Research background

2.1 A parallel network DEA model

A parallel system treats each component as a well-functional unit that operates independently and similarly (Liu et al. 2022; Kao 2009). Suppose a general parallel network system having n DMUs under evaluation and each has q production units with input$X_{ik}^{(p)}, i \in I^{(p)}$, and outputs$Y_{rk}^{(p)}$, $r \varepsilon O^{(p)}$, denoting that the superscript stands for the pth production unit (p = 1, 2, …, q). Under the precondition of identical production elements, the indicators of inputs and outputs, along with their corresponding weights, are uniform across all production units. Liu et al. (2022) argued that imposing a common set of weights on all production units would cause them to operate under a similar mechanism, potentially disregarding the individual strengths of each unit. Consequently, each production process within the parallel structure should be permitted to have its distinct weight scheme. Nevertheless, adopting a unique weight scheme for each production process could result in less distinguishable efficiency results as the number of variables increases. Additionally, maintaining stable values for the importance of inputs and outputs facilitates the negotiation process among DMUs (Zhou et al.

2022). Drawing upon the aforementioned points, a common set of weights is shared among parallel production processes, as depicted in the subsequent model (1):

$$
\begin{aligned}
&\max = \sum_{p=1}^{q} \sum_{r \in O^{(p)}} u_{rk} Y_{rk}^{(p)}, \\
&\text{s.t.} \sum_{p=1}^{q} \sum_{i \in I^{(p)}}^{\cdots} v_{ik} X_{ik}^{(p)} = 1, \\
&\sum_{r \in O^{(p)}} u_{rk} Y_{rj}^{(p)} - \sum_{i \in I^{(p)}}^{\cdots} v_{ik} X_{ik}^{(p)} \le 0, p = 1, \dots q, j = 1, \dots n, \\
&u_r, v_i \ge \varepsilon, i \in I^{(p)}, r \in O^{(p)},
\end{aligned}
\tag{1}
$$

where v_{ik} and u_k indicate the weights of the ith input and rth output for target $\mathrm{DMU_k}$. From model (1), we obtain the self-evaluation optimal efficiency score for $\mathrm{DMU_k}$ through: (a) $E_k^* = \sum_{p=1}^{q} \sum_{r \in O^{(p)}} u_{rk}^* Y_{rk}^{(p)}$; (b) $E_k^{(p)*} = \sum_{r \in O^{(p)}} u_{rk}^* Y_{rk}^{(p)} / \sum_{i \in I^{(p)}} v_{ik}^* X_{ik}^{(p)}$. Note that (a) is equal to the overall efficiency for the parallel structure and the results from (b) calculate the efficiency of each production unit in the system. One main property of a parallel model is that the overall efficiency can be decomposed into the weighted arithmetic average of the efficiencies of each of the constituent production units (Kao 2009). As already proved by Liu et al. (2022), the efficiency of the parallel network system can be evaluated through: $E_k = \sum_{p=1}^{q} \sum_{r \in O^{(p)}} u_{rk} Y_{rk}^{(p)} = \sum_{p=1}^{q} \omega_k^{(p)} E_k^{(p)}$, where $\omega_k^{(p)} = \frac{\sum_{i \in I^{(p)}} v_{ik} X_{ik}^{(p)}}{\sum_{p=1}^{q} \sum_{i \in I^{(p)}} v_{ik} X_{ik}^{(p)}}$ is the relative importance of an individual production unit p. The following section presents the evaluation of a hierarchical-structured network DEA model, which has been empirically demonstrated by Kao (2015) to be equivalent to a parallel structure in terms of measurement.

2.2 Additive aggregation model for a single-stage hierarchical system

We now consider a single-stage hierarchical structure (Kao 2015), as shown in Fig. 1, to illustrate the relationship between upper-level and lower-level sub-units. The overall system which consumes all the flow-in inputs is labeled 0 at the top in Fig. 1 and is then followed by first-level sub-units, (1) and (2). Next, unit (1) flows outputs to its second-level sub-units (1, 1), (1, 2), (1, 3). Similarly, sub-units (2, 1), (2, 2) receive outputs from (2). Lastly, at the third level, sub-units (2, 2, 1), (2, 2, 2) produce final outputs from unit (2, 2).

Denote the ith input and the rth output used by unit P for target $\mathrm{DMU_k}$ as X_{ij}^{P} and Y_{rj}^{P}, where P is a set of indices and $P \varepsilon Q = \{(1,1),(1,2),(1,3),(2,1),(2,2,1),(2,2,2)\}$. Similarly, to production processes in a parallel structure, $i \varepsilon I^P and r \varepsilon O^P$ denote the inputs and outputs employed at unit P, belonging to the respective set of inputs and outputs. Units at higher levels allocate resources to those at the lower level. For example, unit (2, 2) receives outputs from unit (2) and allocates them to units (2, 2, 1) and (2, 2, 2), respectively. The corresponding linear program, developed by Kao (2015), is provided in below:

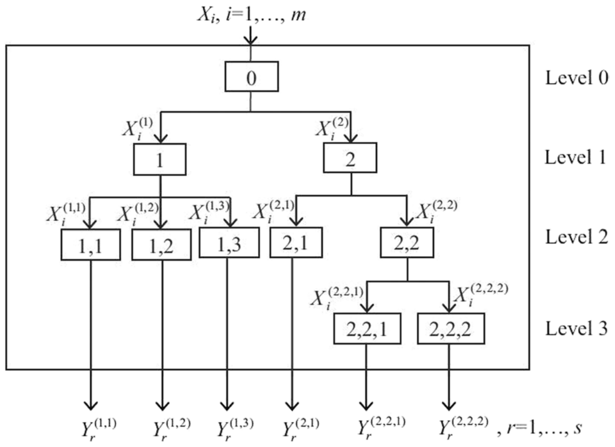

Fig. 1 A single-stage hierarchical system with three levels (Kao 2015)

$$
\begin{aligned}
&\max E_k^{NW} = \sum_{P\in Q}\sum_{r\in O^P} u_r Y_{rk}^P \\
&\text{s.t.} \sum_{P\in Q}\sum_{i\in I^P} v_i X_{ik}^P = 1, \\
&\sum_{r\in O^P} u_r Y_{rj}^P - \sum_{i\in i^P} v_i X_{ij}^P \le 0, j = 1,\dots,n, \\
&Q = \{(1,1),(1,2),(1,3),(2,1),(2,2,1),(2,2,2)\}, \\
&u_r, v_i \ge \varepsilon, i \in I^P, r \in O^P.
\end{aligned}
$$

Note that P considers only the bottom-level units of the integrated hierarchy to avoid redundancy (Kao 2015). These bottom-level units have no lower-level units that take their inputs to produce new outputs, and their production patterns are similar to a parallel structure. Therefore, the formulation for calculating the efficiency of the system and production units within a parallel structure network is also applicable to this situation: $E_k^{P'} = \sum_{r\in O^P} u_{rk} Y_{rk}^P / \sum_{i\in I^P} v_{ik} X_{ik}^P$–the efficiency for the P' sub-unit within this hierarchical structure, equals the weighted sum of total generated outputs divided by the weighted sum of the total consumed inputs in its production process.

To calculate the efficiency of other subordinate units and the whole system, an additive aggregation model combined with an input-proportion strategy is adopted. Based on the input proportion strategy, those units not at the bottom level of the integrated hierarchy can be represented as the sum of weighted efficiency of bottom-level units that receive their outputs, i.e., the efficiency of any non-bottom-level (higher-level) unit $P(2)$, $E_k^{P(2)} = \frac{\sum_{i\in I} p(1) v_{ik} X_{ik}^{P(2)}}{\sum_{P(2)\in B}\sum_{i\in I} P(2) v_{ik} X_{ik}^{(2)}}$, where $P(2)\varepsilon B =$ $P(2)\varepsilon B = \{(2,2),(1),(2),(0)\}$. Another method representing the efficiency of higher-level units within the hierarchical structure measures $E_k^{P(2)} = \sum_{P(1)\in M}\left(\omega_k^{P(1)} / \sum_{P\in M}\omega_k^{P(1)} E_k^{P(1)}\right) = \sum_{P(1)\in M} t^{P(1)} E_k^{P(1)}$, where $P(1)$ ϵ M ={{(1),(2)} ∣ P(2) = (0), {(1,1), (1,2), (1,3)} ∣ P (2) = (1), {(2,1), (2,2)} ∣ P (2) = (2), {(2,2,1), (2,2,2)} ∣ P (2) = (2,2)} and $t^{P(1)}$ satisfies $\sum_{P(1)\in M} t^{P(1)} = 1$. Denoting

$\omega_k^{P(1)} = \frac{\sum_{i \in I^{P(1)}} v_{ik} X_{ik}^{P(1)}}{\sum_{P(1) \in B} \sum_{i \in I^{P(1)}} v_{ik} X_{ik}^{P(1)}}$ is the relative importance of lower-level sub-unit $P(1)$ that receive outputs from upper-level units $P(2)$. For example, the efficiency scores of unit (2, 2, 1) and (2, 2, 2) can be aggregated to the score of higher-level unit (2, 2), which is formulated as: $E_k^{(2,2)} = t \cdot E_k^{(2,2,1)} + (1-t)E_k^{(2,2,2)}$, denoting $t = \omega_k^{(2,2,1)} / (\omega_k^{(2,2,1)} + \omega_k^{(2,2,2)})$, just equaling the ratio of relative importance of P(2,2,1) to the sum of relative importance of all subordinates belonging to P(2,2). A bottom-level or upper-level unit is measured as efficient only when the optimal efficiency equals to 1 ($E_k^P = 1$). The same method is applicable to any other joint unit within the system.

3 Models development

3.1 Maxmin secondary objective model and cross-efficiency

The efficiency scores obtained by model (2) are self-evaluated, in that they do not consider the relationship between DMUs potentially leading to unrealistic input and output weights. Besides, their optimal weighting scheme may be non-unique. To cope with such issues, this paper proposes a combination of an alternative secondary objective model and a CRITIC cross-efficiency approach based on the single-stage hierarchical structure.

Existing methods of the secondary objective model include aggressive, benevolent, neutral and other strategic models (Wu et al. 2021). Under aggressive and benevolent models' framework, the DMU$_\mathrm{k}$ maintains its optimal self-efficiency, while minimizing/maximizing the average efficiency of other DMUs, accordingly. On the contrary, Liu et al. (2022) proposed a neutral secondary model on a parallel network; this does not focus on the relationship of the examined DMU$_\mathrm{k}$ over other DMUs, ensuring the obtained results are not either aggressive or benevolent. Inspired by Liu et al.'s (2022) model, we propose model (3) that corresponds to the single-stage hierarchical structure in Fig. 1:

$$
\begin{aligned}
& \max = \min \left\{ u_r Y_{rk}^P / \sum_{i \in I^P} v_i X_{ik}^P \right\} \\
& \text{s.t.} E_k^{NW*} = \sum_{P \in Q} \frac{\sum_{r \in O^P} u_r Y_{rk}^P}{\sum_{i \in I^P} v_i X_{ik}^P}, E_k^{P*} = \frac{\sum_{r \in O^P} u_r Y_{rk}^P}{\sum_{i \in I^P} v_i X_{ik}^P}, \\
& \sum_{r \in O^P} u_r Y_{rj}^P - \sum_{i \in I^P} v_i X_{ij}^P \le 0, j = 1, \ldots, n, j \ne k, \\
& Q = \{(1,1),(1,2),(1,3),(2,1),(2,2,1),(2,2,2)\}, \\
& u_r, v_i \ge \varepsilon, i \in I^P, r \in O^P.
\end{aligned} \tag{3}
$$

In model (3), $u_r Y_{rk}^P / \sum_{i \in I^P} v_i X_{ik}^P$ is the efficiency of the r th output in unit P of DMU$_k$. E_k^{NW*} and E_k^{P*} are obtained from model (2) to maintain the optimal self-efficiency for DMU$_k$. The objective function of model (3) $u_r Y_{rk}^P / \sum_{i \in I^P} v_i X_{ik}^P$, allows units, within the hierarchical-structured system, to pursue the efficiency that optimizes the lowest boundary. Additionally, the output weights of model (3) are efficient in reducing the values of zero because when max = min $u_r Y_{rk}^P / \sum_{i \in i^P} v_i X_{ik}^P$, $u_r^L (L \neq P)$ must satisfy $u_r^L > u_r^P \geq \varepsilon$.

Our proposed model (3) is non-linear due to its nonlinear objective function and some of its non-linear constraints. By implementing the variable substitution technique (Charnes and Cooper 1962), we build model (4), which calculates the efficiency scores for each of the remaining $(n-1)$ DMUs.

$$
\begin{aligned}
&\max = \min \left\{ u_r Y_{rk}^P \right\} \\
&\text{s.t } E_k^{NW*} = \frac{\sum_{P \in O} \sum_{r \in O^P} u_r Y_{rk}^P}{\sum_{P \in O} \sum_{i \in i^P} v_i X_{ik}^P}, \\
&E_k^{P*} = \sum_{r \in O^P} u_r Y_{rk}^P, \\
&\sum_{i \in I^P} v_i X_{ik}^P = 1, \\
&\sum_{r \in O^P} u_r Y_{rj}^P - \sum_{i \in I^P} v_i X_{ij}^P \leq 0, j = 1, \ldots, n, j \neq k, \\
&E_j^P = \frac{\sum_{r \in O^P} u_r Y_{rj}^P}{\sum_{i \in I^P} v_i X_{ij}^P}, j = 1, \ldots, n, j \neq k, \\
&Q = \{(1,1), (1,2), (1,3), (2,1), (2,2,1), (2,2,2)\}, \\
&u_r, v_i \geq \varepsilon, i \in I^P, r \in O^P.
\end{aligned}
\tag{4}
$$

After totally $2n$ times running of models (2) and (4), the optimal input and output weight sets u_{rk}^*, v_{ik}^* can be obtained. At the optimality of model (4), the cross-efficiency that highlights the peer-evaluation for DMU_j $(j \neq k)$ satisfies $E_{kj}^P = \sum_{r \in O^P} u_r Y_{rj}^P / \sum_{i \in I^P} v_i X_{ij}^P$. To aggregate the cross-efficiency into each DMU_j, we introduce a cross-efficiency aggregation method that is applicable to the efficiency evaluation on a hierarchical structure.

Denote E_{kj}^P as the cross-efficiency score of DMU_j ($j = 1, \ldots$, n) under the weight scheme of (k = 1, …, n). Particularly, model (4) computes the efficiency scores for the remaining n − 1 DMUs by optimising a new objective function. This implies that all DMUs within the criteria system adhere to identical virtual multipliers. It is noticeable that when $j = k$, E_{kj}^P equals the optimal result obtained from model (2) for DMU_k. For convenience, we use p to represent the pth unit in set P. For the pth bottom-level unit of DMU_k, we generate a cross-efficiency matrix containing $n \times n$ elements as illustrated in Table 1.

In the case of a traditional cross-efficiency approach, peer-opinions are equally contributing to the final efficiency for a DMU. Hence, the cross-efficiency result for the pth bottom-level unit of DMU_j can be computed as the following:

Table 1 Cross-efficiency matrix table for the pth bottom-level units of DMU_j

DMU_{kj}	1	2	⋯	n
1	$E_{11}^{(p)}$	$E_{12}^{(p)}$	⋯	$E_{1n}^{(p)}$
2	$E_{21^{(p)}}$	$E_{22}^{(p)}$	⋯	$E_{2n}^{(p)}$
.	.	.	.	.
n	$E_{n1^{(p)}}$	$E_{n2}^{(p)}$	⋯	$E_{nn^{(p)}}$

$$E_j^P = \frac{1}{n}\sum_{k=1}^{n} E_{kj}^P, j = 1, \ldots, n \tag{5}$$

Model (5) computes the cross-efficiency score for the pth bottom-level units of DMUj by aggregating the peer-evaluated opinions from model (4) using an arithmetic average method. In other words, it calculates the cross-efficiency score for a DMU's unit by taking the average of all available alternatives. After obtaining that, the system efficiency can be calculated using the additive aggregation model, in that $E_j^{NW} = \sum_{P=Q} \omega_j^P E_j^P$, where $\omega_j^P = \frac{\sum_{i\in I^P} v_{ij}^* X_{ij}^P}{\sum_{P=Q}\sum_{i\in I^P} v_{ij}^* X_{ij}^P}$ satisfies $\sum_{P=Q} \omega_j^P = 1$, and v_{ij}^* is a self-evaluated optimal input weights matrix, obtained in model (2).

The hierarchical-structured system efficiency in model (2) is optimal but its components' efficiency results may be subjective and non-unique. To further enhance the objectivity and accuracy of efficiency results, a novel combination of the maxmin model (4) and a cross-efficiency technique is proposed, which can additionally measure the efficiency of units at other levels, making the result more informative. The arithmetic average method in aggregating ${E_{kj}}^{\mathrm{P}}$ can lose sight of the weights and does not provide a Pareto optimal solution (Wu et al. 2021; Wang and Wang 2013). To consider the aggregation weights more rationally, this study proposes the CRITIC cross-efficiency approach that is applicable to a single-stage hierarchical-structured system.

3.2 CRITIC cross-efficiency approach

This paper applies the CRITIC method to compute the weights in aggregating peer-evaluated individual cross efficiency scores. CRITIC is deemed as an objective mathematical procedure to derive the level of significance of the criteria involved in the decision-making process. It, additionally, highlights the concept of fairness via the inclusion of the minority opinions, according to Kremantzis et al. (2022a, b); this can ensure a more meritocratic system when measuring and evaluating the performance of DMUs involved in a complex network system. Below, we provide a discussion on how the CRITIC aggregation method could be implemented to a single-stage hierarchical DEA model.

The first step includes the calculation of a set of individual peer-efficiency scores (E_{kj}) that indicate the scores received by jth alternative ($j=1$, …, n), under the kth criterion by DMU_k ($k=1$,…, n) in cross-efficiency evaluation, see in Table 1. The next step refers to the conversion of that matrix into a matrix of relative scores with the element$x_{kj} = \frac{E_k(j)-E_k^{min}}{E_k^{max}-E_k^{min}}, k = 1, \dots, n,$ where $E_k^{max} = max\{E_k(1), E_k(2), \dots, E_k(n)\}$ and $E_k^{min} = min\{E_k(1), E_k(2), \dots, E_k(n)\}$ for criterion k. The normalization method is utilized to eliminate the scale influence on the efficiency table.

Next, the contrast intensity and the conflict, as initially proposed by Diakoulaki et al. (1995), account for the calculation of information entropy. Specifically, the standard deviation representing the contrast intensity, is calculated as the outcome of the normalized relative scores matrix, which is defined as $\sigma = \sqrt{\left[\sum_{j=1}^{n}\left(x_k(j) - \underline{x}_k\right)\right]^2/n}$, where $\underline{x}_k = (\sum_{j=1}^{n} x_k(j))/n$, while conflict $\left(\sum_{j=1}^{n}\left(1 - R_{kj}\right)\right)$ calculates the Spearman ranking correlation coefficients by each criterion, see in Table 2, which contains a symmetric matrix of $n \times n$ elements.

The amount of information, received by criterion k, is denoted by C_k and is determined by the multiplication of the contrast intensity and the conflict:

$$C_k = \sigma_k \sum_{j=1}^{n} \left(1 - R_{kj}\right) \tag{6}$$

The information value of criteria is associated to the proportion of their weights or contributions to the peer-efficiency scores. The higher the information value of C_k, the higher the weight the k criterion will receive. Thereby, the weight of the kth criterion (DMU) in aggregating cross-efficiency scores W_k is formulated as the proportion of its provided CRITIC information value divided by the sum of those of other DMUs, see below:

$$W_k = \frac{C_k}{\sum_{k=1}^{n} C_k} \tag{7}$$

Formula (7) is introduced to determine the weight each criterion should receive for the individual cross-efficiency score of bottom-level subordinate units E_{kj}^{P}, as obtained from model (4). Additionally, it serves as an alternative method for aggregating information compared to the traditional arithmetic average approach, as previously explained in Sect. 3.1. After calculating the weights, the new final cross-efficiency score of p bottom-level unit of DMU j is expressed as:

Table 2 Spearman correlation matrix

DMU_{kj}	1	2	⋯	n
1	R_{11}	R_{12}	⋯	R_{1n}
2	R_{21}	R_{22}	⋯	R_{2n}
n	.	.	.	.

$$e_j^P = \sum_{k=1}^{n} w_k^P E_{kj}^P, j = 1, \ldots, n \tag{8}$$

Note that w_k^P is the weight of the k-th criterion for the P bottom-level subordinate unit. Compared with the arithmetic average cross-efficiency method as shown in formula (5), formula (8) takes into account the relative importance of each criterion, providing a more informative measurement result.

In the case of Fig. 1, CRITIC should run 6 times to calculate $e_j^{(1,1)}, e_j^{(1,2)}, e_j^{(1,3)}, e_j^{(2,1)}, e_j^{(2,2,1)}, e_j^{(2,2,2)}$ for all DMUs. Then, the efficiency of the whole system can be calculated by applying the additive aggregation method as $e_j^{NW} = \sum_{p=Q} \omega_j^{P*} e_j^P$, where $\omega_j^{P*} = \frac{\sum_{i \in I^P} v_{ik}^* X_{ik}^P}{\sum_{P=Q} \sum_{i \in I^P} v_{ik}^* X_{ik}^P}$ is derived from model (2), Q = {(1,1), (1,2), (1,3), (2,1), (2,2,1), (2,2,2)}. Moreover, for better detecting inefficiencies within the whole system, the additive aggregation model for the hierarchical scenario with input proportion strategy, which has been introduced in Sect. 2.2, can be applied to obtain the cross-evaluated efficiency of medium-level subordinate units.

To conclude, traditional arithmetic cross-efficiency aggregation method assigns equal weights to all criteria (opinions), highlighting the majority vote (Kremantzis et al. 2022a, b; Wu et al. 2021). Responding to that issue, the CRITIC method is less influenced from mainstream opinions, while providing adequate representation of the minority opinions (Kremantzis et al. 2022a, b). In Sect. 4, an example of 20 hypothetic universities will be analysed to test the applicability of the suggested cross-efficiency approach.

4 Results and numerical application

This section draws inspiration from Kao's (2015) research, which evaluated the operational effectiveness of several physics departments within 20 different universities. As noted by scholars such as De Witte et al. (2013), Kong and Fu (2012), and Kao and Hung (2008), higher education institutions are commonly recognized for their engagement in three fundamental activities: teaching, research, and enterprise. Moreover, teaching can be subdivided into tasks related to undergraduate and postgraduate education. Seven outputs for each university (Y_1—Graduates-U, Y_2—Credits-U, Y_3—Graduates-G, Y_4—Credits–G, Y_5—Publications, Y_6—Grants, Y_7—Income) have been considered under the single-stage hierarchical network structure with the shared inputs of personnel (X_1) and expenses (X_2) among functions. The network structure of this application is illustrated in Fig. 2. The production units in this three-level hierarchical structure, from top to bottom and left to right, are denoted as U(0), U(1), U(2), U(3), U(1,1), U(1, 2). There are no intermediates, series processes, and feedback in this system; thus, the system efficiency can be decomposed into the efficiency of those production units at the bottom level of the integrated hierarchy.

In addition to evaluating the administration efficiency, another task is associated with the determination of the ratio of inputs distributed to each function. As only

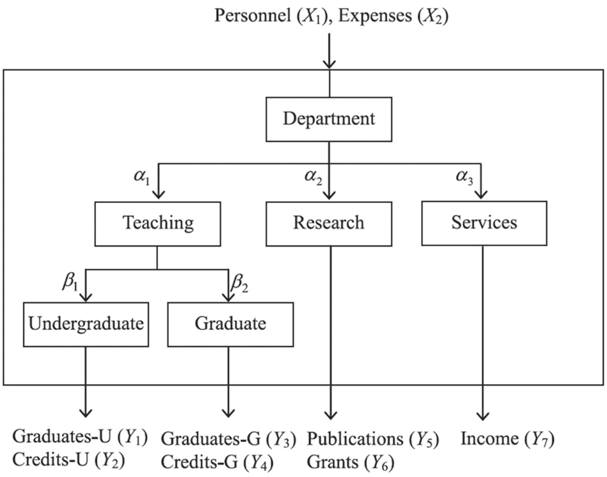

Fig. 2 Single-stage hierarchical structure of the department system (Kao 2015)

the bottom-level production units contribute to the final outcomes Y, $\alpha_1, \alpha_2, \alpha_3$ were used to denote the proportion of inputs for second-level units received from the top, while $\beta_{1,}\beta_2$ are representing the proportions of inputs for Undergraduate and Graduate received from Teaching. Both α and β are freely decided by the decision maker prior to solving the corresponding optimisation model. Following Kao's (2015) spirit on enabling a college department to allocate similar proportions of inputs to its units, we set constraints to control the ratios of α_1/α_2, α_1/α_3, α_2/α_3 lying between [0.5, 2], [1, 4], [1, 4] respectively, while $0.5 \leq \beta_2/\beta_1 \leq 2$. These proportions align with the findings in a comparable study conducted by Kashim et al. (2018). In their research, they employed a hierarchical network framework that featured a broader set of subunits at the second tier. Specifically, they encompassed a PhD component within the Teaching category and two distinct categories, namely, consultation and collaboration activities, within the Service category. Based on the above discussion and the model (4) in Sect. 3.1, we propose the maxmin secondary objective model (9), applied to the structure in Fig. 2, in addition to the self-evaluation model proposed by Kao (2015):

$$
\begin{aligned}
&\max = \min\{u_r Y_{rk}^P\}\text{s.t } E_k^{NW*} = \frac{\sum_{P=O}\sum_{r\in O^P} U_r Y_{rk}^P}{\sum_{P=O}\sum_{i\in i^P} v_i X_{ik}^P}, E_k^{P*} = \sum_{r\in O^P} u_r Y_{rk}^P,\\
&\sum_{i\in I^P} v_i X_{ik}^P = 1, \sum_{r\in O^P} u_r Y_{rj}^P - \sum_{i\in I^P} v_i X_{ij}^P \leq 0, j = 1,\dots,n, j \neq k,\\
&E_j^P = \frac{\sum_{r\in O^P} u_r Y_{rj}^P}{\sum_{i\in I^P} v_i X_{ij}^P}, j = 1,\dots,n, j \neq k, Q = \{(1,1),(1,2),(2),(3)\},\\
&V_i^{(1,1)} = \alpha_1\beta_1 v_i, V_i^{(1,2)} = \alpha_1\beta_2 v_i, V_i^{(2)} = \alpha_2 v_i, V_i^{(3)} = \alpha_3 v_i,\\
&0.5\alpha_2 \leq \alpha_1 \leq 2\alpha_2, \alpha_3 \leq \alpha_1 \leq 4\alpha_3, \alpha_3 \leq \alpha_2 \leq 4\alpha_3, \alpha_1+\alpha_2+\alpha_3 = 1,\\
&0.5\beta_2 \leq \beta_1 \leq 2\beta_2, \beta_1+\beta_2 = 1, u_r, v_i \geq \varepsilon, i\in I^P, r\in O^P.
\end{aligned}
\tag{9}
$$

Note that E_k^{NW*}, E_k^{P*} are the optimal system and subunits' efficiencies calculated from the self-evaluated model (2) proposed by Kao (2015), and V_i stand for the virtual input weights after considering distributed proportions of inputs. The dataset involves 20 university departments and is presented in Table 3. For modelling purposes, we have taken advantage of Python 3.7.6 and the version 2.1 of PuLP as the linear programming library. The experiment ran on a computer with 16 GB RAM.

4.1 Findings and implications

Kao (2015) proposed the self-evaluation model for the single-stage hierarchical-structured network in Fig. 2, that reflects model (2). Table 4 demonstrates the optimal self-efficiency scores of the overall system and its sub-units arranged in a three-level hierarchical structure along with their corresponding rankings across the 20 universities. In particular, columns 4, 6, 10, and 12 present the self-evaluated efficiency scores of all sub-units with their corresponding weights of relative importance, by which they are aggregated into the system performance in column 2 (${E_k}^{\mathrm{NW}}$) as discussed in Sect. 2.2. By the same token, the sub-units of Undergraduate-Teaching (${E_k}^{\mathrm{T-U}}$) and Graduate-Teaching (${E_k}^{\mathrm{T-G}}$) are aggregated to obtain the efficiency score of Teaching department (${E_k}^{\mathrm{T}}$) in column 8. Kao's (2015) model serves as a baseline scenario for comparison with the proposed maxmin secondary objective model (9).

By applying the proposed maxmin secondary objective model (9), some desirable outcomes regarding weight improvement have been obtained. As discussed in Sect. 3.1, model (9) rewards the worst-performing targets for improving unrealistic multipliers (those with a zero value). In total, the optimal multipliers, obtained via solving model (9), generate only 26 zero weights (see Table 5, where the columns 2 & 3 are referring to input weights while the remaining ones the output weights), compared with that of 38 zero weights obtained via solving Kao's (2015) self-evaluation model (see Appendix Table 8). Hence, after applying model (9), there is a decrease in the total number of zero weights, indicating that those less important factors are less likely to be ignored in the evaluation. The optimal multipliers obtained from Table 5 are then used to compute the individual cross-efficiency scores of the general cross-efficiency matrices for Undergraduate-Teaching, Graduate-Teaching, Research, and Service terminal sub-units for all universities.

After totally 40 (20 * 2) times (see Sect. 3.1) running of the self-evaluation hierarchical network model (2) and the proposed maxmin model (9), both optimal self-efficiency and cross-efficiency scores of the overall system and its bottom-level sub-units were obtained, see the optimal multipliers result in Table 5. To aggregate the optimal individual cross-efficiency scores, the traditional arithmetic average cross-efficiency method of Liu et al. (2022) was compared with our proposed CRITIC aggregation cross-efficiency model, see their results in Tables 6 and 7 respectively, which are similarly formatted to those in Table 4. Both aggregation methods adopt the additive aggregation model to obtain the overall network system performance score e_j^{NW} and the middle-level sub-unit (Teaching) score e_j^T (see Sect. 3.2).

Table 3 The numerical application of Kao (2015)

Univ	Personnel (X1)	Expenses (X2)	U-stud. (Y1)	U-hours (Y2)	G-stud. (Y3)	G-hours (Y4)	Publications (Y5)	Grants (Y6)	Income (Y7)
1	35	60	55	25	25	15	70	4	2
2	40	85	60	35	35	20	15	2	1
3	25	50	50	10	25	10	15	1	2
4	30	65	40	20	25	20	25	5	1
5	45	95	50	25	35	20	65	75	3
6	30	100	55	20	65	20	50	35	10
7	15	60	35	5	20	20	10	30	25
8	20	70	50	10	30	15	15	10	10
9	55	140	140	20	315	95	90	80	30
10	60	180	170	20	250	180	160	145	45
11	15	70	50	5	80	30	50	25	30
12	45	110	140	15	100	65	120	80	40
13	20	70	55	5	45	25	15	20	20
14	25	90	60	5	100	50	95	100	25
15	35	110	90	10	120	65	30	30	30
16	30	85	50	5	75	30	20	25	185
17	25	120	65	10	105	55	30	40	15
18	20	65	45	5	45	20	30	15	20
19	60	90	50	5	115	60	70	100	55
20	20	65	50	5	45	20	35	25	85

Table 4 Efficiency and relative importance of system and each sub-unit according to self-evaluation hierarchical model (Kao 2015)

Univ.	E_k^{NW}	Rank	$E_k^{\mathrm{T-U}}$	Rank	$E_k^{\mathrm{T-G}}$	Rank	E_k^{T}	Rank	E_k^{R}	Rank	E_k^{S}	Rank
1	0.790	6	1 (0.667)	1	0.267 (0.333)	19	0.756 (0.286)	8	1 (0.571)	1	0.015 (0.143)	17
2	0.474	19	1 (0.667)	1	0.253 (0.333)	20	0.751 (0.571)	10	0.154 (0.286)	20	0.005 (0.143) (0.142)	20
3	0.479	18	0.867 (0.667)	8	0.380 (0.333)	16	0.705 (0.571)	14	0.257(0.286)	15	0.018 (0.143)	16
4	0.470	20	0.760 (0.667)	15	0.4278 (0.333)	14	0.649 (0.571)	16	0.343(0.286)	14	0.007 (0.143)	19
5	0.558	13	0.632 (0.667)	18	0.3053 (0.333)	17	0.523 (0.286)	19	0.711 (0.571)	7	0.015 (0.143)	18
6	0.564	12	0.917 (0.667)	6	0.4273 (0.333)	15	0.754 (0.571)	9	0.439 (0.286)	11	0.054 (0.143)	15
7	0.600	10	0.776 (0.667)	13	0.6444 (0.333)	8	0.732 (0.571)	11	0.500 (0.286)	9	0.270 (0.143)	5
8	0.486	17	0.945 (0.667)	5	0.3032 (0.333)	18	0.731 (0.571)	12	0.198 (0.286)	19	0.080 (0.143)	14
9	0.722	7	0.817 (0.333)	11	1 (0.667)	1	0.939 (0.571)	2	0.600 (0.286)	8	0.099 (0.143)	12
10	0.791	5	0.772 (0.333)	14	1 (0.667)	1	0.924 (0.444)	3	0.826 (0.444)	6	0.116 (0.111)	11
11	0.869	1	1 (0.667)	1	0.992 (0.333)	3	0.998 (0.444)	1	0.877 (0.444)	5	0.324 (0.111)	3
12	0.848	2	1 (0.667)	1	0.613 (0.333)	9	0.871 (0.444)	5	0.995 (0.444)	4	0.167 (0.111)	6
13	0.510	15	0.855 (0.667)	9	0.473 (0.333)	11	0.727 (0.571)	13	0.251 (0.286)	16	0.159 (0.143)	9
14	0.820	3	0.720 (0.333)	16	0.823 (0.667)	5	0.788 (0.286)	6	1 (0.571)	1	0.162 (0.143)	7
15	0.535	14	0.820 (0.667)	10	0.722 (0.333)	7	0.787 (0.571)	7	0.228 (0.286)	17	0.138 (0.143)	10
16	0.575	11	0.529 (0.667)	19	0.473 (0.333)	10	0.510 (0.333)	20	0.213 (0.333)	18	1 (0.333)	1
17	0.633	8	0.886 (0.667)	7	0.879 (0.333)	4	0.884 (0.571)	4	0.400 (0.286)	12	0.097 (0.143)	13
18	0.493	16	0.716 (0.667)	17	0.438 (0.333)	12	0.623 (0.571)	18	0.398 (0.286)	13	0.161 (0.143)	8
19	0.797	4	0.437 (0.333)	20	0.753 (0.667)	6	0.648 (0.286)	17	1 (0.571)	1	0.281 (0.143)	4
20	0.620	9	0.783 (0.667)	12	0.436 (0.333)	13	0.668 (0.500)	15	0.466 (0.250)	10	0.680 (0.250)	2

Table 5 Optimal multipliers result of the proposed maxmin model (9)

Univ.	v_1	v_2	u_1	u_2	u_3	u_4	u_5	u_6	u_7
1	**0**	6.316	0.538	1.703	0.148	0.396	3.075	0.327	0.415
2	0.223	2.005	**0**	1.952	0.099	0.258	0.521	0.037	0.137
3	**0**	1.324	0.349	0.443	**0**	0.252	0.322	0.034	0.087
4	0.564	1.481	0.081	1.477	0.000	0.318	0.443	0.000	0.110
5	**0**	3.362	**0**	1.537	**0**	0.320	0.000	1.729	0.221
6	0.025	**0**	1.386	4.691	0.394	0.483	1.221	0.000	0.376
7	6.730	0.000	0.575	1.945	0.000	0.427	0.000	0.481	0.156
8	14.068	0.367	1.396	4.071	0.359	0.465	1.157	0.000	0.350
9	**0**	2.813	0.371	0.471	0.264	0.703	0.685	0.073	0.185
10	0.949	2.492	0.287	0.452	0.196	0.560	1.161	0.000	0.144
11	18.680	**0**	1.453	2.072	0.353	0.433	2.185	**0**	0.337
12	**0**	3.384	0.693	0.881	0.123	0.330	1.281	0.136	0.173
13	9.123	0.318	1.212	0.000	0.236	0.312	0.000	0.734	0.232
14	11.946	**0**	0.341	**0**	0.290	0.356	1.796	**0**	0.277
15	4.070	0.106	0.404	1.177	0.104	0.134	0.335	**0**	0.101
16	2.429	0.085	0.188	**0**	0.037	0.048	**0**	0.228	0.144
17	11.159	**0**	0.953	3.225	0.271	0.332	**0**	0.797	0.259
18	6.343	0.165	0.630	1.835	0.161	0.210	0.522	**0**	0.158
19	**0**	4.780	0.358	**0**	0.224	0.600	**0**	2.458	0.314
20	9.783	0.341	1.138	**0**	0.222	0.292	0.724	**0**	0.436

Recalling that in our proposed approach, the weights to aggregate the individual cross-efficiencies into a final cross-efficiency score for each DMU and sub-unit are derived by formula (7) as part of the CRITIC method, see Sect. 3.2. To achieve that, we firstly obtain the matrix of relative scores for each of the four sub-units (Undergraduate-Teaching, Graduate-Teaching, Research, Service), in order to identify the corresponding standard deviations. A symmetric matrix for each of those sub-units is depicted in the Appendix section (Tables 9, 10, 11 and 12). In these matrices, the last two rows indicate the information value and the final weight of each criterion-DMU; these two correspond to formulae (6) & (7), respectively.

The weights derived via formula (7) are used to compute the final cross-efficiency score of each of the four bottom-level sub-unitse_j^{T-U},e_j^{T-G},e_j^R, e_j^S for the 20 universities (see Table 7). For detecting the inefficiency under a hierarchical system, the proposed additive aggregation model is used to aggregate the two Teaching sub-units (e_j^{T-U},e_j^{T-G}) into the Teaching (e_j^T). The latter is also combined with Research (e_j^R) and Services (e_j^S) sub-units to obtain the final cross-efficiency score for the overall system (e_j^{NW}) (see Table 7).

Compared with the results of the self-evaluation model (see Table 4), the proposed CRITIC cross-efficiency model demonstrates a higher level of discriminatory

Table 6 System and sub-unit efficiencies for the arithmetic average cross-efficiency model (Liu et al. 2022)

Univ.	e_j^{NW}	Rank	e_j^{T-U}	Rank	e_j^{T-G}	Rank	e_j^{T}	Rank	e_j^{R}	Rank	e_j^{S}	Rank
1	0.498	8	0.852	2	0.203	20	0.636	8	0.551	5	0.012	18
2	0.390	17	0.843	3	0.215	19	0.634	9	0.094	20	0.005	20
3	0.374	20	0.709	8	0.305	16	0.574	11	0.151	19	0.015	16
4	0.389	19	0.681	9	0.328	15	0.563	14	0.231	15	0.006	19
5	0.429	13	0.574	16	0.244	18	0.464	18	0.515	7	0.013	17
6	0.449	12	0.667	10	0.377	14	0.571	12	0.406	10	0.050	15
7	0.418	14	0.579	15	0.460	9	0.539	16	0.267	13	0.233	4
8	0.389	18	0.712	7	0.272	17	0.565	13	0.176	18	0.110	12
9	0.648	4	0.741	5	0.930	2	0.867	2	0.486	8	0.093	13
10	0.745	2	0.741	4	1	1	0.914	1	0.732	3	0.119	11
11	0.645	5	0.732	6	0.750	3	0.738	4	0.646	4	0.263	3
12	0.689	3	0.862	1	0.552	7	0.758	3	0.753	2	0.155	6
13	0.413	15	0.649	12	0.432	11	0.577	10	0.217	16	0.148	8
14	0.781	1	0.554	18	0.716	4	0.662	6	0.999	1	0.146	9
15	0.465	11	0.654	11	0.679	6	0.662	5	0.237	14	0.133	10
16	0.548	6	0.431	19	0.441	10	0.434	20	0.209	17	1	1
17	0.465	10	0.621	13	0.689	5	0.644	7	0.302	12	0.078	14
18	0.411	16	0.566	17	0.395	12	0.509	17	0.344	11	0.153	7
19	0.473	9	0.281	20	0.539	8	0.453	19	0.550	6	0.207	5
20	0.541	7	0.616	14	0.395	12	0.542	15	0.433	9	0.648	2

power in the evaluation process, that is a unique ranking order of the involved DMUs is attained. For instance, universities 1, 2, 11, and 12 for Undergraduate-Teaching sub-unit are evaluated as relatively efficient with respect to the self-evaluation model and cannot be further distinguished, while both universities 9 & 10 are efficient for Graduate-Teaching sub-unit (see Table 4). At this point, bias triggered by the self-evaluation model is observed, regarding the efficiency score for university 11; this is due to being ranked first (1) and third (0.992) for its Undergraduate and Graduate sub-units respectively, with the aggregated efficiency of its Teaching department ranked first (0.998). This self-evaluated result can be prejudiced when lacking appropriate consideration of peer opinions; this could make the measurement result unfair and in turn less acceptable to DMUs involved in the process. When taking into account the peer-evaluation process, the final cross-efficiency score, for instance, of the Undergraduate-Teaching sub-unit of university 11 is 0.571, ranked in the 7th place (see Table 7). This appears to be more rational considering that its information value was relatively low (1.614) in comparison to the corresponding values of the other evaluators, see Table 9 in the Appendix.

The results of the arithmetic average cross-efficiency model, which coincides with the neutral cross-efficiency approach proposed by Liu et al. (2022) (see Table 6), are compared to the proposed CRITIC aggregation method (see

Table 7 System and sub-unit efficiencies for the proposed CRITIC cross-efficiency model

Univ.	e_j^{NW}	Rank	e_j^{T-U}	Rank	e_j^{T-G}	Rank	e_j^{T}	Rank	e_j^{R}	Rank	e_j^{S}	Rank
1	0.422	9	0.759	2	0.013	20	0.511	7	0.489	7	0.007	18
2	0.275	17	0.716	3	0.026	19	0.486	8	0.001	20	0.000	20
3	0.263	19	0.577	6	0.141	16	0.431	10	0.061	19	0.011	16
4	0.266	18	0.504	9	0.172	15	0.394	12	0.151	15	0.002	19
5	0.353	11	0.365	17	0.064	18	0.264	19	0.490	6	0.008	17
6	0.319	13	0.464	12	0.220	14	0.383	14	0.339	10	0.046	15
7	0.294	15	0.367	15	0.327	9	0.354	16	0.216	13	0.227	4
8	0.256	20	0.540	8	0.090	17	0.390	13	0.091	18	0.069	14
9	0.593	4	0.612	4	0.903	2	0.806	2	0.442	8	0.089	12
10	0.706	2	0.606	5	1	1	0.869	1	0.707	3	0.114	11
11	0.544	5	0.571	7	0.666	3	0.603	4	0.570	4	0.256	3
12	0.633	3	0.788	1	0.446	7	0.674	3	0.728	2	0.151	6
13	0.296	14	0.479	11	0.288	11	0.415	11	0.144	16	0.143	8
14	0.738	1	0.350	18	0.636	4	0.540	5	0.999	1	0.141	9
15	0.360	10	0.486	10	0.596	5	0.523	6	0.163	14	0.128	10
16	0.453	6	0.190	19	0.301	10	0.227	20	0.137	17	1.000	1
17	0.344	12	0.411	14	0.596	6	0.473	9	0.233	12	0.073	13
18	0.280	16	0.365	16	0.242	12	0.324	17	0.267	11	0.148	7
19	0.424	8	0.002	20	0.440	8	0.294	18	0.550	5	0.207	5
20	0.436	7	0.435	13	0.242	12	0.371	15	0.370	9	0.644	2

Table 7). Note that either of the results have been determined via the optimal multipliers of Table 5. Based on Spearman rank correlation coefficient test, it can be concluded that the rankings of the overall systems obtained from the CRITIC (Table 7) and the arithmetic average method (Table 6) are not significantly different, with $s_r = 0.9864$ at the 0.01 significance level. In addition, the efficiency scores of the three models—self-evaluation model (Kao 2015), neutral cross-efficiency model (Liu et al. 2022), and the proposed CRITIC cross-efficiency model—are compared through the line charts in Fig. 3.

It can be clearly seen that the self-evaluation scores are higher than those of the other two models since individual efficiency scores are overestimated excluding the peer opinions. In this scenario, decision-makers confront the challenge of discerning inefficient components. For example, when examining the Research sub-unit of DMU19 in Table 4, it initially shows a self-assessed score of 1 (indicating strong efficiency). However, this score decreases to 0.55 in Table 6 and 7 when incorporating peer opinions. Additionally, the traditional arithmetic method used to aggregate the cross-efficiency matrix lacks comprehensive validation, as pointed out by Soltanifar and Sharafi (2021), who highlight the absence of mechanisms like a voting process. This absence hinders us from obtaining an accurate reflection of DMU's performance. For instance, the Undergraduate sub-unit of

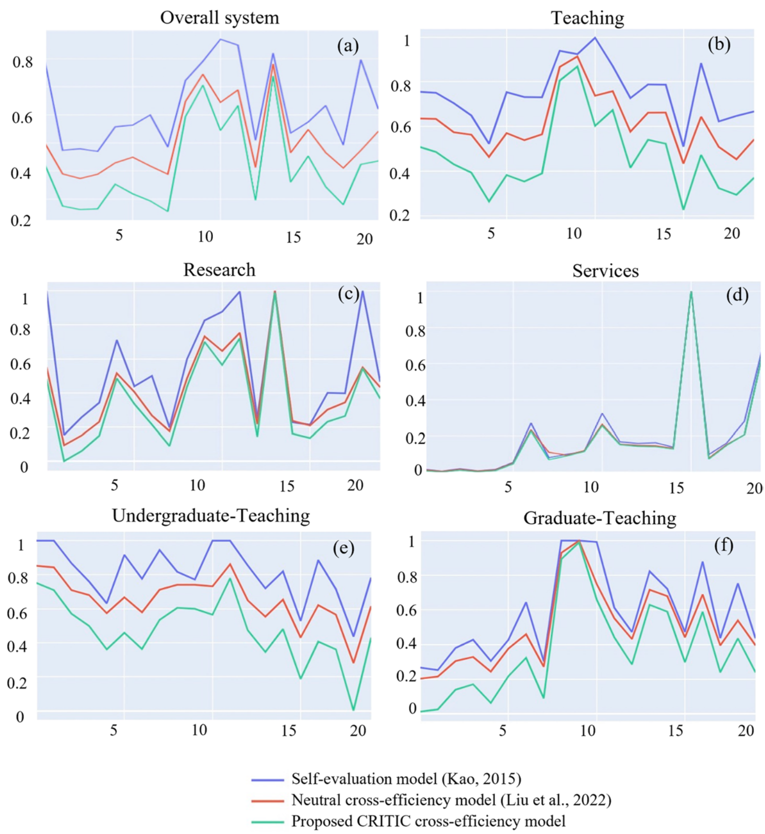

Fig. 3 Comparison of the efficiency scores of 20 Universities for the three models for the **a** overall system, **b** Teaching, **c** Research, and **d** Services units and Teaching sub-units **e** Undergraduate and **f** Graduate

DMU5 has an average cross-efficiency score of 0.574 in Table 6. Nevertheless, a noticeable minority of DMUs 3, 9, 13, 14, 16, 19, and 20 indicates lower scores of 0.117, 0.111, and 0.117 (see Appendix 9), resulting in a CRITIC cross-efficiency of 0.365 in Table 7. Notably, although not universally lower, the Service sub-unit of DMU19 shows an increased score (0.20712) in Table 7 compared to Table 6 (0.2065). Consequently, this empirical case illustrates a prevailing trend of reduced efficiency when compared to both self-evaluation and the arithmetic average cross-efficiency approaches.

The advantageous points of our proposed CRITIC peer-appraisal approach over the arithmetic average cross-efficiency model are presented below. Firstly, the proposed methodology emphasizes a meticulously documented measurement outcome by providing incentives for less-mainstream viewpoints, thereby resulting in a more extensive efficiency range compared to the one obtained through simple arithmetic average, as depicted in Fig. 3 (c,e,f). Secondly, the CRITIC model seamlessly aligns with the configuration of the single-stage hierarchical network, effectively meeting the requirement for accommodating minority peer opinions, while promoting the values of diversity and inclusion (Kremantzis et al. 2022a, b). Moreover, in scenarios where minority opinions are lacking, such as within the Service sub-unit, the final cross-efficiency scores obtained through the CRITIC approach closely approximate the corresponding results of the conventional average method for each unit, see Fig. 3d. Lastly, the enhancement of weight distribution was achieved through the reduction of zero weights, which significantly contributed to fostering greater equity in the outcomes for the universities under investigation. The optimal multipliers derived from the proposed maxmin model (Table 5) have, in addition, been amalgamated with the CRITIC method to reach a higher level of discriminatory power and, in turn, a more meaningful ranking. It is important to note that the final rankings derived from our proposed CRITIC cross-efficiency model exhibit congruence with those obtained from Kao's (2015) and Liu et al.'s (2022) models, as demonstrated in Fig. 3.

5 Conclusions and future research

Hierarchical network structures have garnered significant acceptance among scholars due to their ability to accurately depict internal procedures and systematically measure and evaluate performance. However, these structures have not received considerable attention in the pursuit of fairer results that take into account the perspectives of peer evaluators. In this study, we extend the idea of the single-stage hierarchical network self-evaluation DEA model proposed by Kao (2015) by incorporating a cross-efficiency context; this has been achieved by introducing an original combination of a maxmin secondary objective model and the CRITIC multi-criteria decision-making method that aggregates opposing viewpoints.

Under the prism of this hierarchical system, our proposed maxmin secondary model addresses the issue of non-uniqueness in optimal multipliers obtained from the self-evaluated model (2). By doing so, it effectively diminishes the prevalence of unrealistic weights assigned to input/output factors and enhances the ability to distinguish efficiently the performing DMUs. This study further explains the applicability of the CRITIC cross-efficiency method when applied to a traditional hierarchical-patterned network system, leading to improved fairness, in comparison to the corresponding results obtained from Liu et al.'s (2022) arithmetic average cross-efficiency method. Since our proposed approach has no preference over being benevolent or aggressive, it can be considered as neutral. According to our study's result, the proposed approach is more likely to provide a comprehensive and unique ranking order, due to its nature of supporting diversified and inclusive opinions.

5.1 Managerial implications

In addition to its current application in assessing the relative performance of universities, our proposed cross-efficiency hierarchical network DEA model holds the potential for wide implementation across various domains and sectors. For instance, the resource allocation for different types of service lines for a city transportation enterprise could be explored by the proposed model. Similar to the hierarchical structure of university departments, a city transportation company can have internal services including coach and bus which consume similar inputs (expenditures, seats) to produce dedicated outcomes (miles, passengers, revenue). The same case of transportation system can also take environmental and social production elements into consideration for sustainability assessment by evaluating undesirable environmental outputs (carbon emission), and social outputs (employment, security) in addition to economic elements. It is also possible to facilitate the strategic management in terms of considering the internal mechanisms of the entire system by identifying those inefficient units which require additional support. Finally, our suggested model could be implemented for the evaluation of the sustainable development of airlines, supply chains, and transportation frameworks, with a view to ensuring more acceptable results for the stakeholders involved.

5.2 Future pathways and limitations

The modeling approach presented in this study is under the Constant Returns to Scale (CRS) assumption. This assumption may not hold true in real-world scenarios where processes may incur increasing or decreasing returns to scale; in particular, it has been identified that organisations experience economies of scale, wherein larger-scale hierarchical operations lead to more efficient resource utilisation and lower average costs. This can be extended to a Variable Returns to Scale (VRS) situation as a future pathway, expanding their applicability to other disciplines as well. For instance, VRS applications would be considered advantageous in the agricultural sector, where doubling the required inputs does not necessarily lead to double output. Secondly, the dataset involved in this study only makes use of real positive numbers which may be a restricting factor for cases where the data may be insufficient or include negative values (e.g., the assessment of enterprises with a negative annual net profit) or contain intervals. To this end, fuzzy numbers (Zimmermann 2011) could be used when the probability distribution is unknown. Another future pathway could be the integration of such a peer appraisal setting within more complicated hierarchical network forms, such as the hierarchy with two-stage processes (Zhang and Chen 2019) or the multi-level hierarchy embedded into a general two-stage series structure (Gan et al. 2020) or the multi-function parallel network hierarchical system (Kremantzis et al. 2022a, b).

Appendix

See Tables 8, 9, 10, 11 and 12.

Table 8 Efficiency score and multipliers result of a three-level hierarchical structure, the model by Kao (2015)

Univ	E_k	v_1	v_2	u_1	u_2	u_3	u_4	u_5	u_6	u_7
1	0.790	0	0.017	0.002	0.003	0.000	0.001	0.008	0	0.001
2	0.474	0.001	0.011	0	0.011	0.001	0.001	0.003	0.000	0.001
3	0.479	0	0.020	0.005	0.007	0	0.004	0.005	0.001	0.001
4	0.470	0.005	0.013	0.001	0.013	0	0.003	0.004	0	0.001
5	0.558	0	0.011	0	0.005	0	0.001	0	0.005	0.001
6	0.564	0.033	0	0.003	0.010	0.001	0.001	0.003	0	0.001
7	0.600	0.067	0	0.006	0.019	0	0.004	0	0.005	0.002
8	0.486	0.046	0.001	0.005	0.013	0.001	0.002	0.004	0	0.001
9	0.722	0	0.007	0.001	0.001	0.001	0.000	0.002	0.000	0.001
10	0.791	0.002	0.005	0.001	0.001	0.000	0.001	0.002	0	0.000
11	0.869	0.067	0	0.006	0.000	0.001	0.002	0.008	0	0.001
12	0.848	0	0.009	0.002	0	0.000	0.001	0.003	0.000	0.001
13	0.510	0.045	0.001	0.006	0	0.001	0.002	0	0.004	0.001
14	0.820	0.040	0	0.001	0	0.001	0.001	0.001	0.004	0.001
15	0.535	0.026	0.001	0.003	0.008	0.001	0.001	0.002	0	0.001
16	0.575	0.030	0.001	0.002	0	0.001	0.001	0	0.003	0.002
17	0.633	0.040	0	0.003	0.012	0.001	0.001	0	0.003	0.001
18	0.493	0.046	0.001	0.005	0.013	0.001	0.002	0.004	0	0.001
19	0.797	0	0.011	0.001	0	0.001	0.001	0	0.006	0.001
20	0.620	0.045	0.002	0.005	0	0.001	0.001	0.003	0	0.002

Table 9 Transposed matrix of relative score and weight result for Undergraduate-Teaching sub-unit

DUMs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1.000	1.000	1.000	1.000	1.000	0.859	0.859	0.878	1.000	1.000	0.849	1.000	0.333	0.295	0.878	0.333	0.859	0.878	0.523	0.333
2	0.976	1.000	0.735	1.000	0.986	1.000	1.000	1.000	0.735	0.821	0.987	0.735	0.291	0.267	1.000	0.291	1.000	1.000	0.241	0.291
3	0.416	0.406	0.766	0.464	0.400	0.669	0.669	0.704	0.766	0.764	0.665	0.766	0.519	0.467	0.704	0.519	0.669	0.704	0.635	0.519
4	0.688	0.710	0.464	0.708	0.698	0.739	0.739	0.737	0.464	0.546	0.730	0.464	0.216	0.200	0.737	0.216	0.739	0.737	0.119	0.216
5	0.562	0.584	0.286	0.571	0.575	0.561	0.561	0.554	0.286	0.367	0.553	0.286	0.117	0.111	0.554	0.117	0.561	0.554	0.000	0.117
6	0.392	0.418	0.217	0.440	0.400	0.889	0.889	0.870	0.217	0.340	0.880	0.217	0.410	0.400	0.870	0.410	0.889	0.870	0.032	0.410
7	0.077	0.086	0.077	0.117	0.077	0.701	0.701	0.698	0.077	0.180	0.699	0.077	0.605	0.600	0.698	0.605	0.701	0.698	0.076	0.605
8	0.246	0.255	0.329	0.305	0.242	0.920	0.920	0.925	0.329	0.432	0.915	0.329	0.692	0.667	0.925	0.692	0.920	0.925	0.252	0.692
9	0.261	0.251	0.677	0.329	0.242	0.795	0.795	0.837	0.677	0.714	0.793	0.677	0.745	0.685	0.837	0.745	0.795	0.837	0.635	0.745
10	0.171	0.163	0.560	0.242	0.154	0.851	0.851	0.890	0.560	0.621	0.850	0.560	0.855	0.800	0.890	0.855	0.851	0.890	0.560	0.855
11	0.051	0.053	0.218	0.107	0.044	1.000	1.000	1.000	0.218	0.325	1.000	0.218	1.000	1.000	1.000	1.000	1.000	1.000	0.252	1.000
12	0.257	0.232	1.000	0.350	0.224	0.934	0.934	1.000	1.000	1.000	0.933	1.000	1.000	0.911	1.000	1.000	0.934	1.000	1.000	1.000
13	0.055	0.052	0.305	0.110	0.044	0.741	0.741	0.768	0.305	0.384	0.742	0.305	0.800	0.767	0.768	0.800	0.741	0.768	0.348	0.800
14	0.006	0.006	0.136	0.047	0.000	0.586	0.586	0.604	0.136	0.219	0.587	0.136	0.646	0.627	0.604	0.646	0.586	0.604	0.188	0.646
15	0.110	0.106	0.375	0.167	0.098	0.724	0.724	0.753	0.375	0.445	0.724	0.375	0.735	0.695	0.753	0.735	0.724	0.753	0.391	0.734
16	0.011	0.014	0.045	0.037	0.009	0.333	0.333	0.346	0.045	0.113	0.333	0.045	0.349	0.333	0.346	0.349	0.333	0.346	0.083	0.349
17	0.075	0.088	0.026	0.115	0.077	0.848	0.848	0.825	0.026	0.143	0.846	0.026	0.692	0.707	0.825	0.692	0.848	0.825	0.021	0.692
18	0.065	0.067	0.200	0.109	0.059	0.592	0.592	0.610	0.200	0.278	0.592	0.200	0.592	0.567	0.610	0.592	0.592	0.610	0.222	0.592
19	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.039	0.000
20	0.069	0.067	0.294	0.121	0.059	0.667	0.667	0.693	0.294	0.367	0.667	0.294	0.701	0.667	0.693	0.701	0.667	0.693	0.325	0.701
SD	0.302	0.307	0.301	0.293	0.307	0.231	0.231	0.234	0.301	0.279	0.229	0.301	0.272	0.263	0.234	0.272	0.231	0.234	0.256	0.272
Conflict	10.904	10.925	7.619	10.544	11.006	6.928	6.928	6.388	7.619	7.472	7.043	7.619	12.369	12.725	6.287	12.433	6.928	6.217	9.845	12.369
Information	3.293	3.357	2.295	3.089	3.376	1.597	1.597	1.492	2.295	2.088	1.614	2.295	3.366	3.340	1.468	3.383	1.597	1.452	2.516	3.365
Weight	0.067	0.069	0.047	0.063	0.069	0.033	0.033	0.031	0.047	0.043	0.033	0.047	0.069	0.068	0.030	0.069	0.033	0.030	0.051	0.069

Table 10 Transposed matrix of relative score and weight result for Graduate-Teaching sub-unit

DUMs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0.027	0.016	0.045	0.016	0.045	0.000	0.000	0.000	0.027	0.002	0.000	0.027	0.000	0.000	0.000	0.000	0.000	0.000	0.027	0.000
2	0.013	0.006	0.027	0.009	0.027	0.038	0.028	0.036	0.013	0.000	0.038	0.013	0.036	0.038	0.036	0.036	0.038	0.036	0.013	0.036
3	0.168	0.157	0.211	0.181	0.211	0.105	0.129	0.107	0.168	0.142	0.105	0.168	0.108	0.105	0.107	0.108	0.105	0.107	0.168	0.108
4	0.188	0.178	0.295	0.268	0.295	0.118	0.209	0.121	0.188	0.163	0.118	0.188	0.122	0.118	0.121	0.122	0.118	0.121	0.188	0.122
5	0.060	0.051	0.116	0.094	0.116	0.050	0.084	0.050	0.060	0.042	0.050	0.060	0.050	0.050	0.050	0.050	0.050	0.050	0.060	0.050
6	0.138	0.140	0.096	0.097	0.096	0.320	0.209	0.309	0.138	0.149	0.320	0.138	0.305	0.320	0.309	0.305	0.320	0.309	0.138	0.305
7	0.204	0.205	0.342	0.357	0.342	0.373	0.585	0.361	0.204	0.213	0.373	0.204	0.358	0.373	0.361	0.358	0.373	0.361	0.204	0.358
8	0.000	0.000	0.000	0.000	0.000	0.179	0.125	0.168	0.000	0.009	0.179	0.000	0.165	0.179	0.168	0.165	0.179	0.168	0.000	0.165
9	1.000	1.000	0.591	0.572	0.591	1.000	0.505	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
11	0.420	0.432	0.273	0.292	0.273	0.991	0.611	0.935	0.420	0.464	0.991	0.420	0.919	0.991	0.935	0.919	0.991	0.935	0.420	0.919
12	0.486	0.479	0.479	0.458	0.479	0.417	0.395	0.421	0.486	0.468	0.417	0.486	0.422	0.417	0.421	0.422	0.417	0.421	0.486	0.422
13	0.195	0.196	0.182	0.186	0.182	0.385	0.319	0.373	0.195	0.207	0.385	0.195	0.369	0.385	0.373	0.369	0.385	0.373	0.195	0.369
14	0.521	0.527	0.434	0.445	0.434	0.789	0.611	0.769	0.521	0.544	0.789	0.521	0.763	0.789	0.769	0.763	0.789	0.769	0.521	0.763
15	0.545	0.547	0.479	0.481	0.479	0.676	0.556	0.668	0.545	0.556	0.676	0.545	0.665	0.676	0.668	0.665	0.676	0.668	0.545	0.665
16	0.269	0.269	0.176	0.170	0.176	0.376	0.222	0.371	0.269	0.274	0.376	0.269	0.369	0.376	0.371	0.369	0.376	0.371	0.269	0.369
17	0.359	0.368	0.311	0.333	0.311	0.857	0.689	0.808	0.359	0.395	0.857	0.359	0.793	0.857	0.808	0.793	0.857	0.808	0.359	0.793
18	0.168	0.169	0.119	0.118	0.119	0.339	0.222	0.328	0.168	0.178	0.339	0.168	0.325	0.339	0.328	0.325	0.339	0.328	0.168	0.325
19	0.672	0.639	0.576	0.488	0.576	0.288	0.222	0.300	0.672	0.578	0.288	0.672	0.304	0.288	0.300	0.304	0.288	0.300	0.672	0.304
20	0.168	0.169	0.119	0.118	0.119	0.339	0.222	0.328	0.168	0.178	0.339	0.168	0.325	0.339	0.328	0.325	0.339	0.328	0.168	0.325
Variance	0.290	0.290	0.239	0.237	0.239	0.327	0.255	0.319	0.290	0.289	0.327	0.290	0.317	0.327	0.319	0.317	0.327	0.319	0.290	0.317
Conflict	2.258	2.097	3.475	2.946	3.475	1.998	2.414	2.011	2.244	1.895	1.998	2.244	1.996	1.998	2.011	1.996	1.998	1.998	2.244	1.996
Information	0.654	0.608	0.830	0.700	0.830	0.654	0.615	0.641	0.650	0.548	0.654	0.650	0.632	0.654	0.641	0.632	0.654	0.637	0.650	0.632
Weight	0.050	0.046	0.063	0.053	0.063	0.050	0.047	0.049	0.049	0.042	0.050	0.049	0.048	0.050	0.049	0.048	0.050	0.048	0.049	0.048

Table 11 Transposed matrix of relative score and weight result for Research sub-unit

DUMs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1.000	1.000	1.000	1.000	0.043	0.474	0.019	0.500	1.000	1.000	0.474	1.000	0.020	0.472	0.500	0.020	0.019	0.500	0.043	0.508
2	0.000	0.000	0.000	0.000	0.003	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.003	0.003	0.000	0.003	0.000
3	0.124	0.125	0.124	0.127	0.000	0.066	0.000	0.069	0.124	0.127	0.066	0.124	0.000	0.065	0.069	0.000	0.000	0.069	0.000	0.070
4	0.214	0.217	0.215	0.220	0.052	0.134	0.032	0.139	0.215	0.220	0.134	0.215	0.033	0.133	0.139	0.033	0.411	0.139	0.052	0.141
5	0.578	0.573	0.592	0.534	0.705	0.312	0.411	0.325	0.592	0.534	0.312	0.592	0.431	0.313	0.325	0.431	0.411	0.325	0.705	0.329
6	0.354	0.364	0.360	0.372	0.302	0.377	0.285	0.378	0.360	0.372	0.377	0.360	0.287	0.377	0.378	0.287	0.285	0.378	0.302	0.378
7	0.032	0.029	0.041	0.003	0.440	0.085	0.495	0.080	0.041	0.003	0.085	0.041	0.488	0.087	0.080	0.488	0.495	0.080	0.440	0.078
8	0.048	0.052	0.051	0.054	0.113	0.109	0.116	0.106	0.051	0.054	0.109	0.051	0.116	0.110	0.106	0.116	0.116	0.106	0.113	0.105
9	0.517	0.521	0.527	0.509	0.505	0.368	0.357	0.378	0.527	0.509	0.368	0.527	0.369	0.368	0.378	0.369	0.357	0.378	0.505	0.381
10	0.785	0.798	0.800	0.794	0.720	0.669	0.600	0.679	0.800	0.794	0.669	0.800	0.612	0.669	0.679	0.612	0.600	0.679	0.720	0.682
11	0.570	0.594	0.576	0.635	0.309	0.864	0.411	0.839	0.576	0.635	0.864	0.576	0.397	0.862	0.839	0.397	0.411	0.839	0.309	0.832
12	0.982	0.992	0.994	0.987	0.648	0.669	0.439	0.690	0.994	0.987	0.669	0.994	0.455	0.668	0.690	0.455	0.439	0.690	0.648	0.696
13	0.061	0.062	0.066	0.054	0.244	0.109	0.242	0.106	0.066	0.054	0.109	0.066	0.243	0.110	0.106	0.243	0.242	0.106	0.244	0.105
14	0.980	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
15	0.119	0.121	0.123	0.116	0.232	0.141	0.206	0.140	0.123	0.116	0.141	0.123	0.209	0.141	0.140	0.209	0.206	0.140	0.232	0.140
16	0.083	0.083	0.088	0.072	0.251	0.085	0.200	0.085	0.088	0.072	0.085	0.088	0.205	0.086	0.085	0.205	0.200	0.085	0.251	0.085
17	0.101	0.105	0.107	0.102	0.287	0.241	0.394	0.228	0.107	0.102	0.241	0.107	0.379	0.242	0.228	0.379	0.394	0.228	0.287	0.224
18	0.305	0.315	0.309	0.327	0.193	0.328	0.179	0.329	0.309	0.327	0.328	0.309	0.181	0.328	0.329	0.181	0.179	0.329	0.193	0.330
19	0.700	0.675	0.721	0.585	1.000	0.231	0.411	0.246	0.721	0.585	0.231	0.721	0.440	0.232	0.246	0.440	0.411	0.246	1.000	0.251
20	0.396	0.406	0.403	0.413	0.334	0.401	0.306	0.403	0.403	0.413	0.401	0.403	0.309	0.401	0.403	0.309	0.306	0.403	0.334	0.404
Variance	0.337	0.339	0.341	0.338	0.298	0.272	0.233	0.273	0.341	0.338	0.272	0.341	0.234	0.272	0.273	0.234	0.225	0.273	0.298	0.273
Conflict	4.150	3.911	3.980	3.774	6.437	3.588	6.916	3.432	3.980	3.757	3.606	3.980	6.999	3.576	3.432	6.999	7.275	3.449	6.437	3.454
Information	1.398	1.327	1.356	1.276	1.916	0.976	1.611	0.936	1.356	1.270	0.981	1.356	1.641	0.972	0.936	1.641	1.639	0.940	1.916	0.943
Weight	0.053	0.050	0.051	0.048	0.073	0.037	0.061	0.035	0.051	0.048	0.037	0.051	0.062	0.037	0.035	0.062	0.062	0.036	0.073	0.036

Table 12 Transposed matrix of relative score and weight result for Service sub-unit

DUMs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0.010	0.010	0.010	0.009	0.010	0.005	0.005	0.005	0.010	0.009	0.005	0.010	0.005	0.005	0.005	0.005	0.005	0.005	0.010	0.005
2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.013	0.013	0.013	0.012	0.013	0.009	0.009	0.009	0.013	0.012	0.009	0.013	0.009	0.009	0.009	0.009	0.009	0.009	0.013	0.009
4	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001	0.002	0.002	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001
5	0.009	0.009	0.009	0.009	0.009	0.007	0.007	0.007	0.009	0.009	0.007	0.009	0.007	0.007	0.007	0.007	0.007	0.007	0.009	0.007
6	0.041	0.041	0.041	0.042	0.041	0.050	0.050	0.049	0.041	0.042	0.050	0.041	0.049	0.050	0.049	0.049	0.050	0.049	0.041	0.049
7	0.187	0.190	0.187	0.194	0.187	0.267	0.267	0.260	0.187	0.194	0.267	0.187	0.258	0.267	0.260	0.258	0.267	0.260	0.187	0.258
8	0.061	0.061	0.061	0.062	0.061	0.077	0.077	0.076	0.061	0.062	0.077	0.061	0.076	0.077	0.797	0.076	0.077	0.076	0.061	0.076
9	0.094	0.093	0.094	0.092	0.094	0.085	0.085	0.085	0.094	0.092	0.085	0.094	0.085	0.085	0.085	0.085	0.085	0.085	0.094	0.085
10	0.110	0.110	0.110	0.111	0.110	0.118	0.118	0.117	0.110	0.111	0.118	0.110	0.117	0.118	0.117	0.117	0.118	0.117	0.110	0.117
11	0.193	0.197	0.193	0.202	0.193	0.322	0.322	0.308	0.193	0.202	0.322	0.193	0.304	0.322	0.308	0.304	0.322	0.308	0.193	0.304
12	0.163	0.161	0.163	0.160	0.163	0.141	0.141	0.142	0.163	0.160	0.141	0.163	0.142	0.141	0.142	0.142	0.141	0.142	0.163	0.142
13	0.127	0.128	0.127	0.130	0.127	0.159	0.159	0.156	0.127	0.130	0.159	0.127	0.155	0.159	0.156	0.155	0.159	0.156	0.127	0.155
14	0.123	0.125	0.123	0.126	0.123	0.159	0.159	0.156	0.123	0.126	0.159	0.123	0.155	0.159	0.156	0.155	0.159	0.156	0.123	0.155
15	0.121	0.121	0.121	0.122	0.121	0.135	0.135	0.134	0.121	0.122	0.135	0.121	0.134	0.135	0.134	0.134	0.135	0.134	0.121	0.134
16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
17	0.052	0.054	0.052	0.055	0.052	0.094	0.094	0.089	0.052	0.055	0.094	0.052	0.088	0.094	0.089	0.088	0.094	0.089	0.052	0.088
18	0.137	0.138	0.137	0.139	0.137	0.159	0.159	0.157	0.137	0.139	0.159	0.137	0.157	0.159	0.157	0.157	0.159	0.157	0.137	0.157
19	0.277	0.264	0.277	0.250	0.277	0.145	0.145	0.150	0.277	0.250	0.145	0.277	0.152	0.145	0.150	0.152	0.145	0.150	0.277	0.152
20	0.599	0.603	0.599	0.608	0.599	0.688	0.688	0.681	0.599	0.608	0.688	0.599	0.679	0.688	0.681	0.679	0.688	0.681	0.599	0.679
Variance	0.232	0.232	0.232	0.232	0.232	0.241	0.241	0.240	0.232	0.232	0.241	0.232	0.240	0.241	0.240	0.240	0.241	0.240	0.232	0.240
Conflict	0.470	0.470	0.470	0.470	0.470	0.403	0.403	0.385	0.470	0.470	0.389	0.470	0.385	0.403	0.385	0.385	0.403	0.385	0.470	0.385
Information	0.109	0.109	0.109	0.109	0.109	0.097	0.097	0.093	0.109	0.109	0.094	0.109	0.092	0.097	0.093	0.092	0.097	0.093	0.109	0.092
Weight	0.054	0.054	0.054	0.054	0.054	0.048	0.048	0.046	0.054	0.054	0.047	0.054	0.046	0.048	0.046	0.046	0.048	0.046	0.054	0.046

Declarations

Conflicts of interest The authors declare no conflict of interest.

References

Avkiran NK (2009) Opening the black box of efficiency analysis: An illustration with UAE banks. Omega 37(4):930–941. https://doi.org/10.1016/j.omega.2008.08.001

Angiz MZ, Mustafa A, Kamali MJ (2013) Cross-ranking of Decision Making Units in Data Envelopment Analysis. Appl Math Model 37(1–2):398–405. https://doi.org/10.1016/j.apm.2012.02.038

Azadi M, Yousefi S, Saen RF, Shabanpour H, Jabeen F (2023) Forecasting sustainability of healthcare supply chains using deep learning and network data envelopment analysis. J Bus Res 154:113357

Banker RD, Charnes A, Cooper WW (1984) Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. Manage Sci 30(9):1078–1092. https://doi.org/10.1287/mnsc.30.9.1078

Beullens P, Zaibidi NZ, Jones DF (2012) Goal programming to model human decision making in ultimatum games. Int Trans Oper Res 19(4):599–612. https://doi.org/10.1111/j.1475-3995.2011.00826.x

Bian Y, Hu M, Xu H (2015) Measuring efficiencies of parallel systems with shared inputs/outputs using data envelopment analysis. Kybernetes 44(3):336–352. https://doi.org/10.1108/K-04-2014-0067

Borrás F, Ruiz JL, Sirvent I (2023) Peer evaluation through cross-efficiency based on reference sets. Omega 114:102739

Charnes A, Cooper WW (1962) Programming with linear fractional functionals. Naval Research Logistics Quarterly 9(3–4):181–186. https://doi.org/10.1002/nav.3800090303

Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2(6):429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Chao C-M, Yu M-M, Chen M-C (2010) Measuring the performance of financial holding companies. Serv Ind J 30(6):811–829. https://doi.org/10.1080/02642060701849857

Chou HC, Lu WM, Kweh QL, Tsai CY (2023) Using hierarchical network data envelopment analysis to explore the performance of national research and development organizations. Expert Syst Appl 234:121109

Chu J, Zhu J (2021) Production scale-based two-stage network data envelopment analysis. Eur J Oper Res 294(1):283–294

Contreras I, Lozano S, Hinojosa MA (2021) A DEA cross-efficiency approach based on bargaining theory. Journal of the Operational Research Society 72(5):1156–1167. https://doi.org/10.1080/01605682.2020.1755898

Cui TH, Raju JS, Zhang ZJ (2007) Fairness and Channel Coordination. Manage Sci 53(8):1303–1314. https://doi.org/10.1287/mnsc.1060.0697

De Witte K, Rogge N, Cherchye L, Van Puyenbroeck T (2013) Economies of scope in research and teaching: A non-parametric investigation. Omega 41(2):305–314. https://doi.org/10.1016/j.omega.2012.04.002

Diakoulaki D, Mavrotas G, Papayannakis L (1995) Determining objective weights in multiple criteria problems: The critic method. Comput Oper Res 22(7):763–770. https://doi.org/10.1016/0305-0548(94)00059-H

Ding T, Zhang Y, Zhang D, Li F (2023) Performance evaluation of Chinese research universities: A parallel interactive network DEA approach with shared and fixed sum inputs. Socioecon Plann Sci 87:101582

Dong F, Zhang Y, Zhang X (2020) Applying a data envelopment analysis game cross-efficiency model to examining regional ecological efficiency: Evidence from China. J Clean Prod 267:122031. https://doi.org/10.1016/j.jclepro.2020.122031
Färe R, Grosskopf S (2000) Network DEA. Socioecon Plann Sci 34(1):35–49. https://doi.org/10.1016/S0038-0121(99)00012-9
Gan G, Lee H-S, Lee L, Wang X, Wang Q (2020) Network hierarchical DEA with an application to international shipping industry in Taiwan. Journal of the Operational Research Society 71(6):991–1002. https://doi.org/10.1080/01605682.2019.1603792
Ghasemi, N., Najafi, E., Lotfi, F. H., & Sobhani, F. M. (2020). Assessing the performance of organizations with the hierarchical structure using data envelopment analysis: An efficiency analysis of Farhangian University. Measurement: Journal of the International Measurement Confederation, 156. https://doi.org/10.1016/j.measurement.2020.107609
Goto M, Sueyoshi T (2023) Sustainable development and convergence under energy sector transition in industrial nations: An application of DEA environmental assessment. Socioecon Plann Sci 87:101316. https://doi.org/10.1016/j.seps.2022.101316
Hendrawan BM, Purwanggono B, Suliantoro H (2021) Efficiency measurements of Diponegoro University faculty using the Data Envelopment Analysis hierarchical network model. IOP Conference Series: Materials Science and Engineering 1072(1):012021. https://doi.org/10.1088/1757-899X/1072/1/012021
Kao C (2009) Efficiency measurement for parallel production systems. Eur J Oper Res 196(3):1107–1112. https://doi.org/10.1016/j.ejor.2008.04.020
Kao C (2014) Network data envelopment analysis: A review. Eur J Oper Res 239(1):1–16. https://doi.org/10.1016/j.ejor.2014.02.039
Kao C (2015) Efficiency measurement for hierarchical network systems. Omega 51:121–127. https://doi.org/10.1016/j.omega.2014.09.008
Kao C (2016) Efficiency decomposition and aggregation in network data envelopment analysis. Eur J Oper Res 255(3):778–786. https://doi.org/10.1016/j.ejor.2016.05.019
Kao, C. (2017). Network data envelopment analysis: Foundations and extensions. Springer International Publishing.
Kao C, Hung H (2008) Efficiency analysis of university departments: An empirical study☆. Omega 36(4):653–664. https://doi.org/10.1016/j.omega.2006.02.003
Kong W-H, Fu T-T (2012) Assessing the performance of business colleges in Taiwan using data envelopment analysis and student based value-added performance indicators. Omega 40(5):541–549. https://doi.org/10.1016/j.omega.2011.10.004
Kashim R, Kasim MM, Rahman RA (2017) Measuring efficiency of a university faculty using an extended hierarchical network dea model: A framework. Adv Sci Lett 23(9):9090–9093
Kashim R, Kasim MM, Rosshairy AR (2018) Measuring efficiency of a university faculty using a hierarchical network data envelopment analysis model. Journal of Information and Communication Technology 17(4):569–585
Kounetas K, Androulakis G, Kaisari M, Manousakis G (2023) Educational reforms and secondary school's efficiency performance in Greece: A bootstrap DEA and multilevel approach. Oper Res Int Journal 23(1):9. https://doi.org/10.1007/s12351-023-00764-y
Kremantzis MD, Beullens P, Klein J (2022a) A fairer assessment of DMUs in a generalised two-stage DEA structure. Expert Syst Appl 187:115921. https://doi.org/10.1016/j.eswa.2021.115921
Kremantzis MD, Beullens P, Kyrgiakos LS, Klein J (2022b) Measurement and evaluation of multi-function parallel network hierarchical DEA systems. Socioecon Plann Sci 84:101428. https://doi.org/10.1016/j.seps.2022.101428
Kyrgiakos LS, Kleftodimos G, Vlontzos G, Pardalos PM (2023) A systematic literature review of data envelopment analysis implementation in agriculture under the prism of sustainability. Oper Res Int Journal 23(1):7. https://doi.org/10.1007/s12351-023-00741-5
Lewis HF, Sexton TR (2004) Network DEA: efficiency analysis of organizations with complex internal structure. Comput Oper Res 31(9):1365–1410. https://doi.org/10.1016/S0305-0548(03)00095-9
Liu Y, Wang K (2015) Energy efficiency of China's industry sector: An adjusted network DEA (data envelopment analysis)-based decomposition analysis. Energy 93:1328–1337. https://doi.org/10.1016/j.energy.2015.10.072
Liang L, Wu J, Cook WD, Zhu J (2008) Alternative secondary goals in DEA cross-efficiency evaluation. Int J Prod Econ 113(2):1025–1030. https://doi.org/10.1016/j.ijpe.2007.12.006

Liu P, Zhang Y, Xu H (2022) A neutral cross-efficiency measurement for general parallel production system. Expert Syst Appl 205:117778. https://doi.org/10.1016/j.eswa.2022.117778
Lozano S (2015) A joint-inputs Network DEA approach to production and pollution-generating technologies. Expert Syst Appl 42(21):7960–7968. https://doi.org/10.1016/j.eswa.2015.06.023
Lozano S, Khezri S (2021) Network DEA smallest improvement approach. Omega 98:102140
Lu WM, Kweh QL, Yang KC (2022) Multiplicative efficiency aggregation to evaluate Taiwanese local auditing institutions performance. Ann Oper Res 315(2):1243–1262
Maddahi R, Jahanshahloo GR, Hosseinzadeh Lotfi F, Ebrahimnejad A (2014) Optimizing proportional weights as a secondary goal in DEA cross-efficiency evaluation. International Journal of Operational Research 19(2):234–245
Miškić, S., Stević, Ž., & Marinković, D. (2022). Evaluating the efficiency of a transport company applying an objective-subjective model. International Journal of Management Science and Engineering Management, 1–15. https://doi.org/10.1080/17509653.2022.2101153
Ning Y, Zhang Y, Wang G (2023) An improved DEA prospect cross-efficiency evaluation method and its application in fund performance analysis. Mathematics 11(3):585
Omrani H, Shamsi M, Emrouznejad A, Teplova T (2023) A robust DEA model under discrete scenarios for assessing bank branches. Expert Syst Appl 219:119694. https://doi.org/10.1016/j.eswa.2023.119694
Rezvani, M., & Khazaei, M. (2013). INTERDISCIPLINARY JOURNAL OF CONTEMPORARY RESEARCH IN BUSINESS Prioritization of entrepreneurial marketing dimensions: A case of in higher education Institutions by using entropy.
Ratner SV, Shaposhnikov AM, Lychev AV (2023) Network DEA and Its Applications (2017–2022): A Systematic Literature Review. Mathematics 11(9):2141. https://doi.org/10.3390/math11092141
Sinuany-Stern Z (2023) Foundations of operations research: From linear programming to data envelopment analysis. Eur J Oper Res 306(3):1069–1080. https://doi.org/10.1016/j.ejor.2022.10.046
Sexton TR, Silkman RH, Hogan AJ (1986) Data envelopment analysis: Critique and extensions. New Directions for Program Evaluation 1986(32):73–105. https://doi.org/10.1002/ev.1441
Sharafi H, Lotfi FH, Jahanshahloo GR, Razipour-GhalehJough S (2020) Fair allocation fixed cost using cross-efficiency based on Pareto concept. Asia-Pacific J Oper Res 37(01):1950036
Saen RF, Karimi B, Fathi A (2022) Assessing the sustainability of transport supply chains by double frontier network data envelopment analysis. J Clean Prod 354:131771. https://doi.org/10.1016/j.jclepro.2022.131771
Soltanifar M, Sharafi H (2021) A modified DEA cross efficiency method with negative data and its application in supplier selection. J Comb Optim 43(1):265–296. https://doi.org/10.1007/s10878-021-00765-7
Song L, Liu F (2018) An improvement in DEA cross-efficiency aggregation based on the Shannon entropy. Int Trans Oper Res 25(2):705–714. https://doi.org/10.1111/itor.12361
Sun Y, Wang D, Yang F, Ang S (2023) Efficiency evaluation of higher education systems in China: A double frontier parallel DEA model. Comput Ind Eng 176:108979
Vaz CB, Camanho AS, Guimarães RC (2010) The assessment of retailing efficiency using Network Data Envelopment Analysis. Ann Oper Res 173(1):5–24. https://doi.org/10.1007/s10479-008-0397-z
Vaez-Ghasemi M, Moghaddas Z, Saen RF (2022) Cost efficiency evaluation in sustainable supply chains with marginal surcharge values for harmful environmental factors: A case study in a food industry. Oper Res Int Journal 22(5):5897–5912. https://doi.org/10.1007/s12351-021-00641-6
Wang Y-M, Chin K-S (2010) Some alternative models for DEA cross-efficiency evaluation. Int J Prod Econ 128(1):332–338. https://doi.org/10.1016/j.ijpe.2010.07.032
Wang Y-M, Chin K-S (2011) The use of OWA operator weights for cross-efficiency aggregation. Omega 39(5):493–503. https://doi.org/10.1016/j.omega.2010.10.007
Wang YM, Wang S (2013) Approaches to determining the relative importance weights for cross-efficiency aggregation in data envelopment analysis. Journal of the Operational Research Society 64(1):60–69. https://doi.org/10.1057/jors.2012.43
Wong WP, Wong KY (2007) Supply chain performance measurement system using DEA modeling. Ind Manag Data Syst 107(3):361–381. https://doi.org/10.1108/02635570710734271
Wang M, Huang Y, Li D (2021) Assessing the performance of industrial water resource utilization systems in China based on a two-stage DEA approach with game cross efficiency. J Clean Prod 312:127722. https://doi.org/10.1016/j.jclepro.2021.127722
Wu J, Sun J, Liang L (2021) Methods and applications of DEA cross-efficiency: Review and future perspectives. Frontiers of Engineering Management 8(2):199–211. https://doi.org/10.1007/s42524-020-0133-1

Xu, H., Liu, P., & Xu, K. (2023). Cross-efficiency evaluation method for non-homogeneous parallel network systems with imprecise data. Expert Systems with Applications, 120557.
Yu M-M (2008) Measuring the efficiency and return to scale status of multi-mode bus transit—evidence from Taiwan's bus system. Appl Econ Lett 15(8):647–653. https://doi.org/10.1080/13504850600721858
Yu M-M, Rashid AA, See KF (2022) Developing an innovation capital index of global airlines using a hierarchical data envelopment analysis approach. J Oper Res Soc 73(8):1708–1723. https://doi.org/10.1080/01605682.2021.1923378
Zhang L, Chen K (2019) Hierarchical network systems: an application to high-technology industry in China. Omega 82:118–131. https://doi.org/10.1016/j.omega.2017.12.007
Zhang R, Wei Q, Li A, Ren L (2022) Measuring efficiency and technology inequality of China's electricity generation and transmission system: a new approach of network data envelopment analysis prospect cross-efficiency models. Energy 246:123274. https://doi.org/10.1016/j.energy.2022.123274
Zhang Y, An Y, Wang Y (2023) Evaluation of green industry innovation efficiency based on three-stage DEA model: a case study of Chinese information technology industry. Sustainability 15(8):6368. https://doi.org/10.3390/su15086368
Zhong K, Li C, Wang Q (2021) Evaluation of bank innovation efficiency with data envelopment analysis: from the perspective of uncovering the black box between input and output. Mathematics 9(24):3318
Zhou Z, Ma Y, Pan Y, Zhu Y (2022) Data envelopment analysis models based on decentralized decision making. Comput Ind Eng 170:108318. https://doi.org/10.1016/j.cie.2022.108318
Zhu Q, Li F, Wu J, Sun J (2021) Cross-efficiency evaluation in data envelopment analysis based on the perspective of fairness utility. Comput Ind Eng 151:106926. https://doi.org/10.1016/j.cie.2020.106926
Zhu W, Lin H, Yu Y, Ghiyasi M (2023) Performance measurement of Taiwan hotels with a hierarchical network DEA with shared inputs. Int J Appl Decis Sci 16(4):474–507
Zimmermann HJ (2011). Fuzzy set theory—and its applications. Springer Netherlands. https://doi.org/10.1007/978-94-010-0646-0
Zuo K, Guan J (2017) Measuring the R&D efficiency of regions by a parallel DEA game model. Scientometrics 112(1):175–194. https://doi.org/10.1007/s11192-017-2380-4

Authors and Affiliations

Siwei Xiao[1] · Marios Dominikos Kremantzis[1] · Leonidas Sotirios Kyrgiakos[2] · George Vlontzos[2] · Panos M. Pardalos[3]

✉ Marios Dominikos Kremantzis
marios.kremantzis@bristol.ac.uk

Siwei Xiao
xiaosiwei1006@gmail.com

Leonidas Sotirios Kyrgiakos
lkyrgiakos@uth.gr

George Vlontzos
gvlontzos@agr.uth.gr

Panos M. Pardalos
pardalos@ise.ufl.edu

[1] University of Bristol Business School, University of Bristol, Bristol, UK

[2] Department of Agriculture, Crop Production and Rural Environment, University of Thessaly, 38446 Fytoko, Volos, Greece

[3] Department of Industrial and Systems Engineering, University of Florida, 401 Weil Hall, Gainesville, FL 32611-6595, USA