UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL SCIENCES
SOUTHAMPTON BUSINESS SCHOOL

Essays on the fairer evaluation of units in various network Data Envelopment Analysis structures

by

Marios Dominikos Kremantzis

ORCiD: 0000-0002-9531-404X

A thesis for the degree of
Doctor of Philosophy

June 28, 2022

University of Southampton

Abstract

FACULTY OF SOCIAL SCIENCES
SOUTHAMPTON BUSINESS SCHOOL

Doctor of Philosophy

by Marios Dominikos Kremantzis

This thesis sheds new light on proposing and illustrating the use of several alternative modelling approaches and methodological frameworks to attain fairness in the evaluation outcomes of the decision-making units (DMUs) under exploration, being arranged as a network Data Envelopment Analysis (DEA) system.

The thesis contains three main chapters. Apart from *Chapter 1* (Introduction), in *Chapter 2*, we initially emphasise that in DEA, a variety of approaches have been used in the context of single-stage and basic serial two-stage systems to attain fairness in the evaluation of DMUs. Little work, however, has been done to address this challenge in a generalised two-stage structure featuring additional inputs in the second stage and a proportion of first-stage outputs as final outputs. In this chapter, we argue that in this context, fairness is enhanced by increasing measures related to the discriminatory power and the weighting scheme of the method. We describe a mechanism that gives prominence to a more contemporary concept of fairness, incorporating diversity and inclusion of minority opinions. These aspects have, to our knowledge, not yet received explicit attention in the methodological development of DEA. We propose a novel combination of an additive self-efficiency aggregation model, a minimax secondary goal model, and the CRiteria Importance Through Inter-criteria Correlation (CRITIC) method, in order to promote these aspects of fairness, and thus achieve a better degree of cooperation between the stages of a DMU and among DMUs. The additive aggregation model is chosen over the alternative multiplicative approach for a variety of reasons relating to the emphasis on the intermediate products exchanged and the simplification. The minimax model offers peer evaluation in which each DMU aims to evaluate the worst of the others in the best possible light. Application of the CRITIC method to DEA addresses the aggregation problem within the cross-efficiency concept. Practical applications of this approach could include supporting the determination of training needs in job rotation manufacturing, or evaluation of sustainable supply chains. The chapter includes a description of a numerical experiment, illustrating the approach.

The evaluation of the performance of a 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 spirit, in *Chapter 3*, 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 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; 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.

Many organisations are composed of multiple departments connected either in series or in parallel, which may be further decomposed into a number of functions arranged in a hierarchical structure. Several researchers have successfully used appropriate DEA modelling techniques to assess complex structures. However, to our knowledge, no-one has examined the case of measuring and evaluating a parallel network structure combined with a hierarchical one. *Chapter 4* discusses the development of the novel multi-function parallel system with embedded hierarchical network structures to eliminate this research gap. A linear additive decomposition DEA model and a non-linear multiplicative aggregation DEA model are proposed as alternatives to evaluate the operating performance of such a structure. The system, the sub-systems, and the efficiencies of their internal units, as well as their relationships, are identified. The system efficiency of the additive model is shown to be greater than or equal to that of the multiplicative model. To verify the applicability of our proposed models, we consider a hypothetical example of the evaluation of the performances of several Business Schools across a number of universities. Other envisaged areas of application of our structure could include supporting the evaluation of the supply chain management of a firm, or the determination of most desirable ship design considering maintenance issues. *Chapter 5* summarises the main findings of the PhD thesis, points out limitations of the work, and discusses future research directions.

Keywords Data envelopment analysis; network data envelopment analysis; fairness; cross-efficiency; CRITIC; hierarchical structure; ranking

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Marios Dominikos Kremantzis, declare that this thesis entitled *Essays on the fairer evaluation of units in various network Data Envelopment Analysis structures* and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University.
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
3. Where I have consulted the published work of others, this is always clearly attributed.
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
5. I have acknowledged all main sources of help.
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.
7. Part of this work has been published as:
   a) Kremantzis, M. D., Beullens, P., & Klein, J. (2022). A fairer assessment of DMUs in a generalised two-stage DEA structure. *Expert Systems with Applications*, 115921. [Chapter 2]

   b) Kremantzis, M. D., Beullens, P., & Klein, J. (2022). A ranking framework based on interval self and cross-efficiencies in a two-stage DEA system. *RAIRO - Operations Research*, 56(3), 1293-1319. [Chapter 3]

Signed: Date: 28/06/2022

Acknowledgements

I would like to extend my sincere gratitude to a number of people, without whose help I would have never been able to complete this thesis.

First and foremost, I wish to express how fortunate and honoured I am to work with my two supervisors, Dr. Patrick Beullens and Dr. Jonathan Klein. I wholeheartedly appreciate my main supervisor, Dr. Patrick Beullens. He has taught me all important characteristics that a management scientist should have towards the pursuit of essence through various phenomena. His professional guidance, tremendous support, consistent availability, great patience and encouragement have always pushed me to the frontier and brought in the best in me. This journey will forever remain as the most enjoyable, productive and stimulating life-time experience. The enthusiasm and dedication he possesses for his research was contagious and deeply motivated me to overcome difficulties and challenges in the pursuit of a research goal which I thought was logically impossible to accomplish within almost four years. My sincere thanks go to my supervisor, Dr. Jonathan Klein, who has been greatly and constantly supportive during my thesis writing without reservation. I am profoundly indebted for his unrestricted dedication to the improvement of my critical thinking abilities and research skills as well as the clarification of my research ideas and structures throughout my Ph.D. journey.

I would also like to take this opportunity to thank Prof. Ali Emrouznejad and Dr. Simonov Kusi-Sarpong, for agreeing to serve as external examiner for the evaluation of the performance of this Ph.D thesis. I am most grateful to you and your expertise on the quality improvement of this thesis.

I am hugely grateful to Dr. Konstantinos Katsikopoulos for his constant technical and moral support and encouragement, which have enormously inspired me especially during tough times in my Ph.D. pursuit. I also sincerely thank Professor Martin Kunc who has been generously sharing his research ideas and suggestions with me, which are greatly beneficial of the completion of this thesis.

I would like to thank the Engineering and Physical Sciences Research Council (EPSRC), the BAE Systems and the University of Southampton for the four-year funding support to this thesis. More specifically, the research work has been supported by the EPSRC (Industrial CASE award) under grant 5162031105656. I am also grateful to Mr. Darren Goodman, the Research and Technology Manager from BAE Systems Maritime Services, for his strong moral and practical support. Furthermore, I am so indebted to the University of Southampton Business School's staff who were always ready to offer help whenever I needed them. In particular, my unfailing gratitude goes for Dr. Pingli Li, the director of the Ph.D. programmes and Dr. Rosalind Willis, the faculty director of Graduate School. My sincere thanks to you both.

I also gratefully acknowledge my friends and colleagues (Zeming, Arya, Tony, Leen, Jin, Aygun, Joyce, and Ahmad) who have greatly contributed to my learning and research process. My time at Southampton was full of joy and excitement in large part due to their presence and involvements.

Finally, but by no means least, special mention goes to my mother Mrs Paraskeuh Kremantzi, my father Mr. Anastasios Kremantzis, my brother Mr. Chris Kremantzis and his wife Mrs Koula Valsamidou, my partner Dimitra, as well as my nephew Anastasios, my niece Dionysia, and my friends Christiana, Daphne, Dionysis, Eirini, Fragkiska, Georgia, Ismini, Maria, Panagiotis, and Tasos, for all their love and encouragement. A special thanks to my friend Nikos for his selfless act to support me regarding the use of the Python programming language for the analysis. They are the most important people in my life and I always feel so fortunate to have all of them. I dedicate this thesis to them.

To My Beloved Parents and My Memorable Grandparents

Chapter 1

Introduction

1.1 Research Context & Background

Data Envelopment Analysis (DEA) is a well-established mathematical programming approach, often used in operations management and economics, for measuring and evaluating the relative efficiency of a set of (typically) homogeneous decision making units (DMUs) that make use of multiple inputs to produce multiple outputs **(Cook and Seiford, 2009; Cooper et al., 2011)**. The inputs are typically resources to be minimised and the outputs are products/outcomes to be maximised, leading to a better performance/efficiency measure. When DEA acts as a benchmarking process, then the inputs are usually the less-the-better type of measures and the outputs are the more-the-better type of measures. Under this circumstance, DEA also highlights its nature as a multiple-criteria decision making tool **(Doyle and Green, 1994; Stewart, 1996; Cook et al., 2014)**. In particular, DMUs are the decision alternatives, and the DEA inputs and outputs are the cost and benefit criteria, respectively. DEA is also known as a non-parametric approach, since it does not need to define the mathematical form of the production or operations function. It does not make any specific assumptions with respect to the internal mechanisms of a particular DMU either **(Cooper et al., 2000)**.

DEA was originally developed by **Charnes et al. (1978)** (CCR) for the constant returns-to-scale (CRS) assumption, and extended by **Banker et al. (1984)** to incorporate the variable returns-to-scale assumption (VRS). If a DMU operates under the CRS assumption, then a change in the input level leads to a proportionate change in the output level. If one input or one output has an equal value across the DMUs under consideration, then the CRS assumption is replaced by the VRS assumption. This is due to the fact that the input/output constraint takes the form of a convexity constraint in the model under the CRS assumption **(Cook et al., 2014)**. In general, DEA can be either input- or output-orientated. In the former case, an inefficient DMU becomes efficient via the proportional decrease of its inputs while its outputs proportions are kept fixed. In the latter case, an inefficient DMU becomes efficient via the proportional increase of its outputs while its inputs proportions are held constant **(Kao, 2017)**. To simultaneously consider both orientations, that is to address the input excesses and output shortfalls, **Taleb et al. (2022)** introduced a bi-objective DEA model with undesirable factors and mixed integer restrictions.

Since the seminal work of **Charnes et al. (1978)**, there has been a continuous interest and increased development in the field of DEA. As a result, a considerable amount of published research papers and textbooks have appeared in the DEA literature, including those of **Cook and Zhu (2007)**, **Cooper et al. (2011)**, **Cook et al. (2014)**, and **Emrouznejad and Yang (2018)**. DEA has been widely used in various applications, including energy and environment **(Zhai et al., 2019)**, water resource efficiency **(Liang et al., 2021)**, local governments **(Amatatsu et al., 2012)**, research and development departments **(Wang et al., 2013)**, financial services and banking **(Paradi and Zhu, 2013; Tan**

et al., 2021; Shi et al., 2021; Li et al., 2022), insurance services **(Omrani et al., 2022b)**, supply chain management **(Azadi et al., 2014)**, sports **(Moreno and Lozano, 2014)**, international shipping **(Gan et al., 2019)**, inland transportation **(Stefaniec et al., 2020; Wang et al., 2022)**, hospital efficiency **(Dehnokhalaji et al., 2022; Omrani et al., 2022a)**, higher education **(Ekiz and Tuncer Şakar, 2020; Lee and Johnes, 2021)**, and many more. DEA can facilitate the identification of sources of inefficiency and the classification of DMUs from the most to the least efficient **(Liang et al., 2008; Cook et al., 2014; Halkos et al., 2014)**.

In the conventional DEA, the overall assessment of the target DMU is based on the optimal set of its own most favourable multipliers and on the concept of the efficiency frontier **(Zhu, 2015)**. In this context, "*the efficient DMUs, as defined by DEA, may not necessarily form a production frontier, but rather lead to a best-practice frontier*" **(Cook et al., 2014, p.2)**. In a traditional single-stage DEA structure, it is not known how the inputs (resources) are converted into outputs (outcomes) within the core of a DMU under consideration. The so-called "black box" analysis is assigned to such cases, to depict that a DMU is being treated as a whole unit that measures its relative performance by only considering its exogenous inputs and exogenous outputs **(Kao, 2017)**. As stressed in **Zhu (2020)**, a great number of tasks and performance measures cannot be accommodated in a "black box" DEA structure. The internal operations of a DMU might consist of a number of interrelated and/or independent functions and tasks, depending on the context of the problematic area **(Kao, 2014)**. Ignoring the internal operations of a DMU could lead to misleading outcomes. For instance, while the whole unit could be characterised as efficient, all its constituent sub-stages may be inefficient. This could be better explained by considering an application in the financial sector, where there are two interrelated processes, the capital collection and the profitability due to the investment decisions. By exploring the effect of the information technology (IT) on the efficiency of a bank branch, we might discover that IT is linked to the capital collection but not to profitability. The selection of the appropriate investment opportunities will determine the level of profit. In addition, the traditional concept can lead to a large number of DEA-efficient DMUs **(Ma et al., 2017)**, sharing the top position in the ranking list. Hence, to enable the study of internal operations and easier identify any causes of inefficiency, research has extended DEA models to consider network structures **(Färe and Grosskopf, 2000)**.

According to **Zhu (2020)** and **Charles et al. (2021)**, network DEA is most prominent in the area of big data, highlighting how various performance metrics could be associated via complicated network structures. **Fukuyama and Weber (2021)**, for instance, implemented a network DEA approach under a big data set to optimally reallocate tests for COVID in the United States. The network system differs from the black-box in that it involves more complex structures, thereby leading to a less systematic illustration **(Kao, 2017)**. Furthermore, the traditional single-stage (black-box) DEA model always identifies at least one efficient DMU, whilst a network DEA model might not spot an overall efficient DMU **(Fukuyama and Mirdehghan, 2012; Zhu, 2020)**. In general, a network system converts the inputs forwarded to the system from the external environment to final outputs through the assistance of some intermediate measures. To measure the efficiency of the units and sub-units within a network DEA structure, literature advocates the concepts of the efficiency decomposition and the efficiency aggregation **(Kao, 2016; Kao, 2017; Kao, 2018; Lee, 2021**). The efficiency decomposition (see also **Lozano and Khezri (2021)**) defines the system efficiency taking into account only the exogenous inputs and outputs, and then identifies a mathematical (multiplicative or additive) relationship that associates system and process efficiencies. **Despotis et al. (2016)** argued that the efficiency scores obtained via the additive decomposition method are biased, whereas those obtained via the multiplicative method are not unique. Recently, **Koronakos et al. (2022)** introduced a novel network DEA approach to reach a fair compromise efficiency decomposition. On the other side, the efficiency aggregation defines the system efficiency as a (multiplicative or additive) function of those of its components, involving the intermediate measures **(Kao, 2017)**. **Lu et al. (2020)** proposed a multiplicative efficiency aggregation approach in a two-stage network DEA model into the form of second order cone programming.

In the simplest version of a two-stage tandem system, all inputs used by a DMU feed into a first stage, producing intermediate outputs that all feed into a second stage, producing the final outputs of the entire system. Performance measurement of DMUs in this structure has received a reasonable amount of attention **(Cook et al., 2010b)**. **Kao and Hwang (2008)** proposed that the overall efficiency is decomposed into the product of the efficiencies of the two constituent stages. **Chen et al. (2009)** examined an additive efficiency decomposition approach in which the overall efficiency is decomposed into the weighted sum of the efficiencies of the two divisions, where the weights indicate the degree of importance. While **Kao and Hwang's (2008)** approach was only applicable to CRS situations, **Chen et al. (2009)** showed applicability to both CRS and VRS conditions. **Ang and Chen (2016)** proved that the decomposition weights in **Chen et al.'s (2009)** study are non-increasing in the order of sub-stages. They particularly highlighted that earlier divisions would be assigned higher relative importance, affecting the overall system's efficiency to a greater extent. Based on that, they also demonstrated that the overall and sub-stages' efficiency scores are prone to the impact of the decomposition weights. Some pitfalls that concern the weighted additive efficiency decomposition approach have also been discussed and addressed via an alternative overall efficiency index that reflects the division efficiencies by **(Guo et al., 2017)**. **Wang and Chin (2010c)** proposed alternative DEA models in which the overall system efficiency of the two-stage tandem series process is modeled as a weighted harmonic mean of the efficiencies of the two internal divisions. Which of these methods leads to the best results will depend on the context of application, and while not part of the focus in this Ph.D. thesis, it does form an interesting area for further examination.

Realistic cases extend the two-stage tandem structure into a generalised one in which the first stage additionally generates final outputs, the second stage also produces exogenous inputs, and certain outputs of the second stage are re-utilised by the first (see for example, **Chen and Zhu, 2019; Pérez-González et al., 2021)**. The latter structure can be further illustrated with the use of multiple divisions arranged in a series way **(Despotis et al., 2016)**. **Chen et al. (2010)**, for instance, proposed a model for performance estimation where the stages share non-splittable inputs, while **Zha and Liang (2010)** designed cooperative and non-cooperative models where shared inputs can be freely allocated between the two stages. **Yu and Shi (2014)** examined a two-stage structure with additional inputs in the second stage and part of intermediate products, as final outputs. **Jianfeng (2015)** considered a different two-stage DEA model, in which part of the intermediate products from the first stage convert into an input for the second stage, while the remaining products convert into the final output. Moreover, he classifies the inputs into those that are entirely integrated into one stage and those that are shared between the two stages. **Amirteimoori et al. (2016)** developed a number of additive efficiency models to evaluate the operating performance of DMUs in the form of a two-stage DEA structure with shared resources. **Ma et al. (2017)** proposed a parallel-series hybrid two-stage DEA model. Their model utilised the principles of additive and multiplicative DEA approaches for efficiency measurement and decomposition. **Wang et al. (2020)** constructed a two-stage network DEA model with shared inputs, additional intermediate inputs, and free intermediate outputs to evaluate the technological innovation efficiency of China's high-tech industries. **Chen et al. (2021)** proposed an extended two-stage network DEA approach with shared input resources to measure the operating efficiency of a set of Chinese universities. **Michali et al. (2021)** suggested an additive decomposition network DEA approach with intermediate and undesirable outputs to measure the European railway transport process efficiency. Such systems have mainly a series structure, in that they operate interdependently. Nevertheless, in other types of networks, the internal divisions are largely placed in parallel without impacting one another **(Kao, 2009b; Kao, 2012)**. These networks have been essentially divided into two prevalent systems: the multi-component and the multi-function parallel systems **(Kao, 2017)**. Extensive research has examined such systems in various applications, including the performance evaluation of physics and chemistry departments in UK universities **(Beasley, 1995)**, the assessment of commercial banks in Iran **(Jahanshahloo et al., 2004)**, the maximisation of sales of Portuguese retail stores **(Vaz et al., 2010)**, the impact of coal-fired power plants on pollution-generating processes **(Lozano, 2015)**, and the evaluation of the operational capability of container terminals **(Park et al., 2022)**. In most of the aforementioned studies, scholars examined realistic multi-stage parallel systems with shared/non-shared resources and common/separate outcomes. Although many scholars have emphasised the key assumption of the parallel systems, in that the internal divisions are related to the same type of inputs and outputs (in differing amounts), there are also efforts indicating the development of parallel network DEA models with respect to addressing non-homogeneous parallel divisions' issues **(Du et al., 2015)**. Other special network cases are the hierarchical systems, where the resources of a unit in a certain level are further split to the sub-units of the immediate next level. For instance, **Castelli et al. (2004)** presented single-stage and two-stage hierarchical structures to measure the efficiency of DMUs, assuming that the internal divisions are under the constant returns-to-scale. **Cook and Green (2005)** introduced a DEA model to evaluate the efficiency of a number of power plants that are decomposed into power units within a hierarchical structure. The novelty of their model lay in its nature to consider all hierarchical levels at the same time. **Kao (2015)** also developed a relational model for a single-stage hierarchical structure to measure both the overall system and its divisions' efficiencies simultaneously. He argued that this structure is identical to a parallel system **(Kao, 2009b)**. **Kao (2015)** optimised the efficiency of the overall production system, considering only the constraints corresponding to the terminal divisions. **Li et al. (2020)** focused on the same hierarchical structure by additionally optimising the efficiencies of the terminal divisions. **Chen et al. (2019)** proposed a hierarchical DEA framework to construct an indicator with a view to reevaluating the 2014 Global Food Security Index. **Zhang and Chen (2019)** extended the concept of **Kao (2015)** to a generalised single-stage hierarchical structure wherein all internal units across the different levels can reflect a two-stage tandem system. To examine the relationship between the system and its sub-units, they introduced additive aggregation and multiplicative decomposition DEA models. **Gan et al. (2019)** suggested a general two-stage series process, in which each stage is no longer treated as a black-box, but is further elaborated into a hierarchical structure with multiple layers. **Ghasemi et al. (2020)** assessed the performance of the campuses of the Farhangian University, which are developed as a four-level hierarchical structure. Which network model is most appropriate, is arguably dictated by the area of application, and should consider the interplay between the physical reality of how inputs and outputs connect, and how the decision maker (DM) divides managerial responsibilities across different sections of DMUs.

While network structures have the potential to investigate the sources of inefficiency, two major challenges similar to those in a single-stage production system emerge. The first one concerns the lack of discrimination power **(Mahdiloo et al., 2016; Örkcü et al., 2019)**. This is more likely to occur when the number of DMUs is small relative to the number of inputs and outputs. It is often a problem that coincides with a high number of DEA-efficient DMUs **(Bal et al., 2010; Ghasemi et al., 2014)**. The challenge stems from the self-evaluation concept of (classic) DEA **(Örkcü et al., 2019)**. The inability to discriminate DMUs may lead to a lack of actionable results to be identified by the decision maker.

The second challenge is pertinent to what is referred to in the literature as an 'unrealistic' weighting scheme **(Mahdiloo et al., 2016; Örkcü et al., 2019)**. Indeed, it is allowed for a high relative-importance weight to be assigned to 'less important' inputs or outputs (according to some objective standards), and/or a low weight to significant factors. This choice of weights could turn a DMU into a DEA-efficient unit. When evaluated according to a weighting scheme which would reflect an 'objective' relative significance of various inputs or outputs, this DEA efficient DMU might, in fact, have a worse performance than a seemingly DEA non-efficient DMU **(Wu et al., 2012b; Ghasemi et al., 2014)**. As cited in **Zarei Mahmoudabadi and Emrouznejad (2022)**, there are specific conditions under which the selection of input and output weights requires further considerations. For instance, when the weights allocated to the various known factors are zero or epsilon, or when the decision maker is not compliant with several of the results, then an unbalanced situation can emerge. The common set of weights is one of the many methods introduced in (traditional and network) DEA models to reduce the flexibility in the selection of weight sets assigned to the input and output factors **(Omrani et al., 2019; Mavi et al., 2019; Yu et al., 2021)**. Whether or not a weight scheme is deemed unrealistic has thus to do with the question about whether the degrees of freedom given to the self-evaluating DMUs (the subjective realm) are too high or not in the light of the existence and nature of overarching considerations about areas of consensus across the collective of DMUs (the objective realm). These overarching considerations can either be imposed a priori (e.g., top-down by the decision maker), or revealed during the DEA evaluation process (e.g., through cross-efficiency evaluation).

The aforementioned challenges can impede the attainment of a fair and unique ranking. Fairness can be generally considered as a universal mechanism that corresponds to using a system of evaluation and ranking that is acceptable by the units being assessed **(Fehr et al., 2002; Beullens et al., 2012)**. **Leventhal et al. (1980)** suggested several fairness organisational principles including the consistency, the lack of bias, the use of precise information in decision-making contexts, and the conformity to dominant standards of ethics. According to **Hartmann and Slapnicar (2012)**, the fairness of a project is connected to the uncertainty of its outcomes. As the uncertainty of the project outcomes increases, the efficiency of that project increases too. However, the cost of the uncertainty can outweigh the benefits of fairness at some point. In the DEA context, fairer evaluation outcomes could be achieved by establishing effective rules and guidelines and/or implementing DEA methodologies that everyone can accept and account to fairness in a way everyone can find agreeable. **Banker and Morey (1986)** introduced a fairness approach where the groups are classified by the degree of discrimination power. **Jahanshahloo et al. (2017)** introduced a fairer method for ranking the DMUs under examination; they, particularly, evaluated the optimistic and pessimistic efficiency scores by taking into account the optimistic and pessimistic optimal weights of all DMUs. **Chen et al. (2020)** proposed a fair DEA framework showing that an unprivileged group of DMUs cannot reach the same level of outputs as the privileged group of DMUs, although they both use the same level of inputs. To achieve fairer efficiency scores without creating disparate impact, they imposed additional constraints to the original models. **Yu and Chen (2020)** developed a modelling approach to evaluate production and service efficiency and explore technology biases in the internal divisions of a DMU, by integrating network DEA and the meta-frontier approach. **Radovanović et al. (2021)** integrated fairness into a DEA method through additional constraints regarding the disparate impact of algorithmic decision-making. Finally, **Wu et al. (2021)** investigated how fairness concern influences a DMU by introducing new utility-based network DEA models to analyse the non-cooperative and cooperative modes.

In this Ph.D. thesis, we are interested in methods which aim to ensure appropriate conditions for a fairer evaluation and ranking of the outcomes of the DMUs under exploration. These conditions could lie in a higher level of discriminatory power, a more realistic weight scheme, a greater degree of collaboration between DMUs and sub-stages of a particular DMU, the less mainstream and more diversified profiles and viewpoints within a decision-making context, and/or the more complicated mechanisms and operations within the core of a DMU. In short, we say that we intend to tweak the network DEA methodology to enable fairer efficiency estimations.

There are, broadly speaking, two types of methods for addressing the fairness issues. These methods have been primarily developed for and tested on single-stage DEA models. Methods in the first class incorporate a priori information into the model **(Dyson and Thanassoulis, 1988; Zhu, 1996; Halme et al., 1999)**. Methods in the second class keep the model largely intact but reveal additional principles of consensus via the DEA process. In particular, these methods (often) rely on a generalisation of the process of how each DMU's performance is arrived at, e.g. through: cross-efficiency **(Doyle and Green, 1994)**, interval cross-efficiency **(Yang et al., 2012)**, multi–criteria DEA (MCDEA) **(Li and Reeves, 1999)**, or Nash bargaining game **(An et al., 2017)**. The following paragraphs will introduce some of the DEA concepts related to the second class and explore relevant literature, in order to get a firm grip on understanding how fairness has been attained per case. Nevertheless, it is important to mention that part of the focus in this Ph.D. thesis is to make methodological advancements only in cross-efficiency and interval cross-efficiency concepts (belonging to the second class) within an appropriate network DEA context.

According to the cross-efficiency evaluation concept **(Sexton et al., 1986)**, each DMU is assessed using its own most favourable weights next to the weight profiles of all other DMUs. **Anderson et al. (2002)** emphasised that cross-efficiency greatly improves the probability of obtaining a unique ranking and eliminates the unrealistic weight distribution. A critical drawback of this concept is the non-uniqueness of optimal weights, which leads to the non-uniqueness of cross-efficiencies. This result is in general unfavourable **(Liang et al., 2008)**, indicating that this concept is by itself not sufficiently strong. Indeed, one can imagine game-playing behaviour to become possible in which DMUs will select their self-evaluation multipliers in an aim to, for instance, lower ultimate scores of some of their peers. See also **Wu et al. (2021)** for a review of DEA cross-efficiency methods and applications.

Doyle and Green (1994) were among the first to recommend the adoption of alternative secondary goals in an aim to select unique optimal multipliers. In particular, they introduced an aggressive model and a benevolent model, while the secondary objective functions in **Liang et al. (2008)** reflected the minimisation of total deviation, maximum deviation, and mean absolute deviation from an 'ideal' point i.e. the situation when the self-efficiency score equals 1. Alternative definitions of ideal points were given in **Wang and Chin (2010b)** as to include both DEA efficient and non-efficient units. Stimulated by the concept of the ideal points, **Wu et al. (2016)** proposed extended secondary goal models considering both desirable and undesirable cross-efficiency (CE) targets for the DMUs. Their merit was that the evaluated DMUs have the motivation to accept their ranking since they approach their ideal targets and diverge from the anti-ideal points. In addition, **Wang and Chin (2010a)** introduced a neutral DEA model for CE evaluation, in that the optimal weight profile of each DMU is neither aggressive nor benevolent to the other DMUs. **Ramon et al. (2010)** developed CE models to determine more reasonable weight profiles; they ensured non-zero and least dissimilar weights. **Li et al. (2018)** suggested a game-like iterative algorithm to obtain balanced CE scores. A number of studies have been reported in this direction, such as **Wang et al. (2011)**, **Wu et al. (2012b)**, **Alcaraz et al. (2013)**, **Oukil and Amin, 2015**, **Liu et al. (2017)**, **Shi et al. (2021)**, and **Alcaraz et al. (2022)**. Furthermore, **Kao and Liu (2019)** applied the cross-efficiency to measure the efficiency in two basic network structures, in series and in parallel; for this purpose, they developed an aggressive-based model. **Örkcü et al. (2019)** came up with a neutral-based cross-efficiency model in a two-stage DEA system to quantify the performance of both the overall system and its individual stages. This neutrality was indifferent regarding the impact of the optimal weight sets on the average cross-efficiencies of all other DMUs. **Shao and Wang (2021)** also introduced several novel aggressive, benevolent, and neutral two-stage cross-efficiency DEA evaluation models based on prospect theory.

Overall, the choice of an alternative secondary goal model mainly depends on the developed relationship among DMUs. This relationship could be cooperative (DMUs could work together to attain a shared target) or non-cooperative (DMUs could act in a competitive mode to satisfy their own benefits). As for the first type of relationship, **Troutt (1997)** and later **Liang et al. (2008)** developed a novel secondary goal model, based on the minimisation of the maximum k-inefficiency score within a traditional single-stage DEA structure. In other words, they put emphasis on the designation of the best behaviour of the worst performing DMU, leading to an environment that creates an atmosphere of cooperation. To the best of the authors' knowledge, no-one has yet amended and customised such a model to the specifications of a network structure. In general, far too little literature exists towards the development of a cooperative secondary goal model within a network DEA context.

In addition to the alternative secondary goal model, the aggregation of the individual cross-efficiency scores plays another instrumental role in accommodating fairness in the evaluation outcomes within a DEA context. In DEA literature, several ways of aggregating these scores have been raised, mainly varying between the traditional average and the weighted average methods. Although the arithmetic average method has proven effective in ensuring a credible ranking order in certain cases **(Liang et al., 2008, Wang and Chin, 2010b)**, it loses sight of the weights assigned to individual CE **(Wang and Wang, 2013)**. On top of that, it is not a Pareto optimal solution **(Wu et al., 2011)**. To accommodate these issues, **Wu et al. (2009)** treated the DMUs as players in a cooperative game and the solution of Shapley value was determined to compute the ultimate CE. **Wu et al. (2011)** utilised the Shannon entropy approach, allocating a fixed but different weight to each DMU. **Wu et al. (2012a)** pointed out that this is a problematic condition, since it ignores the primary role of the self-evaluated efficiency of each DMU, located on the leading diagonal of the CE matrix. To this end, they embedded the Shannon entropy into the CE concept by fully considering the association among the self and the peer-evaluation values. **Wang and Chin (2011)** proposed the use of the ordered weighted averaging (OWA) operator weights to fairly allocate the weights to CE in terms of the DM's optimism level. The optimism level is characterized by an orness degree value, which is uncertain and requires DM's subjectivity. By the same token, **Leon et al. (2014)** proposed an aggregation approach based on the induced ordered weighted averaging (IOWA) operators to integrate the DMs' preferences regarding the relative importance weights. **Ruiz and Sirvent (2012)** calculated the CE scores via a weighted average method that reflected the disequilibrium in the DEA weight profiles. More recent works see **Yang et al. (2013)**, **Song et al. (2017)**, **Song and Liu (2018)**, **Wu et al. (2021)**, and **Fu and Li (2022)**.

On the aggregation of the individual cross-efficiencies, existing frameworks **(Wang and Chin, 2011; Wu et al., 2012a; Song and Liu, 2018)** pay close attention to the reasonable allocation of the weights by limiting the range between self and peer-assessment efficiencies. Their direction was probably inspired by the assumption that the opinion of the evaluator for itself is significantly more powerful than the respective opinion of others, yet quite ill-treated since it belongs to the minority; it, thus, does not receive the attention it deserves. The exploration of an aggregation method that will, by design, not only value majority opinions but also well-supported minority opinions is crucial since it will correspond to the more modern mindset of many organisations.

Fairness in the evaluation outcomes has been achieved even via the MCDEA **(Li and Reeves, 1999; Hatami-Marbini and Toloo, 2017)**. As mentioned earlier, MCDEA belongs to those methods that keep the model intact while revealing additional principles of consensus. While classic DEA models pursue solely the maximisation of the efficiency score, in MCDEA other alternative objective functions are used as well. Li and Reeves suggested the following objectives: *(1)* minimisation of the deviation variable, i.e. maximising the efficiency score of the evaluated DMU; *(2)* minimisation of the maximum deviation, which can be called a Chebyshev objective; and *(3)* minimisation of the sum of deviations. By its very nature of being a multi-objective linear programming (MOLP) model, it is typically difficult to identify a global optimal solution **(Li and Reeves, 1999; Bal et al., 2010; Ghasemi et al., 2014; Ghazi and Lotfi, 2022)**. Related research includes the conversion of the MOLP model either into a goal programming (GP) model **(Bal et al., 2010; dos Santos Rubem et al., 2017)** or into a bi-objective weighted MCDEA model **(Ghasemi et al., 2014)**. **Mahdiloo et al. (2016)** adjusted the MCDEA in the context of a two-stage sustainable system in an aim to fix the unrealistic weight scheme and the weak discrimination power.

Furthermore, one of the most 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. However, the performance measurement of a DMU can additionally be supported by the pessimistic viewpoint and in turn by its most unfavourable (optimal) weight sets. Research has consistently shown that considering the optimistic and pessimistic perspectives within a self-appraisal setting simultaneously, can provide more valuable and meaningful insights. **Entani et al. (2002)** were, to our knowledge, among the first to conceive the idea of considering DEA efficiencies from both the best and the worst aspects to obtain an interval efficiency. A drawback of their approach was that only one input and one output were utilised in calculating the worst relative efficiency of each DMU, regardless of the total number of inputs and outputs allocated to each DMU. **Toloo and Tichý (2015)** proposed a multiplier model to identify the maximum efficiency scores and applied the envelopment model to attain the maximum discrimination among efficient DMUs. Based on the ideal and anti-ideal DMUs, **Liu and Wang (2018)** also developed the normalised efficiency metric and then formulated two DEA models to obtain its lower and upper bounds. Numerous studies, in general, exist towards the measurement of the relative efficiency of the target DMU based on its own optimistic and pessimistic weight profiles, simultaneously, within a traditional single-stage DEA structure **(Wang and Luo, 2006; Wu, 2006; Wang and Yang, 2007; Azizi and Ajirlu, 2010;** Azizi and Jahed, 2011; Azizi and Wang, 2013; Chen, 2014; Azizi, 2014; Örkcü et al., 2020). **Badiezadeh et al. (2018)** 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. **Su and Sun (2018)** introduced a new DEA model to measure the efficiency of a multi-stage network supply chain structure with undesirable outputs and dual-role factors. Nevertheless, other relevant studies and applications towards using a double-frontier DEA model within a network system are surprisingly sparse so far.

Finally, research has showcased that the performance measurement of the target DMUs can be attained even with the exploration of the cross-efficiencies in a weight space considering all the weight profiles, within the traditional single-stage DEA structure **(Yang et al., 2012)**. This concept, which was originally introduced by **Yang et al. (2012)**, ensures neutrality (no preference choice between aggressive and benevolent strategies), unique sets of weights, and a unique and meaningful rank. Following the idea of **Yang et al. (2012)**, **Liu (2018)** considered interval cross-efficiencies along with their variances to estimate the variability of the CE intervals and rank the targeted DMUs. In this paper, the author showed how the interval cross-efficiencies will be ranked using the signal-to-noise ratio. Other studies have also been reported in this direction, including **Ramón et al. (2014)**, **Fang and Yang (2019)**, and **Wang et al. (2021)**. Far too little literature, however, exists towards the customisation of such a methodological approach (i.e., the interval cross-efficiency concept) within a network DEA context and its impact thereon.

1.2 General Research Contributions

The general research contributions of this Ph.D. thesis are threefold:

1. To attain fairness in the evaluation and ranking of several competing DMUs in the form of a generalised two-stage DEA structure; this is achieved by increasing measures related to discrimination power and the weight scheme, and by highlighting the concepts of cooperation, diversity, and inclusion.
2. To meaningfully evaluate and rank DMUs within a two-stage tandem DEA structure; this is achieved by considering the optimistic and pessimistic scores in a self-evaluation context as well as the most favourable and unfavourable weight sets of each of the other DMUs in a peer-appraisal setting.
3. To develop a multi-function parallel system with embedded hierarchical network structures and evaluate the operating performance of DMUs with such structure.

1.3 Research Aims & Objectives

A significant number of researchers have paid attention to the measurement of the relative efficiency of DMUs as "black-box" units (whose internal operations are entirely neglected). In reality, however, organisations, entities, departments, premises, and countries (the so-called DMUs) might consist of more complex and various constituent functions and/or operational procedures. Assessing a DMU without considering its internal operations can discourage decision-makers from entering it into the ranking process, which tends to be less fair and balanced. In particular, this may lead to the lack of identifying the sources of inefficiency or the specific internal functions/substages with a highly beneficial footprint. Moreover, the absence of cooperation, the assignment of a smaller relative importance to meaningful factors, and the tie of a number of DMUs (especially in the first place) are a few significant obstacles in pursuing a more meritocratic evaluation and ranking.

The research aim of this Ph.D. thesis is to attain a fairer measurement, evaluation and ranking of the relative performances of DMUs; their structure is not restricted to only considering exogenous inputs to consume and exogenous outputs to produce, but it is further extended to more complex internal operations. In particular, existing and new modelling approaches and methodologies are utilised within self and peer-appraisal settings to meaningfully rank DMUs with either a generalised two-stage structure or a two-stage tandem structure (both structures already exist in the DEA-related literature) or a multi-function parallel network hierarchical system (proposed for the first time in this thesis). The ultimate goal is to gain a valuable insight into not only their network structures' past accomplishments, but also their future developments and aspirations. This Ph.D. thesis sheds light on a number of substantial topics on network analyses that are jointly considerable to academics and practitioners in fields of higher education, logistics, production, banking, and generally in multiple-criteria decision making situations.

Research aims of Chapter 2: to meaningfully evaluate and rank DMUs in the form of a generalised two-stage DEA structure with additional inputs in the second stage and part of intermediate products as final outputs. To attain the former target, we aspire to eliminate the lack of discrimination power, to increase the chances of obtaining a more realistic weight scheme, to highlight the best behaviour of the worst performing DMUs, and to give prominence to the minority opinions.

The research objectives of Chapter 2:

1. To propose an additive self-efficiency aggregation model in a generalised two-stage DEA structure with a view to identifying the strength of each of the constituent stages.
2. To develop a multi-objective minimax secondary goal model within the network structure under consideration, in order to ensure unique optimal multipliers, a more realistic weight scheme, and a better degree of cooperation between substages of a DMU and among DMUs.

3. To apply the CRiteria Importance Through Intercriteria Correlation (CRITIC) multi-criteria decision-making method in DEA to alternatively address the aggregation problem within the cross-efficiency concept via objectively determining the weights assigned to individual cross-efficiencies.

4. To both theoretically and empirically explain how the two pillars of the CRITIC method, the contrast intensity (standard deviation) and the conflict measures, can give voice to the less mainstream opinions, promoting diversity and inclusion.

5. To demonstrate that the CRITIC method is compatible with the minimax secondary goal model proposed for the generalised two-stage DEA structure, obtaining a greater discrimination power among DMUs.

Research aims of Chapter 3: to meaningfully evaluate and rank DMUs in the form of a two-stage tandem DEA structure. To attain this target, each DMU will evaluate itself using its own most favourable and unfavourable weight sets leading to a unified measure. Moreover, each DMU will be further evaluated via the most favourable and unfavourable weight profiles of each of the other DMUs, while keeping the unified self-efficiency measure unchanged.

The research objectives of Chapter 3:

1. To illustrate via several modelling approaches how a DMU can be self-assessed using its own optimistic and pessimistic multipliers, simultaneously, within a two-stage tandem (series) structure.

2. To demonstrate how the two opposing self-efficiency measures are merged to a combined self-efficiency measure via the geometric average.

3. To derive the mathematical relationship among the combined self-efficiency score of the target DMU for the overall system and those for its constituent sub-stages.

4. To extend the interval cross efficiency concept in the two-stage tandem structure, in order to determine the minimum and the maximum individual cross-efficiencies for each DMU and flow, leading to a respective interval peer-efficiency score.

5. To demonstrate that each interval cross efficiency matrix is treated as a multi-criteria decision-making problem and utilise established modelling approaches from the literature to delineate its elements and ultimately rank the DMUs.

Research aims of Chapter 4: to develop a more realistic network DEA structure, according to which each DMU will be decomposed, on a macro level, into a finite number of parallel sub-systems. Each sub-system will further disintegrate, on a micro level, into multiple units with distinctive functions arranged into a hierarchical structure. Several properties of such a system will be analysed.

The research objectives of Chapter 4:

1. To develop a multi-function parallel system with embedded hierarchical network structures in order to broaden the range of options within the network DEA field.
2. To introduce a linear additive decomposition DEA model and a non-linear multiplicative aggregation DEA model to measure and assess the performance of (hypothetical and real) DMUs with such a network structure.
3. To identify the efficiencies of the overall system, its sub-systems, and its internal units at all levels of the hierarchical structure within each sub-system as well as their relationships and properties.
4. To derive the scientific relationship between the system efficiency of the additive decomposition model and the system efficiency of the multiplicative aggregation model.
5. To gauge the operating performance of several Business Schools across a number of hypothetical universities that reflect the proposed parallel network hierarchical DEA structure and then draw compact conclusions about their strengths, weaknesses, and areas of improvement.

1.4 Structure of the Thesis

In this Ph.D. thesis, we have systematised our efforts towards measuring, evaluating, comparing, and ranking, with empirical rigour, the operating performance of units with a network DEA structure under a deterministic multiple-criteria decision making environment. The network systems under consideration are the generalised two-stage structure (Chapter 2), the two-stage tandem (series) structure (Chapter 3), and the newly introduced multi-function parallel system with embedded hierarchical network structures (Chapter 4). These network structures are more flexible in effectively corresponding to the performance evaluation and comprehension of complex and realistic organisation and production systems with independent and interrelated internal processes. The exploration of the past achievements and the more mindful planning of the future business strategy have a significant impact in managers' strategic, tactical, and operational decisions within a firm. In the current thesis, more emphasis is placed on the development of new mathematical (optimisation) models as well as the implementation of consolidated methodologies and techniques from the literature, for a more systematic study of the above-mentioned network DEA structures. The thesis makes a significant effort to ensure appropriate conditions for a fairer evaluation and ranking of the DMUs under examination. A central point which underpins the shared target of fairness in ranking with respect to the three network systems is the investigation of their internal operations and functions. Nevertheless, each of the three systems (in each of the subsequent three main chapters of the thesis) is devoted to a different direction and emphasis by employing different approaches, while addressing the common research question.

Specifically, in Chapter 2, we will discuss how a novel combination of an additive self-efficiency aggregation model, a multi-objective minimax secondary goal model, and the CRITIC method attains a fairer evaluation and a better degree of cooperation between stages of a DMU and among DMUs under consideration. The proposed research framework evaluates the DMUs within a generalised two-stage DEA structure, initially introduced by **Yu and Shi (2014)**, under self and peer-appraisal settings. In particular, the additive self-efficiency aggregation will highlight the size of each division via DMU-specific weights and lead to the most favourable self-efficiency score per DMU and stage. The minimax goal model will address the non-unique optimal multipliers derived via the self-evaluation model. The CRITIC method is the third piece that will alternatively accommodate the aggregation problem within the DEA cross-efficiency and further show how its two main pillars (conflict and contrast intensity) emphasise the more diversified viewpoints.

Chapter 3 is still on the same wavelength with the previous chapter, yet with two major differentiations. In the traditional self-evaluation context, the DMUs will be assessed via their own optimistic and pessimistic weight sets, leading to a combined self-efficiency measure. In the peer setting, each DMU will be evaluated, based on the most favourable and unfavourable weight profiles of each of the other DMUs, while keeping unchanged the combined self-efficiency measure. In this chapter, we will particularly introduce a 7-step methodological framework to ensure more informative and multi-dimensional evaluation outcomes. The first three steps of the framework will highlight how the optimistic and pessimistic DEA models, which are inspired by the studies of **Wang and Luo (2006)** and **Wu (2006)**, are built towards the more realistic two-stage tandem system. The remaining steps of our framework will ensure the peer-evaluation of the DMUs via the customisation of the interval CE method to the specifications of the two-stage tandem structure.

Finally, in Chapter 4, greater emphasis is placed on the development of a novel multi-function parallel network hierarchical DEA system that might better correspond to the reality than the network systems presented in the other two chapters. Its main asset lies in the fact that the selected network scheme (parallel) intertwines with a hierarchical structure. An additive decomposition DEA model and a multiplicative aggregation DEA model are proposed as alternatives to evaluate the operating performance of such a structure. This chapter will showcase how such a new structure will address the weaknesses of the traditional black-box DEA model and the parallel system of **Kao (2009b)**; this will be achieved by improving the level of discriminatory power among efficient DMUs and by measuring the performance scores of not only the overall system and its parallel sub-systems, but also the internal units at all levels of each of the integrated hierarchies.

In conclusion, overall summary of findings and research contributions of the thesis are clearly presented in Chapter 5. In addition, this chapter acknowledges the potential limitations of current work and outlines future research directions.

Chapter 2

A fairer assessment of DMUs in a generalised two-stage DEA structure

2.1 Introduction

A paper based on this chapter has been published, see
Kremantzis, M. D., Beullens, P., & Klein, J. (2021). A fairer assessment of DMUs in a generalised two-stage DEA structure. ***Expert Systems with Applications*****, 115921.**

Data Envelopment Analysis (DEA) is a non-parametric approach for evaluating the performance of decision-making units (DMUs) that use inputs to produce outputs **(Cook et al., 2014)**. DEA was developed by **Charnes et al. (1978)** (CCR) for the constant returns-to-scale assumption. Traditional DEA does not model the internal processes in a DMU. As a result, a relatively large proportion of DMUs emerge as DEA-efficient, without a means to distinguish them **(Ma et al., 2017)**. To enable the study of internal structures, research has extended DEA models to consider network structures **(Färe and Grosskopf, 2000; Tone and Tsutsui, 2009; Kao, 2009; Kao and Hwang, 2011; Wanke and Barros, 2014; Kao, 2014; Zhu, 2015; Guo et al., 2017; Chen and Zhu, 2017; Koronakos et al., 2019; Örkcü et al., 2019; Shi et al., 2021; Koronakos et al., 2022; Qu et al., 2022; Khoveyni and Eslami, 2022; Kiaei and Kazemi Matin, 2022)**. In a two-stage process in particular, inputs used by a DMU feed into a first stage, producing intermediate outputs that feed into a second stage, producing the final outputs of the entire system. Such a structure facilitates the measurement of both the overall system and its individual stages' efficiencies **(Mahdiloo et al., 2016)**.

Measuring the performance can be challenging when inputs and outputs are shared among different processes and are not easily distinguished **(Zha and Liang, 2010)**. **Yu and Shi (2014)** examine a two-stage structure with additional inputs in the second stage and part of intermediate products as final outputs, towards building cooperative and leader-follower models. **Jianfeng (2015)** considers a network DEA model, in which the inputs are classified into those that are entirely integrated into one stage and those that are shared between the two stages. **Ma et al. (2017)** propose a parallel-series hybrid two-stage DEA model utilising the principles of additive and multiplicative efficiency decomposition.

While two-stage DEA models have the potential to increase managerial insight into the sources of inefficiency, two major problems similar to those in a single-stage emerge. The first one concerns the lack of discrimination power due to a high number of efficient DMUs **(Mahdiloo et al., 2016)**. The second challenge relates to an 'unrealistic' weighting scheme. Indeed, it is allowed for high relative-importance weights to be assigned to 'less important' inputs or outputs, and/or low weights to significant factors. This choice of weights could turn a DMU into an efficient unit **(Ghasemi et al., 2014)**.

In this chapter, we are interested in methods which aim to avoid a low degree of discrimination, unrealistic weight schemes, and to use a system of ranking that encourages cooperation by the units being evaluated. While doing so, we also wish to provide a mechanism that gives a voice for minority opinions. This aspect has, to our knowledge, not yet received explicit attention in the methodological development of DEA. In short, we say that we intend to tweak DEA methodology to improve the *fairness*[1] in the evaluation outcomes. We summarize the core literature, relevant to fairness evaluation in DEA, in Figure 2.1.

Among those methods tested towards fairness is the cross-efficiency (CE), which adds peer-evaluation to the self-evaluation principle **(Sexton et al., 1986)**. As stressed by **Anderson et al. (2002)**, CE improves the probability of obtaining a unique ranking. A critical drawback of the CE is the non-uniqueness of optimal weights, which leads to the non-uniqueness of cross-efficiencies. To alleviate this, **Doyle and Green (1994)** recommended the adoption of alternative secondary goals in an aim to select unique optimal multipliers. In particular, they introduced an aggressive and a benevolent model, while the secondary objective functions in **Liang et al. (2008)** reflected the minimisation of total deviation, maximum deviation, and mean absolute deviation from an 'ideal' point. The interested reader could also check **Wang and Chin (2010a), Wang et al. (2011), Wu et al. (2012), Wu et al. (2016)** and **Li et al. (2018)**. The non-uniqueness issue is also critical in a network system. **Kao and Liu (2019)** developed an aggressive CE model to measure the efficiency in two basic network structures. **Örkcü et al. (2019)** 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.

The aggregation of the cross-efficiency scores is another issue in CE. An appropriate aggregation strategy can enable the DMUs to accept their ranking. Although the average method has proven effective in ensuring a credible ranking **(Liang et al., 2008, Wang and Chin, 2010b)**, it loses sight of the weights assigned to scores **(Wang and Wang, 2013)**. To accommodate this issue, **Wu et al. (2011)** utilised the Shannon entropy, allocating a fixed but different weight to each DMU. **Wu et al. (2012a)** highlighted that this is problematic, since it ignores the primary role of the self-evaluated efficiency of each DMU. They, thus, embedded the Shannon entropy into the CE by considering the association among the self and the peer-evaluation values. See also more recent work by **Wang and Chin (2011)**, **Wang and Wang (2013)**, and **Song and Liu (2018)**.

Fairness in the evaluation outcomes has been achieved even via the integration of game theoretic concepts within traditional single-stage and two-stage DEA networks. For instance, **Zhou et al. (2013)** introduced a Nash bargaining game model to obtain a unique efficiency decomposition for the two constituent sub-stages of the centralized model. Their approach leads to a fair context, in that it reflects how the two sub-stages bargain with each other for better efficiencies. **An et al. (2017)** also used Nash bargaining, but introduced a framework for setting fair target values for intermediate products of two-stage systems, so that the two stages are encouraged to collaborate with each other

[1] No attempt is made to give a formal definition of fairness, but aspects which might reasonably be considered to contribute to this are discussed throughout this chapter.

within a pre-agreed range of fair outcomes. **Wu et al. (2016a)** proposed a CE evaluation approach based on Pareto improvement. A merit of their approach is that it always generates a set of Pareto optimal cross-efficiencies for the DMUs. **Li (2017)** introduced a sequence of leader-follower procedures as to ensure a fair evaluation in the sense that it guarantees that the same result is obtained for the second (=follower) stage of a DMU as would be obtained applying the standard DEA model to the second stage independently. A number of studies have been reported in this direction, such as **Yu and Shi (2014)**, **Ma et al. (2014)**, and **Li et al. (2018)**.

	Type of network	Cross-efficiency	Aggregation method	Game approach	Efficiency measurement
Zhou et al. (2013)	two-stage	×	×	Nash bargaining	decomposition
Yu and Shi (2014)	two-stage	×	×	cooperative & leader-follower	×
Ma et al. (2014)	two-stage	centralized	arithmetic average	non-cooperative inspired	decomposition
Wu et al. (2016a)	single-stage	Pareto improvement	arithmetic average	Pareto optimality	n/a
An et al. (2017)	two-stage	×	×	Nash bargaining	×
Li (2017)	two-stage	×	×	cooperative & leader-follower	×
Li et al. (2018)	single-stage	optimal balanced	balanced adjustment	game-like iterative algorithm	n/a
Örkcü et al. (2019)	two-stage	neutral	geometric average	×	×
This Paper	two-stage	minimax	CRITIC	×	aggregation

FIGURE 2.1: Related literature on fairness evaluation in DEA.

In summary, fairness in the evaluation of DMUs has been extensively explored via CE towards single-stage and basic network structures. Nevertheless, when the discussion shifts to more complex structures where inputs and outputs are shared among different processes, there is a limited attention to how to achieve more meaningful results for the DMUs. This intricacy is due to the additional inputs in the second stage obtained from the external environment and the dual role of the intermediate products. There are several enlightening applications, especially in logistics, supply chain, and manufacturing, that could justify the necessity of exploring fairness in the performance evaluation of a generalised two-stage structure. These are discussed in more depth with an example in Section 2.3 and in the implications of Section 2.4.2.

In our paper, we firstly introduce an additive self-efficiency aggregation model that can highlight the strength of each sub-stage and obtain the most favourable efficiency for the DMU overall. Since the optimal set of weights derived from the aggregation model may not be unique, we employ a minimax secondary goal model. The reasons to the adoption of this model are twofold: *(i)* it corresponds to cooperative situations **(Liang et al., 2008b)**, since sub-stages behave benignly, and *(ii)* it is compatible with multi-stage systems where the individual sub-stages pursue mutual cooperation via the maximisation of the overall efficiency **(Yu and Shi, 2014)**. Although other approaches such as the minimisation of the total deviation from the ideal point **(Liang et al., 2008)** were also considered for the exploration of a unique set of optimal multipliers, they were eventually deemed inappropriate for the goals that this study pursues to attain. For this case of minimising the sum of inefficiencies **(Doyle and Green, 1994; Liang et al.,**

2008), each DMU aims to optimise their own benefits while disrespecting the status of the other DMUs. Hence, they act in a non-cooperative environment; we believe that this characteristic is not compatible with the general aims of our paper towards ensuring a fairer evaluation and ranking outcome. The selected multi-objective model is converted using the Compromise Programming methodology as a means to identify a good solution that balances the objectives.

On the aggregation of the individual CE, existing frameworks **(Wang and Chin, 2011; Wu et al., 2012a)** pay attention to the reasonable allocation of the weights by limiting the range between self and peer-assessment efficiencies. This condition may indicate consistency from the perspective of the majority opinion. However, considering that many organisations are moving towards systems of evaluation in which also the opinions of minorities are valued **(Park and DeShon, 2010)**, we introduce an aggregation method that rewards contrast. We rely upon the CRiteria Importance Through Intercriteria Correlation (CRITIC) method **(Diakoulaki et al., 1995)**, an objective method for eliciting weights in multi-criteria problems. With the exception of **He and Ma (2015)**, our paper is the first to apply the CRITIC method in the context of DEA. Its novel function and meaning as deployed in the paper is further described in Section 2.3.3.2, and differences with the above study are discussed in Section 2.4.2. Besides, CRITIC would be compatible with the minimax model introduced herein; this is justified by model's nature to highlight the best behaviour of the worst-performing unit, while the scores of the other better-performing units might decrease.

The remainder of the chapter is organised as follows. Section 2.2 describes the methodological background. In Section 2.3, we develop the alternative modelling approach for the generalised two-stage DEA structure. Section 2.4 illustrates the methods with a numerical example. Section 2.5 presents conclusions and further research.

2.2 Methodological Background

In the typical input-oriented CCR DEA model **(Charnes et al., 1978)**, each DMU_j $(j = 1, 2, ..., n)$ uses m inputs $(i = 1, 2, ..., m)$ to produce s outputs $(r = 1, 2, ..., s)$. 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 $j \in N$. These values are known and non-negative. The multiplier-form model that evaluates the efficiency of the target DMU_k is the following:

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

where μ_{rk}^*, ν_{ik}^* are the rth output and the ith input virtual optimal multipliers, respectively. These are unknown decision variables and they are determined by the linear program. If the set of non-negative optimal multipliers makes the associated objective function equal to 1, then the target DMU_k is called DEA efficient; otherwise, it is called DEA non-efficient.

Two significant challenges of the black-box DEA model, recall the discussion in Section 2.1, are to acquire a unique ranking order of the existing DMUs (dealing with the lack of discrimination power) and to obtain a more realistic weight scheme **(Örkcü et al., 2019)**. They are inter-related and concurrent **(Li and Reeves, 1999)**.

2.2.1 Cross-efficiency concept

A commonly used approach to overcome these inabilities is the cross-efficiency (CE) evaluation, proposed by **Sexton et al. (1986)**. Conventional DEA models provide a self-appraisal of the DMUs, using their own optimal weights **(Örkcü et al., 2019)**. Assume that for model **(1)**, μ_{rk}^*, ν_{ik}^* formulate the optimal set of multipliers. Based on this optimal solution, DMU_k is characterised as efficient if and only if $E_{kk}^* = 1$ **(Charnes et al., 1978)**. Model **(2.1)** needs to be resolved for each DMU (in total n times) to obtain an optimal set of weights for the corresponding DMU. Then by applying the cross-efficiency concept, in which peer-appraisal is the main idea, we evaluate each DMU, considering the weight profiles of all DMUs. In particular, $E_{kj} = \sum_{r=1}^{s} \mu_{rk}^* Y_{rj} / \sum_{i=1}^{m} \nu_{ik}^* X_{ij}$ indicates the individual cross-efficiency of the DMU_j, according to the optimal weighting scheme of DMU_k. A cross-efficiency matrix is a valuable tool for such cases. In this matrix, elements E_{kj} depict the peer-efficiency scores of DMU_j, based on the optimal weights of DMU_k. The diagonal elements of the same matrix indicate the self-efficiency scores of DMU_k. The cross-efficiency score that attributes the final rank of a DMU, is usually estimated by averaging all individual cross-efficiencies of the corresponding DMU which is being evaluated. Thus, $\hat{e}_j = \frac{1}{n} \cdot \sum_{k=1}^{n} E_{kj}\ (j = 1, 2, ..., n)$ **(Anderson et al., 2002)**.

A key difficulty of the CE evaluation is that the optimal weights obtained by model **(2.1)** may not be unique, resulting in the non-uniqueness of cross-efficiency scores and rankings of DMUs. To tackle this difficulty, **Doyle and Green (1994)** proposed the use of aggressive and benevolent models, as alternative secondary goals. Model **(2.2)** is the aggressive. It maximises the performance of the DMU under consideration while minimising the cross-efficiencies of all other DMUs. Model **(2.3)** is the benevolent that ensures the maximisation of the cross-efficiencies of all other DMUs, whilst maintaining the performance of the target DMU.

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

$$Max \quad \sum_{r=1}^{s} \mu_{rk} \Big(\sum_{j=1, j \neq k}^{n} Y_{rj} \Big) \tag{2.3}$$

subject to the same constraints as in model (2.2).

Troutt (1997) and later **Liang et al. (2008)** developed a novel secondary goal, based on the minimisation of the maximum k-inefficiency (or deviation) score. By identifying an optimal set of multipliers that assigns the maximum efficiency score to the DMU with the worst performance, they achieved the reduction of deviations among all the other DMUs. Hence, they presented the following linear programming model, where $\alpha_k^* = 1 - E_{kk}^*$:

$$\begin{aligned}
Min \quad & \theta_k \\
\text{subject to} \quad & \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} + \alpha_j = 0, \forall j, \\
& \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rk} = 1 - \alpha_k^{*}, \\
& \theta_k - \alpha_j \geq 0, \forall j, \\
& \mu_{rk}, \nu_{ik}, \alpha_j, \theta_k \geq 0, \forall\ r, i, j.
\end{aligned} \tag{2.4}$$

Model **(2.4)** corresponds to a cooperative situation towards a single-stage DEA structure. In Section 2.3.2, it will be amended and customised to the specifications of the generalised two-stage structure to accommodate the purposes of our DEA methodology.

2.3 Models Development

Yu and Shi (2014) recommended a DEA structure in which each DMU consists of two sub-stages connected in series, as in Figure 2.2. The initial inputs X_{ij} (where $i = 1, 2, ..., m$) entering stage 1 are converted into intermediate products Z_{dj} (where $d = 1, 2, ..., D$). Part of intermediate products $\alpha_{dj} Z_{dj}$ is consumed during stage 2, and the remaining part $(1 - \alpha_{dj}) Z_{dj}$ is channeled out of the system as final output. α_{dj} is the allocation proportion, dividing this intermediate product into the aforementioned two parts, where $0 \leq \alpha_{dj} \leq 1$. In stage 2, additional inputs X^2_{hj} (where $h = 1, 2, ..., H$) are also supplied from outside. Finally, Y_{rj} (where $r = 1, 2, ..., s$) are the outputs from stage 2 produced for outside.

Note that α_{dj} is pre-specified externally by the decision maker; it is therefore an observed rather than a decision value, that is subjectively designated prior to solving the corresponding mathematical model. This conceptual idea contrasts with the handling of α_{dj} as a variable, according to **Yu and Shi (2014)**. Our decision to illustrate α_{dj} as an observed value determined by the decision maker (externally) and not the model (internally) may represent the reality better, reflecting for example: the market conditions, the contractual requirements, the produced quantity of sub-stage 1, and the alternating requirements and needs of the decision-maker.

To gain a better understanding of the reason we have selected α_{dj} as an observed value, we can refer to a real-life example that clearly describes the two-stage structure (Figure 2.2). A stock-farmer in a cattle farm (DMU) feeds with corn, wheat, and pasture land (initial inputs in stage 1) dairy cows to produce raw milk (intermediate product at the end of stage 1). The farmer ought to decide how much quantity of the produced milk will be further processed (part of intermediate product as input of stage 2) to get butter, cheese, and yoghurt (final output), and how much quantity will be directly allocated to the outside market (remaining intermediate product as final output). Finally, the fungi for the flash pasteurisation of milk could be an additional (exogenous) input of stage 2. In this example, the decision maker i.e. the stock-farmer freely determines beforehand the way to utilise the produced quantity of milk. Evidently, his decision could be influenced by the laws of supply and demand, the production capacity of the cattle farm, and/or the state of health of the cows.

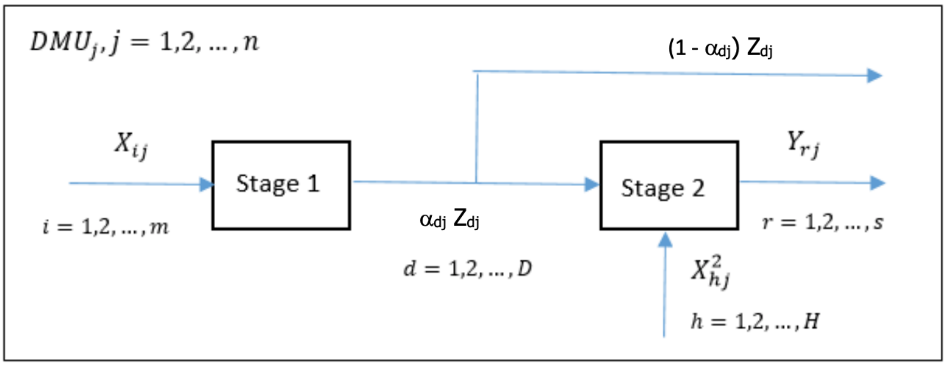

FIGURE 2.2: The *generalised* two-stage structure; **Yu and Shi (2014)**.

2.3.1 Additive efficiency aggregation

The constant-returns-to-scale (CRS) efficiency scores for the target DMU$_k$ can be calculated by the following two CCR models, respective to the first and second stage; they are based upon the CCR model **(Charnes et al., 1978)**:

$$
\begin{aligned}
E_{kk}^{CCR_1} = Max \quad & \frac{\sum_{d=1}^{D} \eta_{dk} Z_{dk}}{\sum_{i=1}^{m} v_{ik} X_{ik}} \\
\text{subject to} \quad & \frac{\sum_{d=1}^{D} \eta_{dk} Z_{dj}}{\sum_{i=1}^{m} v_{ik} X_{ij}} \le 1, \forall j, \\
& \eta_{dk}, v_{ik} \ge 0, \forall\ d, i.
\end{aligned}
\tag{2.5}
$$

$$
\begin{aligned}
E_{kk}^{CCR_2} = Max \quad & \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rk} \ + \sum_{d=1}^{D} \eta_{dk}(1-\alpha_{dk}) Z_{dk}}{\sum_{h=1}^{H} q_{hk} X_{hk}^2 \ + \sum_{d=1}^{D} \eta_{dk}\alpha_{dk} Z_{dk}} \\
\text{subject to} \quad & \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rj} \ + \sum_{d=1}^{D} \eta_{dk}(1-\alpha_{dj}) Z_{dj}}{\sum_{h=1}^{H} q_{hk} X_{hj}^2 \ + \sum_{d=1}^{D} \eta_{dk}\alpha_{dj} Z_{dj}} \le 1, \forall j, \\
& \eta_{dk}, \mu_{rk}, q_{hk} \ge 0, \forall\ d, r, h.
\end{aligned}
\tag{2.6}
$$

Yu and Shi (2014) do not measure $(1-\alpha)Z$ flows as outputs of the second stage, which is a note that merits a comment. This might make sense when part of the outflow of stage 1 is directly forwarded to an outside market without affecting the remaining outflow processed in stage 2. In this way, stage 2 does not need to consider the trade-off and the two sub-stages act rather as being independent. On the other hand, our model **(2.6)** measures the $(1-\alpha)Z$ flows. This conceptual difference justifies our motivation to examine how the outflows of stage 1 (intermediate products) are split into two distinctive instances which interact with one another. As a further reason of the inclusion of the $(1-\alpha)Z$ flows in our study, we draw attention to the commonly used efficiency aggregation method to build our models. As discussed in **Kao (2017)**, in such a case the efficiency of the system is defined as a function of those of the constituent sub-stages. The intermediate products (αZ flows, $(1-\alpha)Z$ flows) should be initially involved in measuring the efficiency of the corresponding sub-stage and then in calculating the overall efficiency. The α value will be freely determined by the decision-maker considering the amount of the respective intermediate measure to be directly forwarded outside of the system and the amount to be further processed in the second sub-stage; the decision-maker's option will reasonably be impacted by the conditions of the market, the availability of the produced quantity, and the laws of supply and demand.

The system efficiency of the DMU_k can be computed from the following CCR model **(2.7)**. Its objective function illustrates the ratio of the aggregate exogenous outputs to that of the aggregate exogenous inputs, considering only the operations of the entire system.

$$
\begin{aligned}
E_{kk}^{CCR} = Max \quad & \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rk} + \sum_{d=1}^{D} \eta_{dk}(1-\alpha_{dk}) Z_{dk}}{\sum_{i=1}^{m} v_{ik} X_{ik} + \sum_{h=1}^{H} q_{hk} X_{hk}^{2}} \\
\text{subject to} \quad & \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rj} + \sum_{d=1}^{D} \eta_{dk}(1-\alpha_{dj}) Z_{dj}}{\sum_{i=1}^{m} v_{ik} X_{ij} + \sum_{h=1}^{H} q_{hk} X_{hj}^{2}} \leq 1, \forall j, \\
& \eta_{dk}, \mu_{rk}, v_{ik}, q_{hk} \geq 0, \forall\, d, r, i, h,
\end{aligned} \tag{2.7}
$$

$\eta_{dk}, \mu_{rk}, v_{ik}, q_{hk}$ correspond to the weights associated with intermediate measure d, output r and inputs i and h, for the DMU_k, respectively. Note that the weights (or multipliers) of the intermediate measures are assumed to be the same for both sub-stages **(Kao and Hwang, 2008)**.

Model **(2.7)** disregards the internal operations of DMUs and treats each DMU as a black box that uses exogenous inputs to produce exogenous outputs. Neglecting the internal operations of DMUs could spur us to results that are not accurate. For instance, while the overall system could be characterised as efficient, one or both of its individual stages may be inefficient. This is one of the main reasons why we need to examine and model the operations of the internal structures for each DMU.

To accommodate the aforementioned issue, we explore the efficiency aggregation method as previously discussed. It is known that it might take either an additive or a multiplicative form depending on the nature of the problem. To **Chen et al. (2009)**, additive efficiency aggregation models are a way of aggregating components in a two-stage structure. This type of aggregation requires the allocation of a relative importance weight to each sub-stage. The weights can be user-specified. They can alternatively be DMU-specific to recognise the strength of each stage as well as the discrepancies between them, and to facilitate the transformation of the non-linear model to a linear one **(Guo**

et al., 2017). As discussed in **Kao (2016)**, the DMU-specific weights will obtain the most favourable efficiency for the system under evaluation. We believe that this might be a reason towards ensuring a fairer and more cooperative environment for the competing DMUs. This approach can also estimate how much more the inputs of the system can be reduced, while ensuring the same level of output production. Finally, it is applicable to both constant and variable returns-to-scale assumptions.

On the other hand, the multiplicative efficiency aggregation method does not require predetermined weights for building the model. Nevertheless, it can put less emphasis on the intermediate products that are being exchanged between the sub-stages of a DMU, whereas a weighted aggregation method does and thus better reflects the level of cooperation between the stages of a DMU. In addition, when it handles a generalised two-stage network structure with exogenous outputs leaving from stage 1 and/or exogenous inputs entering to stage 2, it is extremely nonlinear and cannot be easily converted into a linear model using the Charnes-Cooper transformation. Even the utilisation of a heuristic search method cannot guarantee a global optimal solution **(Chen and Zhu, 2017)**. For the above reasons, this study selects to define the system efficiency as the weighted (arithmetic mean) approach **(Chen et al., 2009)** of its two sub-stage efficiencies.

$$\left(w_k^1 \cdot \frac{\sum_{d=1}^{D} \eta_{dk} Z_{dk}}{\sum_{i=1}^{m} v_{ik} X_{ik}} + w_k^2 \cdot \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rk} + \sum_{d=1}^{D} \eta_{dk}(1-\alpha_{dk}) Z_{dk}}{\sum_{h=1}^{H} q_{hk} X_{hk}^2 + \sum_{d=1}^{D} \eta_{dk}\alpha_{dk} Z_{dk}}\right), \tag{2.8}$$

where w_k^1 and w_k^2 are weights determined by the decision-maker, so that $w_k^1 + w_k^2 = 1$. These weights are not unknown variables, but functions of the optimisation variables. According to **Chen et al. (2009)**, these weights could be DMU-specific, in that they could represent the size/strength of a particular stage. **Guo et al. (2017)** treated the weights of the two sub-stages as parameters (pre-defined by the decision-maker) that can vary between 0 and 1. We can, thus, estimate the overall efficiency of the DMU_k by solving model **(2.9)**.

$$\begin{aligned}
Max \quad & \left(w_k^1 \cdot \frac{\sum_{d=1}^{D} \eta_{dk} Z_{dk}}{\sum_{i=1}^{m} v_{ik} X_{ik}} + w_k^2 \cdot \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rk} + \sum_{d=1}^{D} \eta_{dk}(1-\alpha_{dk}) Z_{dk}}{\sum_{h=1}^{H} q_{hk} X_{hk}^2 + \sum_{d=1}^{D} \eta_{dk}\alpha_{dk} Z_{dk}}\right) \\
\text{subject to} \quad & \frac{\sum_{d=1}^{D} \eta_{dk} Z_{dj}}{\sum_{i=1}^{m} v_{ik} X_{ij}} \leq 1, \forall j, \\
& \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rj} + \sum_{d=1}^{D} \eta_{dk}(1-\alpha_{dj}) Z_{dj}}{\sum_{h=1}^{H} q_{hk} X_{hj}^2 + \sum_{d=1}^{D} \eta_{dk}\alpha_{dj} Z_{dj}} \leq 1, \forall j, \\
& w_k^1 + w_k^2 = 1, \\
& w_k^1, w_k^2, \eta_{dk}, \mu_{rk}, v_{ik}, q_{hk} \geq 0, \forall\ d, r, i, h.
\end{aligned} \tag{2.9}$$

Weights w_k^1 and w_k^2 represent the relative importance of the performances of stages 1 and 2 respectively, divided by the overall performance of the evaluated DMU. A larger weight indicates the corresponding stage's stronger effect on the entire performance of the system. To **Chen et al. (2009)** and **Kao (2016)**, the portions of total resources devoted to each stage could correspond to the relative size of a stage. This is also due to the nature of the models which are input-oriented. Therefore, we define:

$$w_k^1 = \frac{\sum_{i=1}^{m} \nu_{ik} X_{ik}}{\sum_{i=1}^{m} \nu_{ik} X_{ik} \ + \sum_{h=1}^{H} q_{hk} X_{hk}^2 \ + \sum_{d=1}^{D} \eta_{dk} \alpha_{dk} Z_{dk}} \tag{2.10}$$

and

$$w_k^2 = \frac{\sum_{h=1}^{H} q_{hk} X_{hk}^2 \ + \sum_{d=1}^{D} \eta_{dk} \alpha_{dk} Z_{dk}}{\sum_{i=1}^{m} \nu_{ik} X_{ik} \ + \sum_{h=1}^{H} q_{hk} X_{hk}^2 \ + \sum_{d=1}^{D} \eta_{dk} \alpha_{dk} Z_{dk}}. \tag{2.11}$$

Substituting **(2.10)** and **(2.11)** into the objective function of model **(2.9)**, we obtain the following linear fractional programming model:

$$\begin{aligned}
E_{kk}^{CCR} = Max \quad & \frac{\sum_{d=1}^{D} \eta_{dk} Z_{dk} \ + \sum_{r=1}^{s} \mu_{rk} Y_{rk} \ + \sum_{d=1}^{D} \eta_{dk}(1-\alpha_{dk}) Z_{dk}}{\sum_{i=1}^{m} \nu_{ik} X_{ik} \ + \sum_{h=1}^{H} q_{hk} X_{hk}^2 \ + \sum_{d=1}^{D} \eta_{dk} \alpha_{dk} Z_{dk}} \\
\text{subject to} \quad & \frac{\sum_{d=1}^{D} \eta_{dk} Z_{dj}}{\sum_{i=1}^{m} \nu_{ik} X_{ij}} \leq 1, \forall j, \\
& \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rj} \ + \sum_{d=1}^{D} \eta_{dk}(1-\alpha_{dj}) Z_{dj}}{\sum_{h=1}^{H} q_{hk} X_{hj}^2 \ + \sum_{d=1}^{D} \eta_{dk} \alpha_{dj} Z_{dj}} \leq 1, \forall j, \\
& \eta_{dk}, \mu_{rk}, \nu_{ik}, q_{hk} \geq 0, \forall\ d, r, i, h.
\end{aligned} \tag{2.12}$$

By applying the variable substitution technique in **Charnes** and **Cooper (1962)** and by replacing $\eta_{dk}\alpha_{dk} = \phi_{dk}^1$ and $\eta_{dk}(1-\alpha_{dk}) = \phi_{dk}^2$, we introduce the self-evaluation CCR performance score model **(2.13)**, which is equivalent to model **(2.12)**. According to the following (implicitly) linear model, it is possible to measure the performance for each DMU, whose internal structure is illustrated by the two-stage DEA process of Figure 2.2. This relational model estimates the aggregated system efficiency while considering the internal mechanisms of its individual stages.

$$
\begin{aligned}
&E_{kk}^{CCR} = Max \quad \sum_{d=1}^{D} \eta_{dk} Z_{dk} \ + \sum_{r=1}^{s} \mu_{rk} Y_{rk} \ + \sum_{d=1}^{D} \phi_{dk}^{2} Z_{dk} \\
&\text{subject to} \quad \sum_{i=1}^{m} \nu_{ik} X_{ik} \ + \sum_{h=1}^{H} q_{hk} X_{hk}^{2} \ + \sum_{d=1}^{D} \phi_{dk}^{1} Z_{dk} = 1, \\
&\sum_{d=1}^{D} \eta_{dk} Z_{dj} \ - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \forall j, \\
&\sum_{r=1}^{s} \mu_{rk} Y_{rj} \ + \sum_{d=1}^{D} \phi_{dk}^{2} Z_{dj} \ - \sum_{h=1}^{H} q_{hk} X_{hj}^{2} \ - \sum_{d=1}^{D} \phi_{dk}^{1} Z_{dj} \leq 0, \forall j, \\
&\eta_{dk}, \mu_{rk}, \nu_{ik}, q_{hk} \geq 0, \ \ \phi_{dk}^{1} + \phi_{dk}^{2} = \eta_{dk}, \ \ \phi_{dk}^{1} \geq 0, \ \ \phi_{dk}^{2} \geq 0, \ \forall \ d, r, i, h.
\end{aligned}
\tag{2.13}
$$

At the optimality of model **(2.13)**, the system efficiency, the efficiency of sub-stage 1, and the efficiency of sub-stage 2, respectively, are computed as follows:

$$
E_{kk}^{CCR} = (\textstyle\sum_{d=1}^{D} \eta_{dk}^{*} Z_{dk} + \sum_{r=1}^{s} \mu_{rk}^{*} Y_{rk} + \sum_{d=1}^{D} \phi_{dk}^{2*} Z_{dk}) / \\
(\textstyle\sum_{i=1}^{m} \nu_{ik}^{*} X_{ik} \ + \sum_{h=1}^{H} q_{hk}^{*} X_{hk}^{2} \ + \sum_{d=1}^{D} \phi_{dk}^{1*} Z_{dk}),
$$

$$
E_{kk}^{1} = (\textstyle\sum_{d=1}^{D} \eta_{dk}^{*} Z_{dk}) / (\sum_{i=1}^{m} \nu_{ik}^{*} X_{ik}),
$$

$$
E_{kk}^{2} = (\textstyle\sum_{r=1}^{s} \mu_{rk}^{*} Y_{rk} + \sum_{d=1}^{D} \phi_{dk}^{2*} Z_{dk}) / (\sum_{h=1}^{H} q_{hk}^{*} X_{hk}^{2} + \sum_{d=1}^{D} \phi_{dk}^{1*} Z_{dk}).
$$

2.3.2 Proposed cross-efficiency model

Model **(2.13)** searches for the optimal most favourable weights $\eta_{dk}, \mu_{rk}, \nu_{ik}, q_{hk}, \phi_{dk}^{1}, \phi_{dk}^{2}$ to yield an optimistic self-efficiency score for DMU_k. However, this DEA flexibility of the DMU_k in choosing its own weights could sometimes lead to an unrealistically high efficiency score of the corresponding DMU. This results in a lack of discrimination power and therefore in unrealistic weight distribution. Besides, the optimal solution for model **(2.13)** may not be unique, reducing the theoretical value of the potential results **(Mahdiloo et al., 2016)**. A point to focus on in this chapter is the best possible treatment of the limited discriminatory power and the unrealistic weight distribution, for the two-stage structure (Figure 2.2).

To overcome these weaknesses, we apply the cross-efficiency concept in the two-stage DEA structure that we examine. We initially propose an alternative secondary goal model to mainly encounter the shortcoming of the non-unique optimal set of multipliers of model **(2.13)**. This model contains two objective functions (i.e. criteria); each of them represents one of the two stages of the whole system. These two criteria need to be optimised simultaneously.

To advance our multiple criteria-based secondary goal, we have been influenced by the concept of the "minimisation of the maximum k-inefficiency" **(Troutt, 1997; Liang et al., 2008)**. In the minimax model **(2.14)**, there are two independent objective functions. The first objective ($\theta 1$) represents the situation in which we have to minimise the maximum deviation of stage 1 among all DMUs. The second objective ($\theta 2$) illustrates the minimisation of the maximum deviation of stage 2 among all DMUs. There is no former preference order between these criteria. Considering the theoretical framework of the concept, this approach might be proved useful in cooperative situations **(Liang et al., 2008)**. In our case this is vital, as the two stages that constitute the entire system should have the same bargaining power and should cooperate in order to maximise the overall efficiency **(Halkos et al., 2014)**.

$$
\begin{aligned}
&Min && \theta 1\\
&Min && \theta 2\\
&\text{subject to} && \sum_{i=1}^{m} \nu_{ik} X_{ik} + \sum_{h=1}^{H} q_{hk} X_{hk}^2 + \sum_{d=1}^{D} \phi_{dk}^1 Z_{dk} = 1,\\
& && \sum_{d=1}^{D} \eta_{dk} Z_{dk} + \sum_{r=1}^{s} \mu_{rk} Y_{rk} + \sum_{d=1}^{D} \phi_{dk}^2 Z_{dk} = E_{kk}^{CCR^*},\\
& && \sum_{d=1}^{D} \eta_{dk} Z_{dj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} + b_j^1 = 0, \forall j,\\
& && \sum_{r=1}^{s} \mu_{rk} Y_{rj} + \sum_{d=1}^{D} \phi_{dk}^2 Z_{dj} - \sum_{h=1}^{H} q_{hk} X_{hj}^2 - \sum_{d=1}^{D} \phi_{dk}^1 Z_{dj} + b_j^2 = 0, \forall j,\\
& && \theta_1 \geq b_j^1, \forall j,\\
& && \theta_2 \geq b_j^2, \forall j,\\
& && b_j^1, b_j^2, \eta_{dk}, \mu_{rk}, \nu_{ik}, q_{hk} \geq 0, \forall\, j, d, r, i, h,\\
& && \phi_{dk}^1 + \phi_{dk}^2 = \eta_{dk},\ \ \phi_{dk}^1 \geq 0,\ \ \phi_{dk}^2 \geq 0, \forall\, d.
\end{aligned}
\tag{2.14}
$$

In the above model, $E_{kk}^{CCR^*}$ denotes the optimal objective function value of model **(2.13)**. The reason why we are using the restrictions "$\theta_1 \geq b_j^1$" and "$\theta_2 \geq b_j^2$" (where $j = 1, 2, ..., n$) is to set θ_1 as the maximum deviation of stage 1, and θ_2 as the maximum deviation of stage 2. This model pursues to result in an optimal set of weights that will highlight the best behaviour of the worst-performing DMU, underpinning the fairness in the decision-making process.

Model **(2.14)** is a bi-objective programming model that can hardly obtain a global optimal solution. A multi-objective program usually provides a set of non-dominated solutions (see **Li and Reeves, 1999**). The researcher could either apply the objectives interactively (**Mahdiloo et al., 2016**) or identify an alternative way of satisfying the conditions simultaneously. Goal programming has been proposed for optimising all criteria at the same time **(Ghasemi et al., 2014; dos Santos Rubem et al., 2017)**.

We apply the concept of **dos Santos Rubem et al. (2017)** to convert the MOLP model **(2.14)** into a goal programming model. However, given the utopian values assigned to each of the two objective functions (goals), the model should be aligned more closely to Compromise Programming. Moreover, since $b_j^1, b_j^2 \geq 0, \forall\ j$, it follows that $\theta_1, \theta_2 \geq 0$ and thus there is no need to use the negative deviations, d_1^- and d_2^-, in such a model. Actually $\theta_1 = d_1^+$ and $\theta_2 = d_2^+$. Hence, the model is just formulated as follows:

$$\begin{aligned} &Min \quad \theta 1 + \theta 2 \\ &\text{subject to the same constraints as in model (2.14).} \end{aligned} \tag{2.15}$$

Model **(2.15)** is the proposed minimax secondary goal model for the two-stage structure (Figure 2.2) in this chapter and is run under the CRS assumption. It is seeking a particular solution on the Pareto frontier of model **(2.14)** i.e. one with equally weighted deviations. This model can significantly reduce the number of zero weights assigned to the known factors and better discriminate the DEA-efficient DMUs.

2.3.3 Alternative aggregation approach

Recalling the discussion in Section 2.1, we are going to calculate the individual cross-efficiencies, based on the representative optimal weights from model **(2.15)**. In addition, we will determine the cross-efficiencies to get the final ranks of the considered DMUs, based on the CRITIC method.

2.3.3.1 Individual & ultimate cross-efficiencies

Like all DEA models for cross-efficiency evaluation, the proposed secondary model **(2.15)** needs to be solved n times, once for every DMU. There will be n sets of input, intermediate measure and output weights available for cross-efficiency evaluation. According to **Kao and Liu (2019)**, in a series DEA structure as the one we probe in this paper, the discriminatory power is stronger due to the increasing number of restrictions; thus, there are less chances that the optimal set of multipliers derived from the first secondary goal model for each DMU is non-unique. Therefore, we can adopt **Kao and Liu's (2019, p.73)** belief that this optimal set is "*representative enough*" for our analysis. At the optimality of model **(2.15)**, for each DMU_j $(j \neq k)$,

$E_{kj} = (\sum_{d=1}^{D} \eta_{dk}^{*} Z_{dj} \ + \sum_{r=1}^{s} \mu_{rk}^{*} Y_{rj} \ + \sum_{d=1}^{D} \phi_{dk}^{2*} Z_{dj}) /$
$(\sum_{i=1}^{m} \nu_{ik}^{*} X_{ij} \ + \sum_{h=1}^{H} q_{hk}^{*} X_{hj}^{2} \ + \sum_{d=1}^{D} \phi_{dk}^{1*} Z_{dj}),$

$E_{kj}^{1} = (\sum_{d=1}^{D} \eta_{dk}^{*} Z_{dj}) / (\sum_{i=1}^{m} \nu_{ik}^{*} X_{ij}),$

$E_{kj}^{2} = (\sum_{r=1}^{s} \mu_{rk}^{*} Y_{rj} \ + \sum_{d=1}^{D} \phi_{dk}^{2*} Z_{dj}) / (\sum_{h=1}^{H} q_{hk}^{*} X_{hj}^{2} \ + \sum_{d=1}^{D} \phi_{dk}^{1*} Z_{dj}).$

These are referred to as the cross-efficiency values of the DMU_j of the overall system, of stage 1 and of stage 2, according to the optimal weight scheme of DMU_k respectively, and reflect the peer-evaluation of DMU_j.

For each DMU_j, the weighted average cross-efficiency score, produced by the weighted cross-efficiency aggregation is the following:

$$\hat{e}_j = \frac{\sum_{k=1}^{n} w_k \cdot E_{kj}}{\sum_{k=1}^{n} w_k}, \quad \hat{e}_j^1 = \frac{\sum_{k=1}^{n} w_k \cdot E_{kj}^1}{\sum_{k=1}^{n} w_k}, \quad \hat{e}_j^2 = \frac{\sum_{k=1}^{n} w_k \cdot E_{kj}^2}{\sum_{k=1}^{n} w_k}. \tag{2.16}$$

They are called the cross-efficiencies for the overall system, stage 1 and stage 2, respectively. $w_1, ..., w_n$ are the relative importance weights for cross-efficiency aggregation and satisfy the conditions: $w_k \geq 0$ $(k = 1, ..., n)$ and $\sum_{k=1}^{n} w_k = 1$.

2.3.3.2 CRITIC method in DEA

Several subjective and objective weight evaluation methods that have been introduced in the multi-criteria decision-making (MCDM) relevant literature, could obtain the weights in **(2.16)** and solve the aggregation problem. Methods such as the analytic hierarchy process (AHP) and the expert scoring method (Delphi) aim to determine the subjective preference of experts. To overcome the limitation of strong subjectivity and be less dependent on the decision-maker's viewpoint **(Bhadra et al., 2021)**, it is generally suggested to implement objective weighting methods (which are mainly based on the evaluation of the respective data set) for further improvements **(Wu et al., 2019)**. The most widely applied objective methods include the Shannon entropy **(Shao et al., 2020; Pan et al., 2021)**, the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) **(Hwang and Yoon, 1981)**, the VIKOR method **(Amin et al., 2022)**, and the CRITIC method **(Diakoulaki et al., 1995)**.

In this paper, to estimate the weights in **(2.16)**, we apply the CRITIC method, an objective way to determine the relative importance in MCDM situations. This objectivity stems from its formal mathematical procedure and the fact that it is less prone to subjective modifications by a decision-maker. CRITIC considers the evaluation decision-making matrix (in our case the cross-efficiency matrix) to elicit information involved in the evaluation criteria. The emitted information is capable of altering the decision situation and the order of preference. This information delves into two dimensions: the contrast intensity and the conflict among the evaluation criteria **(Diakoulaki et al., 1995)**.

There is a plethora of reasons for selecting the CRITIC method to properly aggregate the individual cross efficiency scores within such a generalised two-stage network DEA structure. Firstly, it has been found that it is more comprehensive and objective than several other well-established objective weighting methods (as cited in **Lu et al., 2022)**. Secondly, several commonly used weight evaluation methods such as AHP, Delphi method, principal component analysis, eigenvalue method, and grey correlation method do not normally explore the internal correlation and the consistency degree among the criteria, resulting in significant deviations from the actual outcomes **(Huang et al., 2018)**. In comparison with the aforementioned methods, the CRITIC method is considered as a more informative option since it examines the variability and conflict among the criteria **(Wu et al., 2020)**. Thirdly, MCDM-relevant literature has shown that entropy-based methods and CRITIC are two of the most commonly used objective weighting methods. As mentioned previously, CRITIC method considers both the contrast intensity and the conflict among the decision criteria. It is also a technique suitable for investigating the trends of individual attributes and the correlations among them. On the other hand, Shannon entropy method only accommodates the contrast intensity **(Peng et al., 2020; Krishnan et al., 2021)** and neglects the interconnections among criteria **(Wu et al., 2018)**.

Below, we will provide an overview of **Diakoulaki et al.'s (1995)** method as we would apply it to single-stage DEA structures; we further explain why it is a sensible tool for promoting fairness, and why it is compatible with the proposed minimax secondary model.

For a finite set A with $j = 1, 2, ..., n$ alternatives and $k = 1, 2, ..., n$ evaluation criteria E_k, the multi-criteria decision making problem is as follows: $Max\ \{E_1(\alpha), E_2(\alpha), ..., E_n(\alpha) \mid \alpha \in A\}$. Initially, we obtain the generalised cross-efficiency matrix (Table 2.1), considering the E_{kj} values for $k, j = 1, 2, ..., n$. See more details of that in Section 2.2.1.

TABLE 2.1: Cross-efficiency matrix; **Doyle and Green (1994)**.

	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}

We proceed to converting the initial cross-efficiency matrix (Table 2.1) into a matrix of relative scores (Table 2.2) with the generic element X_{kj}, where $X_{kj} = (E_k(j) - E_k^{min})/(E_k^{max} - E_k^{min})$. In this mathematical formula, E_k^{max} is equivalent to $max\{E_{k1}, E_{k2}, ..., E_{kn}\}$ and E_k^{min} is equivalent to $min\{E_{k1}, E_{k2}, ..., E_{kn}\}$.

TABLE 2.2: Matrix of relative scores.

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

We generate a vector X_k signifying the scores of all n alternatives $X_k = (X_k(1), X_k(2), ..., X_k(n))$. This vector is characterised by the standard deviation σ_k, which quantifies the contrast intensity of criterion k. Define $\sigma_k = \sqrt{\frac{\sum_{j=1}^{n}(X_k(j) - \hat{X_k})^2}{n}}$, where $\hat{X}_k = \sum_{j=1}^{n} X_k(j)/n$. Then, a symmetric matrix of $n \otimes n$ criteria with R_{kj} elements (Spearman rank correlation coefficients) is constructed (Table 2.3), connecting the rank orders of the elements included in the vector X_k and X_j. Note that, in contrast to the previous two tables, Table 2.3's columns do not list the 'target' DMUs. Instead, each element R_{kj} is a measure of how the degree by which the viewpoint of DMU k as evaluator corresponds to the viewpoint of DMU j as evaluator.

TABLE 2.3: Symmetric matrix.

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

The amount of information C_k emitted by the kth criterion can be determined by multiplying the two measures σ_k (i.e. contrast intensity) and $\sum_{j=1}^{n}(1 - R_{kj})$ (i.e. conflict):

$$C_k = \sigma_k \cdot \sum_{j=1}^{n}(1 - R_{kj}). \tag{2.17}$$

The higher the C_k, the more information we receive from criterion k and the higher its relative importance. Thereby, they define the formula for the weight of criterion k as:

$$w_k = \frac{C_k}{\sum_{l=1}^{n} c_l}. \tag{2.18}$$

The value of the weights w_k ($k = 1, 2, ..., n$) in formula **(2.18)**, can be used to determine the cross-efficiency of DMU_j for the overall system ($\hat{e}_j$), the stage 1 ($\hat{e}_j^1$), and the stage 2 ($\hat{e}_j^2$) in **(2.16)**. CRITIC should, in effect, run three times, based on the investigation of the cross-efficiency matrix of the respective system/stage.

Using the traditional average method, we would assign equal weights $(1/n)$ to everyone's opinion, thus conforming to the majority vote. It would also not matter how diversified or not each of these opinions are. CRITIC, however, emphasises the value of those opinions that are more diversified and less mainstream. In particular, criterion k (here, evaluator DMU_k) will receive more weight if it achieves a wider gap between the best and the worst alternative (here, the target DMUs) in the process of evaluation. This explicitly leads to a higher standard deviation (contrast intensity), implying that its opinion is taken more into account. In other words, the opinion of someone who ranks everyone the same is given less importance, which agrees with the widely accepted viewpoint of **Zeleny (1982)**. This may be justified in the context of DEA, or peer evaluation in general, if the lack of discriminatory signals in the evaluation report of one particular evaluator is believed to represent less reliable information. The only way in which such a viewpoint is able to receive importance would be through the number of evaluators sharing this opinion.

The second feature of the CRITIC method, known as the conflict measure, assigns more weight to the criterion (opinion of evaluator DMU) that puts emphasis on the minority opinion with respect to peer evaluation. The less someone corresponds to a mainstream evaluation profile, the more their opinion is opposed to the majority, the higher their conflict score. This indicates that their opinion will be more valued under these circumstances.

One way to give the application of CRITIC to DEA an interpretation is to say that the CRITIC method infuses a flavour of the 'scientific' approach into a 'political' voting system. Politics is usually in compliance with the majority vote, but in matters of science we often value the most transparent and well-documented opinion. The 'conflict' measure of CRITIC is quite in accordance with the latter viewpoint. However, this analogy is certainly not exact since in science it suffices to have one new opinion that is proven to be correct that can overturn all other opinions (the status quo). CRITIC does not go that far as it does still account for everyone's opinion; the ultimate efficiency measure a DMU receives is still a weighted average.

Another, and perhaps more fruitful interpretation we believe, is that the CRITIC method avoids assigning a too large weight to the majority vote which, by definition, excludes the minority opinion. In this way, it does not let the mass influence too much the public opinion, and in addition, promotes diversity and inclusion. This reflects a contemporary understanding of fairness as an accommodative attitude which is inclusive of a broad variety of legitimate opinion rather than simply mirroring the viewpoint of the majority.

Finally, CRITIC could be compatible with the proposed minimax secondary goal model (see Section 2.3.2), since it rewards contrast intensity. Hence, it is more likely that while the worst-performing DMU attempts to assess itself in its best possible light, the efficiency scores of the other better-performing DMUs might decrease **(Liang et al., 2008)**. Since this situation increases the contrast intensity, our proposed model seems to be an acceptable option to coexist with the CRITIC method.

2.4 Numerical Experiments

This section illustrates the use of the mathematical concepts developed/presented in Section 3 to examine the issue of fairness in DEA context. Our study applies the figures drawn from **Yu and Shi (2014)** for the evaluation of the efficiency of 10 generalised two-stage supply chains of different milk and dairy farm communities. The cattle farms compete with each other, aiming to decide on a sensible allocation of the available raw milk produced. The generalised two-stage DEA structure is considered for this example (see Figure 1), with part of intermediate measures as final outputs and additional inputs in the second stage.

The input resources corn (X_1), wheat (X_2), and pasture land (X_3) are the food of dairy cows consumed by stage 1 to produce raw milk. The raw milk illustrates the intermediate product at the end of stage 1. We distinguish the raw milk between high-fat ($3.5 - 4.5\%$) content and low-fat ($\leq 2.5\%$) content. The former represents the intermediate measure Z_1 and the latter the Z_2. The farmer (i.e. the decision maker) in each community needs to pre-specify how much of these quantities will be further processed in stage 2 and how much will be forwarded to the external environment (i.e. the end-market), as final output. α_{dj} is a proportion, freely determined by the decision-maker, that acts as a regulator of the amount of the dth intermediate measure assigned for processing to stage 2. In this example, we can assume that the stock-farmer has set each α_{dj} equal to 0.7 for simplicity, reflecting market conditions, customer requirements, and updated research surveys; they desire a major proportion of the produced outputs of stage 1 to be further processed in stage 2, while nevertheless channelling a significant quantity as final output. This proportion might consider, for example, the degree to which raw milk contains amino acids, vitamins, minerals, and fatty acids as well as to what extent it is a proper option for those with lactose intolerance, asthma, and allergic conditions. The current observed values of α_{dj} could have been any continuous value between 0 and 1, leading to equally meaningful results. Once the quantity of the respective type of raw milk is processed, the working time for the flash pasteurisation of milk (X_1^2) and the working time for its homogenisation through fine nozzles (X_2^2) will be taken into account. The final (exogenous) outputs will be pasteurised milk (Y_1) and cheese (Y_2). The dataset with the 10 farming communities (DMUs) is summarised in Table 5. 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.

TABLE 2.4: The numerical example of **Yu and Shi (2014)**.

DMU	X_1	X_2	X_3	Z_1	Z_2	X_1^2	X_2^2	Y_1	Y_2
1	9	50	1	20	10	5	8	100	25
2	10	18	10	10	15	7	10	70	20
3	9	30	3	8	20	2	8	96	30
4	8	25	1	20	20	10	10	80	20
5	10	40	5	15	20	5	15	85	15
6	7	35	2	35	10	5	5	90	35
7	7	30	3	10	25	8	10	100	30
8	12	40	4	20	25	4	8	120	10
9	9	25	2	10	10	5	15	110	15
10	10	50	1	20	15	9	10	80	20

2.4.1 Findings

We first consider solving the problem of evaluation and ranking with the classic self-evaluation DEA approach. This serves as a benchmark for comparison with our proposed approach. Table 2.5 exhibits the (non-unique) optimal multipliers from solving the proposed additive self-evaluation two-stage DEA model **(2.13)**, i.e. the basic model without the further model improvements we have introduced in Section 2.3.2. There are 35 zero weights in total, assigned to the respective known factors. The existence of a zero weight indicates that the information of the corresponding known factor is not considered; this prevents us from adhering to the reality. The larger this number of zeros, the more uneven the weight distribution becomes.

Table 2.6 shows the CCR self-efficiency scores and their corresponding rankings of the 10 DMUs for the overall system (E_{kk}^{CCR}), the stage 1 (E_{kk}^{1}), and the stage 2 (E_{kk}^{2}), respectively. Recall that the efficiency scores have been calculated via the optimal weights of model **(2.13)**. DMUs 3,6 and 8 are characterised as DEA-efficient for the overall system, DMUs 1,4,6 and 7 are DEA-efficient for stage 1, and DMUs 3,6,8 and 9 are DEA-efficient for stage 2. Only DMU 6 can be deemed as entirely efficient, since the efficiency of their sub-stages is one. It is evident from the results in Table 2.6 that the DMUs cannot be easily ranked via the self-evaluation method, and from the results in Table 2.5, that this is also based on many flows receiving zero weights and thus not being accounted for.

TABLE 2.5: Optimal multipliers for the proposed self-evaluation model (2.13).

DMU	v_{1k}	v_{2k}	v_{3k}	η_{1k}	η_{2k}	q_{1k}	q_{2k}	μ_{1k}	μ_{2k}
1	0	0	0.3030	0.0152	0	0	0.0606	0.0049	0.0022
2	0	0.0337	0	0.0143	0.0279	0	0	0.0001	0.0087
3	0.0000	0	0	0.0000	0	0.1444	0.0889	0	0.0333
4	0.0513	0.0071	0	0.0109	0.0186	0	0	0.0004	0.0048
5	0.0413	0.0040	0	0.0086	0.0130	0.0064	0.0080	0.0022	0
6	0	0.0140	0	0.0140	0	0	0.0332	0	0.0104
7	0.0592	0.0058	0	0.0123	0.0186	0	0	0.0005	0.0047
8	0	0.0000	0	0.0000	0.0000	0.1000	0.0750	0.0083	0
9	0	0.0217	0.0008	0.0017	0.0254	0.0357	0.0058	0.0034	0
10	0	0	0.5919	0.0278	0.0018	0	0	0.0011	0

We now consider the proposed minimax secondary goal model **(2.15)**. In this manner, we will be able to find flow weights in a cross evaluation approach that exhibit some desirable characteristics. In particular, as discussed in Section 2.3.2, this model will keep the DMU's optimal overall self-efficiency score unchanged, but seeks to minimise the maximum k-inefficiency for each of the stages across all DMUs. Table 2.7 lists the optimal weights from solving model **(2.15)**. The reduction of zero weights compared to the foregoing results of Table 2.5 is noteworthy. In total, there are now only 19 zero weights (compare with 35 in the previous model), improving the weight distribution and providing more balanced results for the evaluated DMUs. The optimal weights from Table 2.7 are subsequently used to calculate the elements of the cross-efficiency matrices for the overall system, stage 1, and stage 2, respectively (see Appendix A, Tables A.1, A.4, and A.7). The latter are the decision-making matrices, whose elements (peer-efficiency scores for each DMU) are found according to the discussion in Section 2.3.3.1.

TABLE 2.6: CCR self-efficiencies for the overall system, stage 1, and stage 2, derived via model (2.13).

DMU	Overall Efficiency E_{kk}^{CCR}	Rank Overall System	Efficiency Stage1 E_{kk}^{1}	Rank Stage1	Efficiency Stage2 E_{kk}^{2}	Rank Stage2
1	0.936	5	**1**	**1**	0.908	5
2	0.909	6	0.924	7	0.886	7
3	**1**	**1**	0.889	8	**1**	**1**
4	0.894	7	**1**	**1**	0.741	8
5	0.690	10	0.676	9	0.709	9
6	**1**	**1**	**1**	**1**	**1**	**1**
7	0.955	4	**1**	**1**	0.890	6
8	**1**	**1**	0.935	6	**1**	**1**
9	0.728	9	0.499	10	**1**	**1**
10	0.844	8	0.985	5	0.640	10

TABLE 2.7: Optimal multipliers for the proposed minimax secondary model (2.15).

DMU	ν_{1k}	ν_{2k}	ν_{3k}	η_{1k}	η_{2k}	q_{1k}	q_{2k}	μ_{1k}	μ_{2k}
1	0.0000	0	0.3030	0.0152	0	0	0.0606	0.0049	0.0022
2	0	0.0337	0	0.0143	0.0279	0	0	0.0001	0.0087
3	0.0000	0.0000	0.0000	0	0.0000	0.0882	0.1029	0.0096	0.0027
4	0.0575	0.0048	0.0089	0.0109	0.0186	0	0.0000	0.0004	0.0048
5	0.0413	0.0040	0	0.0086	0.0130	0.0064	0.0080	0.0022	0.0000
6	0.0561	0.0046	0.0041	0.0111	0.0173	0	0.0087	0.0007	0.0058
7	0.0592	0.0058	0.0000	0.0123	0.0186	0	0	0.0005	0.0047
8	0.0000	0.0000	0	0.0000	0.0000	0.0791	0.0854	0.0082	0.0019
9	0.0000	0.0217	0.0008	0.0017	0.0254	0.0357	0.0058	0.0034	0
10	0.0000	0	0.5919	0.0278	0.0018	0	0	0.0011	0

The next step in our proposed approach is to apply the CRITIC method, see Section 2.3.3.2, to help determine an appropriate weight set for combining the individual cross-efficiency scores into a final cross-efficiency score for each DMU and stage. This technique initially converts the cross-efficiency matrix into a matrix of relative scores for the respective sub-stage, identifying the standard deviation; this indicates the contrast in the viewpoints of the same evaluator DMU_k (see in Appendix A, Tables A.2, A.5, and A.8). It then displays the symmetric matrix for the respective sub-stage, identifying the conflict, the information, and the final weight, for each DMU (see in Appendix A, Tables A.3, A.6, and A.9). Conflict particularly gives voice to the less mainstream opinions of the different evaluators regarding a certain evaluated DMU. An evaluator will be assigned a greater relative importance (final weight) if it provides more valuable information. This information should reward contrast, diversity, and inclusion, in the case of the CRITIC multi-criteria method. As an example, in stage 1, the evaluator (DMU) 1 is assigned the highest final weight (0.119) due to its standard deviation (0.442) and conflict (11.024) measures, which are the highest among their peers. Similarly, in stage 2, the evaluator with the highest final weight (0.128) is DMU3.

Recalling that the weights derived by formula **(2.18)** are used to estimate the final cross-efficiencies in **(2.16)**, see Sections 2.3.3.1 and 2.3.3.2, we display the CRITIC cross-efficiencies for each DMU and stage, in Tables 2.8-2.10. In particular, the CRITIC cross-efficiencies ($\hat{e}_j$) with their respective ranks for the overall system are summarised in the fourth and fifth columns of Table 2.8. The proposed minimax secondary model **(2.15)** evaluated that DMU6 is the most efficient (0.888) and DMU5 has the worst performance (0.532) compared to others; thus, a unique ranking order is achieved. In addition, it can be statistically concluded that the rankings derived from the CRITIC and the traditional average method (third column of table 2.8) are not significantly different based on a Spearman rank correlation coefficient test **(Daniel, 1978)**, with $r_s = 0.94$. This is significant at the 0.01 level (two-tailed).

The CRITIC cross-efficiencies ($\hat{e}_j^1$) with their respective ranks for the stage 1 are exhibited in the fourth and fifth columns of Table 2.9. Model **(2.15)** deemed DMU4 as the most promising DMU (1.000), attaining a unique ranking order once again. The differences between the ranks of CRITIC and average cross-efficiencies (third column of Table 2.9) are also statistically insignificant. With respect to the fourth and fifth columns of Table 2.10 (CRITIC cross-efficiencies and their corresponding ranks for stage 2), DMU3 is located in the first place, with a perfect efficiency score. The dissimilarities with the average cross-efficiency rankings are also negligible based on the Spearman rank correlation test ($r_s = 0.988$). Note that the average cross-efficiencies have been computed following the same reasoning as in CRITIC cross-efficiencies with the sole exception of the method to aggregate the individual cross-efficiencies (see Section 2.2.1). Although their difference is negligible, we consider that the averaging method, privileges the majority vote, and downplays minority opinion by failing to fully respect diversity and the principle of inclusion. CRITIC method fills this gap, assigning more weight to “mavericks” and promoting the modern concept of fairness, as discussed in Section 2.3.3.2.

The CRITIC cross-efficiency scores obtained with our proposed minimax secondary model **(2.15)** are also compared with the geometric average cross-efficiency scores obtained with **Kao and Liu’s (2019)** aggressive-based approach. Note that prior to executing our analysis, we have easily adjusted their model to the specifications of our generalised two-stage DEA structure. The geometric average cross-efficiency scores along with their ranks of the overall system, the stage 1, and the stage 2, are respectively depicted in the sixth and seventh columns of Tables 2.8, 2.9, and 2.10. Correlation analysis suggests that there is a highly strong association between the ranks of these two approaches, as indicated by the correlation values 0.927 (overall system), 0.976 (stage 1), and 0.988 (stage 2), which are significant at the 0.01 level (two-tailed). This can be demonstrated even by the fact that both methods achieve total agreement towards the most desirable unit in all three tables. However, there are a number of points that need to be considered, highlighting the preferability of our method over the other in terms of attaining fairer evaluation results.

Firstly, by solving **Kao and Liu's (2019)** model, we obtained an optimal set of multipliers containing 23 zero weights (as compared with the 19 zero weights of our proposed model). This may indicate a less realistic weight scheme for their method. Secondly, in our minimax model both sub-stages of the generalised two-stage structure have the same bargaining power and improve the overall efficiency. This is conducive to the development of a cooperative situation, where the sub-stages behave altruistically even without having reasons to assume that their cooperation will be returned. This stands in sharp contrast with the aggressive method proposed by **Kao and Liu (2019)**. Although they guaranteed unique cross-efficiencies, they selected a non-cooperative approach, in which DMUs act egoistically with a view to maximising their self-evaluation and downplaying the peer-evaluation. Thirdly, we have managed to acquire a higher absolute cross-efficiency score for each DMU and stage (compared to the respective score in **Kao and Liu's (2019)** results), associated with some performance reward; this is connected with the cooperative role of our model **(2.15)**.

2.4.2 Implications

This example has illustrated the approach proposed in this paper, which is a novel combination of the use of an additive self-efficiency aggregation model, a minimax secondary goal model, and the CRITIC method in order to improve fairness and objectivity in a cross evaluation context for a generalised two-stage DEA system. Firstly, a more sensible weight distribution is obtained via the proposed minimax model **(2.15)** than the basic self-evaluation model **(2.13)** and the aggressive-based model of **Kao and Liu (2019)**, highlighting our successful efforts in obtaining more meaningful rankings. Secondly, the minimax model developed is in addition combined with the CRITIC approach to obtain a greater discrimination power than model **(2.13)** (see Tables 2.8-2.10). Thirdly, on the aggregation of the individual cross-efficiencies, we have compared the traditional average method with the weighted average method, in which the weights are computed via the CRITIC approach. In the former, the opinions of the evaluators are centred around the average (majority) viewpoint. In the latter, more credence and higher inclusion is given to these evaluators that exhibit diversity. These may be desirable characteristics in support of the more modern mindset of many organisations.

Fourthly, it is proven that our proposed minimax model results in higher absolute efficiency scores (than **Kao and Liu's (2019)** scores) connected with some performance prize; this is due to the cooperative nature of the model.

In addition, it is noteworthy that the final rankings obtained are very similar between the three different methods displayed in Tables 2.8-2.10. In practice, however, we think that it is very important for DMUs, when subject to peer evaluation leading ultimately to a ranking, that the methods by which this is achieved are agreeable to modern standards of inclusiveness and diversity and provide an acceptable level of objectivity. We can expect that results are more easily accepted, indeed, if these characteristics are more prominently present in the theoretical foundations of the methods deployed.

As stated in the introduction, the only study having used CRITIC in a DEA context before, seems to be **He and Ma's (2015)**. In that article, CRITIC was used to objectively determine weights used within a DEA collaborative development evaluation model for comparing the internal mechanisms of the regional economy and regional logistics within a 10-year period. Our approach differs in that we use it in the context of peer-evaluation, as an alternative method to address the aggregation problem, in addition to the considering this in the generalised two-stage DEA structure. But more importantly, our study highlights how CRITIC's main components of conflict and contrast intensity can contribute towards a fairer and more diversified cross-efficiency perspective.

As for the possible areas where our study could be applicable, we begin by referring to the manufacturing job shop or to line configurations like clothes manufacturing. In such contexts, the Just-in-Time philosophy takes significantly into account the worker rotations. This practice can eliminate employees' fatigue, encourage their development, and help identifying where they can work best. In this example, it is doable to take day-to-day snapshots (DMUs) of the same factory floor, where the workers are being rotated. In this way, it is possible to measure which of the working stations and settings are (in)efficient and on which days. This will facilitate management towards fairly identifying all those workers that need additional training for certain tasks.

Another promising area could be, for instance, the process of the refinement of the selected cocoa beans into chocolate within a specialized factory. From the first stage, where the cocoa beans are roasted and the cocoa nibs are ground, we mainly obtain cocoa powder. The production manager, in collaboration with the marketing and sales department as well as the outbound logistics manager, will eventually decide on a sensible allocation of the available cocoa powder. On this basis, a proportion of this quantity will be directly forwarded to the outside market for sale, and the remaining will be further blended back with the butter, milk, and liquor in varying quantities, in the second stage, to make different types of chocolate. The main target is to fairly compare the efficiency of several generalised two-stage supply chains of different factory branches or farming communities that make use of cocoa beans from different species of cocoa trees.

As a general ascertainment, it is imperative to improve the processes of efficiency measurement and decision-making under a multi-criteria context and within more advanced network DEA structures; an organisational environment that will promote co-operation, leniency, diversity, and inclusion can result in more effective benchmarking strategies.

TABLE 2.8: Average cross-efficiencies, CRITIC cross-efficiencies, Geometric average cross-efficiencies **(Kao and Liu, 2019)**, and their respective ranks for the overall system.

DMU	Average CE	Ranking	CRITIC CE $\hat{e}_j$	Ranking	Geometric average CE **(Kao and Liu, 2019)**	Ranking
1	0.705	6	0.701	6	0.688	5
2	0.531	9	0.533	9	0.444	10
3	0.760	2	0.747	3	0.699	4
4	0.742	4	0.755	2	0.726	2
5	0.531	10	0.532	10	0.495	9
6	**0.895**	**1**	**0.888**	**1**	**0.878**	**1**
7	0.732	5	0.734	4	0.684	6
8	0.746	3	0.732	5	0.716	3
9	0.567	8	0.567	8	0.566	8
10	0.599	7	0.605	7	0.590	7

TABLE 2.9: Average cross-efficiencies, CRITIC cross-efficiencies, Geometric average cross-efficiencies **(Kao and Liu, 2019)**, and their respective ranks for the stage 1.

DMU	Average CE	Ranking	CRITIC CE $\hat{e}_j^1$	Ranking	Geometric average CE **(Kao and Liu, 2019)**	Ranking
1	0.577	6	0.553	7	0.555	6
2	0.492	9	0.523	9	0.384	10
3	0.575	7	0.596	6	0.509	7
4	**1.000**	**1**	**1.000**	**1**	**1.000**	**1**
5	0.526	8	0.540	8	0.492	8
6	0.828	2	0.814	3	0.889	2
7	0.794	3	0.818	2	0.694	3
8	0.648	4	0.664	4	0.618	5
9	0.412	10	0.421	10	0.411	9
10	0.646	5	0.624	5	0.627	4

TABLE 2.10: Average cross-efficiencies, CRITIC cross-efficiencies, Geometric average cross-efficiencies **(Kao and Liu, 2019)**, and their respective ranks for the stage 2.

DMU	Average CE	Ranking	CRITIC CE $\hat{e}_j^2$	Ranking	Geometric average CE **(Kao and Liu, 2019)**	Ranking
1	0.910	3	0.916	3	0.914	2
2	0.676	7	0.706	7	0.673	7
3	**1.000**	**1**	**1.000**	**1**	**0.997**	**1**
4	0.607	9	0.627	9	0.614	9
5	0.606	10	0.620	10	0.605	10
6	0.924	2	0.921	2	0.912	3
7	0.774	6	0.796	6	0.781	6
8	0.840	4	0.809	5	0.826	4
9	0.818	5	0.837	4	0.815	5
10	0.639	8	0.662	8	0.645	8

2.5 Conclusions & Future Research

Single-stage and the basic serial two-stage DEA systems have fruitfully used various quantitative methods to attain fairness in the evaluation outcomes. Little work, however, has been done addressing the challenge of attaining fairness in a network with more complex interactions among its internal elements. This chapter provides new insight to the generalised two-stage DEA structure of **Yu and Shi (2014)**. We have here proposed a modelling approach for this structure, which promotes fairness among the evaluated DMUs.

In this study, we argue that fairness, or the acceptance of an evaluation and ranking by the different DMUs and their stages, is improved by increasing measures related to the degree of discriminatory power, the weight scheme, and the minority vote. We particularly propose a combination of an additive self-efficiency aggregation model, a multi-objective minimax secondary model, and the CRITIC method in an aim to achieve these aspects of fairness and thus a better degree of cooperation between stages of a DMU and among DMUs. This combination is novel in the DEA literature. Furthermore, the application of the CRITIC method to DEA is by itself novel.

The proposed minimax secondary goal model helps tackle the non-unique optimal multipliers derived from the additive self-evaluation model. The minimax model has the capacity to better discriminate the efficient DMUs than the additive self-evaluation model. In addition, it has significantly eliminated the zero weights assigned to the respective known factors than the additive self-evaluation model and the aggressive-based approach of **Kao and Liu (2019)**.

We have shown in this chapter that the CRITIC method can be applied in DEA to alternatively address the aggregation problem within the DEA cross-efficiency concept. This approach will objectively determine the weights assigned to individual cross-efficiencies to obtain the final cross-efficiencies. By taking into consideration both the contrast intensity and the conflict measures among the DMUs, it manifests the general message of this chapter towards satisfying a more contemporary concept of fairness about diversity and inclusion of minority opinions. Moreover, the proposed minimax model seeks for peer evaluation whereby each peer aims to evaluate the worst of the other players in the best possible light. Its benign and cooperative nature, in conjunction with CRITIC, has the benefit to obtain higher absolute efficiency scores for each DMU and stage than the geometric average efficiencies based on the aggressive method of **Kao and Liu (2019)**. This might be connected with some performance reward, encouraging in a way the DMUs to join the efficiency evaluation and ranking.

In this study, we have proposed an additive self-efficiency aggregation model in the spirit of **Chen et al. (2009)**. This is the basic self-evaluation model without the further improvements introduced in later sections. In such a model, the system efficiency is defined as the weighted arithmetic average of its sub-stages. As for its decomposition weights, **Ang and Chen (2016)** proved that they are non-increasing in the order of sub-stages. Put simply, they highlighted that earlier stages would be assigned higher relative importance, affecting the system's efficiency to a greater extent. Based on that, they also demonstrated that the overall and sub-stages' efficiency scores are prone to the impact of the decomposition weights. We acknowledge this as a limitation of our study, and we believe that a re-definition of the weights, reflecting **Ang and Chen's (2016)** research, could accommodate such an issue. In addition, this chapter could also focus more on the testing of the proposed models and frameworks with empirical data, that is testing their practical value. It would be desirable, for instance, to evaluate the performance of these methods in one of the potential areas described in Section 2.4.2, or other (fair-trade) supply chains.

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 input-oriented DEA models. Another potential path could be the intention to tweak the CRITIC method by focusing perhaps on the level of acceptance of the participants on the final evaluation and ranking scheme obtained. To this end, the conflict measure could be adapted, for example, to fine-tune the impact of opinions with large contrast intensity in relation to their distance to majority opinions. Finally, current research studies the evaluation of the performance of DMUs with a generalised two-stage structure, only when the data are positive real numbers, and the DEA models are based on this condition. In particular in the envisaged areas of application such as sustainable supply chains, datasets can be expected to be incomplete or less accurately described. Future research could thus relax this assumption by allowing the data points to be imprecise and lie in an interval, for example. 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 **(Zimmermann, 2011)** could be used as an alternative option.

Chapter 3

A ranking framework based on interval self and cross-efficiencies in a two-stage DEA system

3.1 Introduction

A paper based on this chapter has been published, see
Kremantzis, M. D., Beullens, P., & Klein, J. (2022). A ranking framework based on interval self and cross-efficiencies in a two-stage DEA system. ***RAIRO - Operations Research*****, 56(3), 1293-1319.**

Data Envelopment Analysis (DEA) is a benchmarking technique for comparing the relative efficiency of a decision-making unit (DMU) with the best observed efficiency **(Charnes et al., 1978)**. The evaluation of a DMU is based on the comparison between the amount of input(s) consumed and the amount of output(s) produced **(Cook et al., 2014)** 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. Hence, in the conventional DEA, the overall assessment of a DMU is based on the optimistic viewpoint **(Zhu, 2015)**. According to these notions, efficiencies are measured up to a maximum of 1. When, from the most optimistic viewpoint, the DMU receives an optimum efficiency score of 1, then it is said to be DEA efficient; otherwise, it is said to be DEA inefficient. On the other hand, if the performance of a DMU is based on the pessimistic viewpoint, then efficiencies are measured within the range of 1 or greater to acquire the worst relative efficiencies of the DMUs. When the DMU is assessed from the angle of the worst relative efficiency, then the following occurs: the DMU that receives an optimum efficiency score of one is called DEA inefficient and the DMUs with score greater than one are called DEA non-inefficient **(Wang and Yang, 2007)**.

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 **(Azizi, 2011)**. 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 (2006)** 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 (2006)** identified a weakness in **Wang and Luo's (2006)** 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 (2007) 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 (2011)**, who suggested a pair of improved bounded models for the target DMU. **Wang et al. (2007)** 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 (2015)** 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 (2017)** 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 (2018)** developed the normalised efficiency metric and then formulated two DEA models to obtain its lower and upper bounds. **Örkcü et al. (2020)** proposed a non-cooperative game like iterative optimistic-pessimistic DEA approach to fully rank the DMUs. **Badiezadeh et al. (2018)** 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. (2018)**, 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 **(Kao and Liu, 2019)**. 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. (2007)**, this score is an effective technique with 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 incremental points make us infer that this framework could be further extended by the use of the cross-efficiency (CE) to ensure fairness in the evaluation outcomes.

The CE concept is based on the peer-evaluation notion **(Sexton et al., 1986)**. As stressed by **Anderson et al. (2002)**, 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 **(Doyle and Green, 1994; Liang et al., 2008; Wu et al., 2016; Li et al., 2018; Zhu et al., 2021; Li et al., 2021)**. The non-uniqueness issue is also critical in a two-stage (network) system **(Ma et al., 2014; Kao and Liu, 2019; Huang et al., 2019; Örkcü et al., 2019; Meng and Xiong, 2021)**. **Kao and Liu (2019)**, for instance, developed an aggressive CE model to measure the efficiency in two basic network structures. **Örkcü et al. (2019)** 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 (1994) 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 **(Wang and Chin, 2010a)**.

To alleviate these deficiencies, **Yang et al. (2012)** 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 chapter 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 3.1, 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.

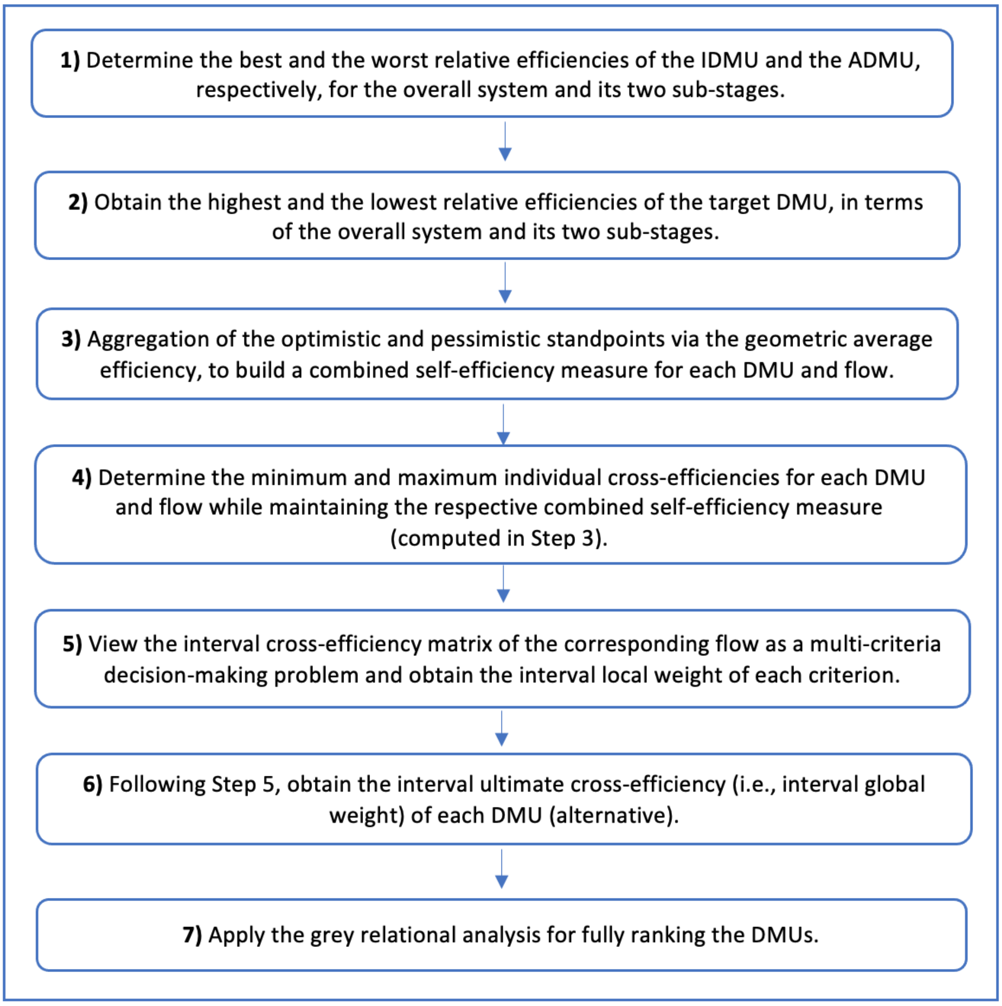

FIGURE 3.1: The proposed framework.

The procedures implemented in the first three steps of our proposed framework (Figure 1) have been applied in several studies (e.g., **Wang et al., 2007)** 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 (2006)** and **Wu (2006)**, 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 chapter views this matrix as a multi-criteria decision-making problem. To solve this problem, we implement the goal programming method of **Wang and Elhag (2007)** 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 (2007). Finally, we apply the grey relational analysis **(Kuo et al., 2008)** 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.

3.2 Methodological Background

We assume that each DMU_j $(j = 1, 2, ..., n)$ uses m inputs $(i = 1, 2, ..., m)$ to produce s outputs $(r = 1, 2, ..., s)$. 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 **(Charnes et al., 1978)**, that assesses the efficiency of the target DMU_k, is illustrated as follows:

$$
\begin{aligned}
E_{kk} = Max \quad & \sum_{r=1}^{s} \mu_{rk} Y_{rk} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \forall j, \\
& \mu_{rk}, \nu_{ik} \geq 0, \forall\, r, i,
\end{aligned}
\tag{3.1}
$$

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.

3.2.1 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 **(Sexton et al., 1986)**. Let μ^*_{rk} and ν^*_{ik} be the optimal set of multipliers of model **(3.1)**. 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 (Table 1) can be a valuable tool to integrate both the peer-efficiency scores E_{kj} $(k, j = 1, 2, ..., n)$ and the self-efficiency scores E_{kk} (in the leading diagonal column). The ultimate cross-efficiency, that attributes the final rank of a DMU, can be defined by averaging all individual cross-efficiencies of the corresponding DMU being evaluated. The ultimate score in this case is $\hat{e}_j = \frac{1}{n} \cdot \sum_{k=1}^{n} E_{kj}, \ \forall j$ **(Angulo-Meza and Lins, 2002)**.

TABLE 3.1: Cross-efficiency matrix; **Doyle and Green (1994)**.

	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}

The existence of multiple optimal weights from model **(3.1)** can deteriorate the theoretical usefulness of the results obtained via the cross-efficiency concept. To tackle this issue, **Doyle and Green (1994)** proposed two opposed secondary goals to choose their weights, favourable or unfavourable, among the optimal solutions. Model **(3.2)** is the aggressive formulation and model **(3.3)** is the benevolent formulation.

$$
\begin{aligned}
Min \quad & \sum_{r=1}^{s} \mu_{rk} \Big(\sum_{j=1, j \neq k}^{n} Y_{rj} \Big) \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_{ik} \Big(\sum_{j=1, j \neq k}^{n} X_{ij} \Big) = 1, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rk} - E^*_{kk} \sum_{i=1}^{m} \nu_{ik} X_{ik} = 0, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \forall\, j,\ j \neq k, \\
& \mu_{rk}, \nu_{ik} \geq 0, \forall r, i,
\end{aligned}
\tag{3.2}
$$

$$Max \quad \sum_{r=1}^{s} \mu_{rk} (\sum_{j=1, j \neq k}^{n} Y_{rj}) \tag{3.3}$$

subject to the same constraints as in model (3.2).

Considering the DEA-related literature, we can draw the conclusion that there is not a well-established methodological approach to guide the DM 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. (2012)** suggested the simultaneous use of the two extreme cases in the context of a single-stage structure.

Model **(3.4)** 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.

$$\begin{aligned}
Min \quad & E_{kj}^{L} = \sum_{r=1}^{s} \mu_{rk} Y_{rj} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_{ik} X_{ij} = 1, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rk} - E_{kk}^{*} \sum_{i=1}^{m} \nu_{ik} X_{ik} = 0, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \forall j; \ j \neq k, \\
& \mu_{rk}, \nu_{ik} \geq 0, \forall \ r, i.
\end{aligned} \tag{3.4}$$

Model **(3.5)** 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.

$$Max \quad E_{kj}^{U} = \sum_{r=1}^{s} \mu_{rk} Y_{rj} \tag{3.5}$$

subject to the same constraints as in model (3.4).

In the above two models, the optimistic self-efficiency score E_{kk}^{*}, derived from model **(3.1)**, 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 $[E_{kj}^{L}, E_{kj}^{U}]$. E_{kj}^{L} is the lower bound and is found from model **(3.4)**, whereas E_{kj}^{U} is the upper bound obtained from model **(3.5)**. Table 3.2 depicts the individual cross-efficiencies as interval numbers, boosting the DM's uncertainty. The elements in the diagonal column of table 3.2 show the special case of the self-efficiency scores, for which $E_{kk} = E^L_{kk} = E^U_{kk}, \forall\, k \in \{1, 2, ..., n\}$.

TABLE 3.2: Interval cross-efficiency matrix; **Yang et al. (2012)**.

	Target	DMU_j		
Evaluator DMU_k	1	2	...	n
1	$[E_{11}, E_{11}]$	$[E^L_{12}, E^U_{12}]$	...	$[E^L_{1n}, E^U_{1n}]$
2	$[E^L_{21}, E^U_{21}]$	$[E_{22}, E_{22}]$	...	$[E^L_{2n}, E^U_{2n}]$
...	...	...	...	...
n	$[E^L_{n1}, E^U_{n1}]$	$[E^L_{n2}, E^U_{n2}]$	...	$[E_{nn}, E_{nn}]$

Models **(3.4)** and **(3.5)** 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 3.3.1, 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 (Section 3.3.2) 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.

3.3 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 $(j = 1, 2, ..., n)$ consumes m inputs $(i = 1, 2, ..., m)$ in the first stage to generate D intermediate products $(d = 1, 2, ..., D)$. The outputs (intermediate measures) of the first stage are converted into inputs in the second stage to produce s final outputs $(r = 1, 2, ..., s)$. 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 $DMU\ j \in N$ **(Kao and Hwang, 2008)**. The above process is illustrated in the exploratory Figure 3.2.

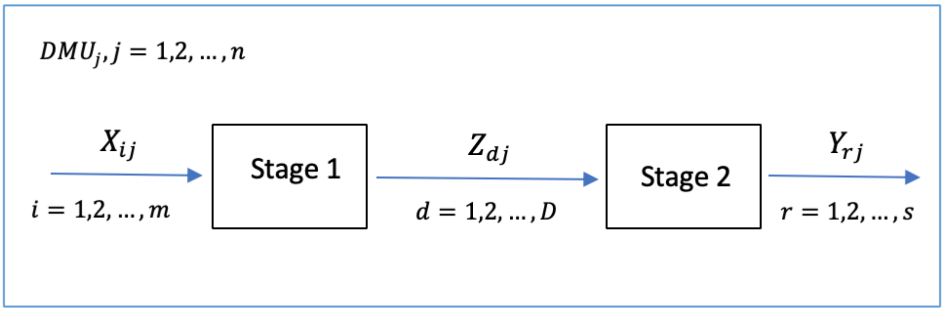

FIGURE 3.2: Typical two-stage tandem DEA structure

According to the relational model of **Kao and Hwang (2008)**, 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 **(3.6)** these operations are described by the constraints, which indicate that the aggregate output can not exceed the aggregate input.

$$
\begin{aligned}
E_k^s = Max \quad & \sum_{r=1}^{s} \mu_{rk} Y_{rk} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1, \\
& \sum_{d=1}^{D} \eta_{dk} Z_{dj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \forall\, j, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{d=1}^{D} \eta_{dk} Z_{dj} \leq 0, \forall\, j, \\
& \mu_{rk}, \nu_{ik}, \eta_{dk} \geq 0, \forall\, r, i, d.
\end{aligned}
\tag{3.6}
$$

At optimality of model **(3.6)**, 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 $E_k^1 = \sum_{d=1}^{D} \eta_{dk}^* Z_{dk} / \sum_{i=1}^{m} \nu_{ik}^* X_{ik}$, and the efficiency of stage 2 as $E_k^2 = \sum_{r=1}^{s} \mu_{rk}^* Y_{rk} / \sum_{d=1}^{D} \eta_{dk}^* Z_{dk}$. It is obvious that the overall efficiency is the product of the efficiencies of the stage efficiencies.

3.3.1 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 **(3.7)** **(Kao and Hwang, 2008)** examines the performance of sub-stage 1:

$$
\begin{aligned}
E_k^1 = Max \quad & \sum_{d=1}^{D} \eta_{dk} Z_{dk} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1, \\
& \sum_{d=1}^{D} \eta_{dk} Z_{dj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \forall\, j, \\
& \nu_{ik}, \eta_{dk} \geq 0, \forall\, i, d.
\end{aligned} \tag{3.7}
$$

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 (2006)**. 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. (2010)**, 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 (2006)** and **Wu's (2006)** models:

$$
\begin{aligned}
E^{IDMU(1)} = Max \quad & \sum_{d=1}^{D} \eta_d Z_d^{max} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_i X_i^{min} = 1, \\
& \sum_{d=1}^{D} \eta_d Z_{dj} - \sum_{i=1}^{m} \nu_i X_{ij} \leq 0, \forall\, j, \\
& \nu_i, \eta_d \geq 0, \forall\, i, d,
\end{aligned} \tag{3.8}
$$

$$
\begin{aligned}
E^{ADMU(1)} = Min \quad & \sum_{d=1}^{D} \eta_d Z_d^{min} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_i X_i^{max} = 1, \\
& \sum_{d=1}^{D} \eta_d Z_{dj} - \sum_{i=1}^{m} \nu_i X_{ij} \leq 0, \forall\, j, \\
& \sum_{d=1}^{D} \eta_d Z_d^{max} - E^{IDMU(1)*} \sum_{i=1}^{m} \nu_i X_i^{min} \geq 0, \\
& \nu_i, \eta_d \geq 0, \forall\, i, d,
\end{aligned} \tag{3.9}
$$

where $E^{IDMU(1)*}$ is the optimal optimistic score of IDMU in terms of sub-stage 1, obtained in model **(3.8)**. Model **(3.9)** ensures that the best relative efficiency of sub-stage 1 is fixed at a value greater than or equal to $E^{IDMU(1)*}$.

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 **(3.6)**. The efficiency of the IDMU for the entire system can be defined as $E^{IDMU(s)} = \sum_{r=1}^{s} \mu_r Y_r^{max} / \sum_{i=1}^{m} \nu_i X_i^{min}$. 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.

$$
\begin{aligned}
E^{IDMU(s)} = Max \quad & \sum_{r=1}^{s} \mu_r Y_r^{max} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_i X_i^{min} = 1, \\
& \sum_{d=1}^{D} \eta_d Z_{dj} - \sum_{i=1}^{m} \nu_i X_{ij} \leq 0, \forall\, j, \\
& \sum_{r=1}^{s} \mu_r Y_{rj} - \sum_{d=1}^{D} \eta_d Z_{dj} \leq 0, \forall\, j, \\
& \mu_r, \nu_i, \eta_d \geq 0, \forall\, r, i, d.
\end{aligned} \tag{3.10}
$$

Similarly, the efficiency of the ADMU for the entire system can be illustrated as $E^{ADMU(s)}$

$= \sum_{r=1}^{s} \mu_r Y_r^{min} / \sum_{i=1}^{m} \nu_i X_i^{max}$. The associated optimisation model is formulated as follows:

$$
\begin{aligned}
E^{ADMU(s)} = Min \quad & \sum_{r=1}^{s} \mu_r Y_r^{min} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_i X_i^{max} = 1, \\
& \sum_{d=1}^{D} \eta_d Z_{dj} - \sum_{i=1}^{m} \nu_i X_{ij} \leq 0, \forall\, j, \\
& \sum_{r=1}^{s} \mu_r Y_{rj} - \sum_{d=1}^{D} \eta_d Z_{dj} \leq 0, \forall\, j, \\
& \sum_{r=1}^{s} \mu_r Y_r^{max} - E^{IDMU(s)*} \sum_{i=1}^{m} \nu_i X_i^{min} \geq 0, \\
& \mu_r, \nu_i, \eta_d \geq 0, \forall\, r, i, d.
\end{aligned}
\tag{3.11}
$$

Model **(3.11)** 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 $E^{IDMU(s)*}$. 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 **(3.12)**, the optimistic relative efficiency of DMU_k for the sub-stage 1 is examined while $E^{IDMU(1)*}$ is kept fixed; it is related to **Wang and Luo's (2006)** framework:

$$
\begin{aligned}
E_k^{IDMU(1)} = Max \quad & \sum_{d=1}^{D} \eta_{dk} Z_{dk} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1, \\
& \sum_{d=1}^{D} \eta_{dk} Z_{dj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \forall\, j, \\
& \sum_{d=1}^{D} \eta_{dk} Z_d^{max} - E^{IDMU(1)*} \sum_{i=1}^{m} \nu_{ik} X_i^{min} = 0, \\
& \nu_{ik}, \eta_{dk} \geq 0, \forall\, i, d
\end{aligned}
\tag{3.12}
$$

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

The overall optimistic efficiency score of DMU_k can be determined as $E_k^{IDMU(s)} = \sum_{r=1}^{s} \mu_{rk} Y_{rk} / \sum_{i=1}^{m} \nu_{ik} X_{ik}$. 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 **(Kao and Hwang, 2008)**. Thus, we propose model **(3.13)**, that maximises the above ratio.

$$
\begin{aligned}
E_k^{IDMU(s)} = Max \quad & \sum_{r=1}^{s} \mu_{rk} Y_{rk} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1, \\
& \sum_{d=1}^{D} \eta_{dk} Z_{dj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \forall\, j, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{d=1}^{D} \eta_{dk} Z_{dj} \leq 0, \forall\, j, \\
& \sum_{d=1}^{D} \eta_{dk} Z_d^{max} - E^{IDMU(1)*} \sum_{i=1}^{m} \nu_{ik} X_i^{min} = 0, \\
& \sum_{r=1}^{s} \mu_{rk} Y_r^{max} - E^{IDMU(2)*} \sum_{d=1}^{D} \eta_{dk} Z_d^{min} = 0, \\
& \mu_{rk}, \nu_{ik}, \eta_{dk} \geq 0, \forall\, r, i, d
\end{aligned}
\tag{3.13}
$$

The fourth and fifth constraints indicate that $E^{IDMU(1)*}$ and $E^{IDMU(2)*}$, respectively, remain unchanged. Let $\nu_k^* = (\nu_{1k}^*, \nu_{2k}^*, ..., \nu_{mk}^*)$, $\eta_k^* = (\eta_{1k}^*, \eta_{2k}^*, ..., \eta_{Dk}^*)$, $\mu_k^* = (\mu_{1k}^*, \mu_{2k}^*, ..., \mu_{sk}^*)$, be an optimal solution to model **(3.13)**. For DMU_k, $E_k^{IDMU(s)} = \sum_{r=1}^{s} \mu_{rk}^* Y_{rk} / \sum_{i=1}^{m} \nu_{ik}^* X_{ik}$, $E_k^{IDMU(1)} = \sum_{d=1}^{D} \eta_{dk}^* Z_{dk} / \sum_{i=1}^{m} \nu_{ik}^* X_{ik}$, and $E_k^{IDMU(2)} = \sum_{r=1}^{s} \mu_{rk}^* Y_{rk} / \sum_{d=1}^{D} \eta_{dk}^* Z_{dk}$, which are referred to as optimistic self-efficiency measures with respect to the overall system and its sub-stages, respectively.

Then, model **(3.14)** evaluates the worst relative efficiency of DMU_k, in terms of sub-stage 1, while the parameter $E^{ADMU(1)*}$ takes the value as determined previously from model **(3.9)**. This model is related to **Wu's (2006)** framework.

$$
\begin{aligned}
E_k^{ADMU(1)} = Min \quad & \sum_{d=1}^{D} \eta_{dk} Z_{dk} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1, \\
& \sum_{d=1}^{D} \eta_{dk} Z_{dj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \forall\, j, \\
& \sum_{d=1}^{D} \eta_{dk} Z_d^{min} - E^{ADMU(1)*} \sum_{i=1}^{m} \nu_{ik} X_i^{max} = 0, \\
& \nu_{ik}, \eta_{dk} \geq 0, \forall\, i, d
\end{aligned}
\tag{3.14}
$$

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

The overall pessimistic score of DMU_k can be determined as $E_k^{ADMU(s)} = \sum_{r=1}^{s} \mu_{rk} Y_{rk} / \sum_{i=1}^{m} \nu_{ik} X_{ik}$ and denotes the product of the pessimistic efficiencies of the DMU_k of the two sub-stages. Thus, we suggest model **(3.15)**, whose purpose is to minimise the above ratio. $E^{ADMU(1)*}$ and $E^{ADMU(2)*}$ are maintained.

$$
\begin{aligned}
E_k^{ADMU(s)} = Min \quad & \sum_{r=1}^{s} \mu_{rk} Y_{rk} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1, \\
& \sum_{d=1}^{D} \eta_{dk} Z_{dj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \forall\, j, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{d=1}^{D} \eta_{dk} Z_{dj} \leq 0, \forall\, j, \\
& \sum_{d=1}^{D} \eta_{dk} Z_d^{min} - E^{ADMU(1)*} \sum_{i=1}^{m} \nu_{ik} X_i^{max} = 0, \\
& \sum_{r=1}^{s} \mu_{rk} Y_r^{min} - E^{ADMU(2)*} \sum_{d=1}^{D} \eta_{dk} Z_d^{max} = 0, \\
& \mu_{rk}, \nu_{ik}, \eta_{dk} \geq 0, \forall\, r, i, d
\end{aligned}
\tag{3.15}
$$

Let $\nu_k^{\sim} = (\nu_{1k}^{\sim}, \nu_{2k}^{\sim}, ..., \nu_{mk}^{\sim})$, $\eta_k^{\sim} = (\eta_{1k}^{\sim}, \eta_{2k}^{\sim}, ..., \eta_{Dk}^{\sim})$, $\mu_k^{\sim} = (\mu_{1k}^{\sim}, \mu_{2k}^{\sim}, ..., \mu_{sk}^{\sim})$, be an optimal solution to model **(3.15)**. For DMU_k, $E_k^{ADMU(s)} = \sum_{r=1}^{s} \mu_{rk}^{\sim} Y_{rk} / \sum_{i=1}^{m} \nu_{ik}^{\sim} X_{ik}$, $E_k^{ADMU(1)} = \sum_{d=1}^{D} \eta_{dk}^{\sim} Z_{dk} / \sum_{i=1}^{m} \nu_{ik}^{\sim} X_{ik}$, and $E_k^{ADMU(2)} = \sum_{r=1}^{s} \mu_{rk}^{\sim} Y_{rk} / \sum_{d=1}^{D} \eta_{dk}^{\sim} Z_{dk}$, 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 $[E_k^{ADMU(s)*}, E_k^{IDMU(s)*}]$ is shaped, where $E_k^{ADMU(s)*}$ (lower bound) represents the worst relative efficiency of DMU_k and $E_k^{IDMU(s)*}$ (upper bound) illustrates the best relative efficiency of DMU_k, obtained via models **(3.15)** and **(3.13)**, 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. (2007)**, to meet this requirement. This is a more comprehensive strategy than either of the two opposing measures, and can better deal with the drawbacks emerged in the efficiency evaluation. In addition, it is not only considered as a ranking index (which is the case with the relative closeness index of the TOPSIS method adopted by **Wang and Luo (2006)** to combine the optimistic and pessimistic efficiencies of a DMU), but also an efficiency measure. **Wang et al. (2007)** also emphasised that the geometric average efficiency requires less computational time to be implemented than the (average) cross efficiency. Let $E_k^{comb(\epsilon)*} = \sqrt{E_k^{ADMU(\epsilon)*} \cdot E_k^{IDMU(\epsilon)*}}$ 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: $E_k^{comb(s)*} = \sqrt{E_k^{ADMU(s)*} \cdot E_k^{IDMU(s)*}} = \sqrt{E_k^{ADMU(1)*} \cdot E_k^{ADMU(2)*} \cdot E_k^{IDMU(1)*} \cdot E_k^{IDMU(2)*}} = \sqrt{E_k^{ADMU(1)*} \cdot E_k^{IDMU(1)*}} \cdot \sqrt{E_k^{ADMU(2)*} \cdot E_k^{IDMU(2)*}} = E_k^{comb(1)*} \cdot E_k^{comb(2)*}$.

The geometric average efficiency is an approachable efficiency measure that leads to a fairer ranking index **(Wang et al., 2007)**. 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 **(Wang et al., 2007)**. Yet, this feature has not been explored in a more complex network structure. To this end, in the next subsection, the archetypal optimistic-pessimistic ranking framework is further extended by the use of the interval CE within a two-stage tandem system, to ensure a more logical ranking order.

3.3.2 Interval cross-efficiencies in basic two-stage structure

In this sub-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 3.3.1. 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 $(k, j = 1, 2, ..., n)$, 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:

$$
\begin{aligned}
E_{kj}^{L(s)} = Min \quad & \sum_{r=1}^{s} \mu_{rk} Y_{rj} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_{ik} X_{ij} = 1, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rk} - E_k^{comb(s)*} \sum_{i=1}^{m} \nu_{ik} X_{ik} = 0, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rk} - E_k^{comb(2)*} \sum_{d=1}^{D} \eta_{dk} Z_{dk} = 0, \\
& \sum_{d=1}^{D} \eta_{dk} Z_{dj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \forall\, j, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{d=1}^{D} \eta_{dk} Z_{dj} \leq 0, \forall\, j, \\
& \mu_{rk}, \nu_{ik}, \eta_{dk} \geq 0, \forall\, r, i, d.
\end{aligned}
\tag{3.16}
$$

In model **(3.16)**, $E_k^{comb(s)*}$ and $E_k^{comb(2)*}$ are the crisp combined self-efficiency measures of the system and the sub-stage 2 for DMU_k, respectively, obtained from Section 3.3.1. The second and third constraint maintain combined system and sub-stage efficiencies for DMUs. Model **(3.16)** 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 $E_{kj}^{L(s)} = \sum_{r=1}^{s} \mu_{rk}^{*} Y_{rj} / \sum_{i=1}^{m} \nu_{ik}^{*} X_{ij}$, $E_{kj}^{L(1)} = \sum_{d=1}^{D} \eta_{dk}^{*} Z_{dj} / \sum_{i=1}^{m} \nu_{ik}^{*} X_{ij}$, and $E_{kj}^{L(2)} = \sum_{r=1}^{s} \mu_{rk}^{*} Y_{rj} / \sum_{d=1}^{D} \eta_{dk}^{*} Z_{dj}$, respectively. By the same token, a benevolent strategy is implemented to construct the following maximisation model:

$$E_{kj}^{U(s)} = Max \quad \sum_{r=1}^{s} \mu_{rk} Y_{rj} \tag{3.17}$$

subject to the same constraints as in model (3.16).

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 $[E_{kj}^{L(\epsilon)}, E_{kj}^{U(\epsilon)}]$, where $E_{kj}^{L(\epsilon)}$ is the lower bound and $E_{kj}^{U(\epsilon)}$ is the upper bound. Therefore, three generalised interval CE matrices (based on the concept of Table 3.2) 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 $E_{jj}^{L(\epsilon)*} = E_{jj}^{U(\epsilon)*} = E_{j}^{comb(\epsilon)*}$ $\forall$ 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.

3.3.3 Interval cross-efficiencies and MCDM context

Each generalised interval CE matrix (see Section 3.3.2) can be treated as a multi-criteria decision making (MCDM) problem with $j = 1, 2, ..., n$ DMUs that act as alternatives. Each DMU_j is assessed considering the weight profile of $k = 1, 2, ..., n$ DMUs that act as criteria. Interestingly, the former intuition is attributed to the novel study of **Cook et al. (2014)**, according to which each DEA-related problem could be viewed as a multi-criteria evaluation problem. This has also been consolidated by **Rakhshan (2017)**, 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 $E_{kj}^{L(\epsilon)}$ and $E_{kj}^{U(\epsilon)}$, 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 $[E_{kj}^{L(\epsilon)}, E_{kj}^{U(\epsilon)}]$. These elements have been obtained in Section 3.3.2. 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 (2007) 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 ($k = 1, 2, ..., n$) 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:

$$\begin{aligned}
\Omega = Min \quad & \sum_{k=1}^{n} (\delta_k^+ + \delta_k^- + \gamma_k^+ + \gamma_k^-) \\
\text{subject to} \quad & (E_L - I)W_U \; - (n-1)W_L \; - \Delta^+ + \Delta^- = 0, \\
& (E_U - I)W_L \; - (n-1)W_U \; - \Gamma^+ + \Gamma^- = 0, \\
& w_k^L \; + \sum_{\omega=1,\omega\neq k}^{n} w_\omega^U \geq 1, \forall\, k, \\
& w_k^U \; + \sum_{\omega=1,\omega\neq k}^{n} w_\omega^L \leq 1, \forall\, k, \\
& W_U - W_L \geq 0, \\
& W_U, W_L, \Delta^+, \Delta^-, \Gamma^+, \Gamma^- \geq 0,
\end{aligned} \tag{3.18}$$

where $\Delta^+ = (\delta_1^+, ..., \delta_n^+)^T$, $\Delta^- = (\delta_1^-, ..., \delta_n^-)^T$, $\Gamma^+ = (\gamma_1^+, ..., \gamma_n^+)^T$, $\Gamma^- = (\gamma_1^-, ..., \gamma_n^-)^T$, $W_U = (w_1^U, ..., w_n^U)^T$, $W_L = (w_1^L, ..., w_n^L)^T$, 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 $E_{kj}^{L(\epsilon)}$ and $E_{kj}^{U(\epsilon)}$ 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 **(3.18)** should, in effect, run three times, based on the investigation of the interval CE matrix of the respective system and stage to compose $[w_k^{L(\epsilon)}, w_k^{U(\epsilon)}]$.

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. (2004)**. 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 (2007)** technique has, to our knowledge, not received attention on eliciting interval local weights from an interval CE matrix. Therefore, this sub-section intends to use their approach to achieve this goal.

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 (2007)** 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:

$$\begin{aligned} E_j^{L.B.(\epsilon)} = Min \quad & \sum_{k=1}^{n} w_k^{(\epsilon)} E_{kj}^{L(\epsilon)} \\ \text{subject to} \quad & \sum_{k=1}^{n} w_k^{(\epsilon)} = 1, \\ & w_k^{L(\epsilon)} \leq w_k^{(\epsilon)} \leq w_k^{U(\epsilon)}, \forall\ k, \end{aligned} \tag{3.19}$$

and

$$E_j^{U.B.(\epsilon)} = Max \quad \sum_{k=1}^{n} w_k^{(\epsilon)} E_{kj}^{U(\epsilon)} \tag{3.20}$$

subject to the same constraints as in model (3.19),

where $w_k^{(\epsilon)}$ is the decision variable of the kth local criterion weight ($k = 1, 2, ..., n$) for $\epsilon = s$ (overall system), 1 (stage 1) or 2 (stage 2). The above pair of LP models **(3.19)-(3.20)** results in the interval global weight for each alternative j ($j = 1, 2, ..., n$), denoted by $[E_j^{L.B.(\epsilon)}, E_j^{U.B.(\epsilon)}]$ for the entire system and its sub-stages. Table 3.3 illustrates the synthesis of the interval cross-efficiencies.

TABLE 3.3: Synthesis of interval cross-efficiencies

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

3.3.4 Grey relational analysis for ranking DMUs

In Section 3.3.3, we obtained an interval ultimate cross-efficiency score for DMU_j $(j = 1, 2, ..., n)$. 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. To this end, alternative techniques have been developed in DEA-related literature.

Wang et al. (2005) proposed the minimax regret approach to compare and rank the efficiency intervals of DMUs; in this method, the DMU with the smallest maximum loss of efficiency is characterised as the most desirable one. **Wang and Yang (2007)** alternatively suggested the Hurwicz criterion approach to deal with such a problem. It is based on the best and the worst relative efficiencies of DMUs. One controversy within the method is the use of a level of optimism α, that is fully determined by the DM. The subjective judgement about the final value of α (DM's attitude towards risk) could alter the result (i.e. final ranking order) at any time, limiting the robustness. **Azizi (2011)** utilised the A-acceptability index, not only to identify the superiority or inferiority of a DMU over another DMU, but also to reveal the amount of satisfaction the DM feels in terms of superiority or inferiority of one interval value over another. **Khalili-Damghani et al. (2015)** took advantage of the well-established Shannon entropy approach to determine the final ranks of DMUs.

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 **Deng (1989)**. It has proved to be a worthy multi-attribute decision-making tool when limited, unreliable, and uncertain information emerges **(Pourmohammadi et al., 2018)**. GRA has fruitfully examined complex interconnections among several factors **(Chang, 2012)** as well as obtained the optimal alternative among several alternatives **(Kuo et al., 2008; Sarraf and Nejad, 2020)**. It has also been discussed as a method that evaluates the changes in the relationship between variables (of homogeneous and non-homogeneous nature) over a period of time **(Zhu and Xu, 2017; Agyemang et al., 2020)**.

There is a plethora of reasons for selecting GRA to compare and rank the interval values within such a created multi-attribute decision-making context. The first is that GRA is generally more applicable when the sample size is small and the distribution pattern unknown **(Wang et al., 2010; Wang et al., 2015; Kaviani et al., 2019)**. Since our study takes into account a limited sample (10 bank branches) and there is no sufficient information on their distribution, it is more justifiable to use GRA. The other reason is related to **Kuo et al.'s (2008)** work, which argued that GRA can solve multi-criteria decision-making problems by aggregating several attribute values into a single value per decision alternative; these attribute values are typically in-commensurable. Our study seems compatible with the latter statement as it makes use of attributes, which we cannot easily assume they are comparable with each other.

In addition, it is important to mention that other equally prevalent multi-criteria decision-making methods may not be suitable for accommodating such an issue. For instance, the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), which was originally introduced by **Hwang and Yoon (1981)**, aims to select the alternative with the shortest distance from the positive ideal solution and the longest distance from the negative ideal solution. GRA is on the same wavelength with TOPSIS method. However, GRA additionally considers different min-max operators, which seem to be more appropriate in our case where the sample size is limited and the distribution pattern is not obvious **(Zhang et al., 2014)**. Another example is the Shannon entropy **(Shannon, 1948)**, which is treated as an uncertainty measure of information. Based on the information entropy, the quality of information derived via a decision-making context can ensure the level of reliability of the respective problem. The entropy measure is normally used to determine the attribute weights, when they are unknown to the decision maker. This limits its applicability, as it cannot be directly implemented to compare and rank interval values of alternatives, such as the ones explored in our study.

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 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 **(Kuo et al., 2008)**. Below, we will provide an overview of the GRA as we would apply this to ranking the 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 3.3.3, 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 $j = 1, 2, ..., n$ alternatives that are assessed by $i = 1, 2$ attributes. Table 3.4 depicts what we described above.

TABLE 3.4: Interval ultimate cross-efficiencies

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

The jth alternative can be expressed as $E_j^{(\epsilon)} = (E_j^{L.B.(\epsilon)}, E_j^{U.B.(\epsilon)})$, 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 $E_j^{\bar{(\epsilon)}} = (E_j^{L.\bar{B}.(\epsilon)}, E_j^{U.\bar{B}.(\epsilon)})$ by use of one of the following equations:

$$E_j^{\bar{i}(\epsilon)} = \frac{E_j^{i(\epsilon)} - min\{E_j^{i(\epsilon)}, \forall\, j\}}{max\{E_j^{i(\epsilon)}, \forall\, j\} - min\{E_j^{i(\epsilon)}, \forall\, j\}}, \ \forall\, j, i, \tag{3.21}$$

$$E_j^{\bar{i}(\epsilon)} = \frac{max\{E_j^{i(\epsilon)}, \forall\, j\} - E_j^{i(\epsilon)}}{max\{E_j^{i(\epsilon)}, \forall\, j\} - min\{E_j^{i(\epsilon)}, \forall\, j\}}, \ \forall\, j, i, \tag{3.22}$$

$$E_j^{\bar{i}(\epsilon)} = \frac{|E_j^{i(\epsilon)} - E_{des}^{i(\epsilon)}|}{max\{E_j^{i(\epsilon)}, \forall\, j\} - min\{E_j^{i(\epsilon)}, \forall\, j\}}, \ \forall\, j, i. \tag{3.23}$$

Equation **(3.21)** is used for the greater-the-better attributes, Equation **(3.22)** is used for the smaller-the-better attributes, and Equation **(3.23)** 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 $(E_j^{i(\epsilon)})$ and the comparability sequence $(E_j^{\bar{i}(\epsilon)})$, which is $\Delta_j^{i(\epsilon)} = |E_j^{i(\epsilon)} - E_j^{\bar{i}(\epsilon)}|, \ \forall\, j, i$. As stressed in **Kuo et al. (2008)**, the reference sequence $E_j^{i(\epsilon)} = (E_j^{1(\epsilon)}, E_j^{2(\epsilon)}) = (1, 1)$.

Then, we compute the grey relational coefficient $\gamma(E_j^{i(\epsilon)}, E_j^{i\bar{(\epsilon)}})$. 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 $\gamma(E_j^{i(\epsilon)}, E_j^{i\bar{(\epsilon)}}) = \frac{\Delta_{min}^{(\epsilon)} + \zeta \cdot \Delta_{max}^{(\epsilon)}}{\Delta_j^{i(\epsilon)} + \zeta \cdot \Delta_{max}^{(\epsilon)}}, \ \forall\, j, i$, where $\Delta_{min}^{(\epsilon)} = min\{\Delta_j^{i(\epsilon)}, \ \forall\, j, i\}$, $\Delta_{max}^{(\epsilon)} = max\{\Delta_j^{i(\epsilon)}, \ \forall\, j, i\}$, and ζ denotes the distinguishing coefficient, $\zeta \in [0, 1]$. The in-built ζ is used to expand or squeeze the range of the grey relational coefficient. In other words, decision makers are offered the opportunity to modify their opinions based on a wide range of values **(Agyemang et al., 2020)**.

Finally, the GRG $\Gamma_j^{(\epsilon)}$, which is the weighted average of the grey relational coefficients, is estimated as $\Gamma_j^{(\epsilon)} = \sum_{i=1}^{2} w_i \cdot \gamma_j^{i(\epsilon)}, \forall\, j$, 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 **(Jahan et al., 2012)**. 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 **(Kuo et al., 2008)**, but also for efficiency evaluation DEA problems as a MCDM context in disguise **(Sarraf and Nejad, 2020)**. 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.

3.4 Numerical Application

This section illustrates the use of the mathematical models presented in Section 3.3 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. (2013)** is used for illustrative purposes. In Table 3.5, ten bank branches of China Construction Bank in Anhui are assessed within the two-stage tandem structure (see Figure 3.1). 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 3.6 exhibits the corresponding scores from solving models **(3.8)-(3.11)**, introduced in Section 3.3.1.

TABLE 3.5: The numerical application of **Zhou et al. (2013)**

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

TABLE 3.6: Highest and lowest relative efficiency scores for the overall system, stage 1, and stage 2

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

Then, models **(3.12)-(3.15)** 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 3.7. Recall that these relative self-efficiency scores indicate their distance from the respective IDMU and ADMU efficiencies, presented in Table 3.6. 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 3.7, no matter what point of view efficiency is measured from, DMU1 is certainly the best unit and DMU10 is the worst unit, in terms of the entire system (second expanded column). Considering stage 1, regardless of the viewpoint, DMU1 and DMU3 are the most and least desirable units, respectively (third expanded column). In stage 2 (fourth expanded column), DMU10 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 ($E_k^{IDMU(s)}$), where the highest score is 0.8132 occurring at DMU1, followed by 0.3490 occurring at DMU6.

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 ($E_k^{comb(s)}$), the stage 1 ($E_k^{comb(1)}$), and the stage 2 ($E_k^{comb(2)}$). In table 3.7, DMU1 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, DMU1 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 DMU1 approximates 0.51, whereas the corresponding rating of DMU2 (in the second place) equals to 0.2733. The geometric average efficiency also indicates that DMU10 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 3.3.1. As an example, the combined self-efficiencies calculated for DMU1 satisfy $0.1267 = 0.5099 * 0.2486$. Since this property is satisfied, every $E_k^{comb(s)}$ is no greater than its corresponding $E_k^{comb(1)}$ and $E_k^{comb(2)}$. Another point to be noted is that most bank branches have a smaller $E_k^{comb(2)}$ than $E_k^{comb(1)}$. Only DMUs 3, 8, and 9 have a smaller $E_k^{comb(1)}$ than $E_k^{comb(2)}$. However, after implementing the Wilcoxon's matched-pairs signed-ranks test **(Daniel, 1978)** 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 DMU3 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 Table 3.7) are also compared with the respective scores (Table 3.8) obtained with **Kao and Hwang's (2008)** approach. As mentioned in Section 3.3, 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 **(3.6)**) 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 3.5, by measuring the efficiencies of the whole process and its constituent sub-stages for the ten DMUs. In Table 3.8, 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 DMU7 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 **(Daniel, 1978)** 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 DMU1 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 **Kao and Hwang, 2008**), 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 3.1, 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 3.7. 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 3.3.2. Tables 3.9-3.11 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 **3.16**) and a benevolent (model **3.17**) strategy; this originally creates an atmosphere of neutrality.

TABLE 3.7: Self-efficiency ratings and ranks of the overall system, the stage 1, and the stage 2, with the proposed method

DMU	$E_k^{IDMU(s)}$	$E_k^{ADMU(s)}$	$E_k^{comb(s)}$	$E_k^{IDMU(1)}$	$E_k^{ADMU(1)}$	$E_k^{comb(1)}$	$E_k^{IDMU(2)}$	$E_k^{ADMU(2)}$	$E_k^{comb(2)}$
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 3.8: Self-efficiency ratings and ranks of the overall system, the stage 1, and the stage 2, with **Kao and Hwang's (2008)** 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)

To make the content of Tables 3.9-3.11 comprehensible to the reader, it should be ideal to present a few examples. In the second column of Table 3.9, DMU1 is assessed based on the weight profile of all other DMUs, except its own weight set. The minimum and maximum individual cross-efficiencies of DMU1 based on the optimal weight scheme of DMU2 are 0.1216 and 0.2371, respectively, for the overall system. In the fifth column of table 3.10, DMU4 is also peer-appraised with respect to the weight profile of all other DMUs. The minimum and maximum individual cross-efficiencies of DMU4 based on the weight profile of DMU10 are 0.1475 and 0.2281, respectively, for sub-stage 1. Table 3.11 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 $E_{jj}^{L(\epsilon)*} = E_{jj}^{U(\epsilon)*} = E_j^{comb(\epsilon)*}$ $\forall$ 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.

Recalling the discussion in Section 3.3.3, we view each interval CE matrix as a MCDM problem. In Tables 3.9-3.11, 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 3.12, with respect to the overall system, stage 1, and stage 2, respectively. The interval weights are obtained via the GP model **(3.18)**, and the interval global weights, according to the pair of optimisation models **(3.19)** and **(3.20)**, as stated in Section 3.3.3.

TABLE 3.9: 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

TABLE 3.10: 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 3.11: 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

There was not any indication that the interval weights in Table 3.12 are not unique. To the best of our knowledge, there are not even any such discussions in the relevant literature to whether these interval weights are unique or not.

For instance, in the second column of the last row of table 3.9, we obtain the interval ultimate CE of DMU1: $[0.1000,\ 0.2611]$, where 0.1000 is the minimum and 0.2611 is the maximum CE score. The minimum score of DMU1 for the overall system is estimated via solving model **(3.19)**. The basic prerequisites of this model are to recognise the minimum individual cross-efficiencies of DMU1 based on the weight profile of all ten DMUs (left side of column 2 of Table 3.9) and the interval weights per criterion for the overall system (column 2 of Table 3.12). The maximum ultimate cross-efficiency of DMU1 for the overall system is estimated via solving model **(3.20)**. The basic prerequisites of this model are to identify the maximum individual cross-efficiencies of DMU1 based on the weight profile of all 10 DMUs (right side of column 2 of Table 3.9) and the interval weights per criterion for the overall system (column 2 of Table 3.12).

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 3.3.4. 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 Tables 3.9-3.11) form the appropriate matrix, as shown in Table 3.4.

TABLE 3.12: 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

The data of performance values of the two attributes are subsequently normalised through the greater-the-better equation **(3.21)**; this choice reflects the necessity of pushing up the peer-efficiency of each DMU. The results are depicted in Appendix B and particularly in the second column of Table B.1 for the overall system, of Table B.2 for stage 1, and of Table B.3 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 three tables in Appendix B. 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 **Deng (1989)**, we have set $\zeta = 0.5$ implying that the DMU has neither an optimistic nor a conservative attitude. The respective results are portrayed in the last column of each of the three tables in Appendix B.

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 3.13, 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 3.13 reflects the improvement of the discriminating power, as compared to the original rank derived from the combined self-efficiency measures in Table 3.7. 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, DMU10 is without a doubt the least desirable unit in all three cases. DMU1 is also considered to be the most promising bank branch for the overall system and stage 1, while DMU3 is the best unit according to stage 2. Generally, one can deduce that the ranking results for all branches (except DMU10) 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 Table 3.13) are also compared with the respective ultimate cross-efficiency ratings (Table 3.14) obtained via the **Kao and Liu's (2019)** 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 3.5. In Table 3.14, 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 DMU6 (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 DMU10 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 DMU8 with a rank difference of 4. All the remaining bank branches have a rank difference of no 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 (2019)** 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 **(Daniel, 1978)**, 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.

TABLE 3.13: 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 Stage1	GRG $\Gamma_j^{(2)}$	Rank Stage2
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 3.14: Peer-efficiency ratings and ranks of the overall system, the stage 1, and the stage 2, with **Kao and Liu's (2019)** 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)

3.5 Conclusions & Future Research

This chapter 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 3.1, 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 (2006)** and **Wu (2006)**, 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 (2007)**, the LP models by **Entani and Tanaka (2007)**, and the GRA of **Kuo et al. (2008)**. 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 **(Kao and Hwang, 2008)**, 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 **(Zhang and Chen, 2019)**. 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 **(Gan et al., 2019)**. Finally, the efficiency evaluation of two-stage (food) supply chains of different factories or farming communities **(Kremantzis et al., 2022)** could also serve the goals of our chapter. 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 chapter 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 **(Saaty and Vargas, 2013)** or the best-worst method **(Rezaei, 2016)**, 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 chapter 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. This chapter has also integrated the best and worst relative efficiencies of each DMU and stage into an overall assessment using the geometric average efficiency; this measure has been adopted after carefully checking the discussions in **Wang et al. (2007)** study. An interesting area for further examination could be the integration of the self-efficiency scores into a single comprehensive measure, using either the (weighted) arithmetic mean or the (weighted) harmonic mean. To the best of our knowledge, such metrics have not been paid attention to comparing and analysing the sensitivity of the integration within a network double-frontier DEA context. Which of these metrics leads to the most meaningful ranking outcome, will be a subject of future in-depth studies. 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 **(Zimmermann, 2011)** could be used as an ideal alternative option. For example, there is a growing body of literature **(Ebrahimnejad, 2012; Hatami-Marbini et al., 2017; Santos Arteaga et al., 2021; Peykani et al., 2022)** 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 **(Hosseinzadeh Lotfi et al., 2010a; Hosseinzadeh Lotfi et al., 2010b; Ebrahimnejad and Hosseinzadeh Lotfi, 2012; Tavana et al., 2018)**.

Chapter 4

Measurement and evaluation of multi-function parallel network hierarchical DEA systems

4.1 Introduction

The purpose of this article is to extend the current literature on the network hierarchical DEA structures. Our study particularly considers the measurement and evaluation of the performance of several parallel processes, wherein each process integrates a multi-function hierarchical structure. Additive decomposition and multiplicative aggregation DEA models are presented and used in a higher education context to investigate the areas of weakness of the considered Business Schools.

An increasing number of companies, especially those with complex organisation charts, are actively engaged in improving their production systems, as a response to the intense market competition. Data envelopment analysis (DEA) has been extensively exploited as an effective performance evaluation technique to gain insight into the past accomplishments and future developments of a decision-making unit (DMU) **(Emrouznejad and Yang, 2018)**. Since the seminal work of **Charnes et al. (1978)**, DEA has been widely used in various contexts, including energy and environment **(Zhou et al., 2008; Zhai et al., 2019)**, local governments **(Amatatsu et al., 2012)**, R & D departments **(Wang et al., 2013)**, financial institutions **(Paradi and Zhu, 2013)**, supply chains **(Azadi et al., 2014)**, sports **(Moreno and Lozano, 2014)**, international shipping **(Gan et al., 2019)**, and inland transportation **(Stefaniec et al., 2020)**. See also **Liu et al. (2013)** for a review of applications.

Traditional DEA approaches put emphasis on evaluating the most favourable efficiency measure of a DMU, only by considering its exogenous inputs and outputs. This is referred to as black-box analysis **(Kao and Hwang, 2008)**. The internal structure of a unit usually consists of several divisions with similar and/or different functions; they may be interrelated, independent, or a mixture of these, depending on the objective of the system **(Kao, 2014)**. Ignoring the internal operations of a DMU could lead to misinterpretation. For instance, while the whole unit could be characterised as efficient, all its individual stages may be inefficient. The traditional concept can also lead to a large relative fraction of DEA-efficient DMUs, without the means to distinguish them **(Ma et al., 2017)**. To enable the study of internal operations, research has extended DEA models to consider network structures **(Färe and Grosskopf, 2000)**.

The network system differs from the black-box in that it involves more complex structures, thereby leading to a less systematic illustration **(Kao, 2017)**. In the two-stage tandem system, all inputs used by a DMU feed into a first stage, producing intermediate outputs that all feed into a second stage, producing the final outputs of the entire system. **Kao and Hwang (2008)** proposed that the system efficiency is decomposed into the product of the efficiencies of the two divisions. Real-world cases, however, may extend the former structure to a general one, in which the first stage additionally generates final outputs, the second stage also produces exogenous inputs, and certain outputs of the second stage are re-utilised by the first. **Lu et al. (2016)** applied this structure to evaluate multiple investment trust corporations.

The above-mentioned systems have a series structure, in that they operate interdependently. In other types of networks, the internal divisions are placed in parallel without impacting one another **(Kao, 2012)**. There are two classes of parallel systems, based on their functions. Multi-component systems involve the assessment of DMUs with multiple divisions of the same function in that they use the same inputs to produce the same outputs. Research into the first class has been conducted to measure the efficiency of forest districts in Taiwan **(Kao, 2009b)**. The second class focuses on the multi-function systems, in which the internal divisions separate their operations by consuming their own inputs, although it is a common practice to also share inputs. Extensive research has examined such systems in various applications **(Beasley, 1995; Vaz et al., 2010; Lozano, 2015)**.

The investigation of the internal composition of a production system enables the improved use of the DEA approach **(Gan et al., 2019)**. However, treating the internal components of a system as black-boxes, continues to be widespread. This paper highlights this issue, considering the context of a parallel system. For example, the department of marketing at university X has two independent functions, teaching and research. If their internal operations are neglected, it cannot identify the potential sources of inefficiency, the way the inputs are further shared, and those layers with a beneficial impact on the respective section. To remedy these issues, each sub-system could be further split into sub-subsystems, and so on, to a reasonable level of detail. In a university department, one may want to identify sources of (in)efficiency down to the level of teaching programmes, for example.

The hierarchical structure has an eminent position in contemporary organisations. It can, inter alia, signify the organisational culture and dynamics, and coordinate the responsibilities of people across several departments and levels. Nevertheless, such a structure has to our knowledge not paid significant attention to exploring the internal operations of a network system, and in particular of a parallel system. In a university, a faculty typically operates as multiple parallel departments, each of which is further hierarchically structured across research, teaching, and enterprise.

Some approaches in DEA have systematically examined the hierarchical structures. **Castelli et al. (2004)** were among the first to propose two distinctive models to measure the performance of single-level and two-level hierarchical structures. This paper, that treated the internal components as independent DMUs, was improved by **Cook and Green (2005)**. The latter developed a model to measure the hierarchical efficiency at all levels simultaneously.

Recently, **Kao (2015)** developed a relational model for a single-stage hierarchical structure to measure both the overall system and its divisions' efficiencies at the same time.

He argued that this structure is identical to a parallel system **(Kao, 2009b)**, in that the system efficiency is decomposed into the weighted arithmetic average of the efficiencies of the units at the bottom level. **Kao (2015)** optimised the efficiency of the overall production system, considering only the constraints corresponding to the terminal divisions. **Li et al. (2020)** focused on the same hierarchical structure by additionally optimising the efficiencies of the terminal divisions as opposed to **Kao (2015)**. **Zhang and Chen (2019)** extended the concept of **Kao (2015)** to a generalised single-stage hierarchical structure wherein all internal units across the different levels can reflect a two-stage tandem system. To examine the relationship between the system and its sub-units, they introduced additive aggregation and multiplicative decomposition DEA models. **Gan et al. (2019)** suggested a general two-stage series process, in which each stage is no longer treated as a black-box, but is further elaborated into a hierarchical structure with multiple layers. They argued that a single-stage hierarchy cannot really correspond to complex production processes. A number of studies have been reported in this direction, such as **Bod'a et al. (2020)** and **Yu et al. (2021)**. We summarize the core literature, relevant to network-hierarchical DEA structures, in Table 1.

TABLE 4.1: Related literature on network-hierarchical DEA systems.

	Type of network	Efficiency measurement	Area of application
Black-box DEA model	SS	n/a	n/a
Kao (2012)	MCP	D	higher-education
Kao (2015)	SSH	D	higher-education
Lozano (2015)	MFP	SBM	pollution generation
Lu et al. (2016)	G2	D	investment trust corporations
Gan et al. (2019)	G2H	D	international shipping industry
Zhang and Chen (2019)	SSHMS	D & A	high-technology
Li et al. (2020)	2LH	$\times$	electric power generation
This Paper	MFPH	D & A	higher-education

D: decomposition, **A**: aggregation, **SBM**: slacks-based measure, **SS**: single-stage system, **MCP**: multi-component parallel system, **SSH**: single-stage hierarchical system, **MFP**: multi-function parallel system, **G2**: general two-stage system, **G2H**: general two-stage system with integrated hierarchies, **SSHMS**: single-stage hierarchical system with integrated multi-stage series processes, **2LH**: two-level hierarchical system, **MFPH**: multi-function parallel system with integrated hierarchies

A real-life organisation is likely to consist of several departments that could be further extended into a number of distinctive tasks, arranged either in sequence or in parallel. To better reflect the reality, we claim that these tasks can be then ordered as multi-layer hierarchical structures. These structures demonstrate that the strategic, tactical, and operational decisions cannot be made across the same level, by the same resources. The above case contributes to a more complex network system with embedded hierarchical structures. **Gan et al. (2019)** adopted such a structure, enabling the initial tasks to be interdependent (i.e., to be connected in series). The primary difference between our proposed system herein, against that of the earlier work of **Gan et al. (2019)** is that we seek to optimise the performance score of a system, in which the departments operate in parallel (i.e., they act independently from one another) incorporating as well a multi-function hierarchical structure. In this paper, we propose an additive decomposition model and a multiplicative aggregation model to measure and evaluate the operating performance of DMUs with a parallel multi-layer multi-function hierarchical structure. The proposed structure seeks to address the weaknesses of the traditional black-box DEA model and the parallel system of **Kao (2012)**; this is attained by improving the level of discriminatory power among efficient DMUs and by measuring the performance scores of not only the overall system and its parallel sub-systems, but also the internal units at all levels of each of the integrated hierarchies. Therefore, our proposed structure is shown to be a more accurate reflection of the entire production/operating process of several large scale organisations.

The remainder of the paper is organised as follows. Section 2 briefly describes the methodological background. Section 3 proposes new models to evaluate DMUs with multi-function parallel network hierarchical structure. Several properties of such a system are also analysed. Section 4 validates the proposed models with an application in the education sector. Finally, Section 5 presents conclusions and further research.

4.2 Methodological Background

In this section, we explore the network nature of two established systems: the parallel with shared inputs, and the single-stage hierarchical structure. This will ease the presentation of the advanced structure and its mathematical models proposed in Section 4.3.

4.2.1 A parallel system with shared inputs

In a real-life application, the core of a production system may be composed of multiple divisions with distinctive functions, operating independently among themselves. Such a system tends to be a more accurate picture of the reality, once joint inputs, shared by a number of divisions, are involved, other than their own inputs. **Beasley (1995)** and **Molinero (1996)** proposed a system with p parallel processes or divisions. In this system, see also Figure 4.1, the X_{ij}, X^S_{lj} and Y_{rj} are the ith dedicated input value, the lth shared input value, and the rth final output value, respectively, of DMU_j $(j = 1, 2, ..., n)$. Let $M = \{1, 2, ..., m\}$, $Q = \{1, 2, ..., q\}$, and $S = \{1, 2, ..., s\}$ be the index sets associated with the dedicated inputs, the shared inputs, and the final outputs, respectively. The division k $(k = 1, 2, ..., p)$ of DMU_j utilises the division-dedicated input i with value $X_{ij}^{(k)}$, $i \in I^{(k)}$, $M = \bigcup_{k \in P} I^{(k)}$ such that $I^{(k)} \cap I^{(i)} = \emptyset$, $\forall\ k, i \in P$, and a proportion $\alpha_l^{(k)}$ of the shared input $l \in Q$ with value X_{lj}^S, to produce the final output r with value $Y_{rj}^{(k)}$, $r \in O^{(k)}$, $S = \bigcup_{k \in P} O^{(k)}$ such that $O^{(k)} \cap O^{(o)} = \emptyset$, $\forall\ k, o \in P$. In such a system, the total division-specific and shared inputs consumed, and total outputs produced by the p divisions of DMU_j are $X_{ij} = \sum_{k=1}^{p} X_{ij}^{(k)}$, $X_{lj}^S = \sum_{k=1}^{p} \alpha_l^{(k)} X_{lj}^S$, and $Y_{rj} = \sum_{k=1}^{p} Y_{rj}^{(k)}$, respectively.

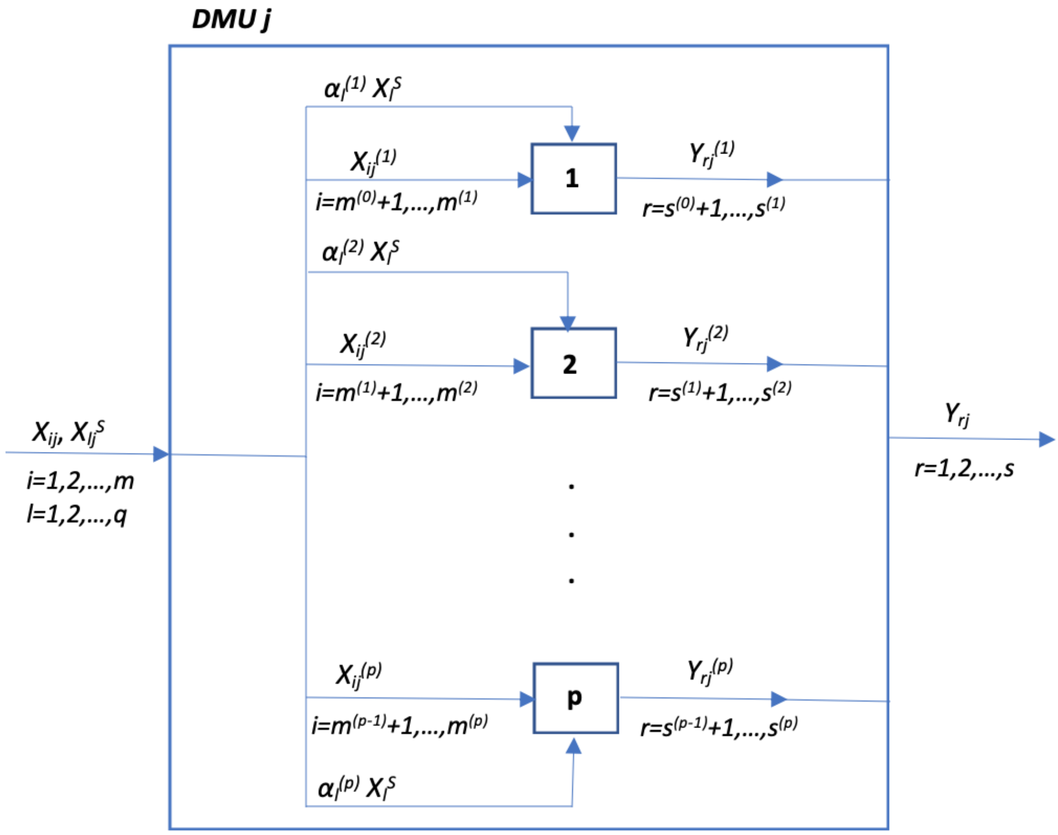

FIGURE 4.1: Parallel system with shared inputs

To measure the performance of the overall system of the target DMU_o, **Beasley (1995)** introduced and later **Kao (2012)** and **Kao (2017)** validated the following model under constant returns to scale:

$$
\begin{aligned}
E_o = Max \quad & \sum_{r=1}^{s} \mu_{ro} Y_{ro} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_{io} X_{io} + \sum_{l=1}^{q} t_{lo} X_{lo}^{S} = 1, \\
& \sum_{r \in O^{(k)}} \mu_{ro} Y_{rj}^{(k)} - (\sum_{i \in I^{(k)}} \nu_{io} X_{ij}^{(k)} + \sum_{l=1}^{q} t_{lo} \alpha_l^{(k)} X_{lj}^{S}) \leq 0, \ \forall\, j,k, \\
& t_{lo}, \nu_{io}, \mu_{ro} \geq \epsilon, \ \forall\, l,i,r,
\end{aligned}
\tag{4.1}
$$

where t_{lo}, ν_{io}, and μ_{ro} are the positive optimal multipliers, and ϵ is an infinitesimal non-Archimedean number. According to the relational model **(1)**, the overall performance score of the evaluated DMU to be maximised is the ratio of the total outputs to that of inputs. It is also required that the aggregation of outputs should not exceed the aggregation of inputs, for every internal division. In such a model, $\alpha_l^{(k)}$ is a parameter that is objectively designated by the decision maker (possibly based on historical data) prior to solving the corresponding mathematical model. It is also ensured that the sum of the proportions of the lth shared input is 1. However, if it is treated as the most favourable value, reflected from the data, then it is additionally essential to involve the following constraints: $L_{lo}^{(k)} \leq \alpha_l^{(k)} \leq U_{lo}^{(k)}$, $\sum_{k=1}^{p} \alpha_l^{(k)} = 1$, and $\alpha_l^{(k)} \geq 0, \ \forall\, l,k$. At optimality, the system efficiency is calculated as $E_o = \sum_{r=1}^{s} \mu_{ro}^{*} Y_{ro} / (\sum_{i=1}^{m} \nu_{io}^{*} X_{io} + \sum_{l=1}^{q} t_{lo}^{*} X_{lo}^{S})$, and the division efficiencies as $E_o^{(k)} = \sum_{r \in O^{(k)}} \mu_{ro}^{*} Y_{ro}^{(k)} / (\sum_{i \in I^{(k)}} \nu_{io}^{*} X_{io}^{(k)} + \sum_{l=1}^{q} t_{lo}^{*} \alpha_l^{(k)} X_{lj}^{S})$. A property of this structure is that the system efficiency equals to the weighted average of its division efficiencies **(Kao, 2009b)**.

To the best of our knowledge, the internal divisions of such a commonly used structure are still treated as black-boxes. This may hinder our efforts to gain further insight into more complex and realistic cases, regarding the activities of a department and the mechanisms behind a core business task.

4.2.2 A single-stage hierarchical structure

A relatively recent network system is that of a hierarchical structure, embedded either in a single-stage or in a general two-stage series network. Its adoption may help the investigation of the operational procedures. As discussed in Section 1, **Kao (2015)** proposed a relational model to evaluate the performance of the overall system and its internal units, reflecting a single-stage hierarchical structure with three levels.

Consider a system with the general hierarchical structure shown in Figure 4.2 **(Kao, 2017)**. The system has q levels and is an extension of the three-level system of **Kao (2015)**. The first level, for example, consists of $p^{(1)}$ divisions, each of which is decomposed into several divisions at the follower level. The kth level $(k = 2, 3, ..., q)$ has a total of $p^{(k)} - p^{(k-1)}$ divisions subordinated to the $p^{(k-1)} - p^{(k-2)}$ divisions at the $(k-1)$th level. Denote $P^{[1]} = \{1, 2, ..., p^{(1)}\}$ and $P^{[k]} = \{p^{(k-1)} + 1, ..., p^{(k)}\}$, as the sets of the divisions in the first and the kth level $(k = 2, 3, ..., q)$, respectively. Moreover, let $S(l)$ be the set of divisions viewed as subordinates to division l. If $S(l) = \emptyset$, then l is referred to as terminal. Let T denote the set of the terminal divisions. They are enabled to generate the outputs, while receiving inputs allocated from their parent unit (the immediate predecessor). On the other hand, the intermediate units i.e. the non-terminal divisions cannot produce outputs themselves, but they can distribute their inputs to their subordinate divisions at the next level.

In such a single-stage system, let X_{ij} and Y_{rj} be the ith input $(i = 1, 2, ..., m)$ and rth output $(r = 1, 2, ..., s)$ for the DMU_j $(j = 1, 2, ..., n)$. Division l distributes its inputs $X_{ij}^{(l)}$, $i \in I^{(l)}$, received by its parent unit, to its subordinate divisions $\xi \in S(l)$, and collects the outputs $Y_{rj}^{(l)}$, $r \in O^{(l)}$ received from its subordinate divisions. Hence, in mathematical terms we have $X_{ij}^{(l)} = \sum_{\xi \in S(l)} X_{ij}^{(\xi)}$ and $Y_{rj}^{(l)} = \sum_{\xi \in S(l)} Y_{rj}^{(\xi)}$.

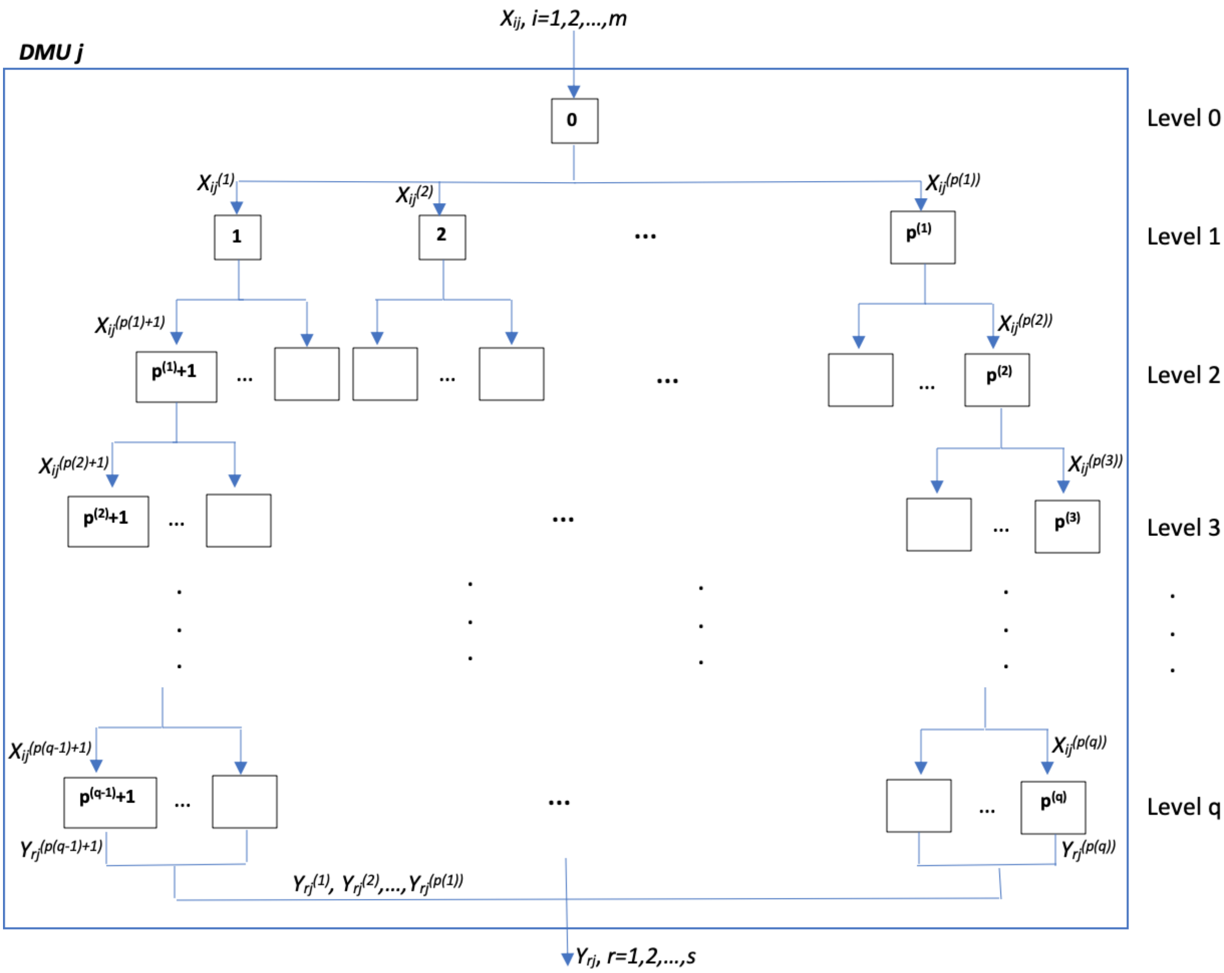

FIGURE 4.2: General single-stage hierarchical structure

Kao (2015) developed the relational input-oriented model **(4.2)** to optimise the efficiency score of the target DMU_o arranged as a multi-layer hierarchical structure within a single-stage system.

$$
\begin{aligned}
E'_o = Max \quad & \sum_{r=1}^{s} \mu_{ro} Y_{ro} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_{io} X_{io} = 1, \\
& \sum_{r \in O^{(l)}} \mu_{ro} Y_{rj}^{(l)} - \sum_{i \in I^{(l)}} \nu_{io} X_{ij}^{(l)} \leq 0, \ l \in T, \ \forall \ j, \\
& \nu_{io}, \mu_{ro} \geq \epsilon, \ \forall \ i, r.
\end{aligned}
\tag{4.2}
$$

To avoid redundancy, Kao highlighted that only the terminal divisions need to be taken into account. In mathematical symbols, $l \in T$. One of the properties of the system is that its overall efficiency is decomposed into the weighted average of those of the terminal divisions. To apply model **(4.2)**, we should ensure that all DMUs have the same hierarchical structure. In particular, for every DMU, a unit at the leader level should have the same number of subordinate units at the follower level, operating different functions.

In Section 4.3.1, the scenario of the integration of such a hierarchical structure into the internal divisions of a parallel system with shared inputs will be thoroughly discussed. This direction can successfully enhance the performance measurement in more complex systems within the production and operations management.

4.3 Models Development

Real-life companies can have a complex corporate structure. The complexity corresponds to their numerous (tangible and intangible) resources, either being interactive or entirely independent, in any department. The utilities of such a structure are to successfully adapt to the constant changes of the internal and external environment, to comply with customers' requirements, and to minimise fixed and variable costs.

There are at least three separate (traditional single-stage and network) production systems, proposed in the DEA-literature, that have intertwined with multi-layer hierarchical structures: *(i)* a three-level multi-function hierarchical structure embedded in the core of a single-stage system **(Kao, 2015)**, *(ii)* a three-level with two-stage processes hierarchical structure embedded in a single-stage system **(Zhang and Chen, 2019)**, and *(iii)* a multi-level hierarchical structure integrated into an operating stage of a general two-stage series system **(Gan et al., 2019)**. In this paper, we extend the above list by considering the case of several parallel processes, wherein each sub-system integrates a multi-function hierarchical structure. The new system is introduced in Figure 4.3.

4.3.1 Parallel-hierarchical network DEA model

Based on the consolidated idea of **Kao (2015)**, the evaluated DMUs should have the same network-hierarchical structure; this can set the basis for a less demanding comparison amongst them. We have combined the ideas developed in Sections 4.2.1 and 4.2.2 into a situation like in Figure 4.3. From the perspective of Figure 4.3, the sub-systems of a system (DMU) must execute different operations, and each sub-system is obliged to have the same function with its counterpart in each of the other DMUs. In addition, the hierarchical structures of the different sub-systems of a DMU may vary in terms of the number and the arrangement of their internal units. However, the hierarchical structure of a certain sub-system of DMU_j $(j = 1, 2, ..., n)$ should be identical with the counterpart structure of the sub-system in each of the remaining DMUs.

On a macro level, the proposed system consists of two successive layers. The external one is associated with the action of retrieving managerial data from the entire system. This examines the overall performance of the DMU under consideration. The system applies m sub-system specific inputs and q shared inputs to generate s final outputs. Subsequently, in the interior part of the system, we detect p sub-systems connected in parallel, that is they are independent among each other and they cannot typically exchange information. A sub-system k $(k = 1, 2, ..., p)$ consumes the dedicated inputs $X_{ij}^{(k)}$, $i \in I^{(k)} \subseteq \{1, 2, ..., m\}$, and the shared inputs $\alpha_l^{(k)} X_{lj}^{S}$ $(l = 1, 2, ..., q)$ to generate the final outputs $Y_{rj}^{(k)}$, $r \in O^{(k)} \subseteq \{1, 2, ..., s\}$. This layer evaluates the performance of each department/task, which is an integral part of the whole system.

On a micro level, in the interior of a sub-system, we identify a three-level multi-function hierarchical structure. The top level 0 (sub-system k) has two subordinate units, labelled (1) and (2), performing distinctive functions, at level 1. Functions (1) and (2) have in rotation three subordinate units (1.1), (1.2), and (1.3), and two subordinate units, (2.1) and (2.2), respectively, at level 2. Only unit (2.2) has two sub-units (2.2.1) and (2.2.2) at the bottom level 3. The internal units (1.1), (1.2), (1.3), (2.1), (2.2.1), and (2.2.2) are characterised as terminal, since they cannot be further broken down into several subordinate units. Note that the hierarchical structure presented herein is indicative and may be subject to modifications, reflecting the respective business environment. The internal unit u $(u = 1, 2, 1.1, 1.2, 1.3, 2.1, 2.2, 2.2.1, 2.2.2)$ of sub-system k of DMU_j allocates the sub-system specific inputs $X_{ij}^{(k)_u}$, $i \in I^{(k)}$, and a proportion $\theta_l^{(k)_u}$ of the lth shared input X_{lj}^{S}, received by its parent unit, to its subordinate units at the follower level, and collects the outputs $Y_{rj}^{(k)_u}$, $r \in O^{(k)}$, received from its subordinate units.

Taking the structure of the above system into account, we obtain the following equalities:

(i) $X_{ij} = \sum_{k=1}^{p} X_{ij}^{(k)} = \sum_{k=1}^{p}(X_{ij}^{(k)_1} + X_{ij}^{(k)_2}) = \sum_{k=1}^{p}(X_{ij}^{(k)_{1.1}} + X_{ij}^{(k)_{1.2}} + X_{ij}^{(k)_{1.3}} + X_{ij}^{(k)_{2.1}} + X_{ij}^{(k)_{2.2}}) = \sum_{k=1}^{p}(X_{ij}^{(k)_{1.1}} + X_{ij}^{(k)_{1.2}} + X_{ij}^{(k)_{1.3}} + X_{ij}^{(k)_{2.1}} + X_{ij}^{(k)_{2.2.1}} + X_{ij}^{(k)_{2.2.2}})$, $\forall\, i, j$,

(ii) $X_{lj}^{S} = \sum_{k=1}^{p} \alpha_l^{(k)_0} X_{lj}^{S} = \sum_{k=1}^{p}(\sum_{k_1=1}^{2} \beta_l^{(k)_{k_1}} \alpha_l^{(k)_0} X_{lj}^{S}) = \sum_{k=1}^{p}(\sum_{k_2=1.1}^{1.3} \gamma_l^{(k)_{k_2}} \beta_l^{(k)_1} \alpha_l^{(k)_0} X_{lj}^{S} + \sum_{k_3=2.1}^{2.2} \gamma_l^{(k)_{k_3}} \beta_l^{(k)_2} \alpha_l^{(k)_0} X_{lj}^{S}) = \sum_{k=1}^{p}(\sum_{k_2=1.1}^{1.3} \gamma_l^{(k)_{k_2}} \beta_l^{(k)_1} \alpha_l^{(k)_0} X_{lj}^{S} + \gamma_l^{(k)_{2.1}} \beta_l^{(k)_2} \alpha_l^{(k)_0} X_{lj}^{S} + \sum_{k_4=2.2.1}^{2.2.2} \delta_l^{(k)_{k_4}} \gamma_l^{(k)_{2.2}} \beta_l^{(k)_2} \alpha_l^{(k)_0} X_{lj}^{S})$, $\forall\, l, k, j$,

(iii) $Y_{rj} = \sum_{k=1}^{p} Y_{rj}^{(k)} = \sum_{k=1}^{p}(Y_{rj}^{(k)_1} + Y_{rj}^{(k)_2}) = \sum_{k=1}^{p}(Y_{rj}^{(k)_{1.1}} + Y_{rj}^{(k)_{1.2}} + Y_{rj}^{(k)_{1.3}} + Y_{rj}^{(k)_{2.1}} + Y_{rj}^{(k)_{2.2}}) = \sum_{k=1}^{p}(Y_{rj}^{(k)_{1.1}} + Y_{rj}^{(k)_{1.2}} + Y_{rj}^{(k)_{1.3}} + Y_{rj}^{(k)_{2.1}} + Y_{rj}^{(k)_{2.2.1}} + Y_{rj}^{(k)_{2.2.2}})$, $\forall\, r, j$.

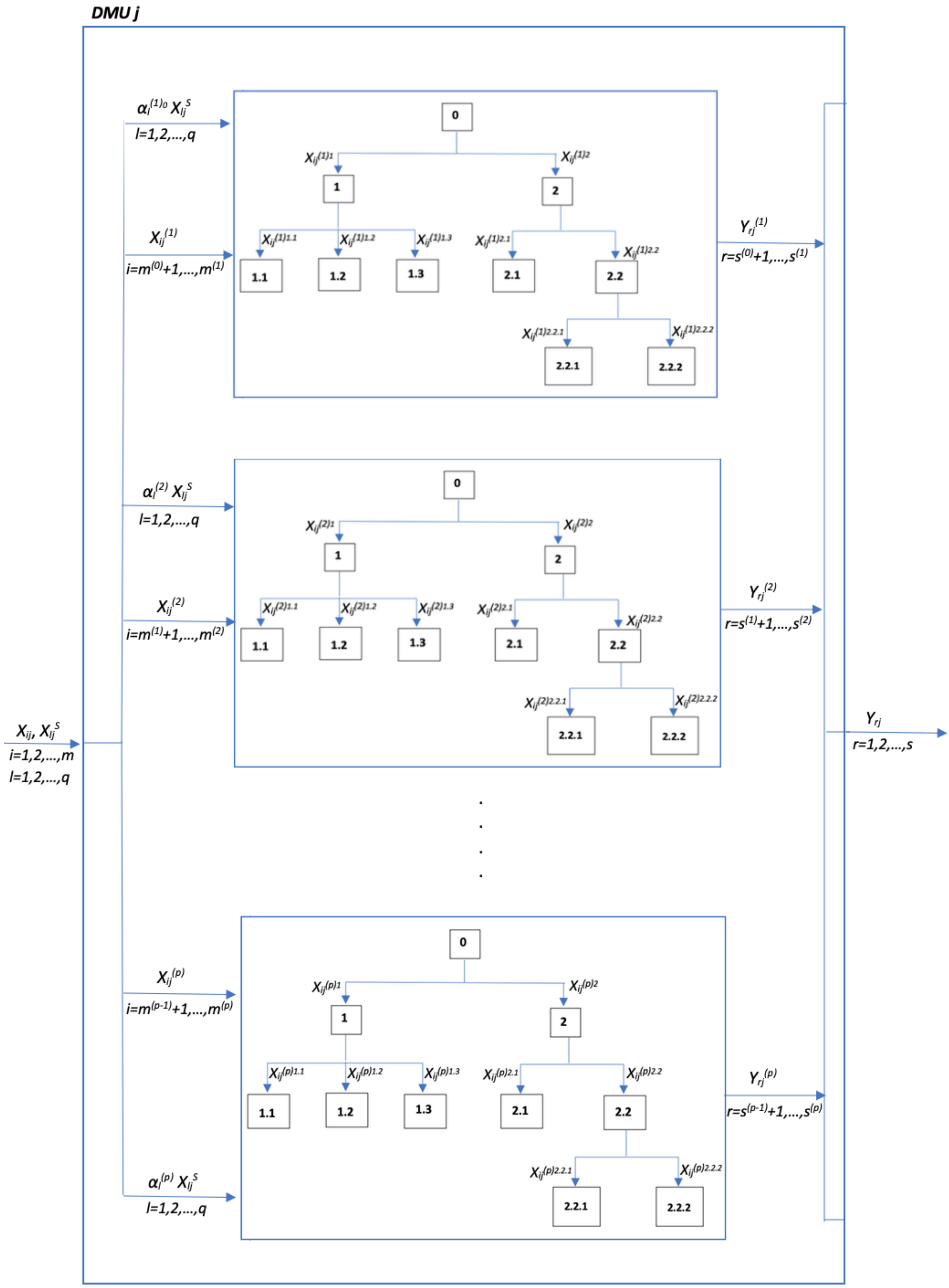

FIGURE 4.3: An embedded hierarchical network structure within a multi-function parallel system

To model the proposed network-hierarchical structure, we adopt two main properties relevant to the network relational model conceptualised by **Kao (2009a)** and **Kao (2015)**. We, firstly, assume that the same factor, either the sub-system dedicated inputs X_{ij}, the shared inputs X^S_{lj} or the outputs Y_{rj}, has the same weight ν_{io}, t_{lo}, and μ_{ro}, respectively, no matter which process (system, sub-system or internal unit of the hierarchy) it corresponds to. This is a common assumption of a relational model within network DEA **(Kao, 2009a)**. Furthermore, the system cannot be handled anymore as a whole unit, but rather as a network with three successive layers, whose operations should be taken into consideration. Therefore, the aggregate output should be less than or equal to the aggregate input for each internal (sub-system or hierarchy) or external (system) process, for each DMU. Our objective function aims to maximise the ratio of the aggregated amount of final outputs to that of the inputs (both the sub-system specific and the shared inputs) for the system, visible from the outside.

The typical ratio-form input-oriented network-hierarchical DEA model under constant returns to scale for DMU_o can be described as follows:

$$
\begin{aligned}
&E_o^{HN} = Max \ \frac{\sum_{r=1}^{s} \mu_{ro} Y_{ro}}{\sum_{i=1}^{m} \nu_{io} X_{io} + \sum_{l=1}^{q} t_{lo} X_{lo}^S} \\
&\text{subject to} \\
&\sum_{r=1}^{s} \mu_{ro} Y_{rj} - (\sum_{i=1}^{m} \nu_{io} X_{ij} + \sum_{l=1}^{q} t_{lo} X_{lj}^S) \leq 0, \ \ \forall j, \\
&\sum_{r\in O^{(k)}} \mu_{ro} Y_{rj}^{(k)} - (\sum_{i\in I^{(k)}} \nu_{io} X_{ij}^{(k)} + \sum_{l=1}^{q} t_{lo} \alpha_l^{(k)0} X_{lj}^S) \leq 0, \ \ \forall j,k, \\
&\sum_{r\in O^{(k)}} \mu_{ro} Y_{rj}^{(k)k_1} - (\sum_{i\in I^{(k)}} \nu_{io} X_{ij}^{(k)k_1} + \sum_{l=1}^{q} t_{lo} \beta_l^{(k)k_1} \alpha_l^{(k)0} X_{lj}^S) \leq 0, \ \ \forall j,k, \ \ k_1 = 1,2, \\
&\sum_{r\in O^{(k)}} \mu_{ro} Y_{rj}^{(k)k_2} - (\sum_{i\in I^{(k)}} \nu_{io} X_{ij}^{(k)k_2} + \sum_{l=1}^{q} t_{lo} \gamma_l^{(k)k_2} \beta_l^{(k)1} \alpha_l^{(k)0} X_{lj}^S) \leq 0, \ \ \forall j,k, \ \ k_2 = 1.1, 1.2, 1.3, \\
&\sum_{r\in O^{(k)}} \mu_{ro} Y_{rj}^{(k)k_3} - (\sum_{i\in I^{(k)}} \nu_{io} X_{ij}^{(k)k_3} + \sum_{l=1}^{q} t_{lo} \gamma_l^{(k)k_3} \beta_l^{(k)2} \alpha_l^{(k)0} X_{lj}^S) \leq 0, \ \ \forall j,k, \ \ k_3 = 2.1, 2.2, \qquad (4.3) \\
&\sum_{r\in O^{(k)}} \mu_{ro} Y_{rj}^{(k)k_4} - (\sum_{i\in I^{(k)}} \nu_{io} X_{ij}^{(k)k_4} + \sum_{l=1}^{q} t_{lo} \delta_l^{(k)k_4} \gamma_l^{(k)2.2} \beta_l^{(k)2} \alpha_l^{(k)0} X_{lj}^S) \leq 0, \ \ \forall j,k, \ \ k_4 = 2.2.1, 2.2.2, \\
&\sum_{k=1}^{p} \alpha_l^{(k)0} = 1, \ \sum_{k_1=1}^{2} \beta_l^{(k)k_1} = 1, \ \sum_{k_2=1.1}^{1.3} \gamma_l^{(k)k_2} = 1, \ \sum_{k_3=2.1}^{2.2} \gamma_l^{(k)k_3} = 1, \ \sum_{k_4=2.2.1}^{2.2.2} \delta_l^{(k)k_4} = 1, \ \ \forall \, l,k, \\
&L_l^{(k,n)} \leq \alpha_l^{(k)0} / \alpha_l^{(n)0} \leq U_l^{(k,n)}, \ \forall \, l, k = 1, ..., p, \ n = 1, ..., p, \ k \neq n, \\
&L_l^{(k)1,2} \leq \beta_l^{(k)1} / \beta_l^{(k)2} \leq U_l^{(k)1,2}, \ \forall \, l,k, \\
&L_l^{(k)2.1,2.2} \leq \gamma_l^{(k)2.1} / \gamma_l^{(k)2.2} \leq U_l^{(k)2.1,2.2}, \ \forall \, l,k, \\
&L_l^{(k)k_2,n_2} \leq \gamma_l^{(k)k_2} / \gamma_l^{(k)n_2} \leq U_l^{(k)k_2,n_2}, \ \forall \, l, k_2 = 1.1, 1.2, 1.3, \ n_2 = 1.1, 1.2, 1.3, \ k_2 \neq n_2, \\
&L_l^{(k)2.2.1,2.2.2} \leq \delta_l^{(k)2.2.1} / \delta_l^{(k)2.2.2} \leq U_l^{(k)2.2.1,2.2.2}, \ \forall \, l,k, \\
&\alpha_l^{(k)0}, \beta_l^{(k)k_1}, \gamma_l^{(k)k_2}, \gamma_l^{(k)k_3}, \delta_l^{(k)k_4} \geq 0, \ \ t_{lo}, \nu_{io}, \mu_{ro} \geq \epsilon, \ \forall \, r, i, l, k, k_1, k_2, k_3, k_4.
\end{aligned}
$$

where ν_{io}, t_{lo}, and μ_{ro} are ensured to be positive, by integrating the small non-Archimedean parameter ϵ. In model **(4.3)**, there are four groups of constraint sets. The first group (first constraint set) reflects the entire system. The second group (second constraint set) is pertinent to the performance of each sub-system k. The third group (from third to sixth constraints sets) illustrates the operations of each of the internal units of the hierarchical structure embedded into sub-system k. For a unit at a certain level, the aggregation of outputs produced by its subordinate units at the follower level should not exceed the aggregation of inputs allocated to it by its parent unit. For example, regarding the unit (2.2) of level 2, it ought to satisfy the constraint $\sum_{r\in O^{(k)}} \mu_{ro} Y_{rj}^{(k)2.2} - (\sum_{i\in I^{(k)}} \nu_{io} X_{ij}^{(k)2.2} + \sum_{l=1}^{q} t_{lo} \gamma_l^{(k)2.2} \beta_l^{(k)2} \alpha_l^{(k)0} X_{lj}^S) \leq 0, \ \ \forall j,k$. The final group (remaining constraints sets) indicates that the proportion of the shared input allocated to the respective internal unit of sub-system k should be treated as a variable. In other words, it seeks for the optimal most favourable value. For instance, with respect to the proportion variable $\delta_l^{(k)_{k_4}}$, two constraint sets are formulated to reflect its dynamics: *(i)* the $\sum_{k_4=2.2.1}^{2.2.2} \delta_l^{(k)_{k_4}} = 1,\ \forall\ l,k$, denotes that the total sum of the proportions of shared resources allocated to the internal units of the third level should be 1, and *(ii)* the $L_l^{(k)_{2.2.1,2.2.2}} \leq \delta_l^{(k)_{2.2.1}}/\delta_l^{(k)_{2.2.2}} \leq U_l^{(k)_{2.2.1,2.2.2}},\ \forall\ l,k$, illustrates that the ratio of the proportions of shared resources in that level is bounded from below by $L_l^{(k)_{2.2.1,2.2.2}}$ and above by $U_l^{(k)_{2.2.1,2.2.2}}$. These are user-specified parameters and typically reflect the requirements of the production.

The constraint corresponding to the system is the sum of those corresponding to the sub-system: $\sum_{r=1}^{s} \mu_{ro} Y_{rj} - (\sum_{i=1}^{m} \nu_{io} X_{ij} + \sum_{l=1}^{q} t_{lo} X_{lj}^S) = \sum_{r=1}^{s} \mu_{ro} \sum_{k=1}^{p} Y_{rj}^{(k)} - (\sum_{i=1}^{m} \nu_{io} \sum_{k=1}^{p} X_{ij}^{(k)} + \sum_{l=1}^{q} t_{lo} \sum_{k=1}^{p} \alpha_l^{(k)_0} X_{lj}^S) = \sum_{k=1}^{p} [\sum_{r=1}^{s} \mu_{ro} Y_{rj}^{(k)} - (\sum_{i=1}^{m} \nu_{io} X_{ij}^{(k)} + \sum_{l=1}^{q} t_{lo} \alpha_l^{(k)_0} X_{lj}^S)]$. The system constraint is thus redundant and can be removed. In addition, the constraint corresponding to the sub-system is the sum of those corresponding to the level 1: $\sum_{r=1}^{s} \mu_{ro} Y_{rj}^{(k)} - (\sum_{i=1}^{m} \nu_{io} X_{ij}^{(k)} + \sum_{l=1}^{q} t_{lo} \alpha_l^{(k)_0} X_{lj}^S) = \sum_{r=1}^{s} \mu_{ro} \sum_{k_1=1}^{2} Y_{rj}^{(k)_{k_1}} - (\sum_{i=1}^{m} \nu_{io} \sum_{k_1=1}^{2} X_{ij}^{(k)_{k_1}} + \sum_{l=1}^{q} t_{lo} \sum_{k_1=1}^{2} \beta_l^{(k)_{k_1}} \alpha_l^{(k)_0} X_{lj}^S) = \sum_{k_1=1}^{2} [\sum_{r=1}^{s} \mu_{ro} Y_{rj}^{(k)_{k_1}} - (\sum_{i=1}^{m} \nu_{io} X_{ij}^{(k)_{k_1}} + \sum_{l=1}^{q} t_{lo} \beta_l^{(k)_{k_1}} \alpha_l^{(k)_0} X_{lj}^S)]$. The sub-system constraint is redundant and may be omitted. By the same token, we identify that the constraints sets regarding the internal units of level 1 and the unit (2.2) of level 2 within each sub-system are additionally redundant and can be deleted. Hence, only the internal operations of the terminal units at the bottom level of the hierarchy for each sub-system should be taken into account.

With regard to the constraint group about proportions, we should also discuss about the removal of $\sum_{k=1}^{p} \alpha_l^{(k)_0} = 1$. In particular, the constraint $\sum_{k_1=1}^{2} \beta_l^{(k)_{k_1}} = 1, \forall\ k, l$, can be transformed into $\sum_{k_1=1}^{2} (\beta_l^{(k)_{k_1}} \alpha_l^{(k)_0} X_{lj}^S) - \alpha_l^{(k)_0} X_{lj}^S = 0, \forall\ k, l, j$. Since the constraint $\sum_{k=1}^{p} \alpha_l^{(k)_0} = 1 \Leftrightarrow \sum_{k=1}^{p} \alpha_l^{(k)_0} X_{lj}^S - X_{lj}^S = 0 \Leftrightarrow \sum_{k=1}^{p} \sum_{k_1=1}^{2} (\beta_l^{(k)_{k_1}} \alpha_l^{(k)_0} X_{lj}^S) - \sum_{k=1}^{p} \alpha_l^{(k)_0} X_{lj}^S = 0 \Leftrightarrow \sum_{k=1}^{p} [\sum_{k_1=1}^{2} (\beta_l^{(k)_{k_1}} \alpha_l^{(k)_0} X_{lj}^S) - \alpha_l^{(k)_0} X_{lj}^S] = 0$, is the sum of those corresponding to $\sum_{k_1=1}^{2} \beta_l^{(k)_{k_1}} = 1$, then it is redundant.

Model **(4.3)** is nonlinear due to its nonlinear objective function and several nonlinear terms, such as $t_{lo}\alpha_l^{(k)_0}$, $t_{lo}\beta_l^{(k)_{k_1}} \alpha_l^{(k)_0}$, and $t_{lo}\gamma_l^{(k)_{k_2}} \beta_l^{(k)_1} \alpha_l^{(k)_0}$. With respect to the objective function, we can assign a value of 1 to the denominator as a constraint, and maximise the value of the numerator. The other nonlinear terms can be linearised by variable transformations as set out below: $t_{lo}\alpha_l^{(k)_0} = \nu_{lo}^{(k)_0}$, $t_{lo}\beta_l^{(k)_{k_1}} \alpha_l^{(k)_0} = b_{lo}^{(k)_{k_1}}$, $t_{lo}\gamma_l^{(k)_{k_2}} \beta_l^{(k)_1} \alpha_l^{(k)_0} = c_{lo}^{(k)_{k_2}}$, $t_{lo}\gamma_l^{(k)_{k_3}} \beta_l^{(k)_2} \alpha_l^{(k)_0} = c_{lo}^{(k)_{k_3}}$, and $t_{lo}\delta_l^{(k)_{k_4}} \gamma_l^{(k)_{2.2}} \beta_l^{(k)_2} \alpha_l^{(k)_0} = d_{lo}^{(k)_{k_4}}$, $\forall\ l, k, k_1, k_2, k_3, k_4$. Thus, we obtain the following linear model **(4.4)**:

$$
\begin{aligned}
&E_o^{HN} = Max \sum_{r=1}^{s} \mu_{ro} Y_{ro} \\
&\text{subject to} \\
&\sum_{i=1}^{m} \nu_{io} X_{io} + \sum_{l=1}^{q} t_{lo} X_{lo}^{S} = 1, \\
&\sum_{r=1}^{s} \mu_{ro} Y_{rj} - (\sum_{i=1}^{m} \nu_{io} X_{ij} + \sum_{l=1}^{q} t_{lo} X_{lj}^{S}) \leq 0, \ \ \forall j, \\
&\sum_{r\in O^{(k)}} \mu_{ro} Y_{rj}^{(k)} - (\sum_{i\in I^{(k)}} \nu_{io} X_{ij}^{(k)} + \sum_{l=1}^{q} \nu_{lo}^{(k)_0} X_{lj}^{S}) \leq 0, \ \ \forall j,k, \\
&\sum_{r\in O^{(k)}} \mu_{ro} Y_{rj}^{(k)_{k_1}} - (\sum_{i\in I^{(k)}} \nu_{io} X_{ij}^{(k)_{k_1}} + \sum_{l=1}^{q} b_{lo}^{(k)_{k_1}} X_{lj}^{S}) \leq 0, \ \ \forall j,k, \ \ k_1 = 1,2, \\
&\sum_{r\in O^{(k)}} \mu_{ro} Y_{rj}^{(k)_{k_2}} - (\sum_{i\in I^{(k)}} \nu_{io} X_{ij}^{(k)_{k_2}} + \sum_{l=1}^{q} c_{lo}^{(k)_{k_2}} X_{lj}^{S}) \leq 0, \ \ \forall j,k, \ \ k_2 = 1.1, 1.2, 1.3, \\
&\sum_{r\in O^{(k)}} \mu_{ro} Y_{rj}^{(k)_{k_3}} - (\sum_{i\in I^{(k)}} \nu_{io} X_{ij}^{(k)_{k_3}} + \sum_{l=1}^{q} c_{lo}^{(k)_{k_3}} X_{lj}^{S}) \leq 0, \ \ \forall j,k, \ \ k_3 = 2.1, 2.2, \\
&\sum_{r\in O^{(k)}} \mu_{ro} Y_{rj}^{(k)_{k_4}} - (\sum_{i\in I^{(k)}} \nu_{io} X_{ij}^{(k)_{k_4}} + \sum_{l=1}^{q} d_{lo}^{(k)_{k_4}} X_{lj}^{S}) \leq 0, \ \ \forall j,k, \ \ k_4 = 2.2.1, 2.2.2, \qquad (4.4) \\
&\sum_{k=1}^{p} \nu_{lo}^{(k)_0} = t_{lo}, \sum_{k_1=1}^{2} b_{lo}^{(k)_{k_1}} = \nu_{lo}^{(k)_0}, \sum_{k_2=1.1}^{1.3} c_{lo}^{(k)_{k_2}} = b_{lo}^{(k)_1}, \forall\, l,k, \\
&\sum_{k_3=2.1}^{2.2} c_{lo}^{(k)_{k_3}} = b_{lo}^{(k)_2}, \sum_{k_4=2.2.1}^{2.2.2} d_{lo}^{(k)_{k_4}} = c_{lo}^{(k)_{2.2}}, \forall\, l,k, \\
&\nu_{lo}^{(n)_0} L_l^{(k,n)} \leq \nu_{lo}^{(k)_0} \leq \nu_{lo}^{(n)_0} U_l^{(k,n)}, \ \forall\, l, k = 1,...,p, \ n = 1,...,p, \ k \neq n, \\
&b_{lo}^{(k)_2} L_l^{(k)_{1,2}} \leq b_{lo}^{(k)_1} \leq b_{lo}^{(k)_2} U_l^{(k)_{1,2}}, \ \forall\, l,k, \\
&c_{lo}^{(k)_{2.2}} L_l^{(k)_{2.1,2.2}} \leq c_{lo}^{(k)_{2.1}} \leq c_{lo}^{(k)_{2.2}} U_l^{(k)_{2.1,2.2}}, \ \forall\, l,k, \\
&c_{lo}^{(k)_{n_2}} L_l^{(k)_{k_2,n_2}} \leq c_{lo}^{(k)_{k_2}} \leq c_{lo}^{(k)_{n_2}} U_l^{(k)_{k_2,n_2}}, \ \forall\, l, k_2 = 1.1, 1.2, 1.3, \ n_2 = 1.1, 1.2, 1.3, \ k_2 \neq n_2, \\
&d_{lo}^{(k)_{2.2.2}} L_l^{(k)_{2.2.1,2.2.2}} \leq d_{lo}^{(k)_{2.2.1}} \leq d_{lo}^{(k)_{2.2.2}} U_l^{(k)_{2.2.1,2.2.2}}, \ \forall\, l,k, \\
&t_{lo}, \nu_{io}, \mu_{ro}, \nu_{lo}^{(k)_0}, b_{lo}^{(k)_{k_1}}, c_{lo}^{(k)_{k_2}}, c_{lo}^{(k)_{k_3}}, d_{lo}^{(k)_{k_4}} \geq \epsilon \ \forall\, r, i, l, k, k_1, k_2, k_3, k_4.
\end{aligned}
$$

After an optimal solution $(t_{lo^*}, \nu_{io^*}, \mu_{ro^*}, \nu_{lo}^{(k)_{0^*}}, b_{lo}^{(k)_{k_1^*}}, c_{lo}^{(k)_{k_2^*}}, c_{lo}^{(k)_{k_3^*}}, d_{lo}^{(k)_{k_4^*}})$ is obtained for DMU_o under the linear model **(4.4)**, the efficiencies of the overall system, its sub-systems, and its internal units at all levels of the hierarchical structure within each sub-system are calculated as follows:

(i) $E_o^{HN} = \sum_{r=1}^{s} \mu_{ro^*} Y_{ro} / (\sum_{i=1}^{m} \nu_{io^*} X_{io} + \sum_{l=1}^{q} t_{lo^*} X_{lo}^{S})$ (overall system efficiency),

(ii) $E_o^{(k)} = \sum_{r\in O^{(k)}} \mu_{ro^*} Y_{ro}^{(k)} / (\sum_{i\in I^{(k)}} \nu_{io^*} X_{io}^{(k)} + \sum_{l=1}^{q} \nu_{lo^*}^{(k)_0} X_{lo}^{S})$, $\forall\, k$ (sub-system k efficiency),

(iii) $E_o^{(k_1,k)} = \sum_{r\in O^{(k)}} \mu_{ro^*} Y_{ro}^{(k)k_1} / (\sum_{i\in I^{(k)}} \nu_{io^*} X_{io}^{(k)k_1} + \sum_{l=1}^{q} b_{lo}^{(k)k_{1*}} X_{lo}^{S}),\ \forall\ k, k_1$ (unit k_1 of level 1 efficiency),

(iv) $E_o^{(k_2,k)} = \sum_{r\in O^{(k)}} \mu_{ro^*} Y_{ro}^{(k)k_2} / (\sum_{i\in I^{(k)}} \nu_{io^*} X_{io}^{(k)k_2} + \sum_{l=1}^{q} c_{lo}^{(k)k_{2*}} X_{lo}^{S}),\ \forall\ k, k_2$ (unit k_2 of level 2 efficiency),

(v) $E_o^{(k_3,k)} = \sum_{r\in O^{(k)}} \mu_{ro^*} Y_{ro}^{(k)k_3} / (\sum_{i\in I^{(k)}} \nu_{io^*} X_{io}^{(k)k_3} + \sum_{l=1}^{q} c_{lo}^{(k)k_{3*}} X_{lo}^{S}),\ \forall\ k, k_3$ (unit k_3 of level 2 efficiency),

(vi) $E_o^{(k_4,k)} = \sum_{r\in O^{(k)}} \mu_{ro^*} Y_{ro}^{(k)k_4} / (\sum_{i\in I^{(k)}} \nu_{io^*} X_{io}^{(k)k_4} + \sum_{l=1}^{q} d_{lo}^{(k)k_{4*}} X_{lo}^{S}),\ \forall\ k, k_4$ (unit k_4 of level 3 efficiency).

4.3.2 Efficiency decomposition

While developing a network DEA model such as the one proposed in this paper, it could be essential to consider the concept of the efficiency decomposition. According to **Kao (2017)**, efficiency decomposition is an approach to measure the system efficiency that utilises exogenous inputs to produce exogenous outputs. It measures system-division efficiencies and then identifies a mathematical relationship that associates them.

As denoted in **Kao (2015)**, when the internal divisions of a system share the available resources, then they are indispensable parts of a parallel structure. In this chapter, there is a parallel hierarchical structure within each operating sub-system and a typical parallel structure among the sub-systems of such a network DEA system. From the perspective of the entire system, its efficiency is decomposed into the weighted arithmetic average of those of the sub-systems, where the weight of the sub-system k is defined as the proportion of the aggregate input consumed by this sub-system in that consumed by all sub-systems (whole system), and $\sum_{k=1}^{p} \omega^{(k)} = 1$:

$$
\begin{aligned}
E_o^{HN} = \sum_{k=1}^{p} \omega^{(k)} E_o^{(k)} &= \sum_{k=1}^{p} \Big(\frac{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)} + \sum_{l=1}^{q} \nu_{lo}^{(k)0} X_{lo}^{S}}{\sum_{i=1}^{m} \nu_{io} X_{io} + \sum_{l=1}^{q} t_{lo} X_{lo}^{S}} \frac{\sum_{r\in O^{(k)}} \mu_{ro} Y_{ro}^{(k)}}{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)} + \sum_{l=1}^{q} \nu_{lo}^{(k)0} X_{lo}^{S}} \Big) \\
&= \frac{\sum_{k=1}^{p} \sum_{r\in O^{(k)}} \mu_{ro} Y_{ro}^{(k)}}{\sum_{i=1}^{m} \nu_{io} X_{io} + \sum_{l=1}^{q} t_{lo} X_{lo}^{S}} = \frac{\sum_{r=1}^{s} \mu_{ro} Y_{ro}}{\sum_{i=1}^{m} \nu_{io} X_{io} + \sum_{l=1}^{q} t_{lo} X_{lo}^{S}}. \qquad (4.5)
\end{aligned}
$$

From the perspective of the hierarchical structure embedded into a sub-system, the efficiency of a unit at level ξ is a weighted average of the ones of the subordinates at level $\xi + 1$, where the respective weight is formulated in a similar approach, as before. Hence, the efficiencies of sub-system k, and the internal units (1), (2), and (2.2) are decomposed as follows:

$$E_o^{(k)} = \sum_{k_1=1}^{2} \omega^{(k_1,k)} E_o^{(k_1,k)} = \sum_{k_1=1}^{2} \Big(\frac{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)k_1} + \sum_{l=1}^{q} b_{lo}^{(k)k_1} X_{lo}^S}{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)} + \sum_{l=1}^{q} \nu_{lo}^{(k)0} X_{lo}^S} \frac{\sum_{r\in O^{(k)}} \mu_{ro} Y_{ro}^{(k)k_1}}{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)k_1} + \sum_{l=1}^{q} b_{lo}^{(k)k_1} X_{lo}^S} \Big) = \frac{\sum_{r\in O^{(k)}} \mu_{ro} Y_{ro}^{(k)}}{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)} + \sum_{l=1}^{q} \nu_{lo}^{(k)0} X_{lo}^S}, \tag{4.6}$$

$$E_o^{(1,k)} = \sum_{k_2=1.1}^{1.3} \omega^{(k_2,k)} E_o^{(k_2,k)} = \sum_{k_2=1.1}^{1.3} \Big(\frac{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)k_2} + \sum_{l=1}^{q} c_{lo}^{(k)k_2} X_{lo}^S}{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)1} + \sum_{l=1}^{q} b_{lo}^{(k)1} X_{lo}^S} \frac{\sum_{r\in O^{(k)}} \mu_{ro} Y_{ro}^{(k)k_2}}{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)k_2} + \sum_{l=1}^{q} c_{lo}^{(k)k_2} X_{lo}^S} \Big) = \frac{\sum_{r\in O^{(k)}} \mu_{ro} Y_{ro}^{(k)1}}{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)1} + \sum_{l=1}^{q} b_{lo}^{(k)1} X_{lo}^S}, \tag{4.7}$$

$$E_o^{(2,k)} = \sum_{k_3=2.1}^{2.2} \omega^{(k_3,k)} E_o^{(k_3,k)} = \sum_{k_3=2.1}^{2.2} \Big(\frac{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)k_3} + \sum_{l=1}^{q} c_{lo}^{(k)k_3} X_{lo}^S}{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)2} + \sum_{l=1}^{q} b_{lo}^{(k)2} X_{lo}^S} \frac{\sum_{r\in O^{(k)}} \mu_{ro} Y_{ro}^{(k)k_3}}{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)k_3} + \sum_{l=1}^{q} c_{lo}^{(k)k_3} X_{lo}^S} \Big) = \frac{\sum_{r\in O^{(k)}} \mu_{ro} Y_{ro}^{(k)2}}{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)2} + \sum_{l=1}^{q} b_{lo}^{(k)2} X_{lo}^S}, \tag{4.8}$$

$$E_o^{(2.2,k)} = \sum_{k_4=2.2.1}^{2.2.2} \omega^{(k_4,k)} E_o^{(k_4,k)} = \sum_{k_4=2.2.1}^{2.2.2} \Big(\frac{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)k_4} + \sum_{l=1}^{q} d_{lo}^{(k)k_4} X_{lo}^S}{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)2.2} + \sum_{l=1}^{q} c_{lo}^{(k)2.2} X_{lo}^S} \frac{\sum_{r\in O^{(k)}} \mu_{ro} Y_{ro}^{(k)k_4}}{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)k_4} + \sum_{l=1}^{q} d_{lo}^{(k)k_4} X_{lo}^S} \Big) = \frac{\sum_{r\in O^{(k)}} \mu_{ro} Y_{ro}^{(k)2.2}}{\sum_{i\in I^{(k)}} \nu_{io} X_{io}^{(k)2.2} + \sum_{l=1}^{q} c_{lo}^{(k)2.2} X_{lo}^S}, \tag{4.9}$$

where $\sum_{k_1=1}^{2} \omega^{(k_1,k)} = 1$, $\sum_{k_2=1.1}^{1.3} \omega^{(k_2,k)} = 1$, $\sum_{k_3=2.1}^{2.2} \omega^{(k_3,k)} = 1$, and $\sum_{k_4=2.2.1}^{2.2.2} \omega^{(k_4,k)} = 1$.

Based on the above decompositions, the network-hierarchical system efficiency, E_o^{HN}, can in effect be decomposed as the weighted average of the ones of the terminal units, belonging to the hierarchical structure of sub-system k:

$$
\begin{aligned}
E_o^{HN} = \sum_{k=1}^{p} \omega^{(k)} E_o^{(k)} = \sum_{k=1}^{p} \omega^{(k)} (\sum_{k_1=1}^{2} \omega^{(k_1,k)} E_o^{(k_1,k)}) = \\
\sum_{k=1}^{p} \omega^{(k)} (\omega^{(1,k)} \sum_{k_2=1.1}^{1.3} \omega^{(k_2,k)} E_o^{(k_2,k)} + \omega^{(2,k)} \sum_{k_3=2.1}^{2.2} \omega^{(k_3,k)} E_o^{(k_3,k)}) = \\
\sum_{k=1}^{p} \omega^{(k)} [(\omega^{(1,k)} \sum_{k_2=1.1}^{1.3} \omega^{(k_2,k)} E_o^{(k_2,k)}) + (\omega^{(2,k)} (\omega^{(2.1,k)} E_o^{(2.1,k)} + \omega^{(2.2,k)} \sum_{k_4=2.2.1}^{2.2.2} \omega^{(k_4,k)} E_o^{(k_4,k)}))] = \\
\sum_{k=1}^{p} \sum_{k_2=1.1}^{1.3} w^{(k_2,k)} E_o^{(k_2,k)} + \sum_{k=1}^{p} w^{(2.1,k)} E_o^{(2.1,k)} + \sum_{k=1}^{p} \sum_{k_4=2.2.1}^{2.2.2} w^{(k_4,k)} E_o^{(k_4,k)},
\end{aligned}
\tag{4.10}
$$

where $w^{(k_2,k)} = \omega^{(k)} \omega^{(1,k)} \omega^{(k_2,k)}$, $w^{(2.1,k)} = \omega^{(k)} \omega^{(2,k)} \omega^{(2.1,k)}$, $w^{(k_4,k)} = \omega^{(k)} \omega^{(2,k)} \omega^{(2.2,k)} \omega^{(k_4,k)}$, and $\sum_{k=1}^{p} \sum_{k_2=1.1}^{1.3} w^{(k_2,k)} + \sum_{k=1}^{p} w^{(2.1,k)} + \sum_{k=1}^{p} \sum_{k_4=2.2.1}^{2.2.2} w^{(k_4,k)} = 1$. According to **(Cook et al., 2010)**, such an additive efficiency decomposition approach enables the measurement of the performance of the system under the assumptions of both constant returns to scale and variable returns to scale.

4.3.3 Efficiency aggregation

Since the proposed model **(4.4)** firstly measures the system and its constituent processes' efficiencies and then seeks for a mathematical relationship (the additive form) that links them, it can be classified as an additive efficiency decomposition model **(Kao, 2018)**. Another known approach for measuring the performance score of a network DEA system is the efficiency aggregation **(Kao, 2017; Zhang and Chen, 2019)**. In such an approach, the internal processes are initially aggregated (either in additive or in multiplicative form) to establish the system efficiency and subsequently to address its performance measurement.

From the perspective of the additive form towards our parallel-hierarchical system, we can easily observe that the efficiency decomposition is identical with the concept of the efficiency aggregation. Nevertheless, if the decision-maker selects to build the system efficiency by aggregating the sub-system efficiencies in a multiplicative way, then we expect that the two theoretical concepts will substantially differ from one another. In this case, we propose the multiplicative efficiency aggregation model **(4.11)** to measure the performance of the system in Figure 4.3:

$$
E_o^{HN'} = Max \prod_{k=1}^{p} E_o^{(k)} = \prod_{k=1}^{p} \left(\frac{\sum_{r \in O^{(k)}} \mu_{ro} Y_{ro}^{(k)}}{\sum_{i \in I^{(k)}} \nu_{io} X_{io}^{(k)} + \sum_{l=1}^{q} \nu_{lo}^{(k)0} X_{lo}^{S}} \right) \tag{4.11}
$$

subject to the constraints of the additive decomposition DEA model.

Model **(4.11)** differs from model **(4.4)** only in terms of its objective function. This illustrates the system efficiency as the product of those of its sub-systems. A limitation of model **(4.11)** is that it is non-linear. We could yet argue that the majority of non-linear solvers run flexibly, ensuring an optimal solution for models with non-linear objective function and linear constraints **(Kao, 2018)**.

We can additionally determine the relationship of the system efficiencies between the additive decomposition model **(4.4)** and the multiplicative aggregation model **(4.11)**, based on the inspirational ideas of **Kao (2018)** and **Zhang and Chen (2019)**:

$$E_o^{HN} = \frac{\sum_{r=1}^{s} \mu_{ro} Y_{ro}}{\sum_{i=1}^{m} \nu_{io} X_{io} + \sum_{l=1}^{q} t_{lo} X_{lo}^{S}} = \sum_{k=1}^{p} \omega^{(k)} E_o^{(k)} \geq \prod_{k=1}^{p} (E_o^{(k)})^{\omega^{(k)}} \geq \prod_{k=1}^{p} E_o^{(k)} \tag{4.12}$$

The first inequality from the left holds, since the weighted arithmetic mean is greater than or equal to the weighted geometric mean. The other inequality is in effect, given that $E_o^{(k)} \leq 1,\ \forall\ k \in \{1, 2, ..., p\}$ and $\sum_{k=1}^{p} \omega^{(k)} = 1$. Hence, the system efficiency of model **(4.4)** is always greater than or equal to that of the multiplicative model **(4.11)**, and this is also confirmed by the numerical application in Section 4.4.

4.4 An illustrative application to higher education

The performance evaluation of the higher education sector has been widely discussed in the literature **(Casu and Thanassoulis, 2006; Kao and Hung, 2008; Kao, 2012; Witte et al., 2013; Kao, 2015; Moncayo–Martínez et al., 2020; Ghasemi et al., 2020)**. **Kao (2012)**, for instance, explored the case of a chemistry and physics university department in UK that consists of two major parallel functions, the teaching and the research. It was said that each university department has a different proportion of resources at its disposal to allocate to teaching and research tasks. In such a parallel production system, the internal parallel divisions are still treated as black-boxes, without enabling the decision-maker to understand and identify the main sources of inefficiency within teaching and research. **Kao (2015)** suggested the measurement and evaluation of a university department in the form of a single-stage hierarchical structure. In their example, the university department under consideration is decomposed into three major functions: the enterprise, the research, and the teaching activities. The latter are further divided into work at the undergraduate and graduate levels. Although **Kao's (2015)** study successfully examined the performance of such a university department with a single-stage hierarchical structure, it did not pay attention to more complicated (parallel) network structures. In reality, a university department (e.g., Business School) could contain multiple parallel sub-departments each of which may consist of a number of internal functions arranged in a multi-layer hierarchical form. To illustrate the effectiveness of our proposed multi-function parallel network hierarchical DEA system, we expand the illustrative application presented in **Kao (2015)** by looking more closely at multiple parallel academic departments with distinctive functions, each of which is further viewed as a hierarchical form, see also Section 4.3.1. An embedded hierarchical structure within a multi-function parallel system has, to our knowledge, not yet been considered in the existing literature, particularly to examine the relative efficiency of the different departments and tasks of a Business School. This study illustrates the proposed models by measuring the operating performance of several Business Schools across a number of hypothetical universities.

Since our target is to better correspond to a real-life scenario, we assume that a Business School can be viewed as a more complicated network system; that is, it contains various departments (Accounting, Banking and Finance, Digital Marketing, Decision Analytics and Risk, Human Resource Management and Organisational Behaviour, Strategy Innovation and Entrepreneurship), that operate independently without affecting each other. Each of those departments performs various academic and managerial functions. For the sake of simplicity, we presume that the Business Schools to be evaluated and compared in this study, have only three departments: Accounting (A), Banking and Finance (B), and Decision Analytics and Risk (D). The internal composition of each department is no longer treated as a black-box but takes into account three main functions: teaching, research, and enterprise. Teaching is further divided into undergraduate and postgraduate teaching activities. These functions are arranged into a multi-layer hierarchical structure. Figure 4.4 illustrates the structure of this parallel network hierarchical system.

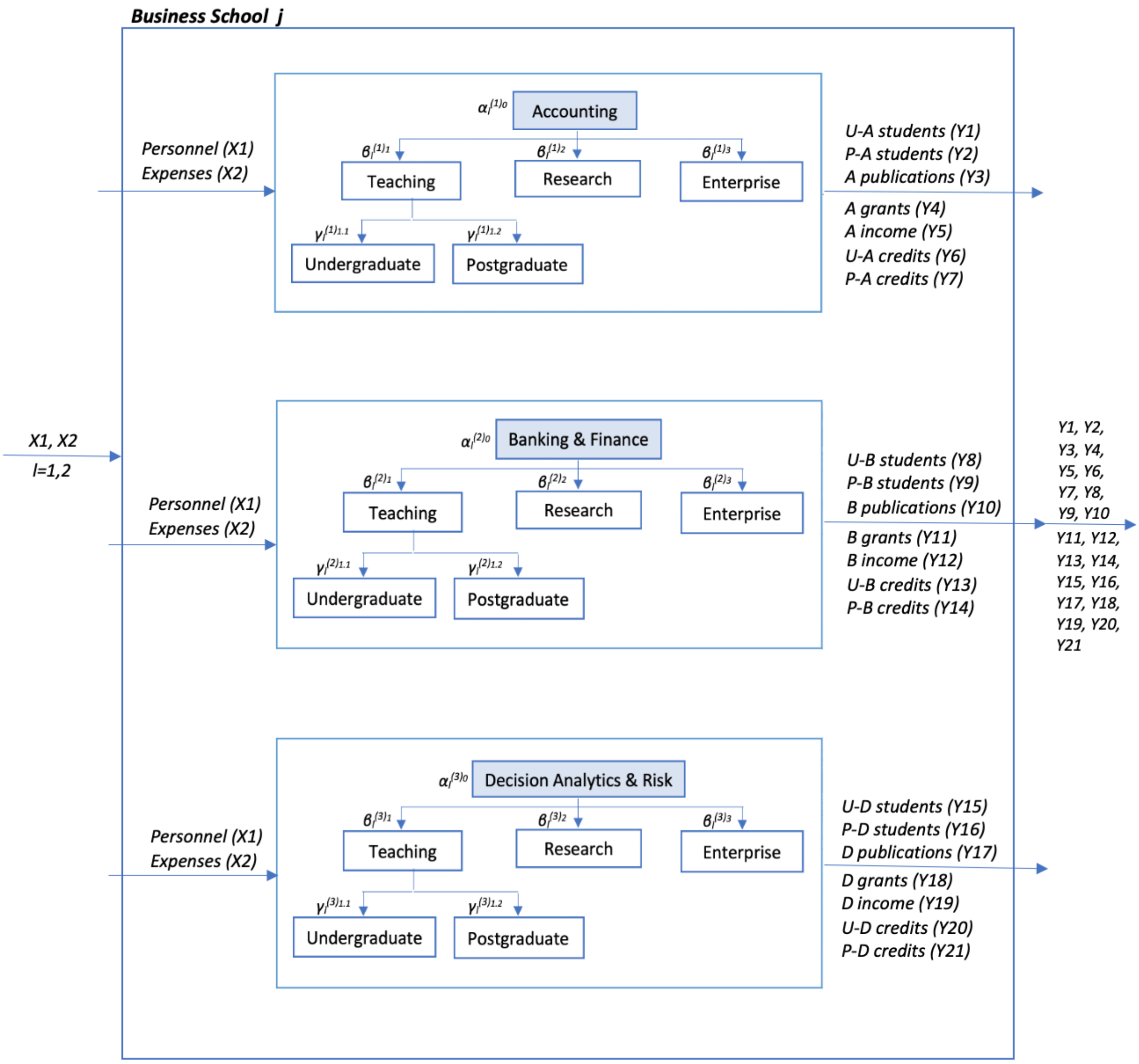

FIGURE 4.4: Embedded hierarchical network structure within a multi-function parallel system - The structure of the hypothetical Business School

In determining departments' and their internal units' efficiencies for the considered Business School, the following two inputs are used: personnel (X_1) and expenses (X_2). The former represents the number of academic and administrative staff and the latter, the general expenditure (e.g., staff salaries, capital investment) and equipment expenditures. With regard to outputs, the following are generated: the number of students (at an undergraduate and postgraduate level) graduating within a year, the credit-hours taught (at an undergraduate and postgraduate level) which are derived by the total number of students attending the unit over all units taught by the department, the total number of publications published by the academic faculty of the particular department within a year, the grants received from government funding councils, and the enterprise income obtained from contractual agreements made between the department and the local businesses with respect to service provision. As for the grants and enterprise income, there is a discussion on **Cook and Zhu (2007)** which suggests that they might be either inputs or outputs. Implicitly, this paper is measuring the Business School performance from the point of view of the University. The University identifies these kinds of income as outputs produced by the Business School. Certainly, they may plough some of that income back into the Business School in the form of salaries and capital investment. However, salaries and investment are already inputs in this example. All the aforementioned outputs are dedicated, that is they are related to different functions within a specific department. In particular, students and credits are associated with teaching, publications and grants with research, and income with enterprise, see also Table 4.2. Note that the purpose of this application is to showcase whether and how the measurement and evaluation of DMUs arranged into a parallel multi-layer multi-function hierarchical structure is attained. The data follows the example of **Kao's (2015)** in spirit, and arguably still is a simplification of most real Business Schools. The study is not intended to represent real Business Schools but instead can help to illustrate how the application of this methodology may help them to identify areas that may benefit from further attention towards improving their performance.

TABLE 4.2: Classification of the outputs to three departments

Outputs	**A**	**B**	**D**
undergraduate students	Y_1	Y_8	Y_{15}
postgraduate students	Y_2	Y_9	Y_{16}
publications	Y_3	Y_{10}	Y_{17}
grants	Y_4	Y_{11}	Y_{18}
income	Y_5	Y_{12}	Y_{19}
undergraduate credits	Y_6	Y_{13}	Y_{20}
postgraduate credits	Y_7	Y_{14}	Y_{21}

Personnel and expenses are shared among the departments and their different functions. They can be distributed using either pre-determined (fixed) proportions or variable proportions. In our scenario, the proportions are treated as variables rather than parameters, as it is difficult to specify instances such as, the amount of time a lecturer dedicates to each function or the amount of money being collected by each of the departments.

In the spirit of **Kao's (2015)** study, we assume that each Business School allocates similar amounts of resources to its three departments; that is, $\alpha_l^{(1)0} \cong \alpha_l^{(2)0} \cong \alpha_l^{(3)0}$, and $\sum_{k=1}^{3} \alpha_l^{(k)0} = 1, l = 1, 2,$

where $\alpha_l^{(k)_0}$ is the proportion of each resource l allocated to department k ($k = 1, 2, 3$). The proportions are expressed in ranges in the form of $0.5 \leq \alpha_l^{(1)_0} / \alpha_l^{(2)_0} \leq 2$, $0.5 \leq \alpha_l^{(1)_0} / \alpha_l^{(3)_0} \leq 2$, and $0.5 \leq \alpha_l^{(2)_0} / \alpha_l^{(3)_0} \leq 2$. Furthermore, every department allocates approximately 40%, 40%, and 20% of each input to the three major functions; that is, $\beta_l^{(k)_1} \cong \beta_l^{(k)_2} \cong 2\beta_l^{(k)_3}$, and $\sum_{k_1=1}^{3} \beta_l^{(k)_{k_1}} = 1$, $l = 1, 2$, and $k = 1, 2, 3$, where $\beta_l^{(k)_{k_1}}$ is the proportion of each resource l of department k allocated to each of these functions ($k_1 = 1, 2, 3$). The proportions are expressed in ranges in the form of $0.5 \leq \beta_l^{(k)_1} / \beta_l^{(k)_2} \leq 2$, $1 \leq \beta_l^{(k)_2} / \beta_l^{(k)_3} \leq 4$, and $1 \leq \beta_l^{(k)_1} / \beta_l^{(k)_3} \leq 4$. We also assume that the teaching function at each department allocates similar amounts of resources to both undergraduate and postgraduate levels; that is, $\gamma_l^{(k)_{1.1}} \cong \gamma_l^{(k)_{1.2}}$, and $\sum_{k_2=1.1}^{1.2} \gamma_l^{(k)_{k_2}} = 1$, $l = 1, 2$, and $k = 1, 2, 3$, where $\gamma_l^{(k)_{k_2}}$ is the proportion of each resource l of department k allocated to each of these levels ($k_2 = 1.1, 1.2$). The proportions are expressed in ranges in the form of $0.5 \leq \gamma_l^{(k)_{1.1}} / \gamma_l^{(k)_{1.2}} \leq 2$. The values of the shared inputs and the dedicated outputs for the evaluation of the Business Schools in twenty hypothetic universities are depicted in Table 4.3.

TABLE 4.3: Data of the Business School in twenty hypothetic universities

DMU	X_1	X_2	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	Y_{11}	Y_{12}	Y_{13}	Y_{14}	Y_{15}	Y_{16}	Y_{17}	Y_{18}	Y_{19}	Y_{20}	Y_{21}
1	20	40	35	60	55	25	25	15	70	4	2	15	70	50	5	80	30	50	25	30	50	95	65
2	55	120	40	85	60	35	35	20	15	2	1	45	110	140	15	100	65	120	80	40	65	110	20
3	25	80	25	50	50	25	10	10	15	1	2	20	70	55	5	45	25	15	20	20	30	40	20
4	70	55	30	65	40	25	20	20	25	5	1	25	90	60	5	100	50	95	100	25	30	30	25
5	80	60	45	95	50	35	25	20	65	75	3	35	110	90	10	120	65	30	30	30	75	120	55
6	75	45	30	100	55	65	20	20	50	35	10	30	85	50	5	75	30	20	25	185	70	80	50
7	75	75	15	60	35	20	5	20	10	30	25	25	120	65	10	105	55	30	40	15	60	45	5
8	20	30	20	70	50	30	10	15	15	10	10	20	65	45	5	45	20	30	15	20	60	60	15
9	30	110	55	140	140	315	20	95	90	80	30	60	90	50	5	115	60	70	100	55	95	55	80
10	85	80	60	180	170	250	20	180	160	145	45	20	65	50	5	45	20	35	25	85	115	60	130
11	12	74	10	28	55	55	34	19	90	101	79	24	5	28	25	20	118	122	20	62	53	64	103
12	125	86	37	40	45	91	20	106	92	43	35	77	72	81	50	72	115	65	113	29	72	103	121
13	113	66	88	101	105	128	42	88	109	81	118	128	14	9	100	107	89	19	123	51	104	123	9
14	74	10	126	5	83	20	22	130	25	116	98	74	57	75	57	111	45	85	119	72	83	32	19
15	121	127	40	127	48	20	123	130	106	91	67	65	130	29	109	81	85	5	19	8	46	55	45
16	23	86	125	70	81	130	83	100	14	61	122	93	112	112	75	22	5	33	36	119	47	85	24
17	32	69	83	56	105	34	56	46	104	100	91	25	25	82	99	81	58	82	34	97	28	119	98
18	118	30	81	110	56	12	15	72	85	87	17	36	122	100	126	72	110	74	44	87	34	126	24
19	72	57	130	77	27	115	30	81	27	89	106	40	23	20	12	43	49	68	76	121	100	30	64
20	17	91	105	6	92	105	47	30	85	51	42	45	26	15	24	36	5	53	44	15	112	6	10

4.4.1 Models from literature

The traditional black-box model has been initially applied to evaluate the operating performance of the Business Schools, as shown in the second column of Table 4.4. This model is simply the typical input-oriented constant returns-to-scale DEA model **(Charnes et al., 1978)** that makes use of the two exogenous shared inputs to produce the twenty-one exogenous dedicated outputs. In other words, it entirely ignores all the internal operations and mechanisms of the system (parallel sub-systems and integrated hierarchies).

According to the second column of Table 4.4, we can easily notice that there are in total 14 efficient DMUs and 6 inefficient ones. The efficient DMUs cannot be readily discriminated and this especially matters, when we consider the problem as a multi-criteria decision-making case. Besides, we are not able to obtain the efficiencies of the constituent departments of the respective Business School.

The model **(4.1)** of **Kao (2012)**, see Section 4.2.1, has also been implemented to evaluate the performance of the Business Schools, as illustrated in columns 3-9 of Table 4.4. The simplification of such a model compared to the proposed models of this paper is that the parallel sub-divisions of the system are treated as black-boxes.

According to column 3 of Table 4.4, there are now 8 DMUs with a perfect efficiency score of one, that is their respective components (departments) are absolutely efficient. There is still, however, the problem with the lack of discrimination of the efficient DMUs that cannot lead us to a unique ranking order. The efficiency scores of the three departments (A, B, D) are given in columns 4, 6, and 8, respectively. The numbers next to each of the efficiency scores (in columns 5, 7, and 9) are the respective weights in the efficiency decomposition. Using the Business School 15 as an example, its efficiency scores for A (0.9997), B (0.5541), and D (0.3353) multiplied by their respective weights of 0.443, 0.309, and 0.248, will obtain the efficiency score of the whole system, 0.6973. Using the information of the efficiency score, we can conclude that DMU's 15 relatively low performance is owed to its weak Decision Analytics and Risk department, whose operations should be improved. Nevertheless, it is not clear to us which constituent functions of that department have the millstone of a heavy burden round their necks.

TABLE 4.4: Efficiency scores of black-box model and **Kao (2012)**

		Kao's (2012) Model						
DMU	Black-box [Rank]	E_o [Rank]	$E_o^{(A)}$	$\omega^{(A)}$	$E_o^{(B)}$	$\omega^{(B)}$	$E_o^{(D)}$	$\omega^{(D)}$
1	1 [1]	1 [1]	1	0.429	1	0.215	1	0.356
2	0.9556 [15]	0.7247 [15]	0.4532	0.237	0.8412	0.433	0.7665	0.330
3	0.7124 [20]	0.5499 [19]	0.4551	0.250	0.6689	0.500	0.4065	0.250
4	0.8845 [17]	0.6596 [17]	0.5015	0.320	0.6915	0.320	0.7715	0.360
5	0.8620 [18]	0.7599 [14]	0.7383	0.337	0.7654	0.283	0.7750	0.380
6	1 [1]	0.9077 [11]	0.9247	0.356	0.6951	0.215	1	0.430
7	0.7200 [19]	0.5339 [20]	0.3111	0.250	0.6903	0.500	0.4441	0.250
8	1 [1]	1 [1]	1	0.252	1	0.499	1	0.249
9	1 [1]	1 [1]	1	0.240	1	0.480	1	0.279
10	1 [1]	0.8961 [12]	1	0.400	0.7496	0.297	0.9026	0.303
11	1 [1]	1 [1]	1	0.496	1	0.248	1	0.255
12	0.8869 [16]	0.6282 [18]	0.5249	0.249	0.3353	0.296	0.8752	0.455
13	1 [1]	0.8415 [13]	0.8911	0.373	0.8199	0.355	0.8016	0.272
14	1 [1]	1 [1]	1	0.403	1	0.202	1	0.395
15	1 [1]	0.6973 [16]	0.9997	0.443	0.5541	0.309	0.3353	0.248
16	1 [1]	1 [1]	1	0.454	1	0.250	1	0.296
17	1 [1]	1 [1]	1	0.250	1	0.250	1	0.500
18	1 [1]	1 [1]	1	0.283	1	0.467	1	0.250
19	1 [1]	0.9197 [10]	0.9973	0.354	0.7184	0.225	0.9623	0.421
20	1 [1]	0.9823 [9]	1	0.400	0.9116	0.200	1	0.400

4.4.2 Results from the parallel-hierarchical network model

To obtain the information regarding the performance of the functions of each department within the considered Business School, we implement the additive decomposition model **(4.4)** and the multiplicative aggregation model **(4.11)**. 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 for model **(4.4)**. As for the non-linear model **(4.11)**, we have implemented the GEKKO which is a Python package for machine learning and optimisation. It is combined with large-scale solvers for non-linear programming models as well. To define the type of the problem, we have used a non-dynamic mode that sets all differential terms to zero to calculate the steady-state conditions. The experiments ran on a computer with 16GB RAM.

These models do not only allow us to discriminate the efficient DMUs, but also to simultaneously calculate the efficiencies of the Business School of interest, its constituent departments, and the functions within the respective department. The results obtained by models **(4.4)** and **(4.11)** are respectively illustrated in Tables 4.5 and 4.6. The second column in each table shows the efficiency of the respective overall system along with its rank. The remaining columns provide the efficiency scores with their respective weights of each sub-system and sub-unit within the sub-system.

The ranks of the efficiency scores of the overall system obtained via our proposed model **(4.4)** are also compared with the overall systems' ranks of black-box and **Kao's (2012)** models. It can be statistically inferred that the ranks are quite similar, and this is verified by the Spearman rank-order correlation test with values 0.678 and 0.924, respectively. These are significant at the 0.01 level (two-tailed). The same situation holds even for our proposed model **(4.11)**. The results of the correlation analysis further validate the underpinning of our model in some way.

Using the Business School 18 as an example, its efficiency score for the Accounting (0.4553) is decomposed into the efficiencies of the teaching (0.7690), research (0.3263), and enterprise (0.2707) multiplied by their respective weights of 0.333, 0.333, and 0.333. Note that teaching efficiency is further decomposed into the efficiencies of the undergraduate (0.3069) and postgraduate (1) levels multiplied by their respective weights, 0.333 and 0.667. By the same token, the efficiency scores of the other two departments, Banking and Finance (0.8045) and Decision Analytics and Risk (0.5141), are identified. Hence, the efficiency scores of the three departments multiplied by their respective weights provide the efficiency of DMU 18, which is 0.6446.

The unsatisfactory performance of Business School 18 is mainly due to the Accounting and secondly to the Decision Analytics and Risk department. If this Business School desires to significantly improve its efficiency, then it should strengthen its contribution to society (enterprise) along with its research activities, as far as the A department is concerned. With regard to D department, particular emphasis should be placed on the enterprise. In summary, this Business School should genuinely pursue constant and long-term synergies with representatives from the public and private sector towards more impactful and effective research and educational actions.

As discussed in Section 4.3.3, the system efficiency of model **(4.4)** should be greater than or equal to the respective one in model **(4.11)**. By comparing the second columns of Table 4.5 and Table 4.6, we validate our initial assumption. This is further bolstered by the fact that the Business School's efficiency is the product of the ones of the three internal departments. For example, DMU's 18 efficiency (0.1946) is obtained by multiplying A's (0.4622), B's (0.8131), and D's efficiency (0.5178). With regard to DMU 18, the promising performance of its B department still has considerable potential for further improvements, through the upgrade of the teaching methods and the training of the teaching staff, to better support postgraduate taught modules. The root cause of the problem, however, is located to the A department that should better adhere to the following guidelines: *(i)* strengthen its contributions to society, and *(ii)* provide greater (financial) incentives to the academic faculty to ensure grants via more powerful research proposals.

TABLE 4.5: Efficiency scores of additive decomposition model (4.4)

DMU	E_o^{HN} [Rank]	$E_o^{(A-U-T)}$	$\omega^{(A-U-T)}$	$E_o^{(A-P-T)}$	$\omega^{(A-P-T)}$	$E_o^{(A-T)}$	$\omega^{(A-T)}$	$E_o^{(A-R)}$	$\omega^{(A-R)}$	$E_o^{(A-E)}$	$\omega^{(A-E)}$	$E_o^{(A)}$	$\omega^{(A)}$
1	0.7713 [4]	0.4768	0.391	1	0.609	0.7952	0.333	0.8827	0.333	0.5472	0.333	0.7417	0.236
2	0.4399 [15]	0.1475	0.333	0.3475	0.667	0.2808	0.333	0.3285	0.333	0.2616	0.333	0.2903	0.231
3	0.3723 [17]	0.1956	0.333	0.4296	0.667	0.3516	0.333	0.4488	0.333	0.1200	0.333	0.3068	0.250
4	0.4164 [16]	0.1452	0.333	0.4085	0.667	0.3207	0.333	0.3056	0.333	0.2255	0.333	0.2839	0.244
5	0.4937 [14]	0.1906	0.333	0.5908	0.667	0.4574	0.333	0.3427	0.333	0.2523	0.333	0.3508	0.244
6	0.5495 [11]	0.1380	0.333	0.7320	0.667	0.5340	0.333	0.4548	0.333	0.2714	0.333	0.4201	0.250
7	0.3459 [19]	0.0778	0.333	0.2911	0.667	0.2200	0.333	0.2197	0.333	0.0464	0.333	0.1620	0.250
8	0.7274 [7]	0.2637	0.333	1	0.667	0.7546	0.333	0.9642	0.333	0.2630	0.333	0.6606	0.250
9	0.6593 [8]	0.7211	0.333	1	0.667	0.9070	0.333	1	0.333	0.1861	0.333	0.6977	0.458
10	0.5628 [10]	0.8729	0.510	1	0.490	0.9352	0.333	1	0.333	0.1565	0.333	0.6973	0.469
11	0.8091 [3]	0.3555	0.336	1	0.664	0.7833	0.333	0.8440	0.333	0.7797	0.333	0.8023	0.250
12	0.3574 [18]	0.3129	0.480	0.4773	0.520	0.3984	0.333	0.2504	0.333	0.1544	0.333	0.2677	0.216
13	0.5081 [13]	0.2134	0.333	0.7266	0.667	0.5555	0.333	0.6052	0.333	0.4880	0.333	0.5496	0.369
14	0.9968 [1]	1	0.667	0.8563	0.333	0.9521	0.333	1	0.333	1	0.333	0.9840	0.200
15	0.3365 [20]	0.2684	0.392	0.4506	0.608	0.3792	0.333	0.1374	0.333	0.8063	0.333	0.4410	0.472
16	0.8730 [2]	1	0.667	0.6394	0.333	0.8798	0.333	0.7451	0.333	1	0.333	0.8750	0.322
17	0.7462 [5]	0.6793	0.377	0.8775	0.623	0.8029	0.333	1	0.333	0.7208	0.333	0.8412	0.476
18	0.6446 [9]	0.3069	0.333	1	0.667	0.7690	0.333	0.3263	0.333	0.2707	0.333	0.4553	0.250
19	0.5130 [12]	0.5508	0.412	0.5241	0.588	0.5351	0.333	0.5363	0.333	0.4162	0.333	0.4959	0.322
20	0.7315 [6]	1	0.667	0.6652	0.333	0.8884	0.333	1	0.333	0.7611	0.333	0.8832	0.500

Table 4.5 Continued:

DMU	$E_o^{(B-U-T)}$	$\omega^{(B-U-T)}$	$E_o^{(B-P-T)}$	$\omega^{(B-P-T)}$	$E_o^{(B-T)}$	$\omega^{(B-T)}$	$E_o^{(B-R)}$	$\omega^{(B-R)}$	$E_o^{(B-E)}$	$\omega^{(B-E)}$	$E_o^{(B)}$	$\omega^{(B)}$
1	0.0848	0.333	1	0.667	0.6949	0.333	0.8968	0.333	0.8115	0.333	0.8010	0.471
2	0.0849	0.333	0.4287	0.667	0.3141	0.333	0.5092	0.333	0.7754	0.333	0.5329	0.463
3	0.0535	0.333	0.3711	0.667	0.2652	0.333	0.6204	0.333	0.4892	0.333	0.4583	0.500
4	0.0429	0.333	0.5566	0.667	0.3853	0.333	0.5723	0.333	0.5014	0.333	0.4863	0.488
5	0.3577	0.333	0.5969	0.667	0.5172	0.333	0.6301	0.333	0.6731	0.333	0.6068	0.489
6	0.1755	0.333	0.3927	0.667	0.3203	0.333	0.6284	0.333	0.3974	0.333	0.4487	0.250
7	0.1325	0.333	0.4818	0.667	0.3654	0.333	0.6274	0.333	0.4470	0.333	0.4799	0.500
8	0.1453	0.333	0.6761	0.667	0.4992	0.333	1	0.333	0.8771	0.333	0.7921	0.500
9	0.3903	0.333	0.9030	0.667	0.7321	0.333	0.6186	0.333	0.3447	0.333	0.5651	0.229
10	0.6364	0.667	0.2361	0.333	0.5030	0.333	0.3011	0.333	0.2900	0.333	0.3647	0.235
11	1	0.336	1	0.664	1	0.333	0.4883	0.333	0.4758	0.333	0.6547	0.250
12	0.2623	0.607	0.2578	0.393	0.2606	0.333	0.4168	0.333	0.4037	0.333	0.3604	0.352
13	0.6235	0.518	0.5960	0.482	0.6102	0.333	0.8125	0.333	0.0481	0.333	0.4903	0.355
14	1	0.667	1	0.333	1	0.333	1	0.333	1	0.333	1	0.400
15	0.4692	0.666	0.2160	0.334	0.3847	0.333	0.4655	0.333	0.1151	0.333	0.3218	0.291
16	0.9943	0.590	1	0.410	0.9967	0.333	1	0.333	1	0.333	0.9989	0.452
17	1	0.580	1	0.420	1	0.333	0.2865	0.333	0.7822	0.333	0.6896	0.238
18	1	0.667	0.3055	0.333	0.7685	0.333	1	0.333	0.6451	0.333	0.8045	0.500
19	0.4950	0.333	0.7000	0.667	0.6316	0.333	0.3556	0.333	0.1488	0.333	0.3787	0.260
20	0.5313	0.337	0.7274	0.663	0.6613	0.333	0.6470	0.333	0.1800	0.333	0.4961	0.250

Table 4.5 Continued:

DMU	$E_o^{(D-U-T)}$	$\omega^{(D-U-T)}$	$E_o^{(D-P-T)}$	$\omega^{(D-P-T)}$	$E_o^{(D-T)}$	$\omega^{(D-T)}$	$E_o^{(D-R)}$	$\omega^{(D-R)}$	$E_o^{(D-E)}$	$\omega^{(D-E)}$	$E_o^{(D)}$	$\omega^{(D)}$
1	1	0.488	0.9247	0.512	0.9614	0.333	0.5586	0.333	0.7214	0.333	0.7471	0.293
2	0.4453	0.496	0.4642	0.504	0.4548	0.333	0.4694	0.333	0.3128	0.333	0.4123	0.306
3	0.3075	0.667	0.1326	0.333	0.2492	0.333	0.3056	0.333	0.2424	0.333	0.2657	0.250
4	0.3510	0.333	0.6143	0.667	0.5265	0.333	0.4900	0.333	0.2125	0.333	0.4097	0.268
5	0.7034	0.667	0.3717	0.333	0.5929	0.333	0.1837	0.333	0.4764	0.333	0.4177	0.267
6	0.5080	0.645	0.4611	0.355	0.4913	0.333	1	0.333	0.5028	0.333	0.6647	0.500
7	0.3368	0.667	0.1442	0.333	0.2726	0.333	0.1723	0.333	0.3402	0.333	0.2617	0.250
8	0.7681	0.667	0.4200	0.333	0.6521	0.333	0.3425	0.333	1	0.333	0.6649	0.250
9	0.3765	0.488	0.4541	0.512	0.4162	0.333	1	0.333	0.5994	0.333	0.6719	0.313
10	0.2493	0.333	0.7345	0.667	0.5727	0.333	0.3407	0.333	0.6064	0.333	0.5066	0.296
11	1	0.667	1	0.333	1	0.333	1	0.333	0.6688	0.333	0.8896	0.500
12	0.5873	0.353	0.6226	0.647	0.6101	0.333	0.3065	0.333	0.2825	0.333	0.3997	0.432
13	0.6513	0.667	0.0873	0.333	0.4633	0.333	0.4339	0.333	0.5291	0.333	0.4754	0.276
14	1	0.656	1	0.344	1	0.333	1	0.333	1	0.333	1	0.400
15	0.3270	0.667	0.1785	0.333	0.2775	0.333	0.0354	0.333	0.1243	0.333	0.1457	0.236
16	0.6904	0.667	0.1607	0.333	0.5138	0.333	1	0.333	0.3410	0.333	0.6183	0.226
17	0.8096	0.483	0.8220	0.517	0.8160	0.333	0.8475	0.333	0.2415	0.333	0.6350	0.286
18	1	0.667	0.4764	0.333	0.8255	0.333	0.5185	0.333	0.1982	0.333	0.5141	0.250
19	0.3354	0.333	0.4810	0.667	0.4325	0.333	0.7108	0.333	0.6858	0.333	0.6097	0.418
20	0.0643	0.337	0.3011	0.663	0.2213	0.333	0.7700	0.333	1	0.333	0.6638	0.250

TABLE 4.6: Efficiency scores of multiplicative aggregation model (4.11)

DMU	$E_o^{HN'}$ [Rank]	$E_o^{(A-U-T)}$	$\omega^{(A-U-T)}$	$E_o^{(A-P-T)}$	$\omega^{(A-P-T)}$	$E_o^{(A-T)}$	$\omega^{(A-T)}$	$E_o^{(A-R)}$	$\omega^{(A-R)}$	$E_o^{(A-E)}$	$\omega^{(A-E)}$	$E_o^{(A)}$	$\omega^{(A)}$
1	0.4452 [4]	0.4816	0.383	1	0.617	0.8016	0.333	0.8830	0.333	0.5435	0.333	0.7427	0.434
2	0.0672 [15]	0.1845	0.333	0.3580	0.667	0.3002	0.333	0.3301	0.333	0.2526	0.333	0.2943	0.409
3	0.0383 [18]	0.1961	0.333	0.4530	0.667	0.3673	0.333	0.4530	0.333	0.1181	0.333	0.3128	0.448
4	0.0589 [16]	0.0962	0.333	0.4599	0.667	0.3387	0.333	0.2363	0.333	0.3013	0.333	0.2921	0.348
5	0.0928 [14]	0.1307	0.333	0.6429	0.667	0.4721	0.333	0.2767	0.333	0.3411	0.333	0.3633	0.321
6	0.1443 [10]	0.0989	0.333	0.8113	0.667	0.5738	0.333	0.4644	0.333	0.3628	0.333	0.4670	0.374
7	0.0209 [20]	0.0825	0.333	0.2889	0.667	0.2201	0.333	0.2261	0.333	0.0452	0.333	0.1638	0.474
8	0.3538 [6]	0.2609	0.333	1	0.667	0.7536	0.333	0.9655	0.333	0.2649	0.333	0.6613	0.402
9	0.2713 [8]	0.7361	0.333	1	0.667	0.9120	0.333	1	0.333	0.1867	0.333	0.6996	0.313
10	0.1324 [11]	0.8872	0.510	1	0.490	0.9425	0.333	1	0.333	0.1565	0.333	0.6997	0.278
11	0.4724 [3]	0.3642	0.333	1	0.667	0.7881	0.333	0.8469	0.333	0.7851	0.333	0.8067	0.461
12	0.0435 [17]	0.2029	0.357	0.5062	0.643	0.3980	0.333	0.3047	0.333	0.1962	0.333	0.2996	0.362
13	0.1292 [12]	0.2148	0.333	0.7287	0.667	0.5574	0.333	0.6051	0.333	0.4870	0.333	0.5498	0.321
14	0.9869 [1]	1	0.667	0.8824	0.333	0.9608	0.333	1	0.333	1	0.333	0.9869	0.483
15	0.0222 [19]	0.1868	0.333	0.4506	0.667	0.3627	0.333	0.1186	0.333	0.8571	0.333	0.4461	0.365
16	0.5456 [2]	1	0.667	0.6477	0.333	0.8826	0.333	0.7495	0.333	1	0.333	0.8774	0.293
17	0.3705 [5]	0.6833	0.393	0.8774	0.607	0.8012	0.333	1	0.333	0.7316	0.333	0.8443	0.360
18	0.1946 [9]	0.2947	0.333	1	0.667	0.7649	0.333	0.3093	0.333	0.3125	0.333	0.4622	0.347
19	0.1178 [13]	0.4716	0.374	0.5242	0.626	0.5046	0.333	0.5671	0.333	0.4384	0.333	0.5033	0.378
20	0.2952 [7]	1	0.667	0.6666	0.333	0.8888	0.333	1	0.333	0.7660	0.333	0.8849	0.272

Table 4.6 Continued:

DMU	$E_o^{(B-U-T)}$	$\omega^{(B-U-T)}$	$E_o^{(B-P-T)}$	$\omega^{(B-P-T)}$	$E_o^{(B-T)}$	$\omega^{(B-T)}$	$E_o^{(B-R)}$	$\omega^{(B-R)}$	$E_o^{(B-E)}$	$\omega^{(B-E)}$	$E_o^{(B)}$	$\omega^{(B)}$
1	0.0850	0.333	1	0.667	0.6950	0.333	0.8965	0.333	0.8126	0.333	0.8014	0.236
2	0.0874	0.333	0.4284	0.667	0.3148	0.333	0.5080	0.333	0.7799	0.333	0.5342	0.288
3	0.0538	0.333	0.3669	0.667	0.2625	0.333	0.6251	0.333	0.4911	0.333	0.4596	0.283
4	0.0429	0.333	0.5646	0.667	0.3907	0.333	0.5725	0.333	0.5014	0.333	0.4882	0.273
5	0.3627	0.333	0.6064	0.667	0.5251	0.333	0.6301	0.333	0.6731	0.333	0.6095	0.324
6	0.1997	0.333	0.4472	0.667	0.3647	0.333	0.5865	0.333	0.4412	0.333	0.4641	0.329
7	0.1334	0.333	0.4879	0.667	0.3697	0.333	0.6281	0.333	0.4470	0.333	0.4816	0.250
8	0.1447	0.333	0.6757	0.667	0.4987	0.333	1	0.333	0.8786	0.333	0.7924	0.310
9	0.3654	0.333	0.9680	0.667	0.7672	0.333	0.6188	0.333	0.3438	0.333	0.5766	0.374
10	0.6871	0.667	0.2489	0.333	0.5410	0.333	0.2954	0.333	0.2754	0.333	0.3706	0.309
11	1	0.617	1	0.383	1	0.333	0.4946	0.333	0.4792	0.333	0.6579	0.309
12	0.2623	0.607	0.2579	0.393	0.2606	0.333	0.4168	0.333	0.4037	0.333	0.3604	0.305
13	0.6235	0.515	0.5965	0.485	0.6104	0.333	0.8214	0.333	0.0480	0.333	0.4932	0.310
14	1	0.493	1	0.507	1	0.333	1	0.333	1	0.333	1	0.267
15	0.4704	0.644	0.2412	0.356	0.3887	0.333	0.4580	0.333	0.1203	0.333	0.3223	0.258
16	0.9966	0.591	1	0.409	0.9980	0.333	1	0.333	1	0.333	0.9993	0.277
17	1	0.580	1	0.420	1	0.333	0.2872	0.333	0.7822	0.333	0.6898	0.381
18	1	0.667	0.3234	0.333	0.7745	0.333	1	0.333	0.6648	0.333	0.8131	0.359
19	0.5073	0.333	0.7152	0.667	0.6459	0.333	0.3574	0.333	0.1484	0.333	0.3839	0.330
20	0.5324	0.333	0.7287	0.667	0.6633	0.333	0.6545	0.333	0.1812	0.333	0.4996	0.311

Table 4.6 Continued:

DMU	$E_o^{(D-U-T)}$	$\omega^{(D-U-T)}$	$E_o^{(D-P-T)}$	$\omega^{(D-P-T)}$	$E_o^{(D-T)}$	$\omega^{(D-T)}$	$E_o^{(D-R)}$	$\omega^{(D-R)}$	$E_o^{(D-E)}$	$\omega^{(D-E)}$	$E_o^{(D)}$	$\omega^{(D)}$
1	1	0.488	0.9301	0.512	0.9642	0.333	0.5586	0.333	0.7214	0.333	0.7481	0.330
2	0.4370	0.530	0.4420	0.470	0.4394	0.333	0.5577	0.333	0.2859	0.333	0.4277	0.303
3	0.3075	0.667	0.1357	0.333	0.2502	0.333	0.3056	0.333	0.2424	0.333	0.2661	0.269
4	0.3510	0.333	0.6236	0.667	0.5327	0.333	0.4940	0.333	0.2125	0.333	0.4131	0.379
5	0.7034	0.667	0.3835	0.333	0.5968	0.333	0.1837	0.333	0.4764	0.333	0.4190	0.355
6	0.4934	0.667	0.4155	0.333	0.4675	0.333	1	0.333	0.5299	0.333	0.6658	0.298
7	0.3540	0.667	0.1604	0.333	0.2894	0.333	0.1609	0.333	0.3460	0.333	0.2655	0.276
8	0.7370	0.667	0.3613	0.333	0.6118	0.333	0.4138	0.333	1	0.333	0.6752	0.289
9	0.3765	0.488	0.4588	0.512	0.4186	0.333	1	0.333	0.5994	0.333	0.6727	0.313
10	0.2493	0.333	0.7526	0.667	0.5848	0.333	0.3407	0.333	0.6064	0.333	0.5106	0.413
11	1	0.572	1	0.428	1	0.333	1	0.333	0.6704	0.333	0.8901	0.230
12	0.4981	0.402	0.6163	0.598	0.5688	0.333	0.3349	0.333	0.3060	0.333	0.4032	0.332
13	0.6513	0.667	0.0896	0.333	0.4640	0.333	0.4365	0.333	0.5291	0.333	0.4765	0.369
14	1	0.403	1	0.597	1	0.333	1	0.333	1	0.333	1	0.250
15	0.3005	0.667	0.1715	0.333	0.2575	0.333	0.0460	0.333	0.1597	0.333	0.1544	0.377
16	0.7348	0.667	0.1521	0.333	0.5406	0.333	1	0.333	0.3262	0.333	0.6222	0.431
17	0.8096	0.483	0.8290	0.517	0.8196	0.333	0.8475	0.333	0.2415	0.333	0.6362	0.258
18	1	0.667	0.4758	0.333	0.8253	0.333	0.5439	0.333	0.1844	0.333	0.5178	0.294
19	0.3354	0.333	0.4818	0.667	0.4330	0.333	0.7108	0.333	0.6858	0.333	0.6099	0.291
20	0.0661	0.333	0.3066	0.667	0.2264	0.333	0.7764	0.333	1	0.333	0.6676	0.417

4.5 Conclusions & Future Research

In our chapter, we have proposed a new multi-function parallel (network) hierarchical structure to more accurately reflect the complex internal mechanisms and procedures of large organisations. These typically consist of multiple departments that could, in turn, be extended into a number of distinctive operational functions, arranged either in series or in parallel or in a hierarchical structure. These components consume and generate resources that can be interactive and/or independent.

The conventional black-box model evaluates a company (system), while ignoring its internal operations. **Kao's (2012)** model evaluates the constituent departments (sub-systems) of a company, which are independent amongst them. However, it still handles the internal structure of each department as a black-box case. **Kao's (2015)** model has successfully considered the internal processes of a single-stage system as a multi-layer hierarchical structure, yet it ignores that each department may have its own complex structure. The above models did not recognise the necessity of assessing a company, in which the network scheme might intertwine with a hierarchical structure. **Gan et al. (2019)** are one of the first to adopt such a notion, enabling the sub-systems to be interdependent. Our chapter presents an alternative to **Gan et al. (2019)**, by proposing an embedded hierarchical network structure within a multi-function parallel system. In our proposed scheme, the constituent sub-systems act independently from one another accommodating another class of problems.

In this chapter, DMUs have a network-hierarchical structure. On a macro level, the external layer is associated with the action of retrieving data from the entire system, whereas the internal layer from each of the sub-systems connected in parallel. On a micro level, that is the interior part of a sub-system, we evaluate the constituent units that form a multi-level multi-function hierarchical structure.

To evaluate the performance of the DMUs with such a structure, we propose an additive decomposition model **(4.4)** and a multiplicative aggregation model **(4.11)**. In both models, we obtain the system, the sub-systems, and their internal units' efficiencies as well as identify their relationship. In particular, the efficiency of a unit at a higher level is the weighted average of those of the subordinates at the immediate lower level; the weight of that unit is the proportion of the input consumed by that subordinate in that consumed by all subordinates. For the additive model **(4.4)**, the overall efficiency is decomposed into the weighted arithmetic average of those of the parallel sub-systems. It can also be expressed as the weighted average of the efficiencies of the terminal units that belong to the hierarchical structure of each sub-system. For the multiplicative model **(4.11)**, the system efficiency is defined as the product of the efficiencies of the constituent sub-systems. We have also proven that the system efficiency of model **(4.4)** is always greater than or equal to the respective one in model **(4.11)**.

The performance measurement and evaluation of several Business Schools across a number of universities illustrates the proposed models. These models allow us to not only discriminate the efficient units, but also to simultaneously calculate the efficiencies of the Business School of interest, its departments, and the functions within the respective department. Hence, decision-makers will be enabled to take certain actions by improving the areas of weakness.

Other areas of application of the proposed structure may include performance evaluation of business functions such as human resources, accounting and finance, marketing, and supply chain. The supply chain management of an organisation, for instance, ensures that goods and services get to customers in the easiest way possible. Such a department could be decomposed into several independent operations such as production, procurement, logistics, and customer service. These operations could be hierarchically divided into people's responsibilities, tasks, and values. The main target is to meaningfully compare the efficiency of several parallel (network) hierarchical supply chains of different factory branches inside and outside the country. Identifying and improving the areas of weakness of the most ineffective supply chains, could reduce operating costs, increase the quality of products, and meet customers' needs.

Another promising area could be the evaluation of the operating performance of a commercial ship. Stakeholders from the shipping industry might be interested in, for instance, exploring the most desirable ship design associated with a valid scenario, in which the maintenance policy is an integral part. The corresponding maintenance policy could be operationalised through the various input and output factors. A ship cannot operate without the effective management of its constituent sub-systems (electrical, diesel propulsion, lube oil, heavy fuel oil, deck) incorporated within its hull and deck. Because of the complex layout of the ship, its overall management system can integrate both the multi-function parallel network and the multi-layer hierarchical structures.

The discussion of both models is under the constant returns to scale assumption. This can be expanded to variable returns to scale situation for the additive model. Another challenge for future research could be the evaluation of a system that requires the integration of a hierarchical structure into other more complex network processes, such as assembly and disassembly, mixed, and dynamic systems **(Cook et al., 2010; Kao, 2016; Kao, 2017)**. It would also be interesting to develop appropriate DEA modelling techniques, which will acknowledge that not all competing DMUs have exactly the same internal structure.

It is also worthwhile to point out that the dataset used in Section 4.4. was based in part on Kao (2015), and has been extended by taking random samples for each of the additional output factors that include integer values in the range of [1 to 330]. The goal of this dataset was to indicate how the theoretical network hierarchical DEA structure is applied to an illustrative example in the higher education sector. Other methods used in the literature aim to develop multiple input-output production frontiers and bring more structure and accuracy in the generation of instances, such as the piecewise Cobb-Douglas and the cubic polynomial production functions **(Banker et al., 1993; Giraleas et al., 2012; Khezrimotlagh, 2022)**.

Among the models proposed in this study to measure the performance of the multi-function parallel network hierarchical system, the multiplicative efficiency aggregation model is the only non-linear network DEA formulation due to its non-linear objective function. Although the majority of non-linear solvers can run flexibly **(Kao, 2018)**, the model is still considered computationally complex and a global optimal solution cannot be easily guaranteed. Alternative algorithms can be used by transforming the model into either a second order cone programming or a semi-definite programming problem, following the spirit of **Chen and Zhu (2017)** and **Kuo et al. (2020)** or **Zhang and Chen (2019)**, respectively. The aforementioned techniques lie in the field of convex optimization, see also **Boyd et al. (2004)** and **Zhu (2020)**.

Finally, current research studies the evaluation of the performance of DMUs with a multi-function parallel network hierarchical structure, only when the data are positive real numbers, and the DEA models are based on this condition. Future research could relax this assumption by allowing the data points (inputs, intermediate measures, and outputs) to be imprecise 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 **(Zimmermann, 2011)** could be used as an alternative option.

Chapter 5

Conclusions

5.1 Overview

This final chapter summarises the main results of the dissertation and outlines the overall research contributions for each of the three main chapters. These are provided in Section 5.2 and Section 5.3, respectively. Additionally, Section 5.4 presents the research limitations of the study. Finally, Section 5.5 identifies potential directions of future research.

5.2 Summary of Findings

The work described in this thesis sheds new light on the use of several alternative modelling approaches and methodological frameworks to attain fairness in the evaluation of DMUs structured as a DEA network. Multiple insightful findings have been produced and they are summarised below.

In Chapter 2, we have initially identified that as the number of zero weights decreases, more information relevant to the known factors are taken into account. This makes the weight distribution significantly less uneven, reflecting the alternating requirements and needs of the stakeholders involved. A more realistic weight distribution has been obtained via the proposed minimax cross-efficiency model for a generalised two-stage DEA structure than our basic additive self-efficiency aggregation model and the aggressive-based model of **Kao and Liu (2019)**.

Moreover, we have shown that the CRITIC multi-criteria decision-making method is compatible with the minimax model, since it rewards contrast. While it is more likely that the worst-performing DMU attempts to assess itself in its best possible light, the efficiency scores of the other better-performing DMUs might decrease. Since this situation increases the contrast, our proposed minimax secondary goal model seems to be an acceptable option to coexist with the CRITIC method. CRITIC and the minimax model have been found to obtain a greater discrimination power than the additive self-efficiency aggregation model.

We have also applied the CRITIC method to determine an appropriate weight set for aggregating the individual cross-efficiencies into a final cross-efficiency score for each DMU and flow. By identifying the standard deviation (one of the technique's main attributes), we indicate the contrast in the viewpoints of an individual evaluator DMU_k. The conflict, which is the other main attribute of the CRITIC method, manages to give voice to the less mainstream viewpoints of the different evaluators regarding the evaluated DMU. An evaluator will be assigned a greater final weight if it provides more valuable information. This information should reward contrast, diversity, and inclusion. CRITIC cross-efficiencies have been compared with the respective final cross-efficiencies obtained via the traditional average method. Although the difference between methods is negligible, we note that the average method, which promotes the majority vote, excludes the minority without fully respecting the degree of diversification of the different opinions. CRITIC method fills this gap, assigning more weight to "mavericks" and promoting the modern concept of fairness.

Finally, we have recognised that in our minimax model both sub-stages of the generalised two-stage structure have the same bargaining power and improve the overall efficiency. This is conducive to the development of a cooperative situation. This stands in contrast with the aggressive method of **Kao and Liu (2019)**. They selected a non-cooperative approach, in which DMUs act egoistically with a view to maximising their self-evaluation and downplaying the peer-evaluation. Following this comparison, we have shown that our proposed minimax (cooperative) model accomplishes a higher absolute cross-efficiency score for each DMU and stage connected with some performance reward than the respective scores of **Kao and Liu (2019)**.

In Chapter 3, we have initially demonstrated that the combined self-efficiency score of the target DMU_k for the overall system is the product of the combined self-efficiency measures of DMU_k for the two sub-stages. In addition, the combined self-efficiency measures obtained with our proposed optimistic-pessimistic evaluation and ranking framework within a two-stage tandem structure, have been compared with the respective scores obtained with **Kao and Hwang's (2008)** approach. Correlation analysis and further comparisons have suggested that there is a very strong association between the ranks of these two approaches. However, our approach is more informative within the self-appraisal context, in that it not only considers the optimistic viewpoint (as in **Kao and Hwang (2008)**), but also the pessimistic viewpoint.

Moreover, since there is no absolute discrimination of some inefficient DMUs considering the combined self-efficiency results at each stage, we have intensified our efforts towards the extension of the optimistic-pessimistic ranking framework by the use of the interval cross-efficiency, in which the respective combined self-efficiency score is embedded. As for the interval cross-efficiency, we have shown that 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 a more reasoned peer-appraisal setting that entails the effects of both the optimistic and pessimistic perspectives.

Furthermore, we have viewed each interval cross-efficiency matrix of the corresponding flow as a multi-criteria decision making problem. To solve this problem, we have initially implemented the goal programming method of **Wang and Elhag (2007)** to obtain the interval local weight of each criterion. To delineate the interval global weight of each alternative, we have then suggested a pair of linear programming models, introduced by **Entani and Tanaka (2007)**. Eventually, we have applied the grey relational analysis tool **(Kuo et al., 2008)** for ranking the interval global weights.

The unique final ranks obtained via the Grey Relational Grades have reflected the improvement of the discriminating power, as compared to the original ranks derived from the combined self-efficiency measures. This practically means that the DMUs, which cannot be fully discriminated by the self-evaluation notion, can be discriminated by the methodologies followed in peer notion. Furthermore, the Grey Relational Grades have been compared with the cross-efficiency ratings obtained via the aggressive-based secondary goal model of **Kao and Liu (2019)**. It has been found that our approach is more multidimensional since it simultaneously considers the most favourable and unfavourable weight sets of each of the other players, while integrating the respective combined self-efficiency measure (and not simply the optimistic self-efficiency). Finally, it has been 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 substages.

In Chapter 4, by applying the traditional black-box model leads to the following: *(i)* all internal operations have been entirely neglected, *(ii)* the efficient DMUs cannot be easily discriminated, and *(iii)* it has been impossible to obtain the efficiencies of the constituent departments within the respective Business School. Furthermore, the results obtained by using **Kao's (2009b)** model, have demonstrated that there is still the problem with the lack of discrimination of the efficient DMUs that cannot lead to a unique and meaningful ranking. It has also been evident that the efficiencies of the main operational functions within the department of a particular hypothetical Business School cannot be computed via **Kao's (2009b)** model.

In addition, the ranks of the efficiency scores of the overall system obtained via the newly proposed additive efficiency decomposition or multiplicative efficiency aggregation models have been compared with the overall systems' ranks of black-box and **Kao's (2009b)** models. It has been statistically inferred that the ranks are quite similar, and this has been verified by the Spearman rank-order correlation test. However, our proposed models have allowed us to not only discriminate the efficient units, but also to simultaneously calculate the efficiencies of the Business School of interest, its departments, and the operational functions within each department. Hence, decision makers will be enabled to take certain actions by strengthening the areas of weakness.

In this chapter, we have additionally drawn other compact conclusions related to the system, the sub-systems, and the efficiencies of their internal units, as well as their type of relationships. As for the additive decomposition model, the system efficiency is decomposed into the weighted arithmetic average of those of the sub-systems, where the weight of the sub-system k is defined as the proportion of the aggregate input consumed by this sub-system in that consumed by all sub-systems. It has also proven that the network-hierarchical system efficiency can be decomposed as the weighted average of the ones of the terminal units, belonging to the hierarchical structure of sub-system k. As for the multiplicative aggregation model, the system efficiency is the product of the efficiencies of its parallel sub-systems (internal departments). Finally, the system efficiency of the additive decomposition model has been confirmed to be greater than or equal to the respective one in the multiplicative aggregation model.

5.3 Research Contributions

This thesis has several appealing contributions to the investigation of systems or organisations composed of a number of independent and/or interdependent internal operations. These contributions have the potential to enhance the fairness and discriminatory power of the evaluation of DMUs with significant internal structure by DEA.

Chapter 2 has identified a novel route to address the challenge of attaining fairness in the evaluation outcomes of the generalised two-stage DEA structure of **Yu and Shi (2014)**. In particular, we have argued that fairness, or the acceptance of an evaluation and ranking by the different DMUs and their constituent flows, is improved by increasing measures related to the degree of discriminatory power, the weight scheme, and the minority vote. As a result, we have proposed a combination of an additive self-efficiency aggregation model, a multi-objective minimax secondary model, and the CRITIC method in an aim to achieve these aspects of fairness and thus a better degree of cooperation between stages of a DMU and among DMUs. This combination is novel in the DEA literature. The application of the CRITIC method to the DEA context is by itself novel.

Chapter 3 has provided new insight into the traditional optimistic-pessimistic evaluation and ranking DEA framework. This methodological framework has been adapted to the specifications of the more realistic two-stage tandem system to better reflect the complex interconnections among its internal sub-systems. There are two salient features that have been explored towards the evaluation of DMUs with such a network structure. As for the first, DMUs have been evaluated based on their own most favourable (optimistic) and unfavourable (pessimistic) optimal multipliers and combined via the geometric average efficiency. As for the second, DMUs have also been assessed in relation to the most favourable and unfavourable weight profiles of each of the other DMUs while integrating the combined self-efficiency measure. The latter statement serves the objectives of the interval cross-efficiency concept.

Chapter 4 has recognised the necessity of assessing a company, in which the network scheme might intertwine with a hierarchical structure. We have, therefore, proposed an embedded hierarchical network structure within a multi-function parallel system. In such a scheme, the constituent sub-systems act independently from one another. This novel network DEA structure can not only discriminate the efficient DMUs, but also calculate simultaneously the efficiencies of the system, its internal parallel sub-systems, and their internal units arranged in a hierarchical form. Their relationships have also been identified. Our ultimate goal, in this chapter, has been to indicate patterns related to the examined system's past achievements, in order to make more concrete future suggestions.

5.4 Research Limitations

While this thesis presents strong results and insightful recommendations and conclusions in ensuring appropriate conditions for a fairer evaluation and ranking of DMUs with a network DEA structure, some limitations and shortcomings are also acknowledged as follows.

In Chapter 2, we have proposed an additive self-efficiency aggregation model in the spirit of **Chen et al. (2009)**. This is the basic self-evaluation model without the further improvements introduced in later chapters. In such a model, the system efficiency is defined as the weighted arithmetic average of its sub-stages. As for its decomposition weights, **Ang and Chen (2016)** proved that they are non-increasing in the order of sub-stages. Put simply, they highlighted that earlier stages would be assigned higher relative importance, affecting the system's efficiency to a greater extent. Based on that, they also demonstrated that the overall and sub-stages' efficiency scores are prone to the impact of the decomposition weights. We acknowledge this as a limitation of our study, and we believe that a re-definition of the weights, reflecting **Ang and Chen's (2016)** research, could accommodate such an issue. In addition, this chapter has successfully managed to consider various factors to attain fairness in the evaluation outcomes of DMUs with a generalised two-stage DEA structure. As a result, our methodological contributions could easily attract the attention of decision-makers in various industries. The utilisation of our proposed techniques in real-world cases and contexts would be essential and interesting.

In Chapter 3, 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 **(Saaty and Vargas, 2013)** or the best-worst method **(Rezaei, 2016)**, to identify in an objective manner the weights. Furthermore, we have 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 to how to identify the DEA-efficient DMUs that constitute the best-practice frontier.

In Chapter 4, the evaluated DMUs should have the same network hierarchical structure. We acknowledge that this requirement will potentially decrease the applicability of the network scheme, since the competing DMUs in a great number of real-life cases do not have exactly the same structure. We believe that this issue can be accommodated in a future research study. Moreover, the proposed multiplicative efficiency aggregation model, that alternatively measures the performance of the multi-function parallel network hierarchical system, is a non-linear network DEA model due to its non-linear objective function. Its solution is considered to be more demanding and computationally complex. In addition, a global optimal solution cannot be easily guaranteed. Instead of solving a series of parametric linear programs, our model could have been transformed into either a second order cone programming or a semi-definite programming problem, following the spirit of **Chen and Zhu (2017)** and **Kuo et al. (2020)** or **Zhang and Chen (2019)**, respectively. The aforementioned techniques lie in the field of convex optimization **(Boyd et al., 2004)** and typically use non-heuristic algorithms (e.g., interior point method) to generate "*feasible approximations and tighter upper bounds on the global optimal solution*" **(Zhu, 2020, p.10)**.

5.5 Future Research Directions

A number of potential directions for future research following on from the work described in this thesis are identified as promising.

Firstly, in Chapter 2, the models were developed under the assumption of the constant returns-to-scale (CRS). A direction for future research could be their advancement to variable returns-to-scale (VRS) input-oriented DEA models. Another potential path could be the intention to tweak the CRITIC method by focusing perhaps on the level of acceptance of the participants on the final evaluation and ranking scheme obtained.

Secondly, in Chapter 3, 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 are positive real numbers, and the DEA models are based on this condition. Future research could relax this assumption by allowing the data points (inputs, intermediate measures, and outputs) to be imprecise 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 **(Zimmermann, 2011)** could be used as an ideal alternative option.

Thirdly, in Chapter 4, the discussion of both the additive decomposition and the multiplicative aggregation models is under the constant returns to scale assumption. These models can be expanded to variable returns to scale situation in order to explore their revised properties. Another challenge for future research could be the measurement and the evaluation of the performance of a system that requires the integration of a hierarchical structure into other more complex and less systematic network processes, such as assembly and disassembly, mixed, and dynamic systems **(Cook et al., 2010; Kao, 2016; Kao, 2017)**. As a final point, it would be promising to associate the network DEA models, developed under this paper's novel network hierarchical structure, with data science **(Zhu, 2020; Shi et al., 2021)**. More specifically, machine learning models such as *k*-nearest neighbours, logistic regression, random forest, classification and regression trees **(Brighton and Mellish, 2002; Jordan and Mitchell, 2015; Katsikopoulos et al., 2021)** could, for instance, accomplish the classification of the DMUs into those in the best-practice frontier and those which are not in this frontier.

Finally, future research could also focus more on the testing of the proposed models and frameworks with empirical data, that is testing their practical value. It would be desirable, for instance, to measure the performance, rank, and select the right mixture of maintenance policies within the shipping industry. We could, alternatively, evaluate and compare the efficiency of potential designs of a particular type of vessel. This could make use of different maintenance policies and technical components to fulfil the requirements of its internal functions arranged in various network DEA formats.

Chapter 6

References

1. Agyemang, M., Kusi-Sarpong, S., Agyemang, J., Jia, F. and Adzanyo, M. (2020). Determining and evaluating socially sustainable supply chain criteria in agri-sector of developing countries: insights from West Africa cashew industry. *Production Planning and Control*, 1-19.

2. Alcaraz, J., Aparicio, J., Monge, J. F. and Ramón, N. (2022). Weight profiles in cross-efficiency evaluation based on hypervolume maximization. *Socio-Economic Planning Sciences*, 101270.

3. Alcaraz, J., Ramón, N., Ruiz, J. L., and Sirvent, I. (2013). Ranking ranges in cross-efficiency evaluations. *European Journal of Operational Research*, 226(3), 516-521.

4. Amatatsu, H., Ueda, T., and Amatatsu, Y. (2012). Efficiency and returns-to-scale of local governments. *Journal of the Operational Research Society*, 63(3), 299-305.

5. Amin, F. U., Dong, Q. L., Grzybowska, K., Ahmed, Z. and Yan, B. R. (2022). A Novel Fuzzy-Based VIKOR–CRITIC Soft Computing Method for Evaluation of Sustainable Supply Chain Risk Management. *Sustainability*, 14(5), 2827.

6. Amirteimoori, A., Kordrostami, S., and Azizi, H. (2016). Additive models for network data envelopment analysis in the presence of shared resources. *Transportation Research Part D: Transport and Environment*, 48, 411-424.

7. An, Q., Chen, H., Xiong, B., Wu, J. and Liang, L. (2017). Target intermediate products setting in a two-stage system with fairness concern. *Omega*, 73, 49-59.

8. Anderson, T.R., Hollingsworth, K. and Inman, L. (2002). The fixed weighting nature of a cross-evaluation model. *Journal of Productivity Analysis*, 17(3), 249-255.

9. Ang, S. and Chen, C.M. (2016). Pitfalls of decomposition weights in the additive multi-stage DEA model. *Omega*, 58, 139-153.

10. Angulo-Meza, L. and Lins, M.P.E. (2002). Review of methods for increasing discrimination in data envelopment analysis. *Annals of Operations Research*, 116(1-4), 225-242.

11. Azadi, M., Shabani, A., Khodakarami, M., and Saen, R. F. (2014). Planning in feasible region by two-stage target-setting DEA methods: An application in green supply chain management of public transportation service providers. *Transportation Research Part E: Logistics and Transportation Review*, 70, 324-338.

12. Azizi, H. (2014). DEA efficiency analysis: A DEA approach with double frontiers. *International Journal of Systems Science*, 45(11), 2289-2300.

13. Azizi, H. (2011). The interval efficiency based on the optimistic and pessimistic points of view. *Applied Mathematical Modelling*, 35(5), 2384-2393.

14. Azizi, H. and Ajirlu, S. F. (2010). Measurement of overall performances of decision-making units using ideal and anti-ideal decision-making units. *Computers and Industrial Engineering*, 59(3), 411-418.

15. Azizi, H. and Jahed, R. (2011). Improved data envelopment analysis models for evaluating interval efficiencies of decision-making units. *Computers and Industrial Engineering*, 61(3), 897-901.

16. Azizi, H. and Wang, Y. M. (2013). Improved DEA models for measuring interval efficiencies of decision-making units. *Measurement*, 46(3), 1325-1332.

17. Badiezadeh, T., Saen, R. F. and Samavati, T. (2018). Assessing sustainability of supply chains by double frontier network DEA: A big data approach. *Computers and Operations Research*, 98, 284-290.

18. Bal, H., Örkcü, H.H. and Çelebioğlu, S. (2010). Improving the discrimination power and weights dispersion in the data envelopment analysis. *Computers and Operations Research*, 37(1), 99-107.

19. Banker, R.D., Charnes, A. and Cooper, W.W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, 30(9), 1078-1092.

20. Banker, R. D. and Morey, R. C. (1986). The use of categorical variables in data envelopment analysis. *Management science*, 32(12), 1613-1627.

21. Banker, R. D., Gadh, V. M. and Gorr, W. L. (1993). A Monte Carlo comparison of two production frontier estimation methods: Corrected ordinary least squares and data envelopment analysis. *European Journal of Operational Research*, 67(3), 332-343.

22. Beasley, J. E. (1995). Determining teaching and research efficiencies. *Journal of the Operational Research Society*, 46(4), 441-452.

23. Beullens, P., Zaibidi, N. Z., and Jones, D. F. (2012). Goal programming to model human decision making in ultimatum games. *International Transactions in Operational Research*, 19(4), 599-612.

24. Bhadra, D., Dhar, N. R. and Salam, M. A. (2021). Sensitivity analysis of the integrated AHP-TOPSIS and CRITIC-TOPSIS method for selection of the natural fiber. *Materials Today: Proceedings*.

25. Boďa, M., Dlouhý, M. and Zimková, E. (2020). Modeling a shared hierarchical structure in data envelopment analysis: An application to bank branches. *Expert Systems with Applications*, 162, 113700.

26. Boyd, S., Boyd, S. P. and Vandenberghe, L. (2004). Convex optimization. Cambridge university press.

27. Brighton, H. and Mellish, C. (2002). Advances in instance selection for instance-based learning algorithms. *Data Mining and Knowledge Discovery*, 6(2), 153-172.

28. Castelli, L., Pesenti, R. and Ukovich, W. (2004). DEA-like models for the efficiency evaluation of hierarchically structured units. *European Journal of Operational Research*, 154(2), 465-476.

29. Casu, B. and Thanassoulis, E. (2006). Evaluating cost efficiency in central administrative services in UK universities. *Omega*, 34(5), 417-426.

30. Chang, A. Y. (2012). Prioritising the types of manufacturing flexibility in an uncertain environment. *International Journal of Production Research*, 50(8), 2133-2149.

31. Charnes, A. and Cooper, W. W. (1962). Programming with linear fractional functionals. *Naval Research logistics quarterly*, 9(3-4), 181-186.

32. Charnes, A., Cooper, W.W. and Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 2(6), 429-444.

33. Charles, V., Gherman, T. and Zhu, J. (2021). Data Envelopment Analysis and Big Data: A Systematic Literature Review with Bibliometric Analysis. *Data-Enabled Analytics*, 1-29.

34. Chen, C., Cook, W. D., Imanirad, R. and Zhu, J. (2020). Balancing Fairness and Efficiency: Performance Evaluation with Disadvantaged Units in Non-homogeneous Environments. *European Journal of Operational Research*, 287(3), 1003-1013.

35. Chen, J. X. (2014). Overall performance evaluation: new bounded DEA models against unreachability of efficiency. *Journal of the Operational Research Society*, 65(7), 1120-1132.

36. Chen, K. and Zhu, J. (2019). Scale efficiency in two-stage network DEA. *Journal of the Operational Research Society*, 70(1), 101-110.

37. Chen, K. and Zhu, J. (2017). Second order cone programming approach to two-stage network data envelopment analysis. *European Journal of Operational Research*, 262(1), 231-238.

38. Chen, P. C., Yu, M. M., Shih, J. C., Chang, C. C. and Hsu, S. H. (2019). A reassessment of the Global Food Security Index by using a hierarchical data envelopment analysis approach. *European Journal of Operational Research*, 272(2), 687-698.

39. Chen, Y., Cook, W.D., Li, N. and Zhu, J. (2009). Additive efficiency decomposition in two-stage DEA. *European Journal of Operational Research*, 196(3), 1170-1176.

40. Chen, Y., Du, J., Sherman, H.D. and Zhu, J. (2010). DEA model with shared resources and efficiency decomposition. *European Journal of Operational Research*, 207(1), 339-349.

41. Chen, Y., Ma, X., Yan, P. and Wang, M. (2021). Operating efficiency in Chinese universities: an extended two-stage network DEA approach. *Journal of Management Science and Engineering*.

42. Cook, W. D. and Green, R. H. (2005). Evaluating power plant efficiency: a hierarchical model. *Computers and Operations Research*, 32(4), 813-823.

43. Cook, W. D. and Seiford, L. M. (2009). Data envelopment analysis (DEA)–Thirty years on. *European Journal of Operational Research*, 192(1), 1-17.

44. Cook, W. D., Zhu, J., Bi, G. and Yang, F. (2010). Network DEA: Additive efficiency decomposition. *European Journal of Operational Research*, 207(2), 1122-1129.

45. Cook, W. D., Liang, L. and Zhu, J. (2010b). Measuring performance of two-stage network structures by DEA: A review and future perspective. *Omega*, 38(6), 423-430.

46. Cook, W. D. and Zhu, J. (2007). Classifying inputs and outputs in data envelopment analysis. *European Journal of Operational Research*, 180(2), 692-699.

47. Cook, W.D., Tone, K. and Zhu, J. (2014). Data envelopment analysis: Prior to choosing a model. *Omega*, 44, 1-4.

48. Cooper, W. W., Seiford, L. M. and Tone, K. (2000). Data envelopment analysis. Handbook on data envelopment analysis, 1-40.

49. Cooper, W. W., Seiford, L. M. and Zhu, J. (2011). Handbook on data envelopment analysis.

50. Daniel, W. W. (1978). Applied nonparametric statistics. Houghton Mifflin.

51. Dehnokhalaji, A., Khezri, S. and Emrouznejad, A. (2022). A box-uncertainty in DEA: A robust performance measurement framework. *Expert Systems with Applications*, 187, 115855.

52. Despotis, D. K., Koronakos, G. and Sotiros, D. (2016). Composition versus decomposition in two-stage network DEA: a reverse approach. *Journal of productivity Analysis*, 45(1), 71-87.

53. Despotis, D. K., Sotiros, D. and Koronakos, G. (2016). A network DEA approach for series multi-stage processes. *Omega*, 61, 35-48.

54. De Witte, K., Rogge, N., Cherchye, L. and Van Puyenbroeck, T. (2013). Economies of scope in research and teaching: a non-parametric investigation. *Omega*, 41(2), 305-314.

55. Diakoulaki, D., Mavrotas, G. and Papayannakis, L. (1995). Determining objective weights in multiple criteria problems: The critic method. *Computers and Operations Research*, 22(7), 763-770.

56. dos Santos Rubem, A.P., de Mello, J.C.C.S. and Meza, L.A. (2017). A goal programming approach to solve the multiple criteria DEA model. *European Journal of Operational Research*, 260(1), 134-139.

57. Doyle, J. and Green, R. (1994). Efficiency and cross-efficiency in DEA: Derivations, meanings and uses. *Journal of the Operational Research Society*, 45(5), 567-578.

58. Du, J., Chen, Y. and Huo, J. (2015). DEA for non-homogenous parallel networks. *Omega*, 56, 122-132.

59. Dyson, R.G. and Thanassoulis, E. (1988). Reducing weight flexibility in data envelopment analysis. *Journal of the Operational Research Society*, 39(6), 563-576.

60. Ebrahimnejad, A. (2012). Cost efficiency measures with trapezoidal fuzzy numbers in data envelopment analysis based on ranking functions: application in insurance organization and hospital. *International Journal of Fuzzy System Applications (IJFSA)*, 2(3), 51-68.

61. Ebrahimnejad, A. and Lotfi, F.H. (2012). Equivalence relationship between the general combined-oriented CCR model and the weighted minimax MOLP formulation. *Journal of King Saud University-Science*, 24(1), 47-54.

62. Ekiz, M. K. and Tuncer Şakar, C. (2020). A new DEA approach to fully rank DMUs with an application to MBA programs. *International Transactions in Operational Research*, 27(4), 1886-1910.

63. Emrouznejad, A. and Yang, G. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences*, 61, 4–8.

64. Entani, T. and Tanaka, H. (2007). Interval estimations of global weights in AHP by upper approximation. *Fuzzy Sets and Systems*, 158(17), 1913-1921.

65. Entani, T., Maeda, Y. and Tanaka, H. (2002). Dual models of interval DEA and its extension to interval data. *European Journal of Operational Research*, 136(1), 32-45.

66. Fang, L. and Yang, J. (2019). An integrated ranking approach using cross-efficiency intervals and the cumulative prospect theory. *Computers and Industrial Engineering*, 136, 556-574.

67. Färe, R. and Grosskopf, S. (2000). Network DEA. *Socio-Economic Planning Sciences*, 34, 35-49.

68. Fehr, E., Fischbacher, U. and Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human nature*, 13(1), 1-25.

69. Fu, Y. and Li, M. (2022). DEA cross-efficiency aggregation based on preference structure and acceptability analysis. *International Transactions in Operational Research*, 29(2), 987-1011.

70. Fukuyama, H. and Mirdehghan, S. M. (2012). Identifying the efficiency status in network DEA. *European Journal of Operational Research*, 220(1), 85-92.

71. Fukuyama, H. and Weber, W. L. (2021). Network DEA and Big Data with an Application to the Coronavirus Pandemic. *In Data-Enabled Analytics* (pp. 175-197). Springer, Cham.

72. Gan, G., Lee, H. S., Lee, L., Wang, X. and Wang, Q. (2019). Network hierarchical DEA with an application to international shipping industry in Taiwan. *Journal of the Operational Research Society*, 1-12.

73. Ghasemi, M.R., Ignatius, J. and Emrouznejad, A. (2014). A bi-objective weighted model for improving the discrimination power in MCDEA. *European Journal of Operational Research*, 233(3), 640-650.

74. Ghasemi, N., Najafi, E., Lotfi, F. H. and Sobhani, F. M. (2020). Assessing the performance of organizations with the hierarchical structure using data envelopment analysis: An efficiency analysis of Farhangian University. *Measurement*, 156, 107609.

75. Ghazi, A. and Lotfi, F. H. (2022). A survey on links between multiple objective decision making and data envelopment analysis. *In Multi-Objective Combinatorial Optimization Problems and Solution Methods* (pp. 29-70). Academic Press.

76. Giraleas, D., Emrouznejad, A. and Thanassoulis, E. (2012). Productivity change using growth accounting and frontier-based approaches–Evidence from a Monte Carlo analysis. *European Journal of Operational Research*, 222(3), 673-683.

77. Guo, C., Shureshjani, R. A., Foroughi, A. A. and Zhu, J. (2017). Decomposition weights and overall efficiency in two-stage additive network DEA. *European Journal of Operational Research*, 257(3), 896-906.

78. Halkos, G.E., Tzeremes, N.G. and Kourtzidis, S.A. (2014). A unified classification of two-stage DEA models. *Surveys in Operations Research and Management Science*, 19(1), 1-16.

79. Halme, M., Joro, T., Korhonen, P., Salo, S. and Wallenius, J. (1999). A value efficiency approach to incorporating preference information in data envelopment analysis. *Management Science*, 45(1), 103-115.

80. Hartmann, F. and Slapničar, S. (2012). The perceived fairness of performance evaluation: The role of uncertainty. *Management Accounting Research*, 23(1), 17-33.

81. Hatami-Marbini, A., Ebrahimnejad, A. and Lozano, S. (2017). Fuzzy efficiency measures in data envelopment analysis using lexicographic multiobjective approach. *Computers & industrial engineering*, 105, 362-376.

82. Hatami-Marbini, A., Saati, S. and Tavana, M. (2010). An ideal-seeking fuzzy data envelopment analysis framework. *Applied Soft Computing*, 10(4), 1062-1070.

83. Hatami-Marbini, A. and Toloo, M. (2017). An extended multiple criteria data envelopment analysis model. *Expert Systems with Applications*, 73, 201-219.

84. He, Y. and Ma, Z. (2015). Study on CRITIC-DEA for Regional Logistics and Regional Economic Coordinated Development Model and Evaluation—A Case Study of Sichuan Province. *Soft Science*, 03.

85. Huang, Y., Wang, Y. M. and Lin, J. (2019). Two-Stage Fuzzy Cross-Efficiency Aggregation Model Using a Fuzzy Information Retrieval Method. *International Journal of Fuzzy Systems*, 21(8), 2650-2666.

86. Huang, Z., Xu, R. H., Yang, R. D. and Liu, W. (2018). Evaluation on equilibrium of passenger flow distribution on urban rail transit network. *J Transp Syst Eng Inform Technol*, 18(3), 139-145.

87. Hwang, C. L. and Yoon, K. (1981). Methods for multiple attribute decision making. In Multiple attribute decision making (pp. 58-191). Springer, Berlin, Heidelberg.

88. Jahan, A., Mustapha, F., Sapuan, S.M., Ismail, M.Y. and Bahraminasab, M. (2012). A framework for weighting of criteria in ranking stage of material selection process. *The International Journal of Advanced Manufacturing Technology*, 58(1-4), 411-420.

89. Jahanshahloo, G.R., Amirteimoori, A.R. and Kordrostami, S. (2004). Multi-component performance, progress and regress measurement and shared inputs and outputs in DEA for panel data: an application in commercial bank branches. *Applied Mathematics and Computation*, 151(1), 1-16.

90. Jahanshahloo, G. R., Sadeghi, J. and Khodabakhshi, M. (2017). Fair ranking of the decision making units using optimistic and pessimistic weights in data envelopment analysis. *RAIRO-Operations Research*, 51(1), 253-260.

91. Jianfeng, M.A. (2015). A two-stage DEA model considering shared inputs and free intermediate measures. *Expert Systems with Applications*, 42(9), 4339-4347.

92. Jordan, M. I. and Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260.

93. Julong, D. (1989). Introduction to grey system theory. *The Journal of Grey System*, 1(1), 1-24.

94. Kao, C. (2009a). Efficiency decomposition in network data envelopment analysis: A relational model. *European Journal of Operational Research*, 192(3), 949-962.

95. Kao, C. (2009b). Efficiency measurement for parallel production systems. *European Journal of Operational Research*, 196(3), 1107-1112.

96. Kao, C. (2012). Efficiency decomposition for parallel production systems. *Journal of the Operational Research Society*, 63(1), 64-71.

97. Kao, C. (2014). Network data envelopment analysis: A review. *European Journal of Operational Research*, 239(1), 1-16.

98. Kao, C. (2015). Efficiency measurement for hierarchical network systems. *Omega*, 51, 121-127.

99. Kao, C. (2016). Efficiency decomposition and aggregation in network data envelopment analysis. *European Journal of Operational Research*, 255(3), 778-786.

100. Kao, C. (2017). Network Data Envelopment Analysis: Foundations and Extensions. 2nd ed. Springer.

101. Kao, C. (2018). Multiplicative aggregation of division efficiencies in network data envelopment analysis. *European Journal of Operational Research*, 270(1), 328-336.

102. Kao, C. and Hwang, S.N. (2008). Efficiency decomposition in two-stage data envelopment analysis: An application to non-life insurance companies in Taiwan. *European Journal of Operational Research*, 185(1), 418-429.

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

104. Kao, C. and Hung, H. T. (2008). Efficiency analysis of university departments: An empirical study. *Omega*, 36(4), 653-664.

105. Kao, C. and Hwang, S. N. (2011). Decomposition of technical and scale efficiencies in two-stage production systems. *European Journal of Operational Research*, 211(3), 515-519.

106. Katsikopoulos, K., Simsek, O., Buckmann, M. and Gigerenzer, G. (2021). Classification in the wild: The art and science of transparent decision making. MIT Press.

107. Kaviani, M. A., Yazdi, A. K., Ocampo, L. and Kusi-Sarpong, S. (2019). An integrated grey-based multi-criteria decision-making approach for supplier evaluation and selection in the oil and gas industry. *Kybernetes*.

108. Khalili-Damghani, K., Tavana, M. and Haji-Saami, E. (2015). A data envelopment analysis model with interval data and undesirable output for combined cycle power plant performance assessment. *Expert Systems with Applications*, 42(2), 760-773.

109. Khezrimotlagh, D. (2022). Simulation designs for production frontiers. *European Journal of Operational Research*.

110. Khodabakhshi, M. and Aryavash, K. (2017). The cross-efficiency in the optimistic–pessimistic framework. *Operational Research*, 17(2), 619-632.

111. Khoveyni, M. and Eslami, R. (2022). Merging two-stage series network structures: A DEA-based approach. *OR Spectrum*, 44(1), 273-302.

112. Kiaei, H. and Kazemi Matin, R. (2022). New common set of weights method in black-box and two-stage data envelopment analysis. *Annals of Operations Research*, 309(1), 143-162.

113. Koronakos, G., Sotiros, D., Despotis, D. K. and Kritikos, M. N. (2022). Fair efficiency decomposition in network DEA: A compromise programming approach. *Socio-Economic Planning Sciences*, 79, 101100.

114. Koronakos, G., Sotiros, D. and Despotis, D. K. (2019). Reformulation of Network Data Envelopment Analysis models using a common modelling framework. *European Journal of Operational Research*, 278(2), 472-480.

115. Koronakos, G., Sotiros, D., Despotis, D. K. and Kritikos, M. N. (2022). Fair efficiency decomposition in network DEA: A compromise programming approach. *Socio-Economic Planning Sciences*, 79, 101100.

116. Kremantzis, M. D., Beullens, P. and Klein, J. (2022). A fairer assessment of DMUs in a generalised two-stage DEA structure. *Expert Systems with Applications*, 187, 115921.

117. Kremantzis, M. D., Beullens, P. and Klein, J. (2022). A ranking framework based on interval self and cross-efficiencies in a two-stage DEA system. *RAIRO - Operations Research*, 56(3), 1293-1319.

118. Krishnan, A. R., Kasim, M. M., Hamid, R. and Ghazali, M. F. (2021). A modified CRITIC method to estimate the objective weights of decision criteria. *Symmetry*, 13(6), 973.

119. Kuo, K. C., Lu, W. M. and Dinh, T. N. (2020). An integrated efficiency evaluation of China stock market. *Journal of the Operational Research Society*, 1-18.

120. Kuo, Y., Yang, T. and Huang, G.W. (2008). The use of grey relational analysis in solving multiple attribute decision-making problems. *Computers and Industrial Engineering*, 55(1), 80-93.

121. Lee, B. L. and Johnes, J. (2021). Using network DEA to inform policy: The case of the teaching quality of higher education in England. *Higher Education Quarterly*.

122. Lee, H. S. (2021). Efficiency decomposition of the network DEA in variable returns to scale: An additive dissection in losses. *Omega*, 100, 102212.

123. León, T., Ramón, N., Ruiz, J. L. and Sirvent, I. (2014). Using Induced Ordered Weighted Averaging (IOWA) Operators for Aggregation in Cross-Efficiency Evaluations. *International Journal of Intelligent Systems*, 29(12), 1100-1116.

124. Leventhal, G. S., Karuza, J. and Fry, W. R. (1980). Beyond fairness: A theory of allocation preferences. *Justice and social interaction*, 3(1), 167-218.

125. Li, F., Wu, H., Zhu, Q., Liang, L. and Kou, G. (2021). Data envelopment analysis cross efficiency evaluation with reciprocal behaviors. *Annals of Operations Research*, 1-38.

126. Li, F., Zhu, Q., Chen, Z. and Xue, H. (2018). A balanced data envelopment analysis cross-efficiency evaluation approach. *Expert Systems with Applications*, 106, 154-168.

127. Li, W., Cook, W. D., Li, Z. and Zhu, J. (2020). Efficiency measurement for hierarchical situations. *Journal of the Operational Research Society*, 1-9.

128. Li, X.B. and Reeves, G.R. (1999). A multiple criteria approach to data envelopment analysis. *European Journal of Operational Research*, 115(3), 507-517.

129. Li, X. (2017). A fair evaluation of certain stage in a two-stage structure: revisiting the typical two-stage DEA approaches. *Omega*, 68, 155-167.

130. Li, Z., Feng, C. and Tang, Y. (2022). Bank efficiency and failure prediction: a nonparametric and dynamic model based on data envelopment analysis. *Annals of Operations Research*, 1-37.

131. Liang, L., Wu, J., Cook, W.D. and Zhu, J. (2008). Alternative secondary goals in DEA cross-efficiency evaluation. *International Journal of Production Economics*, 113(2), 1025-1030.

132. Liang, X., Li, J., Guo, G., Li, S. and Gong, Q. (2021). Evaluation for water resource system efficiency and influencing factors in western China: A two-stage network DEA-Tobit model. *Journal of Cleaner Production*, 328, 129674.

133. Liu, J. S., Lu, L. Y., Lu, W. M. and Lin, B. J. (2013). A survey of DEA applications. *Omega*, 41(5), 893-902.

134. Liu, S. T. (2018). A DEA ranking method based on cross-efficiency intervals and signal-to-noise ratio. *Annals of Operations Research*, 261(1), 207-232.

135. Liu, W. and Wang, Y.M. (2018). Ranking DMUs by using the upper and lower bounds of the normalized efficiency in data envelopment analysis. *Computers and Industrial Engineering*, 125, 135-143.

136. Liu, X., Chu, J., Yin, P., and Sun, J. (2017). DEA cross-efficiency evaluation considering undesirable output and ranking priority: A case study of eco-efficiency analysis of coal-fired power plants. *Journal of Cleaner Production*, 142, 877-885.

137. Lotfi, F.H., Jahanshahloo, G.R., Ebrahimnejad, A., Soltanifar, M. and Mansourzadeh, S.M. (2010a). Target setting in the general combined-oriented CCR model using an interactive MOLP method. *Journal of Computational and Applied Mathematics*, 234(1), 1-9.

138. Lotfi, F.H., Jahanshahloo, G.R., Soltanifar, M., Ebrahimnejad, A. and Mansourzadeh, S.M. (2010b). Relationship between MOLP and DEA based on output-orientated CCR dual model. *Expert Systems with Applications*, 37(6), 4331-4336.

139. Lozano, S. (2015). A joint-inputs Network DEA approach to production and pollution-generating technologies. *Expert Systems with Applications*, 42(21), 7960-7968.

140. Lozano, S. and Khezri, S. (2021). Network DEA smallest improvement approach. *Omega*, 98, 102140.

141. Lu, H., Zhao, Y., Zhou, X. and Wei, Z. (2022). Selection of Agricultural Machinery Based on Improved CRITIC-Entropy Weight and GRA-TOPSIS Method. *Processes*, 10(2), 266.

142. Lu, W. M., Kweh, Q. L. and Yang, K. C. (2020). Multiplicative efficiency aggregation to evaluate Taiwanese local auditing institutions performance. *Annals of Operations Research*, 1-20.

143. Lu, W. M., Liu, J. S., Kweh, Q. L. and Wang, C. W. (2016). Exploring the benchmarks of the Taiwanese investment trust corporations: Management and investment efficiency perspectives. *European Journal of Operational Research*, 248(2), 607-618.

144. Ma, C., Liu, D., Zhou, Z., Zhao, W. and Liu, W. (2014). Game cross efficiency for systems with two-stage structures. *Journal of Applied Mathematics*, 2014, 8.

145. Ma, J., Qi, L. and Deng, L. (2017). Efficiency measurement and decomposition in hybrid two-stage DEA with additional inputs. *Expert Systems with Applications*, 79, 348-357.

146. Mahdiloo, M., Jafarzadeh, A.H., Saen, R.F., Tatham, P. and Fisher, R. (2016). A multiple criteria approach to two-stage data envelopment analysis. *Transportation Research Part D: Transport and Environment*, 46, 317-327.

147. Mavi, R. K., Saen, R. F. and Goh, M. (2019). Joint analysis of eco-efficiency and eco-innovation with common weights in two-stage network DEA: A big data approach. *Technological Forecasting and Social Change*, 144, 553-562.

148. Meng, F. and Xiong, B. (2021). Logical efficiency decomposition for general two-stage systems in view of cross efficiency. *European Journal of Operational Research*.

149. Molinero, C. M. (1996). On the joint determination of efficiencies in a data envelopment analysis context. *Journal of the Operational Research Society*, 47(10), 1273-1279.

150. Michali, M., Emrouznejad, A., Dehnokhalaji, A. and Clegg, B. (2021). Noise-pollution efficiency analysis of European railways: a network dea model. *Transportation Research Part D: Transport and Environment*, 98, 102980.

151. Moncayo–Martínez, L. A., Ramírez–Nafarrate, A. and Hernández–Balderrama, M. G. (2020). Evaluation of public HEI on teaching, research, and knowledge dissemination by Data Envelopment Analysis. *Socio-Economic Planning Sciences*, 69, 100718.

152. Moreno, P. and Lozano, S. (2014). A network DEA assessment of team efficiency in the NBA. *Annals of Operations Research*, 214(1), 99-124.

153. Omrani, H., Alizadeh, A., Emrouznejad, A. and Teplova, T. (2022a). A robust credibility DEA model with fuzzy perturbation degree: An application to hospitals performance. *Expert Systems with Applications*, 189, 116021.

154. Omrani, H., Emrouznejad, A., Shamsi, M. and Fahimi, P. (2022b). Evaluation of insurance companies considering uncertainty: A multi-objective network data envelopment analysis model with negative data and undesirable outputs. *Socio-Economic Planning Sciences*, 101306.

155. Omrani, H., Valipour, M. and Mamakani, S. J. (2019). Construct a composite indicator based on integrating Common Weight Data Envelopment Analysis and principal component analysis models: An application for finding development degree of provinces in Iran. *Socio-Economic Planning Sciences*, 68, 100618.

156. Örkcü, H. H., Özsoy, V. S., Örkcü, M. and Bal, H. (2019). A neutral cross efficiency approach for basic two stage production systems. *Expert Systems with Applications*, 125, 333-344.

157. Örkcü, M., Özsoy, V. S. and Örkcü, H. H. (2020). An optimistic-pessimistic DEA model based on game cross efficiency approach. *RAIRO-Operations Research*, 54(4), 1215-1230.

158. Oukil, A. and Amin, G. R. (2015). Maximum appreciative cross-efficiency in DEA: A new ranking method. *Computers and Industrial Engineering*, 81, 14-21.

159. Pan, B., Liu, S., Xie, Z., Shao, Y., Li, X. and Ge, R. (2021). Evaluating operational features of three unconventional intersections under heavy traffic based on CRITIC method. *Sustainability*, 13(8), 4098.

160. Paradi, J. C., and Zhu, H. (2013). A survey on bank branch efficiency and performance research with data envelopment analysis. *Omega*, 41(1), 61-79.

161. Park, G. and DeShon, R.P. (2010). A multilevel model of minority opinion expression and team decision-making effectiveness. *Journal of Applied Psychology*, 95(5), 824-833.

162. Park, J., Lee, B. K. and Low, J. M. (2022). A two-stage parallel network DEA model for analyzing the operational capability of container terminals. *Maritime Policy and Management*, 49(1), 118-139.

163. Peng, X., Zhang, X. and Luo, Z. (2020). Pythagorean fuzzy MCDM method based on CoCoSo and CRITIC with score function for 5G industry evaluation. *Artificial Intelligence Review*, 53(5), 3813-3847.

164. Pérez-González, A., de Carlos, P. and Alén, E. (2021). An analysis of the efficiency of football clubs in the Spanish First Division through a two-stage relational network DEA model: a simulation study. *Operational Research*, 1-24.

165. Peykani, P., Hosseinzadeh Lotfi, F., Sadjadi, S.J., Ebrahimnejad, A. and Mohammadi, E. (2021). Fuzzy chance-constrained data envelopment analysis: a structured literature review, current trends, and future directions. *Fuzzy Optimization and Decision Making*, 1-65.

166. Pourmohammadi, K., Shojaei, P., Rahimi, H. and Bastani, P. (2018). Evaluating the health system financing of the Eastern Mediterranean Region (EMR) countries

using Grey Relation Analysis and Shannon Entropy. *Cost Effectiveness and Resource Allocation*, 16(1), 1-9.

167. Qu, J., Wang, B. and Liu, X. (2022). A modified super-efficiency network data envelopment analysis: Assessing regional sustainability performance in China. *Socio-Economic Planning Sciences*, 101262.

168. Radovanović, S., Savić, G., Delibašić, B. and Suknović, M. (2021). FairDEA: Removing disparate impact from efficiency scores. *European Journal of Operational Research*.

169. Rakhshan, S.A. (2017). Efficiency ranking of decision-making units in data envelopment analysis by using TOPSIS-DEA method. *Journal of the Operational Research Society*, 68(8), 906-918.

170. Ramón, N., Ruiz, J. L. and Sirvent, I. (2014). Dominance relations and ranking of units by using interval number ordering with cross-efficiency intervals. *Journal of the Operational Research Society*, 65(9), 1336-1343.

171. Ramón, N., Ruiz, J. L. and Sirvent, I. (2010). On the choice of weights profiles in cross-efficiency evaluations. *European Journal of Operational Research*, 207(3), 1564-1572.

172. Rezaei, J. (2016). Best-worst multi-criteria decision-making method: Some properties and a linear model. *Omega*, 64, 126-130.

173. Roll, Y., Cook, W.D. and Golany, B. (1991). Controlling factor weights in data envelopment analysis. *IIE transactions*, 23(1), 2-9.

174. Ruiz, J. L. and Sirvent, I. (2012). On the DEA total weight flexibility and the aggregation in cross-efficiency evaluations. *European Journal of Operational Research*, 223(3), 732-738.

175. Saaty, T. L. and Vargas, L. G. (2013). The analytic network process. In Decision making with the analytic network process (1-40). Springer, Boston, MA.

176. Santos Arteaga, F.J., Ebrahimnejad, A. and Zabihi, A. (2021). A new approach for solving intuitionistic fuzzy data envelopment analysis problems. *Fuzzy Optimization and Modeling Journal*, 2(2), 46-56.

177. Sarraf, F. and Nejad, S. H. (2020). Improving performance evaluation based on balanced scorecard with grey relational analysis and data envelopment analysis approaches: Case study in water and wastewater companies. *Evaluation and Program Planning*, 79, 101762.

178. Sexton, T.R., Silkman, R.H. and Hogan, A.J. (1986). Data envelopment analysis: Critique and extensions. *New Directions for Program Evaluation*, 1986(32), 73-105.

179. Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379-423.

180. Shao, X. and Wang, M. (2021). Two-stage cross-efficiency evaluation based on prospect theory. *Journal of the Operational Research Society*, 1-13.

181. Shao, Y., Luo, Z., Wu, H., Han, X., Pan, B., Liu, S. and Claudel, C. G. (2020). Evaluation of two improved schemes at non-aligned intersections affected by a work zone with an entropy method. *Sustainability*, 12(14), 5494.

182. Shi, H. L., Chen, S. Q., Chen, L. and Wang, Y. M. (2021). A neutral cross-efficiency evaluation method based on interval reference points in consideration of bounded rational behavior. *European Journal of Operational Research*, 290(3), 1098-1110.

183. Shi, X., Emrouznejad, A. and Yu, W. (2021). Overall efficiency of operational process with undesirable outputs containing both series and parallel processes: A SBM network DEA model. *Expert Systems with Applications*, 178, 115062.

184. Shi, Y., Zhu, J., and Charles, V. (2020). Data science and productivity: A bibliometric review of data science applications and approaches in productivity evaluations. *Journal of the Operational Research Society*, 1-14.

185. Song, L. and Liu, F. (2018). An improvement in DEA cross-efficiency aggregation based on the Shannon entropy. *International Transactions in Operational Research*, 25(2), 705-714.

186. Song, M., Zhu, Q., Peng, J. and Gonzalez, E.D.S. (2017). Improving the evaluation of cross efficiencies: A method based on Shannon entropy weight. *Computers and Industrial Engineering*, 112, 99-106.

187. Stefaniec, A., Hosseini, K., Xie, J., and Li, Y. (2020). Sustainability assessment of inland transportation in China: A triple bottom line-based network DEA approach. *Transportation Research Part D: Transport and Environment*, 80, 102258.

188. Stewart, T. J. (1996). Relationships between data envelopment analysis and multicriteria decision analysis. *Journal of the Operational Research Society*, 47(5), 654-665.

189. Su, Y. and Sun, W. (2018). Sustainability evaluation of the supply chain with undesired outputs and dual-role factors based on double frontier network DEA. *Soft Computing*, 22(16), 5525-5533.

190. Sugihara, K., Ishii, H., and Tanaka, H. (2004). Interval priorities in AHP by interval regression analysis. *European Journal of Operational Research*, 158(3), 745-754.

191. Summerfield, N. S., Deokar, A. V., Xu, M., and Zhu, W. (2020). Should drivers cooperate? Performance evaluation of cooperative navigation on simulated road networks using network DEA. *Journal of the Operational Research Society*, 1-16.

192. Taleb, M., Khalid, R., Ramli, R., Ghasemi, M. R. and Ignatius, J. (2022). An integrated bi-objective data envelopment analysis model for measuring returns to scale. *European Journal of Operational Research*, 296(3), 967-979.

193. Tan, Y., Wanke, P., Antunes, J. and Emrouznejad, A. (2021). Unveiling endogeneity between competition and efficiency in Chinese banks: A two-stage network DEA and regression analysis. *Annals of Operations Research*, 306(1), 131-171.

194. Tavana, M., Ebrahimnejad, A., Santos-Arteaga, F.J., Mansourzadeh, S.M. and Matin, R.K. (2018). A hybrid DEA-MOLP model for public school assessment and closure decision in the City of Philadelphia. *Socio-Economic Planning Sciences*, 61, 70-89.

195. Toloo, M. and Tichý, T. (2015). Two alternative approaches for selecting performance measures in data envelopment analysis. *Measurement*, 65, 29-40.

196. Tone, K. and Tsutsui, M. (2009). Network DEA: A slacks-based measure approach. *European Journal of Operational Research*, 197(1), 243-252.

197. Troutt, M.D. (1997). Derivation of the maximin efficiency ratio model from the maximum decisional efficiency principle. *Annals of Operations Research*, 73, 323-338.

198. Vaz, C. B., Camanho, A. S., and Guimarães, R. C. (2010). The assessment of retailing efficiency using network data envelopment analysis. *Annals of Operations Research*, 173(1), 5-24.

199. Wang, C. H., Lu, Y. H., Huang, C. W., and Lee, J. Y. (2013). R&D, productivity, and market value: An empirical study from high-technology firms. *Omega*, 41(1), 143-155.

200. Wang, C. N., Le, T. Q., Yu, C. H., Ling, H. C. and Dang, T. T. (2022). Strategic Environmental Assessment of Land Transportation: An Application of DEA with Undesirable Output Approach. *Sustainability*, 14(2), 972.

201. Wang, L., Chu, J. and Wu, J. (2007). Selection of optimum maintenance strategies based on a fuzzy analytic hierarchy process. *International Journal of Production Economics*, 107(1), 151-163.

202. Wang, Q., Wu, C. and Sun, Y. (2015). Evaluating corporate social responsibility of airlines using entropy weight and grey relation analysis. *Journal of Air Transport Management*, 42, 55-62.

203. Wang, R. T., Ho, C. T. B. and Oh, K. (2010). Measuring production and marketing efficiency using grey relation analysis and data envelopment analysis. *International Journal of Production Research*, 48(1), 183-199.

204. Wang, Y. M. and Elhag, T.M. (2007). A goal programming method for obtaining interval weights from an interval comparison matrix. *European Journal of Operational Research*, 177(1), 458-471.

205. Wang, Y.M. and Chin, K.S. (2010a). A neutral DEA model for cross-efficiency evaluation and its extension. *Expert Systems with Applications*, 37(5), 3666-3675.

206. Wang, Y. M. and Chin, K. S. (2010c). Some alternative DEA models for two-stage process. *Expert Systems with Applications*, 37(12), 8799-8808.

207. Wang, Y.M. and Chin, K.S. (2010b). Some alternative models for DEA cross-efficiency evaluation. *International Journal of Production Economics*, 128(1), 332-338.

208. Wang, Y.M. and Chin, K.S. (2011). The use of OWA operator weights for cross-efficiency aggregation. *Omega*, 39(5), 493-503.

209. Wang, Y.M. and Luo, Y. (2006). DEA efficiency assessment using ideal and anti-ideal decision-making units. *Applied Mathematics and Computation*, 173(2), 902-915.

210. Wang, Y., Pan, J. F., Pei, R. M., Yi, B. W. and Yang, G. L. (2020). Assessing the technological innovation efficiency of China's high-tech industries with a two-stage network DEA approach. *Socio-Economic Planning Sciences*, 71, 100810.

211. Wang, Y., Wu, D., Qian, W. and Li, H. (2021). Cross-efficiency intervals integrated ranking approach based on the generalized Fermat-Torricelli point. *Computers and Industrial Engineering*, 162, 107786.

212. Wang, Y.M. and Wang, S. (2013). Approaches to determining the relative importance weights for cross-efficiency aggregation in data envelopment analysis. *Journal of the Operational Research Society*, 64(1), 60-69.

213. Wang, Y.M. and Yang, J.B. (2007). Measuring the performances of decision-making units using interval efficiencies. *Journal of Computational and Applied Mathematics*, 198(1), 253-267.

214. Wang, Y.M., Chin, K.S. and Luo, Y. (2011). Cross-efficiency evaluation based on ideal and anti-ideal decision-making units. *Expert Systems with Applications*, 38(8), 10312-10319.

215. Wang, Y.M., Chin, K.S. and Yang, J.B. (2007). Measuring the performances of decision-making units using geometric average efficiency. *Journal of the Operational Research Society*, 58(7), 929-937.

216. Wang, Y.M., Greatbanks, R. and Yang, J.B. (2005). Interval efficiency assessment using data envelopment analysis. *Fuzzy Sets and Systems*, 153(3), 347-370.

217. Wanke, P. and Barros, C. (2014). Two-stage DEA: An application to major Brazilian banks. *Expert Systems with Applications*, 41(5), 2337-2344.

218. Wu, D. (2006). A note on DEA efficiency assessment using ideal point: an improvement of Wang and Luo's model. *Applied Mathematics and Computation*, 183(2), 819-830.

219. Wu, D., Wang, N., Yang, Z., Li, C. and Yang, Y. (2018). Comprehensive evaluation of coal-fired power units using grey relational analysis and a hybrid entropy-based weighting method. *Entropy*, 20(4), 215.

220. Wu, D., Wang, Y., Liu, Y. and Wu, J. (2021). DEA cross-efficiency ranking method considering satisfaction and consensus degree. *International Transactions in Operational Research*, 28(6), 3470-3492.

221. Wu, H. W., Zhen, J. and Zhang, J. (2020). Urban rail transit operation safety evaluation based on an improved CRITIC method and cloud model. *Journal of Rail Transport Planning and Management*, 16, 100206.

222. Wu, J., Chu, J., Sun, J. and Zhu, Q. (2016a). DEA cross-efficiency evaluation based on Pareto improvement. *European Journal of Operational Research*, 248(2), 571-579.

223. Wu, J., Chu, J., Sun, J., Zhu, Q. and Liang, L. (2016b). Extended secondary goal models for weights selection in DEA cross-efficiency evaluation. *Computers and Industrial Engineering*, 93, 143-151.

224. Wu, J., Liang, L., and Yang, F. (2009). Determination of the weights for the ultimate cross efficiency using Shapley value in cooperative game. *Expert Systems with Applications*, 36(1), 872-876.

225. Wu, J., Sun, J. and Liang, L. (2012a). DEA cross-efficiency aggregation method based upon Shannon entropy. *International Journal of Production Research*, 50(23), 6726-6736.

226. Wu, J., Sun, J. and Liang, L. (2012b). Cross efficiency evaluation method based on weight-balanced data envelopment analysis model. *Computers and Industrial Engineering*, 63(2), 513-519.

227. Wu, J., Sun, J. and Liang, L. (2021). Methods and applications of DEA cross-efficiency: Review and future perspectives. *Frontiers of Engineering Management*, 8(2), 199-211.

228. Wu, J., Sun, J., Liang, L. and Zha, Y. (2011). Determination of weights for ultimate cross efficiency using Shannon entropy. *Expert Systems with Applications*, 38(5), 5162-5165.

229. Wu, J., Xu, G., Zhu, Q. and Zhang, C. (2021). Two-stage DEA models with fairness concern: Modelling and computational aspects. *Omega*, 105, 102521.

230. Wu, Y., Chen, Z., and Xia, P. (2018). An extended DEA-based measurement for eco-efficiency from the viewpoint of limited preparation. *Journal of Cleaner Production*, 195, 721-733.

231. Wu, Z., Guan, J. and He, J. (2019). An empirical study on the calculation of minimum wage standard—dynamic combination calculation based on objective weight of CRITIC-Entropy Weight Method. *Modern Econ Sci*, 41(03), 103-117.

232. Yang, F., Ang, S., Xia, Q. and Yang, C. (2012). Ranking DMUs by using interval DEA cross efficiency matrix with acceptability analysis. *European Journal of Operational Research*, 223(2), 483-488.

233. Yang, G.L., Yang, J.B., Liu, W.B. and Li, X.X. (2013). Cross-efficiency aggregation in DEA models using the evidential-reasoning approach. *European Journal of Operational Research*, 231(2), 393-404.

234. Yu, M. M., Abdul Rashid, A. and See, K. F. (2021). Developing an innovation capital index of global airlines using a hierarchical data envelopment analysis approach. *Journal of the Operational Research Society*, 1-16.

235. Yu, M. M. and Chen, L. H. (2020). A meta-frontier network data envelopment analysis approach for the measurement of technological bias with network production structure. *Annals of Operations Research*, 287(1), 495-514.

236. Yu, Y. and Shi, Q. (2014). Two-stage DEA model with additional input in the second stage and part of intermediate products as final output. *Expert Systems with Applications*, 41(15), 6570-6574.

237. Yu, Y., Zhu, W., Shi, Q. and Zhuang, S. (2021). Common set of weights in data envelopment analysis under prospect theory. *Expert Systems*, 38(1), e12602.

238. Zeleny, M. (1982). Multiple Criteria Decision Making. McGraw-Hill, New York.

239. Zha, Y. and Liang, L. (2010). Two-stage cooperation model with input freely distributed among the stages. *European Journal of Operational Research*, 205(2), 332-338.

240. Zhai, D., Shang, J., Yang, F. and Ang, S. (2019). Measuring energy supply chains' efficiency with emission trading: A two-stage frontier-shift data envelopment analysis. *Journal of Cleaner Production*, 210, 1462-1474.

241. Zhang, L. and Chen, K. (2019). Hierarchical network systems: An application to high-technology industry in China. *Omega*, 82, 118-131.

242. Zhang, Q., Xing, H., Liu, F. and Huang, Y. (2014). An enhanced grey relational analysis method for interval-valued intuitionistic fuzzy multiattribute decision making. *Journal of Intelligent and Fuzzy Systems*, 26(1), 317-326.

243. Zhou, P., Ang, B. W., and Poh, K. L. (2008). A survey of data envelopment analysis in energy and environmental studies. *European Journal of Operational Research*, 189(1), 1-18.

244. Zhou, Z., Sun, L., Yang, W., Liu, W. and Ma, C. (2013). A bargaining game model for efficiency decomposition in the centralized model of two-stage systems. *Computers and Industrial Engineering*, 64(1), 103-108.

245. Zhu, J. (1996). Data envelopment analysis with preference structure. *Journal of the Operational Research Society*, 47(1), 136-150.

246. Zhu, J. (2015). Data envelopment analysis: a handbook of models and methods (Vol. 221).

247. Zhu, J. (2020). DEA under big data: data enabled analytics and network data envelopment analysis. *Annals of Operations Research*, 1-23.

248. Zhu, Q., Li, F., Wu, J. and Sun, J. (2021). Cross-efficiency evaluation in data envelopment analysis based on the perspective of fairness utility. *Computers and Industrial Engineering*, 151,106926.

249. Zhu, L. and Xu, B. (2017). Grey Relation Analysis on the High-tech Industry and Economic Growth - Empirical Evidence from Yunnan Province of China. *International Journal of Business and Management*, 12(4), 213-213.

250. Zimmermann, H. J. (2011). Fuzzy set theory—and its applications. Springer Science & Business Media.

Appendix A

Supplement to Chapter 2

A.1 Overall system

TABLE A.1: Cross-efficiency Matrix of the overall system for the proposed model (2.15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	0.936	0.461	0.810	0.654	0.691	0.673	0.647	0.801	0.473	0.859
2	0.155	0.909	0.440	0.615	0.640	0.629	0.666	0.433	0.767	0.077
3	0.467	0.771	1.000	0.772	0.794	0.785	0.772	1.000	0.885	0.224
4	0.738	0.891	0.429	0.894	0.837	0.866	0.891	0.420	0.731	0.850
5	0.287	0.604	0.430	0.677	0.690	0.649	0.689	0.431	0.651	0.208
6	0.941	0.767	1.000	0.955	0.939	1.000	0.964	0.975	0.602	0.738
7	0.462	0.851	0.598	0.955	0.905	0.943	0.955	0.588	0.818	0.263
8	0.524	0.675	1.000	0.706	0.827	0.710	0.715	1.000	0.837	0.326
9	0.472	0.605	0.551	0.581	0.669	0.559	0.588	0.553	0.728	0.362
10	0.738	0.503	0.450	0.661	0.644	0.656	0.653	0.442	0.463	0.844

TABLE A.2: Matrix of relative scores for the overall system for the proposed model (2.15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	0.993	0.000	0.668	0.194	0.171	0.259	0.158	0.657	0.024	1.000
2	0.000	1.000	0.019	0.091	0.000	0.158	0.208	0.022	0.721	0.000
3	0.397	0.691	1.000	0.511	0.516	0.514	0.490	1.000	1.000	0.189
4	0.742	0.959	0.000	0.835	0.658	0.697	0.807	0.000	0.635	0.989
5	0.168	0.319	0.003	0.255	0.168	0.204	0.268	0.019	0.446	0.167
6	1.000	0.683	1.000	1.000	1.000	1.000	1.000	0.957	0.329	0.845
7	0.391	0.869	0.297	0.998	0.887	0.870	0.976	0.289	0.841	0.238
8	0.469	0.478	1.000	0.333	0.627	0.342	0.338	1.000	0.886	0.319
9	0.403	0.320	0.214	0.000	0.097	0.000	0.000	0.229	0.627	0.364
10	0.742	0.093	0.036	0.213	0.013	0.220	0.173	0.037	0.000	0.980
Std Deviation	0.332	0.354	0.446	0.374	0.370	0.331	0.361	0.436	0.347	0.397

TABLE A.3: Symmetric Matrix for the overall system for the proposed model (2.15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	0.538	-0.655	0.383	-0.185	-0.163	-0.129	-0.262	0.386	-0.473	0.508
2	-0.536	0.384	-0.360	-0.368	-0.362	-0.353	-0.282	-0.359	0.250	-0.467
3	-0.438	0.238	0.429	-0.219	0.077	-0.211	-0.192	0.449	0.771	-0.608
4	0.109	0.019	-0.745	0.007	-0.290	-0.018	0.026	-0.763	-0.584	0.350
5	-0.160	0.160	-0.621	-0.138	-0.248	-0.183	-0.105	-0.628	-0.118	-0.020
6	0.339	0.107	0.345	0.607	0.541	0.620	0.582	0.335	-0.050	0.247
7	-0.245	0.577	-0.477	0.382	0.258	0.317	0.437	-0.488	0.187	-0.183
8	-0.380	-0.143	0.454	-0.423	-0.124	-0.417	-0.429	0.480	0.562	-0.530
9	-0.103	-0.567	-0.107	-0.861	-0.760	-0.834	-0.877	-0.092	-0.228	-0.040
10	0.610	-0.578	-0.137	-0.128	-0.303	-0.093	-0.189	-0.149	-0.827	0.737
Conflict	10.265	10.459	10.836	11.326	11.375	11.301	11.290	10.829	10.509	10.005
Information	3.412	3.707	4.828	4.236	4.213	3.743	4.081	4.725	3.642	3.967
Final Weight	0.084	0.091	0.119	0.104	0.104	0.092	0.101	0.117	0.090	0.098

A.2 Stage 1

TABLE A.4: Cross-efficiency Matrix of the stage 1 for the proposed model (2.15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	1.000	0.335	0.319	0.526	0.526	0.535	0.526	0.526	0.266	0.970
2	0.050	0.924	0.412	0.516	0.578	0.541	0.578	0.578	1.000	0.052
3	0.133	0.664	0.889	0.666	0.666	0.663	0.666	0.666	0.800	0.145
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.150	0.572	0.557	0.659	0.676	0.669	0.676	0.676	0.613	0.153
6	0.875	0.660	0.398	0.962	1.000	1.000	1.000	1.000	0.413	0.837
7	0.167	0.830	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.182
8	0.250	0.728	0.697	0.743	0.755	0.749	0.755	0.935	0.770	0.254
9	0.250	0.500	0.437	0.449	0.456	0.452	0.456	0.456	0.499	0.250
10	1.000	0.417	0.464	0.602	0.596	0.605	0.596	0.596	0.383	0.985

TABLE A.5: Matrix of relative scores for the stage 1 for the proposed model (2.15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	1.000	0.000	0.000	0.139	0.128	0.150	0.128	0.128	0.000	0.968
2	0.000	0.886	0.137	0.121	0.223	0.162	0.223	0.223	1.000	0.000
3	0.088	0.495	0.837	0.394	0.386	0.385	0.386	0.386	0.728	0.099
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.105	0.357	0.350	0.380	0.404	0.396	0.404	0.404	0.473	0.107
6	0.868	0.488	0.116	0.932	1.000	1.000	1.000	1.000	0.200	0.828
7	0.123	0.744	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.137
8	0.211	0.592	0.555	0.533	0.550	0.542	0.550	0.880	0.686	0.213
9	0.211	0.248	0.174	0.000	0.000	0.000	0.000	0.000	0.318	0.209
10	1.000	0.124	0.213	0.277	0.257	0.280	0.257	0.257	0.160	0.984
Std Deviation	0.442	0.323	0.383	0.377	0.380	0.381	0.380	0.399	0.379	0.429

TABLE A.6: Symmetric Matrix for the stage 1 for the proposed model (2.15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	0.705	-0.673	-0.416	-0.253	-0.297	-0.257	-0.297	-0.316	-0.663	0.703
2	-0.565	0.206	-0.272	-0.471	-0.418	-0.457	-0.418	-0.418	0.236	-0.571
3	-0.661	0.318	0.318	-0.098	-0.083	-0.107	-0.083	-0.092	0.410	-0.658
4	0.285	-0.078	-0.224	-0.491	-0.502	-0.502	-0.502	-0.413	-0.082	0.290
5	-0.621	0.568	0.370	0.309	0.344	0.313	0.344	0.375	0.531	-0.619
6	0.411	0.087	0.059	0.581	0.580	0.592	0.580	0.620	-0.065	0.411
7	-0.570	0.617	0.525	0.456	0.482	0.455	0.482	0.509	0.580	-0.566
8	-0.652	0.517	0.307	0.160	0.193	0.161	0.193	0.346	0.513	-0.650
9	-0.063	-0.434	-0.521	-0.850	-0.850	-0.851	-0.850	-0.892	-0.313	-0.069
10	0.706	-0.688	-0.389	-0.252	-0.300	-0.258	-0.300	-0.321	-0.669	0.704
Conflict	11.024	9.559	10.242	10.908	10.853	10.910	10.853	10.601	9.520	11.025
Information	4.869	3.092	3.921	4.115	4.123	4.162	4.123	4.231	3.610	4.727
Final Weight	0.119	0.075	0.096	0.100	0.101	0.102	0.101	0.103	0.088	0.115

A.3 Stage 2

TABLE A.7: Cross-efficiency Matrix of the stage 2 for the proposed model (2.15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	0.908	1.000	0.810	1.000	1.000	0.969	0.978	0.801	1.000	0.696
2	0.604	0.886	0.440	0.889	0.734	0.802	0.884	0.433	0.610	0.782
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	0.641	0.735	0.429	0.741	0.679	0.710	0.736	0.420	0.547	0.636
5	0.482	0.684	0.430	0.715	0.709	0.617	0.715	0.431	0.703	0.717
6	1.000	1.000	1.000	0.945	0.868	1.000	0.913	0.975	0.937	0.568
7	0.840	0.886	0.598	0.890	0.812	0.876	0.890	0.588	0.670	0.905
8	1.000	0.571	1.000	0.634	0.936	0.645	0.639	1.000	0.926	0.736
9	0.605	0.903	0.551	1.000	1.000	0.762	1.000	0.553	1.000	1.000
10	0.641	0.795	0.450	0.800	0.718	0.750	0.791	0.442	0.592	0.640

TABLE A.8: Matrix of relative scores for the stage 2 for the proposed model (2.15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	0.822	1.000	0.668	1.000	1.000	0.920	0.940	0.657	1.000	0.297
2	0.236	0.735	0.019	0.696	0.172	0.483	0.680	0.022	0.137	0.494
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	0.307	0.382	0.000	0.293	0.000	0.243	0.269	0.000	0.000	0.158
5	0.000	0.263	0.003	0.221	0.096	0.000	0.212	0.019	0.343	0.345
6	1.000	1.000	1.000	0.850	0.590	1.000	0.758	0.957	0.862	0.000
7	0.691	0.735	0.297	0.699	0.415	0.676	0.696	0.289	0.270	0.779
8	1.000	0.000	1.000	0.000	0.801	0.073	0.000	1.000	0.837	0.388
9	0.237	0.775	0.214	1.000	1.000	0.379	1.000	0.229	1.000	1.000
10	0.307	0.523	0.036	0.453	0.124	0.348	0.420	0.037	0.099	0.165
Std Deviation	0.383	0.343	0.446	0.361	0.410	0.372	0.354	0.436	0.419	0.353

TABLE A.9: Symmetric Matrix for the stage 2 for the proposed model (2.15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	-0.263	0.121	-0.215	0.153	-0.073	-0.023	0.158	-0.219	-0.012	0.142
2	-0.307	0.111	-0.540	-0.030	-0.698	0.048	-0.052	-0.562	-0.768	-0.357
3	0.338	-0.063	0.343	-0.040	0.324	0.068	-0.049	0.344	0.360	-0.022
4	-0.029	0.348	-0.260	0.198	-0.376	0.336	0.160	-0.282	-0.451	-0.417
5	-0.723	-0.125	-0.804	-0.029	-0.444	-0.399	0.000	-0.803	-0.576	0.173
6	0.479	0.310	0.526	0.332	0.516	0.395	0.333	0.527	0.527	0.246
7	-0.145	0.235	-0.380	0.051	-0.573	0.225	0.004	-0.405	-0.615	-0.592
8	0.226	-0.095	0.422	0.056	0.682	-0.061	0.081	0.445	0.704	0.345
9	-0.853	-0.268	-0.911	-0.222	-0.679	-0.528	-0.200	-0.913	-0.719	0.036
10	-0.172	0.254	-0.381	0.119	-0.498	0.205	0.090	-0.403	-0.552	-0.338
Conflict	11.450	9.172	12.201	9.410	11.820	9.733	9.476	12.271	12.103	10.785
Information	4.382	3.143	5.437	3.401	4.848	3.616	3.356	5.354	5.076	3.807
Final Weight	0.103	0.074	0.128	0.080	0.114	0.085	0.079	0.126	0.120	0.090

Appendix B

Supplement to Chapter 3

B.1 Overall system and Grey Relational Analysis

TABLE B.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	—	—	—	—

B.2 Stage 1 and Grey Relational Analysis

TABLE B.2: 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	—	—	—	—

B.3 Stage 2 and Grey Relational Analysis

TABLE B.3: 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	—	—	—	—