RAIRO-Oper. Res. 56 (2022) 1293–1319
https://doi.org/10.1051/ro/2022056

RAIRO Operations Research
www.rairo-ro.org

A RANKING FRAMEWORK BASED ON INTERVAL SELF AND CROSS-EFFICIENCIES IN A TWO-STAGE DEA SYSTEM

Marios Dominikos Kremantzis[1,*], Patrick Beullens[2] and Jonathan Klein[3]

Abstract. The evaluation of the performance of a decision-making unit (DMU) can be measured by its own optimistic and pessimistic multipliers, leading to an interval self-efficiency score. While this concept has been thoroughly studied with regard to single-stage systems, there is still a gap when it is extended to two-stage tandem structures, which better correspond to a real-world scenario. In this paper, we argue that in this context, a meaningful ranking of the DMUs is obtained; this outcome simultaneously considers the optimistic and pessimistic viewpoints within the self-appraisal context, and the most favourable and unfavourable weight sets of each of the other DMUs in a peer-appraisal setting. We initially extend the optimistic-pessimistic Data Envelopment Analysis (DEA) models to the specifications of such a two-stage structure. The two opposing self-efficiency measures are merged to a combined self-efficiency measure *via* the geometric average. Under this framework, the DMUs are further evaluated in a peer setting *via* the interval cross-efficiency (CE). This methodological tool is applied to evaluate the target DMU in relation to the most favourable and unfavourable weight profiles of each of the other DMUs, while maintaining the combined self-efficiency measure. We, thus, determine an interval individual CE score for each DMU and flow. By treating the interval CE matrix as a multi-criteria decision making problem and by utilising several well-established approaches from the literature, we delineate its remaining elements; we show how these lead us to a meaningful ultimate ranking of the DMUs. A numerical example about the efficiency evaluation of ten bank branches in China illustrates the applicability of our modelling approaches.

Mathematics Subject Classification. 90B10, 90C05, 90C90, 91B06.

Received July 1, 2021. Accepted April 20, 2022.

1. Introduction

Data Envelopment Analysis (DEA) is a benchmarking technique for comparing the relative efficiency of a Decision Making Unit (DMU) with the best observed efficiency [7]. The evaluation of a DMU is based on the comparison between the amount of input(s) consumed and the amount of output(s) produced [8] by DMUs.

One of the undeniably attractive features of DEA is its weight flexibility. This allows each DMU to be allocated its most favourable set of weights to be assigned to inputs and outputs for determining its relative efficiency.

Keywords. Data envelopment analysis, network, interval self-efficiency, interval cross-efficiency, ranking.

[1] School of Management, University of Bristol, Bristol, UK.

[2] Mathematical Sciences, Southampton Business School and CORMSIS, University of Southampton, Southampton, UK.

[3] Southampton Business School and CORMSIS, University of Southampton, Southampton, UK.

*Corresponding author: marios.kremantzis@bristol.ac.uk

Hence, in the conventional DEA, the overall assessment of a DMU is based on the optimistic viewpoint and on the concept of the efficiency frontier [56]. On the other hand, if the performance of a DMU is based on the pessimistic viewpoint, then an inefficiency frontier is defined.

Optimistic and pessimistic perspectives illustrate two extreme cases for each DMU. Taking only one scenario into account limits the examination of the performance of a unit. The obtained results might be unreasonable [3]. Therefore, it is thought to be valuable to consider the two distinctive efficiencies together.

Research on exploring both aspects of viewing the efficiency of a DMU within a single-stage structure is relatively extensive. Wang and Luo [48] evaluated each DMU in terms of the optimistic and pessimistic viewpoint, by introducing an input-oriented virtual Ideal DMU (IDMU) and an output-oriented virtual Anti-ideal DMU (ADMU). The two separate efficiencies were combined into the Relative Closeness (RC) index to obtain a unique ranking order. Wu [51] identified a weakness in Wang and Luo's [48] paper dealing with the ADMU for DEA modelling. Wu argued that it is inconsistent to aggregate an input-oriented IDMU and an output-oriented ADMU into the RC index.

Wang and Yang [49] proposed an alternative way of measuring the performance of DMUs. The efficiencies of DMUs are measured within the range of an interval, in which the upper bound is 1 and the lower bound equals to the performance of a virtual ADMU, which is the worst among all DMUs. This approach, which only considers the performance of the lower bound, was extended by Azizi and Jahed [4], who suggested a pair of improved bounded models for the target DMU. Wang *et al.* [50] combined optimistic and pessimistic efficiencies into a geometric average efficiency to measure the overall performance of a DMU. The geometric average efficiency was deemed effective, as it was simultaneously an efficiency measure and a ranking index. Toloo and Tichy [45] proposed a multiplier model to identify the maximum efficiency scores and applied the envelopment model to attain the maximum discrimination among efficient DMUs . Khodabakhshi and Aryavash [23] used a double frontier DEA procedure to introduce a new cross-efficiency method; the merit of their approach lied on the non-use of any alternative secondary goal. Based on the ideal and anti-ideal DMUs, Liu and Wang [29] developed the normalised efficiency metric and then formulated two DEA models to obtain its lower and upper bounds. Örkcü *et al.* [35] proposed a non-cooperative game like iterative optimistic-pessimistic DEA approach to fully rank the DMUs. Badiezadeh *et al.* [5] were, to our knowledge, the first to conceive the idea of considering optimistic-pessimistic DEA models under a network DEA context to evaluate the performance of a sustainable supply-chain management.

With the exception of Badiezadeh *et al.* [5], the majority of the existing studies on the double frontier DEA models are concerned with a system handled as a whole unit, ignoring its internal structure. Several studies illustrate that this condition might produce misleading results [22]. In reality, systems can be composed of two sub-stages operating interdependently. In this paper, we will extend our selected optimistic-pessimistic ranking procedure to a two-stage tandem system to not only measure the efficiency of the overall system and its individual stages' efficiencies; thus, the stage that causes inefficiencies can be identified.

The optimistic and pessimistic self-efficiency scores can be unified *via* the geometric average efficiency. As shown in Wang *et al.* [50], this score has a better discriminating power than either of the opposing efficiencies. Yet, this feature has not been explored in a network environment, implying the possible existence of a non-unique ranking. It also considers the effects of the optimistic and pessimistic standpoints only within the self-appraisal context. The integration of the geometric average score in a peer-appraisal context would contribute to the assessment of a DMU in terms of the weight sets of other players, leading to a more logical ranking. These points make us infer that this framework could be further extended by the use of the cross-efficiency (CE) to ensure more fairness in the evaluation outcomes.

The CE concept is based on the peer-evaluation notion [42]. As stressed by Anderson *et al.* [1], CE improves the probability of obtaining a unique ranking. A shortcoming of the CE is the non-uniqueness of DEA optimal weights, leading to the non-uniqueness of cross-efficiencies. Remedial actions have been suggested towards the adoption of secondary goals in an aim to select unique optimal multipliers [10,26–28,52,57]. The non-uniqueness issue is also critical in a two-stage (network) system [18,22,32–34]. Kao and Liu [22], for instance, developed an aggressive CE model to measure the efficiency in two basic network structures. Örkcü *et al.* [34] came up with a neutral CE model in a two-stage system, which is indifferent to the preference choice between the aggressive and benevolent formulations.

Doyle and Green [10] introduced an aggressive and a benevolent secondary goal model to remedy the non-uniqueness of the optimal weights. The former ensures the minimisation and the latter the maximisation of the cross-efficiencies of all other DMUs, whilst both maintaining the optimistic self-efficiency of the target DMU. The use of any formulation of the two may be subject to an individual judgement, possibly leading to an irrational selection of either model. There is also no confirmation that these formulations will result in the same ranking or that their optimal set of multipliers are unique [46].

To alleviate these deficiencies, Yang *et al.* [53] suggested the "interval CE" for the exploration of the cross-efficiencies in a weight space considering all the weight profiles, within the single-stage DEA structure. In such a peer-appraisal setting, the base DMU is assessed regarding the most unfavourable and favourable weight profiles of each of the other DMUs. The aggressive and benevolent models of this process were, however, keeping only the optimistic self-efficiency value of each DMU fixed.

In summary, this paper adapts an optimistic-pessimistic DEA approach in the light of the two-stage tandem system, in order to then support the interval CE method in such a network system. Using the proposed framework as shown in Figure 1, a meaningful evaluation and ranking of the considered DMUs is attained. Decision makers will be enabled to simultaneously consider: (i) both the optimistic and the pessimistic viewpoints within the self-appraisal context, and (ii) the most favourable and unfavourable weight sets of each of the other DMUs in a peer-appraisal setting. We believe that the combination of the methods that compose our framework has not been considered before in the literature; in our view, this could lead to a meaningful ranking in addition to it being adjusted to a two-stage tandem DEA structure.

The procedures implemented in the first three steps of our proposed framework (Fig. 1) have been applied in several studies (*e.g.*, [50]) that focus on double frontier DEA models to evaluate DMUs in a self-appraisal context in a single-stage structure. As for these steps, our study differs in that our optimistic-pessimistic DEA models, which are inspired by the studies of Wang and Luo [48] and Wu [51], are built towards the two-stage tandem (network) system.

The remaining steps of the proposed framework pursue to support the peer-evaluation of the considered DMUs *via* the customisation of the interval CE method to the specifications of the two-stage tandem structure while embedding the respective combined self-efficiency measure (that considers the effects of both opposing standpoints). To rank the DMUs in the interval CE matrix of the corresponding flow, this paper views this matrix as a multi-criteria decision-making problem. To solve this problem, we implement the goal programming method of Wang and Elhag [47] to obtain the interval local weight of each criterion. To delineate the interval global weight of each alternative, we suggest a pair of linear programming models, introduced by Entani and Tanaka [13]. Finally, we apply the grey relational analysis [25] for ranking the interval global weights. To our knowledge, the aforementioned well-established approaches have not been previously considered for extracting valuable information from an interval CE matrix. We have also shown that our proposed framework offers a more informative assessment of the units under consideration than particular existing methods in network DEA-relevant literature.

The remainder of the paper is organised as follows. Section 2 shortly describes the preliminaries and the methodological background. Section 3 proposes the framework to meaningfully rank DMUs. Section 4 illustrates the methods with a numerical example. Section 5 presents conclusions and further research.

2. Methodological background

We assume that each DMU_j $(j = 1, 2, \ldots, n)$ uses m inputs $(i = 1, 2, \ldots, m)$ to produce s outputs $(r = 1, 2, \ldots, s)$. Let X_{ij} be the input value of $i \in M$ for DMU $j \in N$ and Y_{rj} be the output value of $r \in S$ for DMU $j \in N$. We estimate the optimistic self-efficiency for each DMU, based on determining an optimal set of the most favourable input and output weights. The conventional input-oriented CCR DEA model [7], that assesses

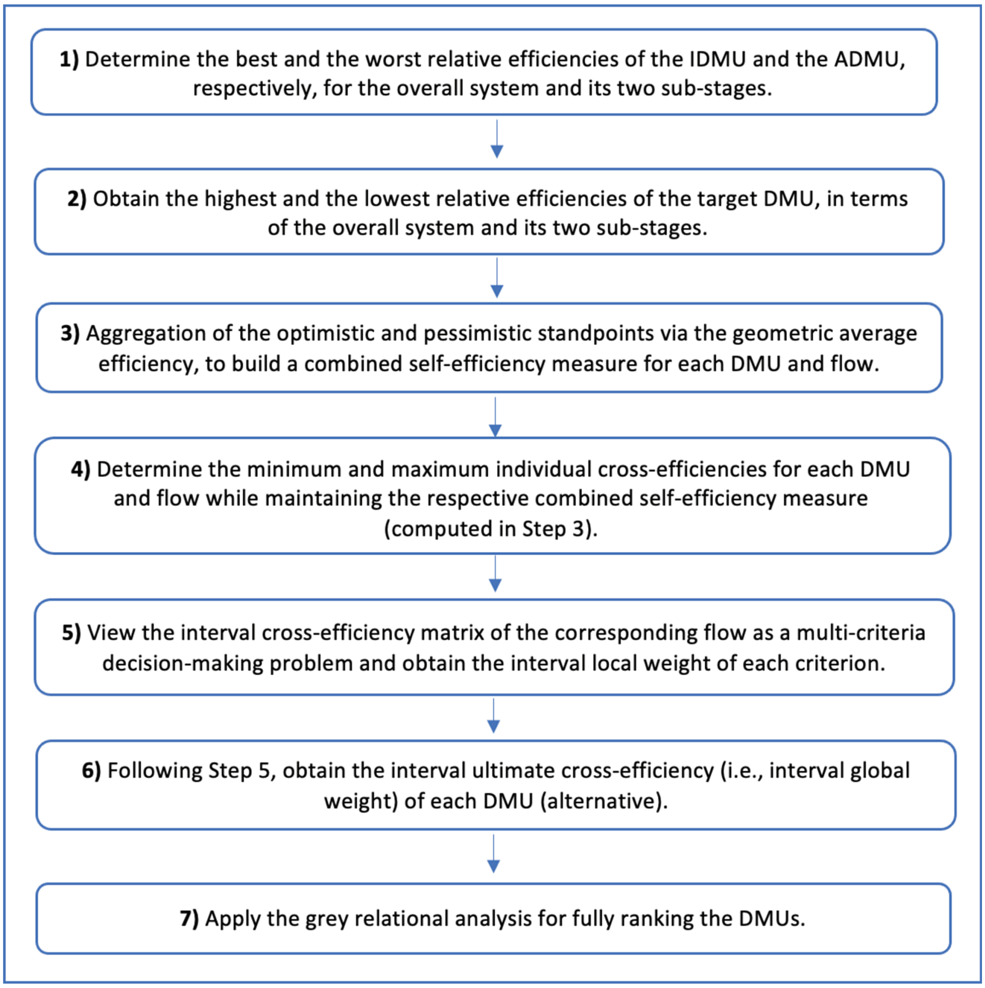

FIGURE 1. The proposed framework.

the efficiency of the target DMU_k, is illustrated as follows:

$$\begin{aligned}
E_{kk} = \text{Max} \ & \sum_{r=1}^{s} \mu_{rk} Y_{rk} \\
\text{subject to} \ & \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \qquad \forall j, \\
& \mu_{rk}, \nu_{ik} \geq 0, \qquad \forall r, i,
\end{aligned} \tag{2.1}$$

where μ_{rk}, ν_{ik} are the rth output and the ith input weights for DMU_k, respectively. If the optimal (optimistic) self-efficiency $E^*_{kk} = 1$, then DMU_k is called DEA efficient; otherwise it is said to be DEA inefficient.

TABLE 1. Cross-efficiency matrix [10].

		Target	DMU_j	
Evaluator DMU_k	1	2	...	n
1	E_{11}	E_{12}	...	E_{1n}
2	E_{21}	E_{22}	...	E_{2n}
...	...	...	...	...
n	E_{n1}	E_{n2}	...	E_{nn}

2.1. Cross-efficiency & interval cross-efficiency in single-stage structures

A significant challenge of the conventional single-stage DEA model, is to distinguish the efficient DMUs and thus to acquire a unique ranking of the DMUs. A potential remedy to overcome this inability is the implementation of the CE concept [42]. Let μ_{rk}^* and ν_{ik}^* be the optimal set of multipliers of model (2.1). Then, $E_{kk}^* = \sum_{r=1}^{s} \mu_{rk}^* Y_{rk}$ is the optimal self-efficiency score of DMU_k and reflects its desire to be assessed only on the basis of its own most favourable weights. On the other hand, CE, in which peer-appraisal is the main notion, evaluates each DMU, considering the weight profiles of all DMUs. The ratio $E_{kj} = \sum_{r=1}^{s} \mu_{rk}^* Y_{rj} / \sum_{i=1}^{m} \nu_{ik}^* X_{ij}$ denotes the individual cross-efficiency of DMU_j, based on the optimal weight scheme of DMU_k. A CE matrix (Tab. 1) can be a valuable tool to integrate both the peer-efficiency scores E_{kj} $(k, j = 1, 2, \ldots, n)$ and the self-efficiency scores E_{kk} (in the leading diagonal column). The ultimate cross-efficiency can be defined by averaging all individual cross-efficiencies of the corresponding DMU being evaluated. The ultimate score in this case is $\hat{e}_j = \frac{1}{n} \cdot \sum_{k=1}^{n} E_{kj},\ \forall j$ [2].

The existence of multiple optimal weights from model (2.1) can deteriorate the theoretical usefulness of the results obtained *via* the cross-efficiency concept. To tackle this issue, Doyle and Green [10] proposed two opposed secondary goals to choose their weights, favourable or unfavourable, among the optimal solutions.

Considering the DEA-related literature, there is not a well-established methodological approach to guide the decision-maker in reasonably selecting either the benevolent or the aggressive strategy. In addition, the selection of either the former or the latter model might not provide the same ranking or a unique optimal set of weights. To overcome these obstacles, Yang *et al.* [53] suggested the simultaneous use of the two extreme cases in the context of a single-stage structure.

Model (2.2) is an aggressive-based model to obtain an optimal set of multipliers and thus to identify the minimum individual cross-efficiency value of DMU_j based on DMU_k.

$$
\begin{aligned}
\text{Min } E_{kj}^{\mathrm{L}} &= \sum_{r=1}^{s} \mu_{rk} Y_{rj} \\
\text{subject to } & \sum_{i=1}^{m} \nu_{ik} X_{ij} = 1, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rk} - E_{kk}^* \sum_{i=1}^{m} \nu_{ik} X_{ik} = 0, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, && \forall j;\ j \neq k, \\
& \mu_{rk}, \nu_{ik} \geq 0, && \forall r, i.
\end{aligned} \tag{2.2}
$$

Table 2. Interval cross-efficiency matrix [53].

		Target	DMU_j	
Evaluator DMU_k	1	2	...	n
1	$[E_{11}, E_{11}]$	$\left[E_{12}^{\mathrm{L}}, E_{12}^{\mathrm{U}}\right]$	...	$\left[E_{1n}^{\mathrm{L}}, E_{1n}^{\mathrm{U}}\right]$
2	$\left[E_{21}^{\mathrm{L}}, E_{21}^{\mathrm{U}}\right]$	$[E_{22}, E_{22}]$	...	$\left[E_{2n}^{\mathrm{L}}, E_{2n}^{\mathrm{U}}\right]$
...	...	...	...	...
n	$\left[E_{n1}^{\mathrm{L}}, E_{n1}^{\mathrm{U}}\right]$	$\left[E_{n2}^{\mathrm{L}}, E_{n2}^{\mathrm{U}}\right]$	...	$[E_{nn}, E_{nn}]$

Model (2.3) is a benevolent-based model to obtain an optimal set of weights and thus to determine the maximum individual cross-efficiency of DMU_j based on DMU_k.

$$\begin{aligned}&\text{Max} \quad E_{kj}^{\mathrm{U}} = \sum_{r=1}^{s} \mu_{rk} Y_{rj}\\&\text{subject to the same constraints as in model (2.2).}\end{aligned} \tag{2.3}$$

In the above two models, the optimistic self-efficiency score E_{kk}^{*}, derived from model (2.1), remains fixed; this keeps one of the basic properties of the traditional CE concept intact. Overall, in this peer-evaluation procedure an interval individual cross-efficiency score of DMU_j in terms of DMU_k is formed and lies in the range $\left[E_{kj}^{\mathrm{L}}, E_{kj}^{\mathrm{U}}\right]$. E_{kj}^{L} is the lower bound and is found from model (2.2), whereas E_{kj}^{U} is the upper bound obtained from model (2.3). Table 2 depicts the individual cross-efficiencies as interval numbers, boosting the DM's uncertainty. The elements in the diagonal column of Table 2 show the special case of the self-efficiency scores, for which $E_{kk} = E_{kk}^{\mathrm{L}} = E_{kk}^{\mathrm{U}}, \forall k \in \{1, 2, \ldots, n\}$.

Models (2.2) and (2.3) make use of unfavourable and favourable multipliers, respectively, to identify the individual cross-efficiencies towards the single-stage structure. In either case, only the optimistic self-efficiency measure is involved to accommodate their purpose.

In Section 3.1, a combined self-efficiency score is obtained indicating the merger of the optimistic and pessimistic self-efficiencies. That score is embedded to the adjusted cross-efficiency models (Sect. 3.2) to explore the effect of both opposing viewpoints. The above-mentioned processes are part of a broader framework presented herein to reasonably rank DMUs towards the two-stage tandem structure.

3. Models development

The exploration of the internal processes taking place in the core of a DMU sets the foundation for the transition from a single-stage to a two-stage DEA structure. Each DMU_j ($j = 1, 2, \ldots, n$) consumes m inputs ($i = 1, 2, \ldots, m$) in the first stage to generate D intermediate products ($d = 1, 2, \ldots, D$). The outputs (intermediate measures) of the first stage are converted into inputs in the second stage to produce s final outputs ($r = 1, 2, \ldots, s$). Let X_{ij} be the input value of $i \in M$, Z_{dj} be the intermediate product of $d \in D$, and Y_{rj} be the output value of $r \in S$, for DMU $j \in N$ [21]. The above process is illustrated in the exploratory Figure 2.

According to the relational model of Kao and Hwang [21], to measure the performance of the overall system it is necessary to consider not only its operations, but also the operations of its individual sub-stages. In model (3.1) these operations are described by the constraints, which indicate that the aggregate output can not exceed the aggregate input.

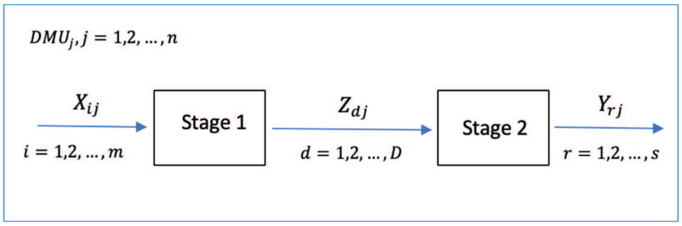

FIGURE 2. Typical two-stage tandem DEA structure.

$$
\begin{aligned}
E_k^s = \text{Max } & \sum_{r=1}^{s} \mu_{rk} Y_{rk} \\
\text{subject to } & \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1, \\
& \sum_{d=1}^{D} \eta_{dk} Z_{dj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \qquad \forall j, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{d=1}^{D} \eta_{dk} Z_{dj} \leq 0, \qquad \forall j, \\
& \mu_{rk}, \nu_{ik}, \eta_{dk} \geq 0, \qquad \forall r, i, d.
\end{aligned} \tag{3.1}
$$

At optimality of model (3.1), the system efficiency is estimated as $E_k^s = (\sum_{r=1}^{s} \mu_{rk}^* Y_{rk})/(\sum_{i=1}^{m} \nu_{ik}^* X_{ik})$, the efficiency of stage 1 as $E_k^1 = \left(\sum_{d=1}^{D} \eta_{dk}^* Z_{dk}\right)/(\sum_{i=1}^{m} \nu_{ik}^* X_{ik})$, and the efficiency of stage 2 as $E_k^2 = (\sum_{r=1}^{s} \mu_{rk}^* Y_{rk})/\left(\sum_{d=1}^{D} \eta_{dk}^* Z_{dk}\right)$. It is obvious that the overall efficiency is the product of the efficiencies of the stage efficiencies.

3.1. Optimistic & pessimistic models in basic two-stage structure

The above model can set the basis for the exploration of the optimistic and pessimistic self-efficiencies and, in turn, their integration into a geometric average efficiency score within the two-stage tandem system.

Sub-stage 1 consumes inputs to generate intermediate products. The following input-oriented CCR model (3.2) [21] examines the performance of sub-stage 1:

$$
\begin{aligned}
E_k^1 = \text{Max } & \sum_{d=1}^{D} \eta_{dk} Z_{dk} \\
\text{subject to } & \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1, \\
& \sum_{d=1}^{D} \eta_{dk} Z_{dj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \qquad \forall j, \\
& \nu_{ik}, \eta_{dk} \geq 0, \qquad \forall i, d.
\end{aligned} \tag{3.2}
$$

With reference to sub-stage 1 of a basic two-stage DEA structure, two fundamental concepts, the IDMU and the ADMU, are introduced, following the principles of Wang and Luo [48]. IDMU is a hypothetical DMU that utilises the least amount of inputs to generate the most intermediate products. An ADMU, on the other side, uses the most inputs to produce the least intermediate products. The IDMU can be expressed with the vectors $(X^{\min}, Z^{\max})$, where $X_i^{\min} = \min_k\{X_{ik}\}$ and $Z_d^{\max} = \max_k\{Z_{dk}\}$, $\forall\, i, d$. The ADMU can be determined with the vectors $(X^{\max}, Z^{\min})$, where $X_i^{\max} = \max_k\{X_{ik}\}$ and $Z_d^{\min} = \min_k\{Z_{dk}\}$, $\forall\; i, d$. As stressed in

Hatami-Marbini *et al.* [16], the performance of the IDMU cannot be worse than any of the actual DMUs, and the performance of the ADMU cannot be better than that of the worst performing actual DMU.

The best and worst relative efficiency scores in terms of sub-stage 1 can be defined by the following two CCR models, respective to the IDMU and the ADMU; they are related to Wang and Luo [48] and Wu's [51] models:

$$\begin{aligned} E^{\mathrm{IDMU}(1)} = \mathrm{Max}\ & \sum_{d=1}^{D} \eta_d Z_d^{\max} \\ \text{subject to}\ & \sum_{i=1}^{m} \nu_i X_i^{\min} = 1, \\ & \sum_{d=1}^{D} \eta_d Z_{dj} - \sum_{i=1}^{m} \nu_i X_{ij} \leq 0, && \forall j, \\ & \nu_i, \eta_d \geq 0, && \forall i, d, \end{aligned} \tag{3.3}$$

$$\begin{aligned} E^{\mathrm{ADMU}(1)} = \mathrm{Min}\ & \sum_{d=1}^{D} \eta_d Z_d^{\min} \\ \text{subject to}\ & \sum_{i=1}^{m} \nu_i X_i^{\max} = 1, \\ & \sum_{d=1}^{D} \eta_d Z_{dj} - \sum_{i=1}^{m} \nu_i X_{ij} \leq 0, && \forall j, \\ & \sum_{d=1}^{D} \eta_d Z_d^{\max} - E^{\mathrm{IDMU}(1)*} \sum_{i=1}^{m} \nu_i X_i^{\min} \geq 0, \\ & \nu_i, \eta_d \geq 0, && \forall i, d, \end{aligned} \tag{3.4}$$

where $E^{\mathrm{IDMU}(1)*}$ is the optimal optimistic score of IDMU in terms of sub-stage 1, obtained in model (3.3). Model (3.4) ensures that the best relative efficiency of sub-stage 1 is fixed at a value greater than or equal to $E^{\mathrm{IDMU}(1)*}$.

By the same token, we establish the definitions as well as formulate the appropriate optimisation models for the IDMU and the ADMU, regarding sub-stage 2 of the basic two-stage structure. Note that sub-stage 2 focuses on the consumption of intermediate products for the generation of the final outputs.

The next stage concerns the determination of the optimistic and pessimistic efficiency scores of the IDMU and the ADMU, respectively, in terms of the overall system. The reference model is the relational two-stage DEA model (3.1). The efficiency of the IDMU for the entire system can be defined as $E^{\mathrm{IDMU}(s)} = \left(\sum_{r=1}^{s} \mu_r Y_r^{\max}\right) / \left(\sum_{i=1}^{m} \nu_i X_i^{\min}\right)$. The factor weights μ_r and ν_i are assigned to the rth output and the ith input, respectively. We thus construct the following LP model that aims to maximise the efficiency of the IDMU.

$$\begin{aligned} E^{\mathrm{IDMU}(s)} = \mathrm{Max}\ & \sum_{r=1}^{s} \mu_r Y_r^{\max} \\ \text{subject to}\ & \sum_{i=1}^{m} \nu_i X_i^{\min} = 1, \\ & \sum_{d=1}^{D} \eta_d Z_{dj} - \sum_{i=1}^{m} \nu_i X_{ij} \leq 0, && \forall j, \\ & \sum_{r=1}^{s} \mu_r Y_{rj} - \sum_{d=1}^{D} \eta_d Z_{dj} \leq 0, && \forall j, \\ & \mu_r, \nu_i, \eta_d \geq 0, && \forall r, i, d. \end{aligned} \tag{3.5}$$

The efficiency of the ADMU for the entire system can be illustrated as $E^{\text{ADMU}(s)} = \left(\sum_{r=1}^{s} \mu_r Y_r^{\min}\right) / \left(\sum_{i=1}^{m} \nu_i X_i^{\max}\right)$. The associated optimisation model is formulated as follows:

$$
\begin{aligned}
E^{\text{ADMU}(s)} = \text{Min } & \sum_{r=1}^{s} \mu_r Y_r^{\min} \\
\text{subject to } & \sum_{i=1}^{m} \nu_i X_i^{\max} = 1, \\
& \sum_{d=1}^{D} \eta_d Z_{dj} - \sum_{i=1}^{m} \nu_i X_{ij} \leq 0, && \forall j, \\
& \sum_{r=1}^{s} \mu_r Y_{rj} - \sum_{d=1}^{D} \eta_d Z_{dj} \leq 0, && \forall j, \\
& \sum_{r=1}^{s} \mu_r Y_r^{\max} - E^{\text{IDMU}(s)*} \sum_{i=1}^{m} \nu_i X_i^{\min} \geq 0, \\
& \mu_r, \nu_i, \eta_d \geq 0, && \forall r, i, d.
\end{aligned} \tag{3.6}
$$

Model (3.6) aims to minimise the pessimistic efficiency measure of the ADMU, while keeping the optimistic efficiency of the IDMU for the overall system no less than $E^{\text{IDMU}(s)*}$. It should be noted that the second and third sets of constraints in both models imply that the overall efficiency of DMU cannot exceed 1.

The next point to focus on in this paper is the examination of the highest and the lowest relative efficiency of each DMU, considering their self-evaluation. In model (3.7), the optimistic relative efficiency of DMU_k for the sub-stage 1 is examined while $E^{\text{IDMU}(1)*}$ is kept fixed; it is related to Wang and Luo's [48] framework:

$$
\begin{aligned}
E_k^{\text{IDMU}(1)} = \text{Max } & \sum_{d=1}^{D} \eta_{dk} Z_{dk} \\
\text{subject to } & \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1, \\
& \sum_{d=1}^{D} \eta_{dk} Z_{dj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, && \forall j, \\
& \sum_{d=1}^{D} \eta_{dk} Z_d^{\max} - E^{\text{IDMU}(1)*} \sum_{i=1}^{m} \nu_{ik} X_i^{\min} = 0, \\
& \nu_{ik}, \eta_{dk} \geq 0, && \forall i, d.
\end{aligned} \tag{3.7}
$$

In the same manner, we construct the counterpart model for measuring the highest relative efficiency of DMU_k for the sub-stage 2, considering $E^{\text{IDMU}(2)*}$ as the fixed parameter.

The overall optimistic efficiency score of DMU_k can be determined as $E_k^{\text{IDMU}(s)} = \left(\sum_{r=1}^{s} \mu_{rk} Y_{rk}\right) / \left(\sum_{i=1}^{m} \nu_{ik} X_{ik}\right)$. It is clear that this measure is the product of the optimistic efficiencies of the DMU_k of the two sub-stages, adopting the principle of the multiplicative efficiency decomposition approach [21]. Thus, we propose model (3.8), that maximises the above ratio.

$$
\begin{aligned}
E_k^{\text{IDMU}(s)} = \text{Max } & \sum_{r=1}^{s} \mu_{rk} Y_{rk} \\
\text{subject to } & \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1,
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{d=1}^{D}\eta_{dk}Z_{dj}-\sum_{i=1}^{m}\nu_{ik}X_{ij}\leq 0, && \forall j,\\
&\sum_{r=1}^{s}\mu_{rk}Y_{rj}-\sum_{d=1}^{D}\eta_{dk}Z_{dj}\leq 0, && \forall j,\\
&\sum_{d=1}^{D}\eta_{dk}Z_{d}^{\max}-E^{\mathrm{IDMU(1)}*}\sum_{i=1}^{m}\nu_{ik}X_{i}^{\min}=0, &&\\
&\sum_{r=1}^{s}\mu_{rk}Y_{r}^{\max}-E^{\mathrm{IDMU(2)}*}\sum_{d=1}^{D}\eta_{dk}Z_{d}^{\min}=0, &&\\
&\mu_{rk},\nu_{ik},\eta_{dk}\geq 0, && \forall r,i,d. \qquad (3.8)
\end{aligned}
$$

The fourth and fifth constraints indicate that $E^{\mathrm{IDMU(1)}*}$ and $E^{\mathrm{IDMU(2)}*}$, respectively, remain unchanged. Let $\nu_k^*=(\nu_{1k}^*,\nu_{2k}^*,\ldots,\nu_{mk}^*)$, $\eta_k^*=(\eta_{1k}^*,\eta_{2k}^*,\ldots,\eta_{Dk}^*)$, $\mu_k^*=(\mu_{1k}^*,\mu_{2k}^*,\ldots,\mu_{sk}^*)$, be an optimal solution to model (3.8). For DMU_k, $E_k^{\mathrm{IDMU}(s)}=(\sum_{r=1}^{s}\mu_{rk}^*Y_{rk})/(\sum_{i=1}^{m}\nu_{ik}^*X_{ik})$, $E_k^{\mathrm{IDMU(1)}}=\left(\sum_{d=1}^{D}\eta_{dk}^*Z_{dk}\right)/(\sum_{i=1}^{m}\nu_{ik}^*X_{ik})$, and $E_k^{\mathrm{IDMU(2)}}=(\sum_{r=1}^{s}\mu_{rk}^*Y_{rk})/\left(\sum_{d=1}^{D}\eta_{dk}^*Z_{dk}\right)$, which are referred to as optimistic self-efficiency measures with respect to the overall system and its sub-stages, respectively.

Then, model (3.9) evaluates the worst relative efficiency of DMU_k, in terms of sub-stage 1, while the parameter $E^{\mathrm{ADMU(1)}*}$ takes the value as determined previously from model (3.4). This model is related to Wu's [51] framework.

$$
\begin{aligned}
E_k^{\mathrm{ADMU(1)}}=\ &\mathrm{Min}\ \sum_{d=1}^{D}\eta_{dk}Z_{dk}\\
\text{subject to}\ &\sum_{i=1}^{m}\nu_{ik}X_{ik}=1,\\
&\sum_{d=1}^{D}\eta_{dk}Z_{dj}-\sum_{i=1}^{m}\nu_{ik}X_{ij}\leq 0, && \forall j,\\
&\sum_{d=1}^{D}\eta_{dk}Z_{d}^{\min}-E^{\mathrm{ADMU(1)}*}\sum_{i=1}^{m}\nu_{ik}X_{i}^{\max}=0,\\
&\nu_{ik},\eta_{dk}\geq 0, && \forall i,d. \qquad (3.9)
\end{aligned}
$$

Similarly, we formulate the counterpart model for measuring the lowest relative efficiency of DMU_k for the sub-stage 2, considering $E^{\mathrm{ADMU(2)}*}$ as the unchanged parameter.

The overall pessimistic score of DMU_k can be determined as $E_k^{\mathrm{ADMU}(s)}=(\sum_{r=1}^{s}\mu_{rk}Y_{rk})/(\sum_{i=1}^{m}\nu_{ik}X_{ik})$ and denotes the product of the pessimistic efficiencies of the DMU_k of the two sub-stages. Thus, we suggest model (3.10), whose purpose is to minimise the above ratio. $E^{\mathrm{ADMU(1)}*}$ and $E^{\mathrm{ADMU(2)}*}$ are maintained.

$$
\begin{aligned}
E_k^{\mathrm{ADMU}(s)}=\ &\mathrm{Min}\ \sum_{r=1}^{s}\mu_{rk}Y_{rk}\\
\text{subject to}\ &\sum_{i=1}^{m}\nu_{ik}X_{ik}=1,\\
&\sum_{d=1}^{D}\eta_{dk}Z_{dj}-\sum_{i=1}^{m}\nu_{ik}X_{ij}\leq 0, && \forall j,
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{d=1}^{D} \eta_{dk} Z_{dj} \leq 0, && \forall j,\\
&\sum_{d=1}^{D} \eta_{dk} Z_d^{\min} - E^{\mathrm{ADMU}(1)*} \sum_{i=1}^{m} \nu_{ik} X_i^{\max} = 0, &&\\
&\sum_{r=1}^{s} \mu_{rk} Y_r^{\min} - E^{\mathrm{ADMU}(2)*} \sum_{d=1}^{D} \eta_{dk} Z_d^{\max} = 0, &&\\
&\mu_{rk}, \nu_{ik}, \eta_{dk} \geq 0, && \forall r, i, d.
\end{aligned} \tag{3.10}
$$

Let $\nu_k^{\sim} = (\nu_{1k}^{\sim}, \nu_{2k}^{\sim}, \ldots, \nu_{mk}^{\sim})$, $\eta_k^{\sim} = (\eta_{1k}^{\sim}, \eta_{2k}^{\sim}, \ldots, \eta_{Dk}^{\sim})$, $\mu_k^{\sim} = (\mu_{1k}^{\sim}, \mu_{2k}^{\sim}, \ldots, \mu_{sk}^{\sim})$, be an optimal solution to model (3.10). For DMU_k, $E_k^{\mathrm{ADMU}(s)} = (\sum_{r=1}^{s} \mu_{rk}^{\sim} Y_{rk})/(\sum_{i=1}^{m} \nu_{ik}^{\sim} X_{ik})$, $E_k^{\mathrm{ADMU}(1)} = \left(\sum_{d=1}^{D} \eta_{dk}^{\sim} Z_{dk}\right)/(\sum_{i=1}^{m} \nu_{ik}^{\sim} X_{ik})$, and $E_k^{\mathrm{ADMU}(2)} = (\sum_{r=1}^{s} \mu_{rk}^{\sim} Y_{rk})/\left(\sum_{d=1}^{D} \eta_{dk}^{\sim} Z_{dk}\right)$, which are referred to as pessimistic self-efficiency measures with respect to the overall system and its constituent parts, respectively.

Consequently, in a two-stage DEA structure, a self-efficiency interval is formulated for each DMU under consideration, both for the overall system and its constituent stages. For instance, considering the overall system, an efficiency interval denoted by $\left[E_k^{\mathrm{ADMU}(s)*}, E_k^{\mathrm{IDMU}(s)*}\right]$ is shaped, where $E_k^{\mathrm{ADMU}(s)*}$ (lower bound) represents the worst relative efficiency of DMU_k and $E_k^{\mathrm{IDMU}(s)*}$ (upper bound) illustrates the best relative efficiency of DMU_k, obtained *via* models (3.10) and (3.8), respectively.

There is a clear need to integrate both optimistic and pessimistic self-efficiency measures to provide an overall assessment of the performance of each DMU in a two-stage DEA process. This study adopts the geometric average efficiency measure, proposed and verified by Wang *et al.* [50], to meet this requirement. Let $E_k^{\mathrm{comb}(\epsilon)*} = \sqrt{E_k^{\mathrm{ADMU}(\epsilon)*} \cdot E_k^{\mathrm{IDMU}(\epsilon)*}}$ be the combined self-efficiency measure of DMU_k, where $\epsilon = s$ (overall system) or 1 (sub-stage 1) or 2 (sub-stage 2). We easily prove that the combined self-efficiency score of DMU_k for the overall system is the product of the combined self-efficiency measures of DMU_k for the two sub-stages: $E_k^{comb(s)*} = \sqrt{E_k^{\mathrm{ADMU}(s)*} \cdot E_k^{\mathrm{IDMU}(s)*}} = \sqrt{E_k^{\mathrm{ADMU}(1)*} \cdot E_k^{\mathrm{ADMU}(2)*} \cdot E_k^{\mathrm{IDMU}(1)*} \cdot E_k^{\mathrm{IDMU}(2)*}} = \sqrt{E_k^{\mathrm{ADMU}(1)*} \cdot E_k^{\mathrm{IDMU}(1)*}} \cdot \sqrt{E_k^{\mathrm{ADMU}(2)*} \cdot E_k^{\mathrm{IDMU}(2)*}} = E_k^{comb(1)*} \cdot E_k^{comb(2)*}$.

The geometric average efficiency is an approachable efficiency measure that leads to a fairer ranking index [50]. However, we should consider that it sheds light on the effects of the optimistic and pessimistic standpoints only within the self-appraisal context. In other words, each DMU is assessed, based on its own most favourable and unfavourable weights, without considering the weight scheme of each of the other DMUs. This score also ensures a better discriminating power than either of the optimistic and pessimistic efficiencies [50]. Yet, this feature has not been explored in a more complex network structure. To this end, in the next section, the double frontier DEA models are further extended by the use of the interval CE within a two-stage tandem system, to ensure a more logical ranking order.

3.2. Interval cross-efficiencies in basic two-stage structure

In this section, we will propose the customisation and simultaneous use of the traditional aggressive and benevolent secondary models in the context of the basic two-stage DEA structure with combined self-efficiencies, obtained in Section 3.1. Their purpose is the determination of the minimum and maximum individual cross-efficiencies of DMU_j, with respect to the optimal weight scheme of DMU_k $(k, j = 1, 2, \ldots, n)$, respectively. A fruitful aspect we believe, is the integration of the combined self-efficiency score for the corresponding system/stage within the CE process. This is irrespective of the type of multipliers, favourable for a benevolent or unfavourable for an aggressive strategy, that are used to capture the cross-efficiencies.

We initially adopt an aggressive strategy to establish the following minimisation model:

$$\begin{aligned}
E_{kj}^{L(s)} = \text{Min} \ & \sum_{r=1}^{s} \mu_{rk} Y_{rj} \\
\text{subject to} \ & \sum_{i=1}^{m} \nu_{ik} X_{ij} = 1, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rk} - E_{k}^{\text{comb}(s)*} \sum_{i=1}^{m} \nu_{ik} X_{ik} = 0, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rk} - E_{k}^{\text{comb}(2)*} \sum_{d=1}^{D} \eta_{dk} Z_{dk} = 0, \\
& \sum_{d=1}^{D} \eta_{dk} Z_{dj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \qquad \forall j, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{d=1}^{D} \eta_{dk} Z_{dj} \leq 0, \qquad \forall j, \\
& \mu_{rk}, \nu_{ik}, \eta_{dk} \geq 0, \qquad \forall r, i, d.
\end{aligned} \tag{3.11}$$

In model (3.11), $E_{k}^{\text{comb}(s)*}$ and $E_{k}^{\text{comb}(2)*}$ are the crisp combined self-efficiency measures of the system and the sub-stage 2 for DMU_k, respectively, obtained from Section 3.1. The second and third constraint maintain combined system and sub-stage efficiencies for DMUs. Model (3.11) pursues to minimise the cross-efficiency value of DMU_j under the condition that the combined self-efficiency scores for the overall system and its constituent parts remain unchanged. At optimality, the minimum individual cross-efficiencies of DMU_j based on DMU_k ($j \neq k$) for the overall system, the stage 1, and the stage 2, are determined as $E_{kj}^{L(s)} = (\sum_{r=1}^{s} \mu_{rk}^{*} Y_{rj})/(\sum_{i=1}^{m} \nu_{ik}^{*} X_{ij})$, $E_{kj}^{L(1)} = \left(\sum_{d=1}^{D} \eta_{dk}^{*} Z_{dj}\right)/(\sum_{i=1}^{m} \nu_{ik}^{*} X_{ij})$, and $E_{kj}^{L(2)} = (\sum_{r=1}^{s} \mu_{rk}^{*} Y_{rj})/\left(\sum_{d=1}^{D} \eta_{dk}^{*} Z_{dj}\right)$, respectively. By the same token, a benevolent strategy is implemented to construct the following maximisation model:

$$\begin{aligned}
E_{kj}^{U(s)} = \text{Max} \quad & \sum_{r=1}^{s} \mu_{rk} \cdot Y_{rj} \\
& \text{subject to the same constraints as in model (3.11).}
\end{aligned} \tag{3.12}$$

This model seeks to maximise the cross-efficiency of DMU_j given that the combined self-efficiency measures are kept fixed for the overall system and its sub-stages. Similarly, we define the maximum individual cross-efficiencies of DMU_j for the system and its stages.

In terms of ϵ, where $\epsilon = s$ (overall system), 1 (stage 1) or 2 (stage 2), for DMU_j, its cross-efficiency rated by DMU_k lies in $\left[E_{kj}^{\text{L}(\epsilon)}, E_{kj}^{\text{U}(\epsilon)}\right]$, where $E_{kj}^{\text{L}(\epsilon)}$ is the lower bound and $E_{kj}^{\text{U}(\epsilon)}$ is the upper bound. Therefore, three generalised interval CE matrices (based on the concept of Tab. 2) are shaped for the n DMUs, in regard to the overall system, the stage 1, and the stage 2, respectively. The diagonal column in each of these matrices demonstrates the special case in which $E_{jj}^{\text{L}(\epsilon)*} = E_{jj}^{\text{U}(\epsilon)*} = E_{j}^{\text{comb}(\epsilon)*}$ $\forall j$, where $\epsilon = s$, 1 or 2.

The recently created interval CE matrices can be viewed as MCDM problems. Taking that into consideration, we will set the scene for the determination of the interval local weights of criteria and the interval global weights of alternatives (ultimate cross-efficiencies) to fully rank the DMUs, in a basic two-stage DEA structure.

3.3. Interval cross-efficiencies and MCDM context

Each generalised interval CE matrix (see Sect. 3.2) can be treated as a multi-criteria decision making (MCDM) problem with $j = 1, 2, \ldots, n$ DMUs that act as alternatives. Each DMU_j is assessed considering the weight profile of $k = 1, 2, \ldots, n$ DMUs that act as criteria. Interestingly, the former intuition is attributed to the novel study of Cook *et al.* [8], according to which each DEA-related problem could be viewed as a multi-criteria evaluation problem. This has also been consolidated by Rakhshan [37], who argues that the combination of the MCDM and the DEA tools could mitigate their drawbacks when applied as stand-alone techniques.

Our primary target is to estimate the interval ultimate cross-efficiency scores, which are the interval global weights for the evaluated DMUs. To this end, our approach is twofold as it requires not only the local weights of alternatives with respect to a certain criterion, but also the local weights of criteria. The former are the elements $E_{kj}^{\mathrm{L}(\epsilon)}$ and $E_{kj}^{\mathrm{U}(\epsilon)}$, which act as lower-level and upper-level local weights of alternative j in reference to criterion k for $\epsilon = s, 1$ or 2, respectively, and overall compose $\left[E_{kj}^{\mathrm{L}(\epsilon)}, E_{kj}^{\mathrm{U}(\epsilon)}\right]$. These elements have been obtained in Section 3.2. The latter illustrates the local weight of criterion k, that is manifested as an interval value with lower bound w_k^{L} and upper bound w_k^{U}. The existence of this interval value is due to dealing with two diametrically opposed strategies for the overall system and its constituent stages.

Wang and Elhag [47] suggest a goal programming (GP) method to elicit normalised interval local weights from an interval comparison matrix. In our scenario, the interval CE matrix is committed to undertaking the role of the interval comparison matrix. Their method captures the lower and upper limits of the local weight of criterion k ($k = 1, 2, \ldots, n$) without ignoring the interval individual cross-efficiencies and the potential existence of uncertainty. We will provide their optimisation model as we would apply this within the basic 2-stage series structure:

$$
\begin{aligned}
\Omega = \mathrm{Min}\ & \sum_{k=1}^{n} \left(\delta_k^+ + \delta_k^- + \gamma_k^+ + \gamma_k^-\right) \\
\text{subject to}\ & (E_{\mathrm{L}} - I)W_{\mathrm{U}} - (n-1)W_{\mathrm{L}} - \Delta^+ + \Delta^- = 0, \\
& (E_{\mathrm{U}} - I)W_{\mathrm{L}} - (n-1)W_{\mathrm{U}} - \Gamma^+ + \Gamma^- = 0, \\
& w_k^{\mathrm{L}} + \sum_{\omega=1,\omega\neq k}^{n} w_\omega^{\mathrm{U}} \geq 1, && \forall k, \\
& w_k^{\mathrm{U}} + \sum_{\omega=1,\omega\neq k}^{n} w_\omega^{\mathrm{L}} \leq 1, && \forall k, \\
& W_{\mathrm{U}} - W_{\mathrm{L}} \geq 0, \\
& W_{\mathrm{U}}, W_{\mathrm{L}}, \Delta^+, \Delta^-, \Gamma^+, \Gamma^- \geq 0,
\end{aligned} \tag{3.13}
$$

where $\Delta^+ = \left(\delta_1^+, \ldots, \delta_n^+\right)^T$, $\Delta^- = \left(\delta_1^-, \ldots, \delta_n^-\right)^T$, $\Gamma^+ = \left(\gamma_1^+, \ldots, \gamma_n^+\right)^T$, $\Gamma^- = \left(\gamma_1^-, \ldots, \gamma_n^-\right)^T$, $W_{\mathrm{U}} = \left(w_1^{\mathrm{U}}, \ldots, w_n^{\mathrm{U}}\right)^T$, $W_{\mathrm{L}} = \left(w_1^{\mathrm{L}}, \ldots, w_n^{\mathrm{L}}\right)^T$, I is a $n \otimes n$ unit matrix whose elements on the diagonal are 1, and E_{L} and E_{U} are the minimum and maximum individual cross-efficiency matrices, whose elements are in the form of $E_{kj}^{\mathrm{L}(\epsilon)}$ and $E_{kj}^{\mathrm{U}(\epsilon)}$ respectively. The deviation vectors $\Delta^+, \Delta^-, \Gamma^+, \Gamma^-$, that appear in the first two constraint sets, pursue to eliminate the uncertainty and connect the lower level criteria W_{L} with the upper level criteria W_{U}. The third and fourth sets of constraints ensure the normalisation of the local interval weights, whereas the fifth constraint set determines their lower and upper bounds. Model (3.13) should, in effect, run three times, based on the investigation of the interval CE matrix of the respective system and stage to compose $\left[w_k^{\mathrm{L}(\epsilon)}, w_k^{\mathrm{U}(\epsilon)}\right]$.

Their approach might make sense in our study for two reasons. It has a greater scope for action due to its compatibility with any interval comparison matrix, and involves less constraints than other methods such as that of Sugihara *et al.* [43]. This enables it as an easier-to-use method for the DM. The fewer number of constraints was owed to its practice, putting more emphasis on the matrix as a whole rather than on each element individually. Wang and Elhag's [47] technique has, to our knowledge, not received attention on eliciting interval local weights from an interval CE matrix. Therefore, this section intends to use their approach to achieve this goal.

Table 3. Synthesis of interval cross-efficiencies.

		Target	DMU$_j$	
Evaluator DMU$_k$	1	2	...	n
1 $\left[w_1^{\mathrm{L}(\epsilon)}, w_1^{\mathrm{U}(\epsilon)}\right]$	$\left[E_{11}^{\mathrm{L}(\epsilon)}, E_{11}^{\mathrm{U}(\epsilon)}\right]$	$\left[E_{12}^{\mathrm{L}(\epsilon)}, E_{12}^{\mathrm{U}(\epsilon)}\right]$	...	$\left[E_{1n}^{\mathrm{L}(\epsilon)}, E_{1n}^{\mathrm{U}(\epsilon)}\right]$
2 $\left[w_2^{\mathrm{L}(\epsilon)}, w_2^{\mathrm{U}(\epsilon)}\right]$	$\left[E_{21}^{\mathrm{L}(\epsilon)}, E_{21}^{\mathrm{U}(\epsilon)}\right]$	$\left[E_{22}^{\mathrm{L}(\epsilon)}, E_{22}^{\mathrm{U}(\epsilon)}\right]$	...	$\left[E_{2n}^{\mathrm{L}(\epsilon)}, E_{2n}^{\mathrm{U}(\epsilon)}\right]$
...	...	...	...	...
n $\left[w_n^{\mathrm{L}(\epsilon)}, w_n^{\mathrm{U}(\epsilon)}\right]$	$\left[E_{n1}^{\mathrm{L}(\epsilon)}, E_{n1}^{\mathrm{U}(\epsilon)}\right]$	$\left[E_{n2}^{\mathrm{L}(\epsilon)}, E_{n2}^{\mathrm{U}(\epsilon)}\right]$	...	$\left[E_{nn}^{\mathrm{L}(\epsilon)}, E_{nn}^{\mathrm{U}(\epsilon)}\right]$
Ultimate cross-efficiencies	$\left[E_1^{\mathrm{L.B.}(\epsilon)}, E_1^{\mathrm{U.B.}(\epsilon)}\right]$	$\left[E_2^{\mathrm{L.B.}(\epsilon)}, E_2^{\mathrm{U.B.}(\epsilon)}\right]$	...	$\left[E_n^{\mathrm{L.B.}(\epsilon)}, E_n^{\mathrm{U.B.}(\epsilon)}\right]$

Taking the interval local weight for each criterion k and the interval local weight of each alternative j with respect to criterion k into account, we determine the interval ultimate cross-efficiencies for the alternatives. We recommend using the practical method of Entani and Tanaka [13] that is based on a pair of linear programming (LP) models. Their approach treats the local weights of criteria as decision variables to be optimised and intends to determine the global weights for each DMU. The pair of LP models is described as follows:

$$\begin{aligned} E_j^{\mathrm{L.B.}(\epsilon)} = \mathrm{Min}\ & \sum_{k=1}^{n} w_k^{(\epsilon)} E_{kj}^{\mathrm{L}(\epsilon)} \\ \text{subject to}\ & \sum_{k=1}^{n} w_k^{(\epsilon)} = 1, \\ & w_k^{\mathrm{L}(\epsilon)} \leq w_k^{(\epsilon)} \leq w_k^{\mathrm{U}(\epsilon)}, \qquad \forall k, \end{aligned} \tag{3.14}$$

and

$$\begin{aligned} E_j^{\mathrm{U.B.}(\epsilon)} = \mathrm{Max}\quad & \sum_{k=1}^{n} w_k^{(\epsilon)} E_{kj}^{\mathrm{U}(\epsilon)} \\ & \text{subject to the same constraints as in model (3.14),} \end{aligned} \tag{3.15}$$

where $w_k^{(\epsilon)}$ is the decision variable of the kth local criterion weight ($k = 1, 2, \ldots, n$) for $\epsilon = s$ (overall system), 1 (stage 1) or 2 (stage 2). The above pair of LP models (3.14) and (3.15) results in the interval global weight for each alternative j ($j = 1, 2, \ldots, n$), denoted by $\left[E_j^{\mathrm{L.B.}(\epsilon)}, E_j^{\mathrm{U.B.}(\epsilon)}\right]$ for the entire system and its sub-stages. Table 3 illustrates the synthesis of the interval cross-efficiencies.

3.4. Grey relational analysis for ranking DMUs

In Section 3.3, we obtained an interval ultimate cross-efficiency score for DMU$_j$ ($j = 1, 2, \ldots, n$). It is apparent that there is a significant need to identify a simple yet efficient ranking approach for comparing and ranking different DMUs, whose performance is expressed in the form of interval values. In this study, the Grey Relational Analysis (GRA) is applied to obtain a unique ranking order for the DMUs, whose ultimate cross-efficiencies are illustrated within certain boundaries, and thus to determine the most desirable alternative. GRA is based on the grey system theory proposed by Julong [20]. It has proved to be a worthy methodological tool when uncertain information emerges. GRA has fruitfully examined complex interconnections among several factors [6] as well as obtained the optimal alternative among several alternatives [25, 41].

GRA consists of four main steps: grey relational generating, reference sequence definition, grey relational coefficient calculation, and grey relational grade (GRG) calculation. In a first step, GRA translates the existing

TABLE 4. Interval ultimate cross-efficiencies.

DMU_j	1 (L.B.)	2 (U.B.)
1	$E_1^{\text{L.B.}(\epsilon)}$	$E_1^{\text{U.B.}(\epsilon)}$
2	$E_2^{\text{L.B.}(\epsilon)}$	$E_2^{\text{U.B.}(\epsilon)}$
...	...	...
n	$E_n^{\text{L.B.}(\epsilon)}$	$E_n^{\text{U.B.}(\epsilon)}$

performance of all alternatives into comparability sequence. According to the comparability sequence, an ideal target sequence (reference sequence) is defined in the reference sequence definition (second step). In a third step, a grey relational coefficient is calculated to illustrate the distance between the comparability and the reference sequence. In a final step, the GRG between the reference and every comparability sequence is calculated, based on the grey relational coefficient. If the comparability sequence of an alternative has the highest grey relational grade, then this alternative is deemed as the most desirable one [25]. Below, we will provide an overview of the GRA as we would apply this to ranking interval ultimate cross-efficiencies.

To start with, we collect the data to be evaluated from the mathematical viewpoint. The interval ultimate cross-efficiency scores, defined in Section 3.3, are gathered into a $n \otimes 2$ matrix, setting out the appropriate conditions for translating the DMUs into alternatives and the two extreme cases (lower bound, upper bound) into criteria. Hence, we form another MCDM problem with $j = 1, 2, \ldots, n$ alternatives that are assessed by $i = 1, 2$ attributes. Table 4 depicts what we described above.

The jth alternative can be expressed as $E_j^{(\epsilon)} = \left(E_j^{\text{L.B.}(\epsilon)}, E_j^{\text{U.B.}(\epsilon)}\right)$, where $E_j^{i(\epsilon)}$ is the ultimate cross-efficiency of attribute i of alternative j and where $\epsilon = s$ (overall system), 1 (stage 1) or 2 (stage 2). The term $E_j^{(\epsilon)}$ is translated into the comparability sequence $E_j^{\bar{(\epsilon)}} = \left(E_j^{\text{L.}\bar{\text{B}}.(\epsilon)}, E_j^{\text{U.}\bar{\text{B}}.(\epsilon)}\right)$ by use of one of the following equations:

$$E_j^{\bar{i}(\epsilon)} = \frac{E_j^{i(\epsilon)} - \min\left\{E_j^{i(\epsilon)}, \forall\, j\right\}}{\max\left\{E_j^{i(\epsilon)},\ j\right\} - \min\left\{E_j^{i(\epsilon)}, \forall\, j\right\}}, \qquad \forall j, i, \tag{3.16}$$

$$E_j^{\bar{i}(\epsilon)} = \frac{\max\left\{E_j^{i(\epsilon)}, \forall\, j\right\} - E_j^{i(\epsilon)}}{\max\left\{E_j^{i(\epsilon)}, \forall\, j\right\} - \min\left\{E_j^{i(\epsilon)}, \forall\, j\right\}}, \qquad \forall j, i, \tag{3.17}$$

$$E_j^{\bar{i}(\epsilon)} = \frac{\left|E_j^{i(\epsilon)} - E_{des}^{i(\epsilon)}\right|}{\max\left\{E_j^{i(\epsilon)}, \forall\, j\right\} - \min\left\{E_j^{i(\epsilon)}, \forall\, j\right\}}, \qquad \forall j, i. \tag{3.18}$$

Equation (3.16) is used for the greater-the-better attributes, equation (3.17) is used for the smaller-the-better attributes, and equation (3.18) is used for the closer-to-the-desired-value- $E_{des}^{i(\epsilon)}$ -the-better.

We proceed to calculating the grey relation distance between the reference sequence $\left(E_j^{i(\epsilon)}\right)$ and the comparability sequence $\left(E_j^{\bar{i}(\epsilon)}\right)$, which is $\Delta_j^{i(\epsilon)} = \left|E_j^{i(\epsilon)} - E_j^{\bar{i}(\epsilon)}\right|$, $\forall\, j, i$. As stressed in Kuo *et al.* [25], the reference sequence $E_j^{i(\epsilon)} = \left(E_j^{1(\epsilon)}, E_j^{2(\epsilon)}\right) = (1, 1)$.

Then, we compute the grey relational coefficient $\gamma\left(E_j^{i(\epsilon)}, E_j^{\bar{i}(\epsilon)}\right)$. It is used to determine how close $E_j^{i(\epsilon)}$ is to $E_j^{\bar{i}(\epsilon)}$. The larger the coefficient, the closer $E_j^{i(\epsilon)}$ and $E_j^{\bar{i}(\epsilon)}$ are. Let $\gamma\left(E_j^{i(\epsilon)}, E_j^{\bar{i}(\epsilon)}\right) = \frac{\Delta_{\min}^{(\epsilon)} + \zeta \cdot \Delta_{\max}^{(\epsilon)}}{\Delta_j^{i(\epsilon)} + \zeta \cdot \Delta_{\max}^{(\epsilon)}}$, $\forall j, i$ where

TABLE 5. The numerical application of Zhou *et al.* [55].

DMU	X_1	X_2	X_3	Z_1	Z_2	Y_1	Y_2
1	0.526	0.478	0.385	49.917	5.461	34.990	0.843
2	0.713	1.236	0.555	37.495	4.083	20.601	0.486
3	0.443	0.446	0.342	20.985	0.690	8.633	0.129
4	0.638	1.248	0.457	45.051	1.724	9.235	0.302
5	0.575	0.705	0.404	38.163	2.249	12.017	0.314
6	0.432	0.645	0.401	30.168	2.335	13.813	0.377
7	0.510	0.724	0.371	26.539	1.342	5.096	0.145
8	0.322	0.336	0.233	16.124	0.489	5.980	0.093
9	0.423	0.668	0.347	22.185	1.177	9.235	0.200
10	0.256	0.342	0.159	13.436	0.406	2.533	0.006

$\Delta^{(\epsilon)}_{\min} = \min\left\{\Delta^{i(\epsilon)}_j,\ \forall\, j, i\right\}$, $\Delta^{(\epsilon)}_{\max} = \max\left\{\Delta^{i(\epsilon)}_j,\ \forall\, j, i\right\}$, and ζ denotes the distinguishing coefficient, $\zeta \in [0, 1]$. ζ is used to expand or squeeze the range of the grey relational coefficient.

Finally, the GRG $\Gamma^{(\epsilon)}_j$, which is the weighted average of the grey relational coefficients, is estimated as $\Gamma^{(\epsilon)}_j = \sum_{i=1}^{2} w_i \cdot \gamma^{i(\epsilon)}_j, \forall j$, where w_i is the weight of the criterion i and can be more prone to subjective modifications by a DM. Nevertheless, it is possible to delineate it with the use of an objective method [19]. Besides, $\sum_{i=1}^{2} w_i = 1$. We should emphasise that GRG only ranks the alternatives; thus, it is not an efficiency measure. The DMU with the highest GRG is placed first.

To conclude, GRA is considered as an efficient ranking tool not only for traditional MCDM problems [25], but also for efficiency evaluation DEA problems as a MCDM context in disguise [41]. Nevertheless, GRA has, to our knowledge, not yet received explicit attention on ranking interval values and, in particular, interval ultimate cross-efficiencies within an interval CE matrix. Hence, this section has aspired to attain this target, in the light of a meaningful prioritisation of the DMUs.

4. NUMERICAL APPLICATION

This section illustrates the use of the mathematical models presented in Section 3 to meaningfully evaluate and rank the DMUs. There are two salient factors that evaluate each DMU within the two-stage tandem structure herein: (i) the optimistic and pessimistic efficiency scores within a self-evaluation context, and (ii) the most favourable and unfavourable weight sets of each of the other DMUs, in a peer-appraisal setting that integrates the combined self-efficiency measure.

The numerical example drawn from Zhou *et al.* [55] is used for illustrative purposes. In Table 5, ten bank branches of China Construction Bank in Anhui are assessed within the two-stage tandem structure (see Fig. 2). The employee (X_1), the fixed assets (X_2), and the expenses (X_3) are the input resources of the first stage to be consumed to produce the intermediate products; the credit (Z_1) and the inter-bank loans (Z_2). The latter are used as inputs in the second stage to generate the final outputs; the loan (Y_1) and the profit (Y_2). For modelling, running, and analysing our data, we have utilised the programming language Python 3.7.6 and in particular the version 2.1 of PuLP as the free linear programming library. The experiment ran on a computer with 16GB RAM.

In our framework, we first consider determining the best and worst relative efficiencies of the IDMU and the ADMU, respectively, for the overall system and its individual stages. Table 6 exhibits the corresponding scores from solving models (3.3)–(3.6), introduced in Section 3.1.

Then, models (3.7)–(3.10) are used to obtain the highest and the lowest relative efficiency scores of the target DMU_k in terms of the overall system, the stage 1, and the stage 2. These scores are given in Table 7. Recall that these relative self-efficiency scores indicate their distance from the respective IDMU and ADMU efficiencies,

TABLE 6. Highest and lowest relative efficiency scores for the overall system, stage 1, and stage 2.

$E^{\mathrm{IDMU}(1)}$	2.41405
$E^{\mathrm{ADMU}(1)}$	0.05162
$E^{\mathrm{IDMU}(2)}$	10.92813
$E^{\mathrm{ADMU}(2)}$	0.00550
$E^{\mathrm{IDMU(s)}}$	2.41405
$E^{\mathrm{ADMU(s)}}$	0.00469

presented in Table 6. These scores are also accompanied by the combined self-efficiency ratings for each DMU and system/stage. The numbers in parentheses illustrate the rankings of the corresponding bank branches for each type of efficiency measure.

In Table 7, no matter what point of view efficiency is measured from, DMU 1 is certainly the best unit and DMU 10 is the worst unit, in terms of the entire system (second expanded column). Considering stage 1, regardless of the viewpoint, DMU 1 and DMU 3 are the most and least desirable units, respectively (third expanded column). In stage 2 (fourth expanded column), DMU 10 is deemed as the least promising unit. However, there is no correspondence between the optimistic and pessimistic perspectives regarding the best unit. Notably, none of the 10 bank branches perform efficiently in both stages and viewpoints. This is, for instance, seen in the non-efficient overall optimistic self-efficiency scores $\left(E_k^{\mathrm{IDMU}(s)}\right)$, where the highest score is 0.8132 occurring at DMU 1, followed by 0.3490 occurring at DMU 6.

The next focal point of the framework is the geometric aggregation of the optimistic and pessimistic perspectives, to build a combined self-efficiency measure for each DMU, with respect to the system $\left(E_k^{\mathrm{comb}(s)}\right)$, the stage 1 $\left(E_k^{\mathrm{comb}(1)}\right)$, and the stage 2 $\left(E_k^{\mathrm{comb}(2)}\right)$. In Table 7, DMU 1 has the best performance among all units, reflecting the two opposed standpoints. This is completely verified by the consistent results of the overall system and the stage 1. Nevertheless, regarding stage 2, there is a significant inconsistency between the optimistic and the pessimistic efficiency. In detail, DMU 1 receives a moderate rating (0.8132) with respect to the optimistic aspect. This rating is compensated by its exceptional pessimistic performance (0.0760). The overall performance of bank branch 1 is also grievously higher than the corresponding performance of all others. For instance, in stage 1 the combined self-efficiency score of DMU 1 approximates 0.51, whereas the corresponding rating of DMU 2 (in the second place) equals to 0.2733. The geometric average efficiency also indicates that DMU 10 has the worst performance in terms of the overall system and the stage 2.

The combined self-efficiencies calculated for every DMU, satisfy the sound mathematical property that the overall system combined self-efficiency score is the product of the two sub-stages, as discussed in Section 3.1. As an example, the combined self-efficiencies calculated for DMU 1 satisfy $0.1267 = 0.5099 * 0.2486$. Since this property is satisfied, every $E_k^{\mathrm{comb}(s)}$ is no greater than its corresponding $E_k^{\mathrm{comb}(1)}$ and $E_k^{\mathrm{comb}(2)}$. Another point to be noted is that most bank branches have a smaller $E_k^{\mathrm{comb}(2)}$ than $E_k^{\mathrm{comb}(1)}$. Only DMUs 3, 8, and 9 have a smaller $E_k^{\mathrm{comb}(1)}$ than $E_k^{\mathrm{comb}(2)}$. However, after implementing the Wilcoxon's matched-pairs signed-ranks test [9] we found that there is not sufficient evidence to say that the average efficiency measures of these two sub-stages are not equal. This may be due to the limited sample under examination. In addition, it is noteworthy that the difference between ratings and ranks of the combined self-efficiency measures in all stages is quite significant for several bank branches. For instance, the rank of DMU 3 for the overall system, the stage 1, and the stage 2, is 8, 10, and 2, respectively, indicating that at least 6 ranks difference exists. A large difference may reveal the source that causes the inefficiency of the overall system. For example, DMUs 2 and 5 perform satisfactorily in stage 1 (as compared to stage 2) whereas DMUs 3 and 8 perform satisfactorily in stage 2 (as compared to stage 1). Decomposing the overall system combined self-efficiency score into the product of its two sub-stages, may assist the respective bank branch in identifying the sub-stage that triggers inefficiency.

The combined self-efficiency measures obtained with our proposed method (see the respective columns of Tab. 7) are also compared with the respective scores (Tab. 8) obtained with Kao and Hwang's [21] approach. As mentioned in Section 3, the latter approach aims to explore the efficiency decomposition in a two-stage production process by taking into consideration the series relationship of the two sub-stages. Their relational model (see model (3.1)) was found to be reliable in terms of measuring overall and division efficiencies along with the better identification of the causes of inefficiency. Our study has applied their relational model to further analyse and validate the dataset provided in Table 5, by measuring the efficiencies of the whole process and its constituent sub-stages for the ten DMUs. In Table 8, the self-efficiency scores along with their ranks of the overall system, the stage 1, and the stage 2, are depicted in the second, third, and fourth column, respectively. The rankings of the two models with respect to the overall system are quite similar, showing that the largest difference is 1 occurring at the bank branches 2, 3, and 8. The rankings of the two models with respect to sub-stage 1 are also quite close to each other. In the latter case, the largest difference occurs at DMU 7 with a rank difference of 4. The second largest difference occurs at DMUs 9 and 10 with a rank difference of 2. For the remaining 7 bank branches, their rank differences are less than 2. The rank differences look very similar even with the case of sub-stage 2. Correlation analysis suggests that there is a highly strong association between the ranks of these two approaches, as indicated by the Spearman coefficients [9] of 0.985 (overall system), 0.806 (stage 1), and 0.841 (stage 2), which are significant at the 0.01 level (two-tailed). This can be demonstrated even by the fact that both our method and Kao and Hwang's method identify DMU 1 as the best performer. However, our approach is more informative within the self-appraisal context, in that it not only considers the most favourable self-efficiency scores (as in [21]), but also the most unfavourable ones to obtain a more accurate and less misleading overall assessment for each DMU and flow. As a result, it puts emphasis on both sides of the same coin simultaneously. The above points further validate the rationale of our approach.

As discussed in Section 1, the geometric average efficiency is an easy-to-use measure with a good discriminating power amongst the evaluated DMUs. However, it may not be sufficiently strong in terms of leading to a unique ranking in this two-stage process. As a matter of fact, there is no absolute discrimination of some inefficient DMUs considering the combined self-evaluation results at each stage, presented in Table 7. In particular, in the overall system the DMUs 2 and 6 tied (0.0528) in the second place. Similarly, at stage 2 the DMUs 3 and 6 also tied (0.2094), sharing the second place. In such results, each DMU is self-assessed ignoring the weight profile of each of the other DMUs. Embedding the geometric average score into a peer context, would possibly contribute to a more comprehensive ranking. To this end, the proposed framework is further extended by the use of the interval CE.

The next step in our proposed approach concerns the implementation of the interval CE towards the evaluated network structure, as discussed in Section 3.2. Tables 9–11 showcase the interval individual cross-efficiencies of DMU_j based on the optimal weight scheme of DMU_k for the overall system, the stage 1, and the stage 2, respectively. In this case, each DMU is evaluated considering simultaneously an aggressive (model (3.11)) and a benevolent (model (3.12)) strategy; this originally creates an atmosphere of neutrality.

To make the content of Tables 9–11 comprehensible to the reader, it should be ideal to present a few examples. In the second column of Table 9, DMU 1 is assessed based on the weight profile of all other DMUs, except its own weight set. The minimum and maximum individual cross-efficiencies of DMU 1 based on the optimal weight scheme of DMU 2 are 0.1216 and 0.2371, respectively, for the overall system. In the fifth column of Table 10, DMU 4 is also peer-appraised with respect to the weight profile of all other DMUs. The minimum and maximum individual cross-efficiencies of DMU 4 based on the weight profile of DMU 10 are 0.1475 and 0.2281, respectively, for sub-stage 1. Table 11 determines in a similar manner the individual cross-efficiencies for each DMU, for the sub-stage 2. The diagonal leading column in each of these three matrices demonstrates the special case in which $E_{jj}^{\mathrm{L}(\epsilon)*} = E_{jj}^{\mathrm{U}(\epsilon)*} = E_{j}^{\mathrm{comb}(\epsilon)*}\ \forall j$, where $\epsilon = s$ (overall system), 1 (stage 1) or 2 (stage 2). These are the combined self-efficiency scores. Clearly, the property of maintaining the combined self-efficiency measure for each DMU is satisfied both for the overall system and its individual stages; this accomplishes our efforts towards a more reasoned peer-appraisal setting that entails the effects of the optimistic and pessimistic viewpoints.

TABLE 7. Self-efficiency ratings and ranks of the overall system, the stage 1, and the stage 2, with the proposed method.

DMU	$E_k^{\text{IDMU(s)}}$	$E_k^{\text{ADMU(s)}}$	$E_k^{\text{comb(s)}}$	$E_k^{\text{IDMU(1)}}$	$E_k^{\text{ADMU(1)}}$	$E_k^{\text{comb(1)}}$	$E_k^{\text{IDMU(2)}}$	$E_k^{\text{ADMU(2)}}$	$E_k^{\text{comb(2)}}$
1	0.8132 (1)	0.0197 (1)	0.1267 (1)	1.0000 (1)	0.2600 (1)	0.5099 (1)	0.8132 (6)	0.0760 (1)	0.2486 (1)
2	0.3255 (3)	0.0086 (2)	**0.0528** (2)	0.5186 (2)	0.1441 (5)	0.2733 (2)	0.6277 (8)	0.0595 (2)	0.1933 (6)
3	0.1398 (9)	0.0058 (6)	0.0284 (8)	0.1421 (10)	0.1298 (10)	0.1358 (10)	0.9838 (2)	0.0446 (5)	**0.2094** (2)
4	0.2450 (5)	0.0038 (8)	0.0307 (6)	0.2655 (5)	0.1731 (3)	0.2144 (5)	0.9227 (3)	0.0222 (8)	0.1432 (8)
5	0.2886 (4)	0.0062 (4)	0.0423 (4)	0.3927 (4)	0.1818 (2)	0.2672 (3)	0.7350 (7)	0.0341 (7)	0.1584 (7)
6	0.3490 (2)	0.0080 (3)	**0.0528** (2)	0.4101 (3)	0.1549 (4)	0.2521 (4)	0.8509 (5)	0.0515 (3)	**0.2094** (2)
7	0.1454 (8)	0.0030 (9)	0.0208 (9)	0.2549 (6)	0.1425 (7)	0.1906 (6)	0.5706 (9)	0.0208 (9)	0.1090 (9)
8	0.1476 (7)	0.0055 (7)	0.0285 (7)	0.1476 (9)	0.1372 (8)	0.1423 (9)	1.0000 (1)	0.0402 (6)	0.2005 (5)
9	0.2141 (6)	0.0061 (5)	0.0363 (5)	0.2389 (7)	0.1362 (9)	0.1804 (7)	0.8963 (4)	0.0451 (4)	0.2011 (4)
10	0.0133 (10)	0.0029 (10)	0.0062 (10)	0.1796 (8)	0.1438 (6)	0.1607 (8)	0.0739 (10)	0.0204 (10)	0.0389 (10)

TABLE 8. Self-efficiency ratings and ranks of the overall system, the stage 1, and the stage 2, with Kao and Hwang's [21] method.

DMU	E_k^s (Rank)	E_k^1 (Rank)	E_k^2 (Rank)
1	1.00 (1)	1.00 (1)	1.00 (1)
2	0.43 (3)	0.55 (2)	0.79 (8)
3	0.29 (7)	0.29 (9)	1.00 (1)
4	0.30 (6)	0.32 (7)	0.94 (4)
5	0.35 (4)	0.43 (4)	0.83 (7)
6	0.54 (2)	0.54 (3)	1.00 (1)
7	0.18 (9)	0.28 (10)	0.62 (9)
8	0.28 (8)	0.31 (8)	0.92 (5)
9	0.33 (5)	0.38 (5)	0.85 (6)
10	0.17 (10)	0.37 (6)	0.47 (10)

TABLE 9. Interval cross-efficiencies for the overall system.

DMU	1	2	3	4	5	6	7	8	9	10
1	0.1267 0.1267	0.0282 0.0550	0.0207 0.0371	0.0127 0.0381	0.0294 0.0449	0.0370 0.0690	0.0121 0.0226	0.0198 0.0357	0.0215 0.0415	0.0011 0.0221
2	0.1216 0.2371	0.0528 0.0528	0.0225 0.0613	0.0234 0.0397	0.0382 0.0597	0.0489 0.0785	0.0182 0.0269	0.0223 0.0563	0.0347 0.0438	0.0017 0.0234
3	0.0970 0.1738	0.0244 0.0667	0.0284 0.0284	0.0108 0.0498	0.0250 0.0587	0.0314 0.0854	0.0103 0.0295	0.0261 0.0300	0.0203 0.0463	0.0016 0.0178
4	0.1018 0.3039	0.0407 0.0691	0.0175 0.0802	0.0307 0.0307	0.0354 0.0707	0.0437 0.0888	0.0182 0.0292	0.0184 0.0737	0.0268 0.0573	0.0014 0.0307
5	0.1192 0.1819	0.0374 0.0585	0.0205 0.0480	0.0183 0.0366	0.0423 0.0423	0.0489 0.0677	0.0174 0.0220	0.0216 0.0441	0.0285 0.0442	0.0015 0.0225
6	0.0966 0.1801	0.0354 0.0568	0.0175 0.0475	0.0181 0.0369	0.0329 0.0455	0.0528 0.0528	0.0164 0.0219	0.0173 0.0437	0.0269 0.0407	0.0013 0.0243
7	0.1161 0.2161	0.0407 0.0600	0.0199 0.0571	0.0218 0.0349	0.0397 0.0502	0.0498 0.0664	0.0208 0.0208	0.0210 0.0524	0.0305 0.0453	0.0016 0.0240
8	0.1012 0.1825	0.0267 0.0675	0.0270 0.0310	0.0118 0.0473	0.0273 0.0557	0.0343 0.0864	0.0112 0.0282	0.0285 0.0285	0.0221 0.0468	0.0017 0.0176
9	0.1105 0.2136	0.0437 0.0551	0.0222 0.0507	0.0194 0.0415	0.0347 0.0538	0.0469 0.0708	0.0166 0.0246	0.0220 0.0466	0.0363 0.0363	0.0017 0.0216
10	0.0357 0.6610	0.014 0.1913	0.0099 0.1081	0.0062 0.1326	0.0117 0.1666	0.0135 0.2448	0.0053 0.0798	0.0100 0.1033	0.0104 0.1327	0.0062 0.0062
ultim. CE	0.1000 0.2611	0.0323 0.0755	0.0197 0.0579	0.0158 0.0508	0.0305 0.0673	0.0394 0.0953	0.0140 0.0316	0.0198 0.0540	0.0246 0.0554	0.0015 0.0227

Recalling the discussion in Section 3.3, we view each interval CE matrix as a MCDM problem. In Tables 9–11, the ten DMUs (alternatives) located in their last 11 columns, are evaluated by the weight profiles of the ten DMUs (criteria) presented in their first column. To designate the interval global weights (interval ultimate cross-efficiencies) in the last row of each of these matrices, it is required to determine the interval weight per criterion except the known interval individual cross-efficiencies. To start with, the interval weight of each criterion is determined in the second, third, and fourth column of Table 11, with respect to the overall system, stage 1,

Table 10. Interval cross-efficiencies for the stage 1.

DMU	1	2	3	4	5	6	7	8	9	10
1	0.5099 0.5099	0.1472 0.2811	0.0689 0.0764	0.0615 0.1353	0.1422 0.2002	0.1615 0.2654	0.0826 0.1299	0.0648 0.0752	0.0785 0.1366	0.0530 0.0915
2	0.4956 0.9466	0.2733 0.2733	0.0743 0.1279	0.1143 0.1399	0.1867 0.2640	0.2161 0.2998	0.1255 0.1533	0.0724 0.1203	0.1258 0.1458	0.0757 0.0984
3	0.9056 0.9999	0.2887 0.4994	0.1358 0.1358	0.1211 0.2537	0.2793 0.3751	0.3170 0.4715	0.1623 0.2435	0.1277 0.1410	0.1543 0.2426	0.1044 0.1716
4	0.8073 1.0000	0.4186 0.2893	0.1147 0.2665	0.2144 0.2144	0.3103 0.3785	0.3311 0.3713	0.2057 0.2407	0.1191 0.2653	0.1928 0.2223	0.1259 0.2174
5	0.6804 0.9578	0.2765 0.3911	0.0967 0.1295	0.1156 0.1845	0.2672 0.2672	0.2790 0.3692	0.1552 0.1796	0.1004 0.1217	0.1475 0.1900	0.0995 0.1222
6	0.4923 0.8093	0.2336 0.3241	0.0738 0.1094	0.0977 0.1659	0.1855 0.2454	0.2521 0.2521	0.1247 0.1593	0.0720 0.1028	0.1246 0.1493	0.0752 0.1122
7	0.7478 1.0000	0.2889 0.4148	0.1062 0.1828	0.1547 0.1985	0.2834 0.3151	0.3067 0.3917	0.1906 0.1906	0.1100 0.1773	0.1786 0.2015	0.1150 0.1343
8	0.9640 0.9999	0.2887 0.5365	0.1370 0.1496	0.1310 0.2559	0.2898 0.3785	0.3227 0.5066	0.1706 0.2465	0.1423 0.1423	0.1615 0.2606	0.1164 0.1731
9	0.6732 1.0000	0.2889 0.3915	0.1009 0.1858	0.1568 0.2005	0.2536 0.3173	0.3097 0.3378	0.1705 0.1924	0.0984 0.1803	0.1804 0.1804	0.1029 0.1356
10	0.2473 0.6610	0.1239 0.2484	0.1170 0.1958	0.1475 0.2281	0.1802 0.2651	0.1433 0.2834	0.1364 0.1816	0.1317 0.1961	0.1218 0.1858	0.1607 0.1607
ultim. CE	0.6083 0.7509	0.2172 0.3407	0.0892 0.1257	0.1018 0.1754	0.2013 0.2661	0.2260 0.3277	0.1245 0.1705	0.0880 0.1230	0.1195 0.1731	0.0808 0.1240

Table 11. Interval cross-efficiencies for the stage 2.

DMU	1	2	3	4	5	6	7	8	9	10
1	0.2486 0.2486	0.1918 0.1957	0.3007 0.4856	0.2078 0.2820	0.2072 0.2246	0.2294 0.2600	0.1473 0.1743	0.3056 0.4745	0.2739 0.3044	0.0225 0.2417
2	0.2454 0.2504	0.1933 0.1933	0.3030 0.4795	0.2052 0.2842	0.2046 0.2263	0.2265 0.2620	0.1455 0.1757	0.3080 0.4686	0.2760 0.3006	0.0227 0.2387
3	0.1072 0.1738	0.0848 0.1336	0.2094 0.2094	0.0897 0.1964	0.0896 0.1564	0.0993 0.1811	0.0636 0.1214	0.2045 0.2128	0.1316 0.1908	0.0157 0.1042
4	0.1261 0.3039	0.0973 0.2389	0.1526 0.3012	0.1432 0.1432	0.1140 0.1867	0.1320 0.2393	0.0885 0.1213	0.1551 0.2779	0.1390 0.2580	0.0114 0.1413
5	0.1752 0.1899	0.1352 0.1495	0.2119 0.3710	0.1588 0.1988	0.1584 0.1584	0.1753 0.1833	0.1126 0.1229	0.2154 0.3626	0.1931 0.2326	0.0159 0.1847
6	0.1963 0.2225	0.1515 0.1753	0.2375 0.4348	0.1861 0.2228	0.1775 0.1856	0.2094 0.2094	0.1319 0.1377	0.2414 0.4249	0.2164 0.2726	0.0178 0.2165
7	0.1553 0.2161	0.1409 0.1447	0.1878 0.3122	0.1410 0.1762	0.1403 0.1595	0.1625 0.1696	0.1090 0.1090	0.1909 0.2959	0.1711 0.2251	0.0141 0.1787
8	0.1050 0.1825	0.0925 0.1258	0.1972 0.2075	0.0905 0.185	0.0943 0.1473	0.1065 0.1706	0.0661 0.1143	0.2005 0.2005	0.1373 0.1797	0.0148 0.1021
9	0.1641 0.2136	0.1512 0.1408	0.2207 0.2732	0.1237 0.2070	0.1368 0.1697	0.1515 0.2095	0.0973 0.1280	0.2243 0.2585	0.2011 0.2011	0.0165 0.1596
10	0.1444 1.0000	0.1132 0.7699	0.0847 0.5521	0.0422 0.5815	0.0649 0.6285	0.0943 0.8637	0.0395 0.4397	0.0764 0.5268	0.0858 0.7140	0.0389 0.0389
ultim. CE	0.1606 0.2970	0.1310 0.2247	0.2055 0.3782	0.1354 0.2522	0.1349 0.2235	0.1540 0.2722	0.0977 0.1644	0.2071 0.3656	0.1770 0.2904	0.0168 0.1773

and stage 2, respectively. The interval weights are obtained *via* the GP model (3.13), and the interval global weights, according to the pair of optimisation models (3.14) and (3.15), as stated in Section 3.3.

For instance, in the second column of the last row of Table 9, we obtain the interval ultimate CE of DMU 1: [0.1000, 0.2611], where 0.1000 is the minimum and 0.2611 is the maximum CE score. The minimum score of DMU 1 for the overall system is estimated *via* solving model (3.14). The basic prerequisites of this model are to recognise the minimum individual cross-efficiencies of DMU 1 based on the weight profile of all ten DMUs (left side of column 2 of Tab. 9) and the interval weights per criterion for the overall system (column 2 of Tab. 12). The maximum ultimate cross-efficiency of DMU 1 for the overall system is estimated *via* solving model (3.17). The basic prerequisites of this model are to identify the maximum individual cross-efficiencies of DMU 1 based on the weight profile of all 10 DMUs (right side of column 2 of Tab. 9) and the interval weights per criterion for the overall system (column 2 of Tab. 12).

The final step of our methodological approach seeks for a unique and reasonable prioritisation of the interval ultimate cross-efficiencies *via* the established GRA, as discussed in Section 3.4. This step continues to allow the DMUs, located in the columns of the interval CE matrices mentioned above, to act as alternatives and to be assessed by two attributes; the first attribute concerns the minimum (worst condition) and the second attribute is pertinent to the maximum (best condition) ultimate CE of each DMU, towards the corresponding system/stage. The interval ultimate cross-efficiencies (last row of each of the Tabs. 9–11) form the appropriate matrix, as shown in Table 4.

The data of performance values of the two attributes are subsequently normalised through the greater-the-better equation (3.16); this choice reflects the necessity of pushing up the peer-efficiency of each DMU. The results are depicted in the second column of Appendix A for the overall system, of Appendix B for stage 1, and of Appendix C for stage 2. The grey relational distance calculation is also utilised to measure the distance between the reference sequence and the comparability sequence (normalised values), see the third column of each of the appendices. In addition, we compute the grey relational coefficient to explore how close the reference and the comparability sequences are. In this formula, the value of ζ may affect the size of the correlation degree distribution interval, thereby affecting the results of the correlation analysis. The value of ζ can be determined considering the DMU's tendency towards optimism-pessimism. Following [21], we have set $\zeta = 0.5$ implying

Table 12. Interval weights per criterion for the overall system, the stage 1, and the stage 2.

Criteria (DMUs)	Interval weights per criterion (overall system)		Interval weights per criterion (stage 1)		Interval weights per criterion (stage 2)	
1	0.0000	0.1111	0.4502	0.4726	0.0134	0.1198
2	0.0000	0.1111	0.0286	0.0510	0.0133	0.1197
3	0.0000	0.1111	0.0506	0.0730	0.0016	0.1080
4	0.0000	0.1111	0.0479	0.0703	0.0023	0.1088
5	0.0000	0.1111	0.0390	0.0613	0.0068	0.1132
6	0.0000	0.1111	0.0283	0.0507	0.0091	0.1156
7	0.0000	0.1111	0.0433	0.0657	0.0052	0.1117
8	0.0000	0.1111	0.0538	0.0761	0.0017	0.1081
9	0.0000	0.1111	0.0394	0.0618	0.0060	0.1125
10	0.0000	0.1111	0.0170	0.0394	0.0000	0.0887

that the DMU has neither an optimistic nor a conservative attitude. The respective results are portrayed in the last column of each of the appendices.

The GRG and the rank for each DMU with respect to the overall system, the stage 1, and the stage 2, are illustrated in the second, third, and fourth column of Table 13, respectively. It is important to make two remarks about the process of obtaining the GRG: firstly, the relative importance weights of the two performance attributes were assumed to be equal ($w_1 = w_2 = 0.5$) illustrating that the two extremes are of the same importance, and secondly, the GRG is just an index that only captures the rank rather than an efficiency measure. The unique final rank in Table 13 reflects the improvement of the discriminating power, as compared to the original rank derived from the combined self-efficiency measures in Table 7. This practically means that the non-dominated bank branches, which cannot be fully discriminated by the self-evaluation notion, can be discriminated by the methodologies followed in peer notion. In detail, DMU 10 is without a doubt the least desirable unit in all three cases. DMU 1 is also considered to be the most promising bank branch for the overall system and stage 1, while DMU 3 is the best unit according to stage 2. Generally, one can deduce that the ranking results for all branches (except DMU 10) are not consistent and may show a higher degree of uncertainty and inefficiency in specific stages.

The GRG grades obtained with our proposed framework (see Tab. 13) are also compared with the respective ultimate cross-efficiency ratings (Tab. 14) obtained *via* the Kao and Liu's [21] approach. In their study, they applied the concept of cross-evaluation to measure the efficiency of basic (parallel & series) network structures. Their proposed aggressive-based secondary goal model was particularly able to decompose the cross-efficiency score of the overall system into the product of those of the internal sub-stages for the series structure. Our study has applied their aggressive-based model under the two-stage tandem series structure and the peer-appraisal setting to further analyse the dataset provided in Table 5. In Table 14, the peer-efficiency scores along with their ranks of the overall system, the stage 1, and the stage 2, are respectively presented in the second, third, and fourth column. Firstly, we have noticed that the multiplicative mathematical relationship between the overall system and its sub-stage efficiencies is indeed satisfied. For example, the ultimate cross-efficiency score of DMU 6 (0.446) is equal to the product of its sub-stage 1 (0.574) and sub-stage 2 (0.778) efficiencies. Secondly, the rankings of the two methods with respect to the overall system and the stage 1 are not significantly different based on a Spearman rank order correlation test with statistics of 0.964 and 0.830, respectively. These are significant at the 0.01 level (two-tailed). However, it is worthwhile to mention that DMU 10 has a difference of 3 ranks in terms of the evaluation of stage 1. Thirdly, as for the stage 2, the rankings from the two methods are not so close. The bank branch 2 is the extreme case with a rank difference of 6. The second largest difference occurs at DMU 8 with a rank difference of 4. All the remaining bank branches have a rank difference of no

TABLE 13. Grey Relational Grade and ranks of the overall system, the stage 1, and the stage 2.

DMU	GRG $\Gamma_j^{(s)}$	Rank overall system	GRG $\Gamma_j^{(1)}$	Rank stage 1	GRG $\Gamma_j^{(2)}$	Rank stage 2
1	1.0000	**1**	1.0000	**1**	0.6199	4
2	0.4062	3	0.4181	2	0.4830	8
3	0.3750	6	0.3356	8	0.9916	**1**
4	0.3655	8	0.3477	7	0.5145	6
5	0.3978	4	0.3931	4	0.4886	7
6	0.4333	2	0.4171	3	0.5719	5
7	0.3530	9	0.3519	5	0.3992	9
8	0.3729	7	0.3349	9	0.9471	2
9	0.3811	5	0.3512	6	0.6543	3
10	0.3333	10	0.3335	10	0.3403	10

TABLE 14. Peer-efficiency ratings and ranks of the overall system, the stage 1, and the stage 2, with Kao and Liu's [22] method.

DMU	System CE (Rank)	Stage 1 CE (Rank)	Stage 2 CE (Rank)
1	1.000 (1)	1.000 (1)	1.000 (1)
2	0.416 (3)	0.534 (4)	0.780 (2)
3	0.239 (7)	0.315 (10)	0.760 (4)
4	0.251 (6)	0.488 (5)	0.513 (8)
5	0.330 (4)	0.573 (3)	0.576 (7)
6	0.446 (2)	0.574 (2)	0.778 (3)
7	0.160 (9)	0.420 (6)	0.381 (9)
8	0.238 (8)	0.336 (9)	0.710 (6)
9	0.297 (5)	0.392 (8)	0.756 (5)
10	0.072 (10)	0.397 (7)	0.182 (10)

more than 3. Statistically, this situation is even further validated by the Spearman coefficient of 0.503, which implies a moderate association between the rankings of the two methods. Finally, Kao and Liu's [22] approach only considers the most unfavourable weight sets of each of the other DMUs, while keeping the optimistic self-efficiency score constant. However, our study is more multi-dimensional since it simultaneously takes into account the most favourable and unfavourable weight sets of each of the other players, while integrating the respective combined self-efficiency measure.

Finally, it can be statistically inferred that the rankings of the DMUs obtained from the combined self-efficiency measures (self-appraisal), and the grey relational grades after showing peer-appraisal, are similar with respect to the overall system and its sub-stages. As an example, for the overall system, according to the Spearman correlation test [9], the $r_s = 0.948$. This indicates that under the significance level of 0.01, there is a strong positive association between the ranking values of the DMUs obtained by the two separate conditions (self-appraisal & peer-appraisal), confirming the validity of our framework. Exceptions are considered the DMUs 1, 6, and 8 within the evaluation of the second sub-stage, where there is a larger rank difference of 3. This could be justified by the nature of the self-appraisal setting to let each bank branch to be evaluated based only on its own (favourable and unfavourable) standpoint, while the peer-appraisal setting expects the bank branches to be evaluated from the (favourable and unfavourable) standpoint of all branches.

5. Conclusions & future research

This paper has provided new insight into the attainment of a meaningful and unique ranking of DMUs under a two-stage tandem (network) structure. In particular, it extends the selected optimistic-pessimistic DEA models into the two-stage tandem system, to then complement the interval CE method within such a system. Decision makers are offered with the chance of evaluating the performance of the DMUs by considering: (i) the optimistic and pessimistic self-efficiency scores, and (ii) the most favourable and unfavourable weight profiles of each of the other DMUs in a peer-appraisal setting. In this study, we have introduced a 7-step methodological approach, as shown in Figure 1, which combines existing methods from the literature in a novel way. This approach supports the aforementioned conditions and ensures more multi-dimensional evaluation outcomes.

The procedures implemented in the first three steps of our framework indicate how the optimistic and pessimistic DEA models, which are inspired by the studies of Wang and Luo [48] and Wu [51], are built towards the more realistic two-stage tandem system that better reflects the complex interconnections among its internal sub-systems. The DMUs are initially evaluated, based on their own most favourable (optimistic) and unfavourable (pessimistic) optimal multipliers, and then are aggregated into a combined self-efficiency measure *via* the geometric average.

The remaining steps of our framework ensure the peer-evaluation of the DMUs *via* the customisation of the interval CE method to the specifications of the two-stage tandem structure while keeping the combined self-efficiency measure unchanged. To rank all DMUs in the interval CE matrix of the corresponding flow, the study introduces an alternative novel use of the GP method of Wang and Elhag [47], the LP models by Entani and Tanaka [13], and the GRA of Kuo *et al.* [25]. The combination of such well-established techniques for extracting valuable insights from an interval CE matrix has not been considered before. This combination underpins the wider MCDM context to which the elements of the interval CE matrix belong.

We envisage that our study could be applicable in several areas. In the non-life insurance industry [21], for example, operations consist of the insurance service and the capital investment. Customers pay direct written and reinsurance premiums, which are then invested in a portfolio to earn underwriting profit. Another promising area would be the evaluation of the performance of the high-technology industry that is decomposed into the technology development and the economic application [54]. In this two-stage tandem network, raw data and knowledge are processed into technological achievements, which are then transformed into economic development. A third application connects our study's methodological framework with the operational activities of the international shipping industry; these could be divided into the supervision of the ship dispatching management and the control of the working time in the port [15]. Finally, the efficiency evaluation of two-stage (food) supply chains of different factories or farming communities [24] could also serve the goals of our paper. For instance, the process of the refinement of selected cocoa beans into milk/dark chocolate and the production of black tea through withering, fermentation, drying, and sieving across a number of specialised factory branches could further highlight the importance of our evaluation and ranking framework.

This paper treats the two sub-stages of a DMU equally. In reality, however, there might be a certain degree of leader-follower relationship between the upstream and downstream of a particular DMU. We acknowledge this as a limitation of our study and we believe that the introduction of relative weights for the different stages when calculating overall efficiency could accommodate such an issue. In addition, one of the main steps of the grey relational analysis methodology, used to rank the interval ultimate cross-efficiencies within an interval cross-efficiency matrix, is the calculation of the grey relational grade. It is defined as the weighted average of the grey relational coefficients, where the weight of the respective criterion is subjectively determined by the decision maker. To better reflect the reality, we would have taken advantage of an existing powerful multi-criteria decision-making method, such as the analytic network process [39] or the best-worst method [38], to identify in an objective manner the weights. We have also recognised that the grey relational grade is just an index that can only capture the rank rather than an efficiency measure. In other words, there is no sufficient information that would allow the identification of the DEA-efficient DMUs that constitute the best-practice frontier. However, we acknowledge that the GRA technique has not received attention on ranking interval cross-efficiencies within an interval CE matrix and, thus, our paper has worked towards this direction. Finally, further study could focus on the testing of the proposed models and frameworks with empirical data. In the shipping industry, for example, it could be deployed to compare the efficiency of potential designs of a particular type of vessel, including the selection of the right mixture of maintenance policies.

The models in this study were developed under the assumption of the constant returns-to-scale. A direction for future research could be their advancement to variable returns-to-scale DEA models. In addition, current research studies the evaluation of the performance of several DMUs with a two-stage tandem structure in a self and in a peer-appraisal setting, only when the data (*i.e.*, the input and output factors) are accurate and unambiguous, and the DEA models are based on this condition. Future research could relax this assumption by allowing the data points to be imprecise (*e.g.*, to be expressed as linguistic terms) and lie in an interval. Other cases to be investigated concern missing data or intervals, where some values are more likely to occur over other values. In the latter case, since there is no information of the probability distributions, fuzzy numbers and mathematical operations [58] could be used as an ideal alternative option. For example, there is a growing body of literature [11,17,36,40] surrounding the development of novel fuzzy DEA approaches and models characterised by intuitionistic fuzzy data, applied possibility, necessity, credibility, general fuzzy measures, and/or trapezoidal fuzzy numbers. Some of these models were solved with the aid of either a linear programming with an intuitionistic fuzzy objective function and an alphabetical technique, a chance-constrained programming, a lexicographic multi-objective linear programming, or a fuzzy linear programming. The network double-frontier DEA models introduced in this study could be adjusted to the specifications of such an uncertain (fuzzy) environment adopting the most suitable formulation and solution techniques.

Finally, it would be worthwhile to adjust the modelling approaches, introduced in our study, ensuring that they will be taking into consideration the decision maker's preferences. Relevant literature has already focused on this aspect by combining DEA and multiple-objective linear programming [12,30,31,44].

Appendix A. Overall system and grey relational analysis

Table A.1. Normalisation of data, calculation of grey relational distance and grey relational coefficient for the overall system.

DMU	Normalisation of data		Grey relational distance		Grey relational coefficient	
	C1	C2	C1	C2	C1	C2
1	1.0000	1.0000	0.0000	0.0000	1.0000	1.0000
2	0.3132	0.2217	0.6868	0.7783	0.4213	0.3911
3	0.1851	0.1477	0.8149	0.8523	0.3803	0.3697
4	0.1456	0.1181	0.8544	0.8819	0.3692	0.3618
5	0.2941	0.1872	0.7059	0.8128	0.4146	0.3809
6	0.3847	0.3046	0.6153	0.6954	0.4483	0.4183
7	0.1267	0.0376	0.8733	0.9624	0.3641	0.3419
8	0.1861	0.1313	0.8139	0.8687	0.3806	0.3653
9	0.2349	0.1372	0.7651	0.8628	0.3952	0.3669
10	0.0000	0.0000	1.0000	1.0000	0.3333	0.3333
Reference value	1.0000	1.0000	–	–	–	–

Appendix B. Stage 1 and grey relational analysis

Table B.1. Normalisation of data, calculation of grey relational distance and grey relational coefficient for the stage 1.

DMU	Normalisation of data		Grey relational distance		Grey relational coefficient	
	C1	C2	C1	C2	C1	C2
1	1.0000	1.0000	0.0000	0.0000	1.0000	1.0000
2	0.2585	0.3467	0.7415	0.6533	0.4028	0.4335
3	0.0159	0.0043	0.9841	0.9957	0.3369	0.3343
4	0.0397	0.0834	0.9603	0.9166	0.3424	0.3529
5	0.2284	0.2278	0.7716	0.7722	0.3932	0.3930
6	0.2753	0.3259	0.7247	0.6741	0.4083	0.4259
7	0.0827	0.0756	0.9173	0.9244	0.3528	0.3510
8	0.0135	0.0000	0.9865	1.0000	0.3364	0.3333
9	0.0732	0.0797	0.9268	0.9203	0.3504	0.3520
10	0.0000	0.0015	1.0000	0.9985	0.3333	0.3337
Reference value	1.0000	1.0000	–	–	–	–

Appendix C. Stage 2 and grey relational analysis

Table C.1. Normalisation of data, calculation of grey relational distance and grey relational coefficient for the stage 2.

DMU	Normalisation of data		Grey relational distance		Grey relational coefficient	
	C1	C2	C1	C2	C1	C2
1	0.7556	0.6199	0.2444	0.3801	0.6717	0.5681
2	0.6000	0.2820	0.4000	0.7180	0.5555	0.4105
3	0.9915	1.0000	0.0085	0.0000	0.9833	1.0000
4	0.6230	0.4104	0.3770	0.5896	0.5701	0.4589
5	0.6204	0.2765	0.3796	0.7235	0.5685	0.4087
6	0.7209	0.5041	0.2791	0.4959	0.6418	0.5021
7	0.4250	0.0000	0.5750	1.0000	0.4651	0.3333
8	1.0000	0.9409	0.0000	0.0591	1.0000	0.8942
9	0.8417	0.5893	0.1583	0.4107	0.7596	0.5491
10	0.0000	0.0602	1.0000	0.9398	0.3333	0.3473
Reference value	1.0000	1.0000	–	–	–	–

Acknowledgements. The work was supported by the Engineering and Physical Sciences Research Council (Industrial CASE award) under grant 5162031105656, and BAE Systems.

References

[1] T.R. Anderson, K. Hollingsworth and L. Inman, The fixed weighting nature of a cross-evaluation model. *J. Prod. Anal.* **17** (2002) 249–255.

[2] L. Angulo-Meza and M.P.E. Lins, Review of methods for increasing discrimination in data envelopment analysis. *Ann. Oper. Res.* **116** (2002) 225–242.

[3] H. Azizi, The interval efficiency based on the optimistic and pessimistic points of view. *Appl. Math. Modell.* **35** (2011) 2384–2393.

[4] H. Azizi and R. Jahed, Improved data envelopment analysis models for evaluating interval efficiencies of decision-making units. *Comput. Ind. Eng.* **61** (2011) 897–901.

[5] T. Badiezadeh, R.F. Saen and T. Samavati, Assessing sustainability of supply chains by double frontier network DEA: a big data approach. *Comput. Oper. Res.* **98** (2018) 284–290.

[6] A.Y. Chang, Prioritising the types of manufacturing flexibility in an uncertain environment. *Int. J. Prod. Res.* **50** (2012) 2133–2149.

[7] A. Charnes, W.W. Cooper and E. Rhodes, Measuring the efficiency of decision-making units. *Eur. J. Oper. Res.* **2** (1978) 429–444.

[8] W.D. Cook, K. Tone and J. Zhu, Data envelopment analysis: prior to choosing a model. *Omega* **44** (2014) 1–4.

[9] W.W. Daniel, Applied Nonparametric Statistics. Houghton Mifflin, Boston (1978).

[10] J. Doyle and R. Green, Efficiency and cross-efficiency in DEA: derivations, meanings and uses. *J. Oper. Res. Soc.* **45** (1994) 567–578.

[11] A. Ebrahimnejad, Cost efficiency measures with trapezoidal fuzzy numbers in data envelopment analysis based on ranking functions: application in insurance organization and hospital. *Int. J. Fuzzy Syst. App. (IJFSA)* **2** (2012) 51–68.

[12] A. Ebrahimnejad and F.H. Lotfi, Equivalence relationship between the general combined-oriented CCR model and the weighted minimax MOLP formulation. *J. King Saud Univ.-Sci.* **24** (2012) 47–54.

[13] T. Entani and H. Tanaka, Interval estimations of global weights in AHP by upper approximation. *Fuzzy Sets Syst.* **158** (2007) 1913–1921.

[14] T. Entani, Y. Maeda and H. Tanaka, Dual models of interval DEA and its extension to interval data. *Eur. J. Oper. Res.* **136** (2002) 32–45.

[15] G. Gan, H.S. Lee, L. Lee, X. Wang and Q. Wang, Network hierarchical DEA with an application to international shipping industry in Taiwan. *J. Oper. Res. Soc.* **71** (2020) 991–1002.

[16] A. Hatami-Marbini, S. Saati and M. Tavana, An ideal-seeking fuzzy data envelopment analysis framework. *Appl. Soft Comput.* **10** (2010) 1062–1070.

[17] A. Hatami-Marbini, A. Ebrahimnejad and S. Lozano, Fuzzy efficiency measures in data envelopment analysis using lexicographic multiobjective approach. *Comput. Ind. Eng.* **105** (2017) 362–376.

[18] Y. Huang, Y.M. Wang and J. Lin, Two-stage fuzzy cross-efficiency aggregation model using a fuzzy information retrieval method. *Int. J. Fuzzy Syst.* **21** (2019) 2650–2666.

[19] A. Jahan, F. Mustapha, S.M. Sapuan, M.Y. Ismail and M. Bahraminasab, A framework for weighting of criteria in ranking stage of material selection process. *Int. J. Adv. Manuf. Technol.* **58** (2012) 411–420.

[20] D. Julong, Introduction to grey system theory. *J. Grey Syst.* **1** (1989) 1–24.

[21] C. Kao and S.N. Hwang, Efficiency decomposition in two-stage data envelopment analysis: an application to non-life insurance companies in Taiwan. *Eur. J. Oper. Res.* **185** (2008) 418–429.

[22] C. Kao and S.T. Liu, Cross efficiency measurement and decomposition in two basic network systems. *Omega* **83** (2019) 70–79.

[23] M. Khodabakhshi and K. Aryavash, The cross-efficiency in the optimistic–pessimistic framework. *Oper. Res.* **17** (2017) 619–632.

[24] M.D. Kremantzis, P. Beullens and J. Klein, A fairer assessment of DMUs in a generalised two-stage DEA structure. *Expert Syst. App.* **187** (2022) 115921.

[25] Y. Kuo, T. Yang and G.W. Huang, The use of grey relational analysis in solving multiple attribute decision-making problems. *Comput. Ind. Eng.* **55** (2008) 80–93.

[26] F. Li, Q. Zhu, Z. Chen and H. Xue, A balanced data envelopment analysis cross-efficiency evaluation approach. *Expert Syst. App.* **106** (2018) 154–168.

[27] F. Li, H. Wu, Q. Zhu, L. Liang and G. Kou, Data envelopment analysis cross efficiency evaluation with reciprocal behaviors. *Ann. Oper. Res.* **302** (2021) 173–210.

[28] L. Liang, J. Wu, W.D. Cook and J. Zhu, Alternative secondary goals in DEA cross-efficiency evaluation. *Int. J. Prod. Econ.* **113** (2008) 1025–1030.

[29] W. Liu and Y.M. Wang, Ranking DMUs by using the upper and lower bounds of the normalized efficiency in data envelopment analysis. *Comput. Ind. Eng.* **125** (2018) 135–143.

[30] F.H. Lotfi, G.R. Jahanshahloo, A. Ebrahimnejad, M. Soltanifar and S.M. Mansourzadeh, Target setting in the general combined-oriented CCR model using an interactive MOLP method. *J. Comput. Appl. Math.* **234** (2010) 1–9.

[31] F.H. Lotfi, G.R. Jahanshahloo, M. Soltanifar, A. Ebrahimnejad and S.M. Mansourzadeh, Relationship between MOLP and DEA based on output-orientated CCR dual model. *Expert Syst. App.* **37** (2010) 4331–4336.

[32] C. Ma, D. Liu, Z. Zhou, W. Zhao and W. Liu, Game cross efficiency for systems with two-stage structures. *J. Appl. Math.* **2014** (2014) 8.

[33] F. Meng and B. Xiong, Logical efficiency decomposition for general two-stage systems in view of cross efficiency. *Eur. J. Oper. Res.* **294** (2022) 622–632.

[34] H.H. Örkcü, V.S. Özsoy, M. Örkcü and H. Bal, A neutral cross efficiency approach for basic two stage production systems. *Expert Syst. App.* **125** (2019) 333–344.

[35] H.H. Örkcü, V.S. Özsoy, M. Örkcü and H. Bal, An optimistic-pessimistic DEA model based on game cross efficiency approach. *RAIRO: Oper. Res.* **54** (2020) 1215–1230.

[36] P. Peykani, F. Hosseinzadeh Lotfi, S.J. Sadjadi, A. Ebrahimnejad and E. Mohammadi, Fuzzy chance-constrained data envelopment analysis: a structured literature review, current trends, and future directions. *Fuzzy Optim. Decision Making* **21** (2022) 197–261

[37] S.A. Rakhshan, Efficiency ranking of decision making units in data envelopment analysis by using TOPSIS-DEA method. *J. Oper. Res. Soc.* **68** (2017) 906–918.

[38] J. Rezaei, Best-worst multi-criteria decision-making method: some properties and a linear model. *Omega* **64** (2016) 126–130.
[39] T.L. Saaty and L.G. Vargas, The analytic network process. In: Decision Making with the Analytic Network Process. Springer, Boston, MA (2013) 1–40.
[40] F.J. Santos Arteaga, A. Ebrahimnejad and A. Zabihi, A new approach for solving intuitionistic fuzzy data envelopment analysis problems. *Fuzzy Optim. Model. J.* **2** (2021) 46–56.
[41] F. Sarraf and S.H. Nejad, Improving performance evaluation based on balanced scorecard with grey relational analysis and data envelopment analysis approaches: case study in water and wastewater companies. *Eval. Program Planning* **79** (2020) 101762.
[42] T.R. Sexton, R.H. Silkman and A.J. Hogan, Data envelopment analysis: critique and extensions. *New Directions Program Eval.* **1986** (1986) 73–105.
[43] K. Sugihara, H. Ishii and H. Tanaka, Interval priorities in AHP by interval regression analysis. *Eur. J. Oper. Res.* **158** (2004) 745–754.
[44] M. Tavana, A. Ebrahimnejad, F.J. Santos-Arteaga, S.M. Mansourzadeh and R.K. Matin, A hybrid DEA-MOLP model for public school assessment and closure decision in the City of Philadelphia. *Soc. Econ. Planning Sci.* **61** (2018) 70–89.
[45] M. Toloo and T. Tichý, Two alternative approaches for selecting performance measures in data envelopment analysis. *Measurement* **65** (2015) 29–40.
[46] Y.M. Wang and K.S. Chin, A neutral DEA model for cross-efficiency evaluation and its extension. *Expert Syst. App.* **37** (2010) 3666–3675.
[47] Y.M. Wang and T.M. Elhag, A goal programming method for obtaining interval weights from an interval comparison matrix. *Eur. J. Oper. Res.* **177** (2007) 458–471.
[48] Y.M. Wang and Y. Luo, DEA efficiency assessment using ideal and anti-ideal decision-making units. *Appl. Math. Comput.* **173** (2006) 902–915.
[49] Y.M. Wang and J.B. Yang, Measuring the performances of decision-making units using interval efficiencies. *J. Comput. Appl. Math.* **198** (2007) 253–267.
[50] Y.M. Wang, K.S. Chin and J.B. Yang, Measuring the performances of decision-making units using geometric average efficiency. *J. Oper. Res. Soc.* **58** (2007) 929–937.
[51] D. Wu, A note on DEA efficiency assessment using ideal point: an improvement of Wang and Luo's model. *Appl. Math. Comput.* **183** (2006) 819–830.
[52] J. Wu, J. Chu, J. Sun, Q. Zhu and L. Liang, Extended secondary goal models for weights selection in DEA cross-efficiency evaluation. *Comput. Ind. Eng.* **93** (2016) 143–151.
[53] F. Yang, S. Ang, Q. Xia and C. Yang, Ranking DMUs by using interval DEA cross efficiency matrix with acceptability analysis. *Eur. J. Oper. Res.* **223** (2012) 483–488.
[54] L. Zhang and K. Chen, Hierarchical network systems: an application to high-technology industry in China. *Omega* **82** (2019) 118–131.
[55] Z. Zhou, L. Sun, W. Yang, W. Liu and C. Ma, A bargaining game model for efficiency decomposition in the centralized model of two-stage systems. *Comput. Ind. Eng.* **64** (2013) 103–108.
[56] J. Zhu, Data Envelopment Analysis: A Handbook of Models and Methods. Vol. 221. Springer (2015).
[57] Q. Zhu, F. Li, J. Wu and J. Sun, Cross-efficiency evaluation in data envelopment analysis based on the perspective of fairness utility. *Comput. Ind. Eng.* **151** (2021) 106926.
[58] H.J. Zimmermann, Fuzzy Set Theory – And Its Applications. Springer Science & Business Media (2011).