UNIVERSITY OF TECHNOLOGY SYDNEY

Performance Measurement of Cloud Service Suppliers and Cloud Supply Chain

by Majid Azadi

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of A/Prof. Farookh Khadeer Hussain as principal supervisor, Dr. Fahimeh Ramezani as co-supervisor, Professor Ali Emrouznejad as external supervisor and Professor Reza Farzipoor Saen as external supervisor

University of Technology Sydney
Faculty of Engineering and Information Technology

February 2021

Certificate of Authorship/Originality

I, Majid Azadi, declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Computer Science / Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature:

Production Note:
Signature removed prior to publication.

Date: 1/02/2021

Dedication

To my beloved parents and siblings for their encouragement which helped my dreams come true.

Acknowledgment

I would like to express my appreciation to Associate Professor Farookh Khadeer Hussain, who supervised this Ph.D. program, for his knowledgeable comments, valuable support, appropriate guidance during difficult times, and beneficial suggestions along the way. I want to thank Farookh for his wonderful scientific and spiritual support during this journey and his willingness to allow my research to follow my interests. I am also grateful to Dr. Fahimeh Ramezani, my co-supervisor, for her knowledgeable suggestions and valuable technical advice throughout my PhD study. I would also like to express my special thanks to Professor Ali Emrouznejad and Professor Reza Farzipoor Saen for their generous support and letting me work beside their amazing research team.

I would like to thank my parents, the first ones who taught me, for their encouragement and support despite the geographical distance. Pursuing a PhD would not have been possible without their love and assistance. I express my gratitude to my siblings, all my friends and colleagues at the University of Technology Sydney for their help and valuable comments during my study.

I also want to express my appreciation to the Faculty of Engineering and Information Technology and the Centre for Artificial Intelligence at the University of Technology Sydney for their support during this research. This research was also supported by the International Research Scholarship (IRS) and UTS President's Scholarship (UTSP) funded by the University of Technology Sydney.

I would like to thank all of my friends especially Dr. Amir Shabani, Dr. Mahdi Saki and Dr. Hadi Hosseini for their great friendship and generous support during my PhD.

Last, but not least, a special thank you to Mrs. Michele Mooney for helping me to identify and correct grammar and syntax problems in my publications.

Abstract

Over the past decade, cloud computing has received much consideration from both practitioners and academics. Nonetheless, the rapid increase in the industrial use-cases of cloud computing and the sharp increase in the number of cloud service providers (CSPs) have resulted in many challenges in performance measurement and the selection of the optimal CSPs according to quality of service (QoS) requirements. To date, there is no reliable approach for the performance assessment of CSPs and quantitative models are rarely used to support such decisions. In addition, the existing approaches to CSP performance measurement suffer from several limitations and drawbacks such as requiring complex calculations, being effort-intensive and being time-consuming. Furthermore, the existing approaches are unable to find slight differences between CSPs in a cloud marketplace owing to the high level of competition. These limitations are major obstacles to applying the existing approaches to assess CSPs. To address these issues in the existing literature, the objective of this study is to develop performance evaluation models based on the Data Envelopment Analysis (DEA) approach that can act as a proxy for many conventional performance assessment problems in cloud computing. The obtained results show that the models proposed in this study are extremely effective in measuring the efficiency of CSPs. Furthermore, the proposed models can deal with different types of data simultaneously in the efficiency measurement of CSPs and cloud supply chains, giving decision makers the ability to estimate the performance of CSPs efficiently.

Keywords: Cloud services, Cloud service providers (CSPs), Cloud supply chain, Data Envelopment Analysis (DEA), Network DEA, Performance measurement and selection.

Contents

List of Figures

List of Tables

Chapter 1 INTRODUCTION

1.1 BACKGROUND

Advances in virtualization techniques and the construction of numerous large commodity data centres around the world have resulted in a new approach to computing referred to as cloud computing, which has become an important topic of research and development (Ghanbari et al. 2012). Cloud computing services have become a new paradigm for delivering computing resources in the last few years. Unlike conventional approaches to the provision of storage, computation and network resources to meet the customer's needs, cloud services provide customers with on-demand services that are available over a network. The major types of cloud computing services are Software as a Service (SaaS), Infrastructure as a Service (IaaS) and Platform as a Service (PaaS). For each of these services, the cloud-based service delivery model offers many benefits including economies of scale, quick bug fixes, potential cost-savings through the "pay-as-you-go" model, and the fast deployment of new features (Atmaca et al. 2015; Elgendy et al. 2018). Due to these benefits, increasingly, business organizations are hosting their applications on cloud infrastructures in order to save huge investment or up-front costs (Kumar & Agarwal 2014).

Over the last few years, the rapid and significant developments in the cloud computing area have provided customers with a wide range of cloud services (Hussain, Chun & Khan 2020). Cloud services are available to customers through networks from data centres operated by cloud service providers (CSPs). By using these services, the customers do not need to provision or deploy their own resources or allocate IT staff to manage the service. Cloud computing services also provide customers with multiple deployment models such as public, private, hybrid, community and distributed infrastructures (Duan 2017). According to Erl (2005), Service-Oriented Architecture (SOA) plays a major role in provisioning cloud services by encapsulating computational resources.

With the increasing technological advances in the area of cloud computing and networking, a wide spectrum of cloud services has been adopted as an integral part in the information systems (IS) of many organizations; consequently, the performance measurement of CSPs has an important influence on the performance of organizations. Cloud service customers need to have a thorough understanding of the performance of CSPs in order to meet their IS requirements. Cloud computing service customers also need effective and powerful approaches to evaluate the performance of CSPs to make correct and appropriate decisions on CSP selection and composition. The performance evaluation of CSPs using appropriate and powerful methods can help customers choose the most appropriate CSP while at the same time, it helps CSPs to identify resources which do not meet the needs of their customers in order to improve their performance in the marketplace.

Because of the numerous business benefits offered by cloud services, many corporations are interested in building applications on the cloud infrastructure to have an agile and flexible business. However, with the growth of cloud services, it has become increasingly difficult to decide which CSP is able to meet their quality of service (QoS) requirements (Choudhury et al. 2012). In fact, a large number of similar services with different sets of features, dissimilar performance levels and highly competitive prices are offered by other CSPs. Consequently, a critical challenge for cloud customers is how to evaluate a CSP's efficiency and identify the optimal CSPs to meet their needs based on their priorities.

One of the major issues with the existing approaches for measuring and benchmarking the performance of CSPs is that none of the approaches can discern the differences between CSPs in a crowded and highly competitive marketplace. Moreover, most existing methods suffer from complex calculations, being effort-intensive, time-consuming, and having ranking irregularities (Azadi et al. 2019; Ghosh, Ghosh & Das 2014; Huang, Hsu & Tzeng 2012; Rajarajeswari & Aramudhan 2015; Sun et al. 2019). In addition, all weights given to quality of service (QoS) criteria by managers and experts for performance assessment and selection of CSPs are subjective measures, which results in an unfair assessment. Moreover, no study addresses the performance of CSPs in the context of an entire supply chain, where multiple services interact to achieve a business objective or goal. Finally, the current methods ignore undesirable outputs, integer-valued data, and stochastic data – such as latency, the number of security certifications, services prices, which can lead to inaccurate results.

To address these shortcomings, this study proposes and develops a number of advanced data envelopment analysis (DEA) models for performance evaluation and the selection of the optimal CSPs. The usefulness and applicability of the proposed models have also been evaluated using a real data set of globally recognized CSPs. Furthermore, to the best of our knowledge, for the first time the concept of cloud supply chain and its structure is defined. A decision support system (DSS) is designed that accurately evaluates the efficiency of multiple cloud service providers in the supply chain.

The rest of this chapter is set out as follows. Sections 1.2, 1.3 and 1.4 contain the research problems, objectives and expected contributions respectively. Section 1.5 describes the research methodology and plan. Section 6 outlines the thesis structure followed by the publications of this study in section 1.7.

1.2 RESEARCH PROBLEMS

The main motivations of this study and the research questions are presented in this section.

(1) CSPs performance evaluation and selection refers to the process of assessing and approving potential CSPs based on QoS requirements (Garg, Versteeg & Buyya 2013). The main objective of this process is to ensure a portfolio of the optimal CSP is available for use. The evaluation and monitoring of CSP performance can ensure contract compliance, reduce costs, mitigate risk and drive continuous improvement (Choudhury et al. 2012). However, the rapid development of cloud computing and the sharp increase in the number of CSPs have resulted in many challenges in the suitability and selection of the optimal CSPs according to the QoS requirements. Although considerable research has been undertaken to solve the problem of evaluating and selecting CSPs, such as the utilization of multiple-criteria decision-making (MCDM) or multiple-criteria decision analysis (MCDA) (Alabool & Mahmood 2016; Aruna & Aramudhan 2016; Ataş & Gungor 2014; Chan & Chieu 2010; Garg, Versteeg & Buyya 2013; Menzel et al. 2015; Rajarajeswari & Aramudhan 2015; Singh & Sidhu 2017), the existing methods need further improvement.

(2) Based on the literature, most CSP evaluation performance and selection problems are solved using MCDM or MCDA methods such as analytic hierarchy process (AHP), grey relational analysis (GRA) and analytic network process (ANP). Nonetheless, these approaches suffer from complex calculations, being effort-intensive, time-consuming, and having subjective weight and ranking irregularities (Sun et al. 2019). Therefore, advanced methods and frameworks for evaluating and selecting CSPs need to be developed and applied to decrease the risk of selecting an inappropriate CSP.

(3) The rapid development of cloud computing and the sharp increase in the number of CSPs have resulted in most CSPs being highly competitive in terms of performance and price for cloud services. Hence, it is difficult for the existing approaches to differentiate between CSPs which have a very similar rank. Furthermore, all the weights given to the QoS criteria by managers and experts for the evaluation and selection of CSPs are subjective measures, which results in an unfair assessment. As a result, more objective approaches with high discrimination power are needed to solve or address performance evaluation and selection problems.

(4) To date, considerable research has been undertaken to solve the problem of evaluating the efficiency of cloud service providers. However, no study addresses the efficiency of providers in the context of an entire supply chain, where multiple services interact to achieve a business objective or goal. Moreover, the current models for performance evaluation and selection problems ignore undesirable factors, integer-valued data, and stochastic data – such as latency, security certifications, and services prices, which can lead to inaccurate results. Furthermore, none of the existing approaches to evaluate and select a CSP is able to provide customers with an optimal CSP composition given their QoS priorities, such as cost or availability in a cloud supply chain.

Based on the aforementioned issues, we identify the following as our research questions:

1. Research Question 1. How can we develop a model with high discrimination power that has the ability to differentiate between CSPs even if they have been given the same rank, or are rated the same using the current methods?

2. Research Question 2. How can we develop a model for the performance evaluation of CSPs with high discrimination power that is more objective compared to other models and can deal with undesirable factors?

3. Research Question 3. How can we effectively evaluate the efficiency of CSPs in a cloud supply chain in separate stages of the chain and overall using an integrated model and based on QoS indicators while providing a composition of CSPs for different services?

4. Research Question 4. How can we develop a model to evaluate the efficiency of the cloud supply chain in the presence of undesirable factors?

5. Research Question 5. How can we develop a model to evaluate the efficiency of a cloud supply chain in the presence of both undesirable factors and integer-valued data simultaneously?

6. Research Question 6. How can we develop a model to evaluate the efficiency of a cloud supply chain in the presence of undesirable factors, integer-valued and stochastic data simultaneously?

7. How can we extend the model developed in response to Research Question 5 to ensure a more objective performance evaluation of the cloud supply chain and which has high discrimination power?

1.3 RESEARCH OBJECTIVE

Based on the research problems, this study has seven primary research objectives as follows:

Research Objective 1: *To rank CSPs and to be able to accurately discriminate between CSPs.*

The first research objective corresponds to research question 1. To achieve this research objective, we propose three new ranking models with different properties based on the extensions of the Enhanced Russell Model (ERM) while considering the distances to two special artificial decision-making units namely ideal decision-making units (IDMU) and anti-ideal DMU (AIDMU). The first model is developed based on the distance to IDMU; the distance to AIDMU is considered to develop the second ranking model, and the third model is designed based on the distances to IDMU and AIDMU. There are several advantages to the proposed ranking models. First, they consider both pessimistic and optimistic scenarios of data envelopment analysis (DEA), so they are more equitable than methods that are based on only one of these scenarios. The second strength of this approach is its discrimination power, enabling it to provide the complete ranking for all CSPs. The proposed method can help customers to choose the most appropriate CSP while at the same time, it helps software developers to identify inefficient CSPs in order to improve their performance in the marketplace. Third, the proposed models can easily be computerized, enabling them to serve as decision-making tools to assist decision makers. Furthermore, we propose a network DEA model to increase discrimination power in the performance measurement of CSPs.

The methods proposed to meet this research objective will be applied to a prototype and a real data set to evaluate and select the optimal CSPs according to QoS requirements. Moreover, to demonstrate the advantages of the proposed method, it is compared to other similar methods.

Research Objective 2: *To develop a model that considers both undesirable factors (outputs) and weight restrictions with a high discrimination to evaluate and select the optimal CSPs.*

The second research objective corresponds to research question 2. In the performance evaluation and benchmarking of CSPs, there are some undesirable variables such as latency that can affect the evaluation results. Moreover, there may be problems with the decision makers’ value judgments in the evaluation and benchmarking problems of CSPs (Kumar & Agarwal 2014). To be more precise, in the performance evaluation of CSPs, different weights are given by different decision makers to performance evaluation criteria, which results in incorrect evaluation (Halabi & Bellaiche 2017; Kumar & Agarwal 2014; Saen 2010b). Value judgments can reflect known information about how the factors used by the CSPs behave, and/or "accepted" beliefs or preferences as to the relative worth of inputs, outputs or even CSPs (Saen 2009, 2010b). Furthermore, as previously mentioned, the existing approaches for the performance measurement and benchmarking of CSPs are unable to discern differences between CSPs in the crowded and highly competitive cloud marketplace. Therefore, there is a need to develop a model to deal with both undesirable factors (outputs) and decision makers' value judgments with a high discrimination power to evaluate and select the optimal CSPs. To achieve this research objective, we propose a new model based on a super-efficiency DEA model for the benchmarking and selection of CSPs in the presence of both undesirable and weight restrictions.

To achieve this research objective the following steps are proposed:

Step 1: Provide a list of potential CSPs in the market.

Step 2: Determine inputs, outputs and undesirable factors of the model based on QoS requirements such as price, availability, and latency.

Step 3: Consider decision makers' comments in the performance evaluation and selection of CSPs based on the QoS criteria.

Step 4) Develop the new model in the presence of both undesirable factors (outputs) and weight restrictions.

Step 5) Validate the new model using a real data set.

The working of our method is explained in detail in Chapter 4.

The optimal CSPs have the highest efficiency scores compared to other CSPs in the presence of weight restrictions and undesirable factors (outputs) such as latency.

Research Objective 3: *To apply the network DEA model for the performance evaluation of the cloud supply chain*

This objective corresponds to research question 3. We apply the network DEA model to evaluate the efficiency of a cloud supply chain provided by CSPs where multiple cloud computing services interact to achieve a business objective or goal. A cloud supply chain network consists of IaaS, PaaS and SaaS stages respectively. By using the network DEA model, the efficiencies for each sub-unit in addition to the efficiency for the entire cloud supply chain are computed. In the cloud supply chain model, each CSP is considered as a decision-making unit (DMU) with its own set of inputs $(x_1 \dots x_n)$ and outputs $(y_1 \dots y_n)$. Each cloud computing service (IaaS, PaaS and SaaS) is assigned a weight by the user based on its importance in determining cloud supply chain efficiency $(w_1 \dots w_n)$. In addition, there is at least one link between two sub-units in the model. These links are classified as intermediate.

To achieve this research objective, the following steps are proposed:

Step 1: Construct the cloud supply chain.

Step 2: Determine the decision-making variables (i.e., the input, the intermediate, and the output variables) based on QoS criteria.

Step 3: Apply a rigorous network DEA model to the performance evaluation of cloud supply chains.

Step 4: Determine the scope of the problem, i.e., the number of stages (services) in the supply chain that need to be considered given the customer's priorities.

Step 5: Select and apply the relevant and rigorous network DEA model based on the number of stages and the type of decision-making variables.

Step 6: Analyse the results of the evaluation.

Step 7: Recommend the highest-ranking CSPs.

The working of our method is explained in detail in Chapter 5.

Research Objective 4: *To develop a model to evaluate the efficiency of a cloud supply chain in the presence of undesirable factors*

This research objective corresponds to research question 4. In a cloud supply chain consisting of IaaS, PaaS and SaaS, there might be some undesirable factors. These factors can exist in each stage of the chain as inputs, outputs or intermediates. In addition to this, undesirable factors may change the performance evaluation results of a cloud supply chain. Thus, we need to develop a model for the performance measurement of cloud supply chains so that it can deal with undesirable factors in the whole of the chain. Moreover, the effect of undesirable factors on the performance measurement results of a cloud supply chain needs to be analysed and compared with those identified in research objective 3.

The working of our method is explained in Chapter 5.

Research Objective 5: *To develop a model to evaluate the efficiency of cloud supply chains in the presence of both undesirable factors and integer-valued data.*

This research objective corresponds to research question 5. In the performance measurement of cloud supply chains, as well as undesirable factors, there might be some QoS criteria such as CPU, memory and storage that only take integer values. The efficiency evaluation results of a cloud supply chain also can be inaccurate if this condition is not considered in the performance evaluation process. One way to deal with integer-valued variables is to round the obtained results from performance evaluation models to the nearest integer values. However, it has been proven that this simple rounding approach can result in misleading evaluation results (Chen et al. 2012). Therefore, to carry out the performance evaluation of cloud supply chains, we need to develop a performance evaluation model that can deal with both undesirable factors and integer-valued variables.

Table 1.1 The factors used for the efficiency measurement of CSPs

Factors	Factor type	Functional/non-functional
Central Processing Unit (CPU)	Integer	Functional
Memory	Real	Functional
Storage	Real	Functional
Data transfer	Real	Functional
Latency	Real	Functional
Availability	Ratio	Non-functional
The number of security certifications	Integer	Non-functional
Price	Real	Non-functional

The working of our method is explained in detail in Chapter 5.

Research Objective 6: *To develop a comprehensive model to evaluate the efficiency of cloud supply chains in the presence of undesirable factors, integer-valued data and stochastic data.*

This research objective corresponds to research question 6. As well as undesirable factors and integer-valued data, many observations in the real world are stochastic. Consequently, the resulting efficiencies are stochastic as well (Kao & Liu 2009). A case in point is the stock market, in which some observations, such as price, change dramatically due to uncertainty in the environment. To deal with stochastic data in the performance evaluation of cloud supply chains, we develop a comprehensive model considering undesirable factors, integer-valued data and stochastic data.

To achieve this research objective, the following steps are proposed:

Step 1: Begin.

Step 2: Construct the cloud supply chain.

Step 3: Determine the decision-making variables (i.e., the input, the intermediate and the output variables).

Step 4: Build three separate DEA ranking models to consider the undesirable, integer, and stochastic variables.

Step 5: Integrate the three models from the previous step into one model that considers all three variables.

Step 6: Identify the three different variable types.

Step 7: Determine the scope of the problem, i.e., the number of stages (services) in the supply chain that need to be considered given the customer's priorities.

Step 8: Select the relevant ranking models based on the number of stages and the type of decision-making variables.

Step 9: Analyse the results of the evaluation.

Step 10: Recommend the highest-ranking cloud supply chain provided by CSPs.

The working of our method is explained in detail in Chapter 5.

Research Objective 7: *To extend the model developed in response to objective 5 to evaluate the cloud supply chain more objectively and with higher discrimination power.*

This research objective corresponds to research question 7. In the performance evaluation of the model related to research objective 5, the performance of each cloud supply chain is compared with virtual ones. To undertake a more accurate evaluation, we first propose a new model that compares the performance of each cloud supply chain with actual ones using the free disposal hull (FDH) model. Nonetheless, in such conditions, the problem that may occur is that the discrimination power of the proposed model for the performance measurement of cloud supply chains decreases. To address this issue, we develop the proposed model using the super efficiency concept. By doing so, not only will we have a more objective model for the performance evaluation of cloud supply chains but also the model will benefit from a high discrimination power to identify inefficient resources in each stage of the cloud supply chain. In addition, it provides a better composition of cloud service providers.

The working of our method is explained in detail in Chapter 6.

Objective 1

- **To rank CSPs and to be able to accurately discriminate between CSPs:**
  - Develop a new model to compute the distance of each DMU to IDMU.
  - Develop a new model to compute the distance of each DMU to AIDMU.
  - Develop a model based on mixed aspects of IDMU and AIDMU.
  - Propose a network DEA model to increase discrimination power in performance measurement of CSPs.
  - Evaluate and test the developed methods.
  - Benchmark our proposed and developed methods against other similar methods in the literature.

Objective 2

- **To develop a model to consider both undesirable factors (outputs) and weight restrictions with a high discrimination to evaluate and select the optimal CSPsl CSPs:**
  - Propose a a new super-efficiency Russell directional distance function DEA model by considering undesirable outputs.
  - Develop the proposed model by considering decision makers' subjective judgements.
  - Evaluate and test the developed methods.
  - Benchmark the proposed model against other similar models in the literature.

Objective 3

- **To propose a network DEA model for performance evaluation of cloud supply chain:**
- Create a network structure of cloud services.
- Identify inputs, intermediates and outputs of each stage based on QoS indicators.
- Model and measure the cloud supply chain efficiency.

Objective 4

- **To develop a model to evaluate the efficiency of cloud supply chain in the presence of undesirable factors:**
  - Identify undesirable factors in the cloud supply chain as inputs, outputs or intermediates.
  - Model undesirable factors in the cloud supply chain.
  - Measure the efficiency of the cloud supply chain by considering undesirable factors.
  - Evaluate and test the developed methods.

Objective 5

- **To develop a model to evaluate the efficiency of cloud supply chain in the presence of both undesirable factors and integer-valued data:**
  - Extend concepts and definitions of network DEA to make them consistent with integer-valued data.
  - Model cloud supply chain efficiency with respect to undesirable factors and integer-valued data.
  - Evaluate the developed method and computing efficiency score of overall and division efficiency of cloud supply chain.

Objective 6

- **To develop a comprehensive model to evaluate the efficiency of cloud supply chain in the presence of undesirable factors, integer-valued data and stochastic data:**
- Extend the concepts and definitions of network DEA to make them consistent with chance constrained programming.
- Model and measure the cloud supply chain efficiency with respect to undesirable factors, integer-valued and stochastic data.
- Evaluate and test the developed methods.
- Benchmark the proposed model against other similar models in the literature in response to research objective 3-5.

Objective 7

- **To develop the model mentioned in research objective 5 used for performance evaluation of cloud supply chain while is more objective than it and benefits from high discrimination power:**
- Develop the proposed model in research objective 5 based on free disposal hull concept.
- Incorporate super-efficiency DEA model in the developed model.
- Evaluate and test the developed methods.
- Benchmark the proposed model against other similar models in the literature.

Figure 1.1 Research Objectives

1.4 RESEARCH CONTRIBUTIONS

The aim of this study is to develop novel performance measurement models based on the data envelopment analysis (DEA) approach to the efficiency evaluation and selection of CSPs. On the basis of the developed DEA-based models, an efficiency evaluation framework is presented which is applicable for the evaluation and selection of CSPs. The proposed models assist cloud service customers to mitigate the risk of CSP selection under different conditions and suggest an optimal cloud service composition to cloud service customers.

Based on the research objectives and goals, this study contributes the following research innovations:

1) **Develop a performance evaluation system for CSPs:** This study proposes a performance measurement system to evaluate and select the optimal CSPs based on one of the rigorous Operations Research techniques. The approach applied in this system can deal with conflicting criteria in QoS such as price and quality or security used for CSP selection. In addition, the models proposed in this study can deal with different types of data, integer-valued and stochastic, in the performance measurement and selection of CSPs.
2) **Develop novel models for performance measurement and selection of the optimal CSPs:** The existing methods for assessing CSPs and cloud service performance suffer from many problems such as being effort-intensive, being time-consuming, having subjective weights, and ranking irregularities (Huang, Hsu & Tzeng 2012) (Huang, Hsu & Tzeng (2012); Kumar & Agarwal (2014);(Supriya, Sangeeta & Patra 2016). To obviate a number of barriers in the performance measurement and selection of CSPs, several novel models are developed. Some of the advantages of these models are as follows:

- The proposed models do not demand weights from the decision makers.
- The proposed models consider both undesirable factors and weight restrictions simultaneously in the performance evaluation and selection of CSPs.

- The proposed models are more objective than other approaches for the performance measurement and benchmarking of CSPs.
- For the first time, these models are proposed and applied for the evaluation and benchmarking of CSPs.
- The proposed models can easily be computerized, enabling them to serve as a decision-making tool to assist decision makers.

3) **Define and design a cloud supply chain system:** In this research for the first time, the concept of a cloud supply chain system is defined and the related structure is designed based on reasonable interactions in cloud computing environment where multiple services interact together to achieve a business objective or goal.
4) **Evaluate the performance of cloud supply chain divisionally[1] and overall:** Cloud supply chain activities include providing computing infrastructure, software development platforms, and software to the end customer. In a cloud supply chain, IaaS is often provided to PaaS suppliers; PaaS suppliers deliver their services to SaaS suppliers; and all services can be delivered to cloud service customers. In this study for the first time, the performance of a cloud supply chain is evaluated using an advanced performance evaluation model. The proposed model can evaluate the efficiency of the chain divisionally and overall, and it can also deal with all types of variables that might be involved in the performance evaluation of a cloud supply chain. In addition to this, the proposed model is suitable for uncertain conditions, which can help managers and decision makers to make the right decisions. Furthermore, the model proposed in this study is able to provide customers with an optimal CSP composition given their QoS priorities, which has not been considered in the literature to date.
5) **Increasing customers' satisfaction and managerial implications:** The models proposed in this study can increase clouds customers' satisfaction by allowing them to make an optimal decision with respect to their objectives, constraints and preferences. Moreover, given that the initial investment in cloud computing services can be both costly and time-consuming, the performance measurement techniques proposed in this research can serve as appropriate decision support system tools for

[1] Each of the cloud supply chain stages is considered as a division.

managers and decision makers. Finally, the models proposed in this study can be applied with minor modifications to other problems related to performance measurement in different areas including health care, supply chain management, banking and education.

1.5 RESEARCH METHODOLOGY

Research methodology is the "collection of problem solving methods governed by a set of principles and a common philosophy for solving targeted problems" (Gallupe 2007). Several research methodologies such as case studies, field studies, design research, field experiments, laboratory experiments, surveys, and action research have been proposed and applied in the domain of information systems. The methodology of this research is planned according to the practice of design research (Niu, Lu & Zhang 2009), which has been proposed and applied in information systems.

1.5.1 DESIGN SCIENCE RESEARCH METHODOLOGY

As shown in Figure 2, the design research methodology includes five stages as follows (Niu, Lu & Zhang 2009):

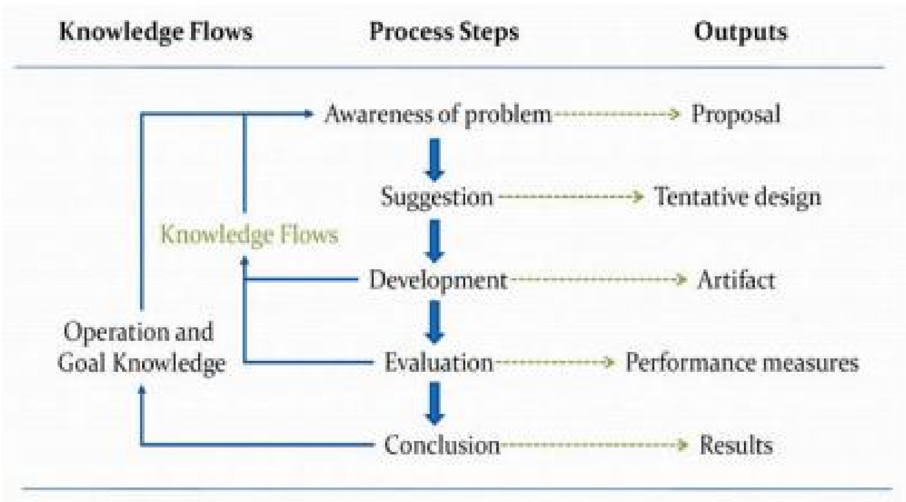

Figure 1.2 The design research methodology

I. **Awareness of the problem**: In this first step, the limitations of the existing applications are analysed and significant research problems are acknowledged. The research problems reflect a gap between the existing applications and the expected status. Research problems can be identified from different sources: industry experiences, observations on practical applications and literature reviews. A clear definition of the research problem provides a focus for the research throughout the development process. The output of this phase is a research proposal for new research effort.

II. **Suggestion**: This phase follows the identification of research problems where a tentative design is suggested. The tentative design describes what the prospective artefacts will be and how they can be developed. Suggestion is a creative process during which new concepts, models and functions of artefacts are demonstrated. The tentative design resulting from this step is usually one part of the research proposal.

III. **Development**: This phase considers the implementation of the suggested tentative design artefacts. The techniques for implementation will be based on the artefact to be constructed. The implementation itself can be simple and does not need to involve novelty; the novelty is primarily in the design not the construction of the artefact. The development process is often an iterative process in which an initial prototype is first built and then evolves as the researcher gains a deeper comprehension of research problems. Thus, the output of the suggestion step is also the feedback of the first step, whereby the research proposal can be revised. This step includes the following sub-steps to create the prototype (Niu, Lu & Zhang 2009): a) planning, b) analysis, c) design, d) development, e) testing, f) implementation, and g) maintenance.

IV. **Evaluation**: This phase considers the evaluation of the implemented artefacts. The artefacts' performance can be evaluated according to criteria defined in the research proposal and the suggested design. The evaluation results, which might not meet expectations, are fed back to the first two steps. Accordingly, the proposal and design might be revised and the artefacts might be improved.

V. **Conclusion**: This is the final phase of a design research effort. Typically, it is the result of satisfaction with the evaluation results of the developed artefacts. However, there are still deviations in the behaviour between the suggested proposal and the artefacts that are actually developed. A design research effort concludes as long as the developed artefacts are considered ‘good enough’ wherein the anomalous behaviour may well serve as the subject of further research.

1.5.2 RESEARCH PLAN

With respect to the designed general methodology of research, the research plan of this study consists of the following steps:

Step 1. Select a topic

The choice of a research topic can arise from personal interest, from observation, or from the literature describing previous theory and research in the area, from social concern or as the outcome of some currently popular issues. The topic of this research was chosen from the previous literature and research and also the author’s observation and experience in the process industry.

Step 2. Review the literature

Irrespective of the reason for choosing a topic, a literature review of the previous research in the topic area is an essential component of any research process. To undertake a comprehensive literature review on the topic, a large number of journal articles were searched, selected and categorized. Then, the existing literature in the related areas was retrieved and critically reviewed. The results of the literature review help us to identify a number of key research gaps in the cloud computing area.

Step 3. Finalize research problems

The identification of the research gaps leads to the definition of the specific research questions for this current study. The research questions are directly addressed in this research study. As the research questions grew clearer and more definite, more studies in the literature closely related to the research questions were reviewed. The existing work is compared to the desirable expectations and the gaps and limitations are identified. Based on the work done so far, the research questions and sub-questions are identified.

Step 4. Design a performance measurement framework

After defining the research problems, the need to design a performance measurement framework for the evaluation and selection CSPs is recognized. To develop a CSP evaluation and selection framework to cover the research problems in this study, the following specifications are considered:

1. There are a number of CSPs providing cloud computing services in a highly competitive market.
2. There are different types of data in the performance evaluation and selection of appropriate CSPs.
3. The existence of decision makers' subjective judgments in CSPs evaluation and selection.
4. The existence of uncertain environments in the cloud computing market.
5. There are multiple cloud computing services which interact to achieve a business objective or goal.

For step 3 (to finalize research problems) and the specifications in step 4 (to design a performance measurement framework), a number of reliable performance evaluation models need to be developed under different conditions and scenarios so that each of them can meet the customers' expectations of receiving high quality cloud computing services with respect to their objectives, constraints and preferences. The defined goals and expected functions in the proposed models are as follows:

1. Increase the power discrimination in CSP performance evaluation.
2. Minimize decision makers' subjective judgments in CSP performance evaluation and the selection process.
3. Minimize the impact of undesirable factors, integer-valued and stochastic data in CSP performance measurement.
4. Evaluate a chain of cloud services provided by CSPs.
5. Deal with all types of influential data in evaluating cloud supply chains using a unified model.

Step 5. Determine performance evaluation requirements of CSPs

To evaluate the efficiency of CSPs, the following items are required:

1. Prepare a list of potential CSPs.

2. Determine the QoS indicators.
3. Determine the scope of the problem.
4. Identify the type of decision-making variables.
5. Select an appropriate approach for performance evaluation.
6. Develop the performance evaluation model with respect to the objectives, constraints and decision variables type.

Step 6. Develop an enhanced Russell Model (ERM) to increase discrimination power between CSPs

We develop an enhanced Russell Model (ERM) to increase the discrimination power in capturing and measuring the performance measurement of the CSPs. The proposed model has the ability to rank CSPs in an optimal way where slight differences between CSPs are considered/recognised. The new ERM considers ideal DMU, anti-ideal DMU and mixed ideal and anti-ideal DMU aspects at the same time in the performance evaluation process. Moreover, this method is based on both pessimistic and optimistic viewpoints, which leads to a more equitable performance evaluation. Therefore, this method enables decision makers to include some of their preferences in the ranking process.

Step 7. Develop a new super-efficiency Russell directional distance function model[2] for CSP performance evaluation and selection in the presence of undesirable factors and weight restrictions

To evaluate and select the optimal CSPs, a rigorous model based on the super-efficiency model is developed. The developed model can not only rank CSPs with a high discrimination power, it also can deal with both undesirable factors and weight restrictions in the performance evaluation process. In addition, the proposed model can minimize the impact of decision makers' subjective judgments in the CSP performance measurement process.

[2] Super-efficiency Russell directional distance function is one of DEA models to increase discrimination power in performance measurement of decision-making units (DMUs).

Step 8. Propose a rigorous model for performance evaluation of cloud supply chains

When a system consists of several components operating interdependently, ignoring the operations within a component may result in misleading efficiency measurements (Kao 2016). Hence, the operations of the components need to be considered when measuring performance in a network structure. To evaluate the efficiency of a cloud supply chain, we propose a rigorous model which is able to consider the internal structure of the chain in the evaluation process. To do so, we propose a two-stage network DEA model in this study.

Step 9. Develop a chance-constrained two-stage network DEA model for the efficiency evaluation of cloud supply chains in the presence of undesirable factors, integer-valued data and stochastic data

Owing to the presence of different types of data in cloud supply chain efficiency evaluation, the model proposed in the previous section needs to be modified and developed. The most common type of data in the efficiency evaluation of cloud supply chains is undesirable data. Therefore, we first develop the two-stage network DEA model in the presence of undesirable data. Another important data type in the efficiency evaluation of a cloud supply chain is integer-valued data such as CPUs. We then develop the model by deriving a DEA production possibility set (PPS) that satisfies the minimum extrapolation principle under our refined set of axioms. In addition to do this, a mixed integer linear programming formula for computing the efficiency scores of the overall and divisional efficiency of cloud supply chains is presented. As well as undesirable factors and integer-valued data, many observations in the real world are imprecise/stochastic. Consequently, the resulting efficiencies are stochastic as well (Kao & Liu 2009). A case in point is the stock market, in which some observations, such as price, change dramatically due to uncertainty in the environment. Thus, in this step, we develop the two-stage network DEA model using chance-constrained programming (CCP) to deal with stochastic data in the performance evaluation of cloud supply chains. Finally, the deterministic equivalent of the proposed model is presented using a quadratic program.

Step 10. Develop a performance measurement model using a non-convex technology and super-efficiency technique

In this stage, in order to obtain more accurate efficiency results in the performance evaluation of cloud supply chains, first we develop the proposed model in step 8 using non-convex technology. The developed model uses the free disposal hull (FDH) technique which computes the efficiency of cloud supply chains based on actual observations of performance. Then, to distinguish the performance evaluation results of the cloud supply chain, we develop the model using the super-efficiency technique. Compared to the model developed in step 8, the new model is more objective, so it benefits from high discrimination power. In addition, the developed model can provide a more optimal composition of CSPs for cloud service customers.

Step 11. Data gathering

In this step, we need a data set to demonstrate the applicability of the proposed models. The data set for this research is gathered from different resources such as reports, cloud computing experts, sales employees and websites. The data set is related to the QoS data of leading companies such as Amazon Web Services (AWS), Microsoft Azure, and IBM SoftLayer.

Step 12. Test, implement and evaluate the proposed models

In this step, the proposed models are evaluated using the data set. Furthermore, where applicable, the developed models will be benchmarked against other similar models in the literature. Based on the obtained results, our proposed methodology may have unexpected results so we should use feedback and revise our methodology for revision to reach suitable results. This step is part of step 4 (Evaluation) of the design science research methodology framework.

Step 13. Write up the thesis

Writing up the PhD thesis concludes the research.

1.6. THESIS STRUCTURE

This thesis comprises 7 chapters as shown in Figure 1.3. Chapter 1 presents the research background, challenges, objectives, contributions, methodology, and the structure of the thesis. Chapter 2 provides the preliminary information on cloud computing, such as cloud architecture and services, physical topology, resources, middleware and components, and also reviews the literature related to the performance evaluation approaches of CSPs, DEA and network DEA models. Chapter 3 presents the developed models for the performance measurement of CSPs. Chapters 4, 5 and 6 present novel models for the performance evaluation of cloud supply chains. In should be noted that in Chapters 4, 5 and 6 we respectively describe the testing, evaluation and implementation of the proposed models using some numerical examples as well as the real data set. Chapter 7 presents the conclusion and future research directions of this research.

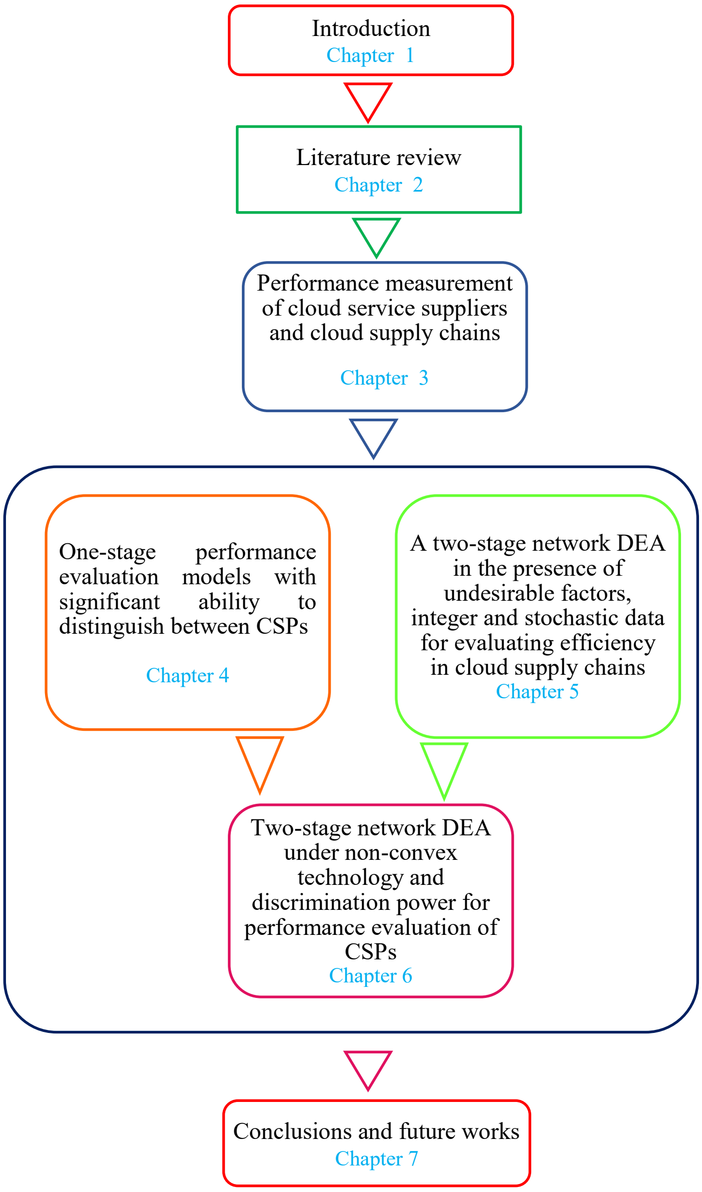

Figure 1.3 Thesis structure

1.7. PUBLICATIONS RELATED TO THIS RESEARCH

Some chapters of this thesis are based on articles that have been submitted and accepted by peer-reviewed journals during my Ph.D. candidature. The details of the articles are as follows:

PEER-REVIEWED INTERNATIONAL JOURNAL PAPERS

1. Azadi, M., Izadikhah, M., Ramezani, F., and Khadeer Hussain, F. ‘A mixed ideal and anti-ideal DEA model: An application to evaluate cloud service providers’, *IMA Journal of Management Mathematics* (JCR **Tier Q1** Journal, Impact Factor: 1.277, 5 year Impact Factor, 1.218, **Accepted**).

2. Azadi, M., Emrouznejad, A. Ramezani, F., and Khadeer Hussain, F. ‘Efficiency measurement of cloud service providers using network data envelopment analysis’, *IEEE Transactions on Cloud Computing*, (JCR **Tier Q1** Journal, Impact Factor: 5.967, Article Influence Score, 4.349, **Accepted**).
3. Azadi, M., Farzipoor Saen., R., and Khadeer Hussain, F. ‘Benchmarking cloud service providers: An extension and application of data envelopment analysis’, Annals of Operations Research (JCR **Tire Q1** Journal, Impact Factor: 2.284, **Submitted and under peer review**).
4. Azadi, M., Toloo, M., Ramezani, F., and Khadeer Hussain, F. Evaluating efficiency in cloud supply chains: a two-stage network data envelopment analysis, *IEEE Transactions on Services Computing*, (JCR **Tire Q1** Journal, Impact Factor: 5.707, Article Influence Score, 1.331, **Submitted and under peer review**).
5. Azadi, M., Farzipoor Saen., R., and Khadeer Hussain, F. ‘Developing a novel SBM network DEA model under non-convex technology with both integer and undesirable outputs: A case study in selecting cloud service providers’, *Expert Systems with Applications: An International Journal*, JCR **Tire Q1** Journal, Impact Factor: 4.292, 5 year Impact Factor, 4.577, (**Submitted and under peer review**).

Chapter 2 LITERATURE REVIEW

2.1 INTRODUCTION

To obtain a better understanding of this thesis, this chapter explains important background and preliminary information regarding cloud computing, cloud layer architecture and services, service-level agreement and quality of service in sections 2.2 to 2.4. Sections 2.5 and 2.6 discuss cloud service providers (CSPs) and the performance measurement of CSPs, respectively. Sections 2.7 to 2.16 review the literature on data envelopment analysis, the implementation of DEA in the cloud environment, discrimination power in data envelopment analysis, undesirable outputs, integer-valued data, chance-constrained DEA, network DEA, free disposal hull (FDH), supply chain and quality of service. Section 2.17 discusses the research gaps in the literature.

2.2 CLOUD COMPUTING

The term 'cloud' was introduced for the first time in the area of information and communication technologies (ICT) early in the 1990s when virtual private network (VPN) services were established for data communications (Jadeja & Modi 2012). Ranjan, Benatallah & Wang (2011) discussed a cloud as a network of data centres over the entire globe, where each centre consists of thousands of computers working together that can perform the functions of software on a personal or business computer by providing user access to platforms, powerful applications and services delivered over the Internet.

Cloud computing provides a network-based environment to consumers, which paves the way for shared resources and calculations irrespective of location (Subramanian & Jeyaraj 2018). According to a definition provided by The National Institute of Standards and Technology (NIST), cloud computing is "a template for providing suitable and when needed access to the internet, to a collective pool of programmable grids, storage, servers, software, and amenities that can be rapidly emancipated, with little communication and supervision from the provider" (Mell & Grance 2011). Therefore, cloud computing provides the users with large pools of resources in a transparent way as well as a mechanism to manage the resources so that the users can access it ubiquitously and without incurring an unnecessary performance overhead. Mahmoud & Xia (2019) and Menzel et al. (2015) summarized the compelling features of cloud computing as follows:

1. *No up-front investment.* Cloud computing uses a pay-as-you-go pricing model. A CSP does not need to invest in the infrastructure to start gaining a benefit from cloud computing. It simply rents resources from the cloud based on its needs and pays for the usage.
2. *Lowering operating cost.* Resources in a cloud environment can be quickly allocated and de-allocated on demand. Thus, a CSP no longer needs to provision capacities based on the peak load. This provides huge savings since resources can be released to save on operating costs when service demand is low.
3. *Highly scalable.* Infrastructure providers pool a large amount of resources from data centres and make them easily accessible. A CSP can easily expand its service to a large scale in order to handle a rapid increase in service demands (e.g., flash-crowd effect). This model sometimes is called surge computing (Armbrust et al. 2009).
4. *Easy access.* Services hosted in the cloud are generally web-based. Therefore, they are easily accessible through a variety of devices with Internet connections. These devices not only include desktop and laptop computers, but also cell phones and PDAs.
5. *Reducing business risks and maintenance expenses.* By outsourcing the service infrastructure to the cloud, a CSP shifts its business risks (such as hardware failures) to infrastructure providers, who often have better expertise and are better equipped to manage these risks. Furthermore, a CSP can cut down the hardware maintenance and the associated staff training costs.

In a cloud environment, users can access the services provided by CSPs only using an Internet connection. Some examples of cloud computing are social networking services, online backup services, and personal data services. In addition to this, cloud computing includes online applications, such as those which are offered by Microsoft and Google online services. Furthermore, hardware services such as mirrored websites, redundant servers and Internet-based clusters are other examples of CSPs (Devi, Gupta & Choudhary 2014).

A private cloud is cloud infrastructure operated solely for a single organization, whether managed internally or by a third-party, and is hosted either internally or externally (Mell & Grance 2011). A private cloud project needs important engagement for virtualizing the computing environment and needs the organization to re-assess decisions regarding existing resources. Although it is able to improve business, every step in the project raises security issues that should be tackled to prevent serious vulnerabilities (Haff 2009). Unlike in a public cloud (such as IBM, Google Cloud, Oracle, Microsoft Azure and Amazon Web Services) where several layers may be offered by multiple providers, in a private cloud the entire stack is controlled by a single provider and so it has access and control over the various infrastructure, applications, and middlewares simultaneously (Ghanbari et al. 2012). A hybrid cloud is a cloud computing environment which uses a combination of on-premises, private cloud, community cloud and public cloud services with orchestration between public and private platforms. A hybrid cloud provides businesses with greater flexibility and more data deployment options by allowing workloads to move between private and public clouds as computing needs and costs change (Varia 2008). The public and private clouds in a hybrid cloud arrangement are distinct and independent elements. This allows organizations to store protected or privileged data on a private cloud while retaining the ability to leverage computational resources from the public cloud to run applications that rely on this data. This keeps data more secure because sensitive data are not stored long-term on the public cloud component (Moltó, Caballer & De Alfonso 2016).

Figure 2.1 illustrates the various cloud deployment models.

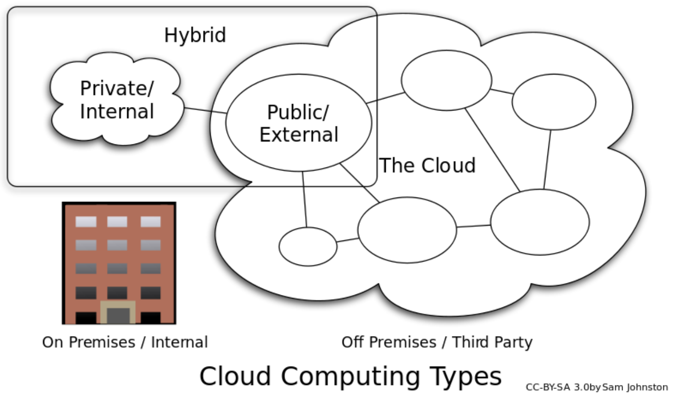

Figure 2.1 Deployment models types

2.3 CLOUDS LAYERS ARCHITECTURE AND SERVICES

As Ramezani (2016) discussed, the cloud is viewed as a layered architecture where services of a higher layer can be composed of services of the underlying layer. The reference model proposed by (Buyya, Pandey & Vecchiola 2009) explains the role of each layer in an integrated architecture. A core middleware manages physical resources and virtual machines (VM) are deployed on top of them. Furthermore, it provides the required features (e.g., accounting, billing, service level agreement (SLA) management, QoS negotiation and execution management) by offering multi-tenant pay-as-you-go services. Cloud development environments are built on top of infrastructure services for offering application development and deployment capabilities; at this level, various programming models, libraries, application programming interfaces (APIs), and mashup editors enable the creation of a range of business, Web, and scientific applications. Once deployed in the cloud, these applications can be consumed by end users (Buyya, Broberg & Goscinski 2011).

According to service-oriented architecture, cloud-computing providers provide their customer with three standard cloud computing services. These are Infrastructure as a Service, Platform as a Service, and Software as a Service. Figure 2.2 shows the cloud computing services.

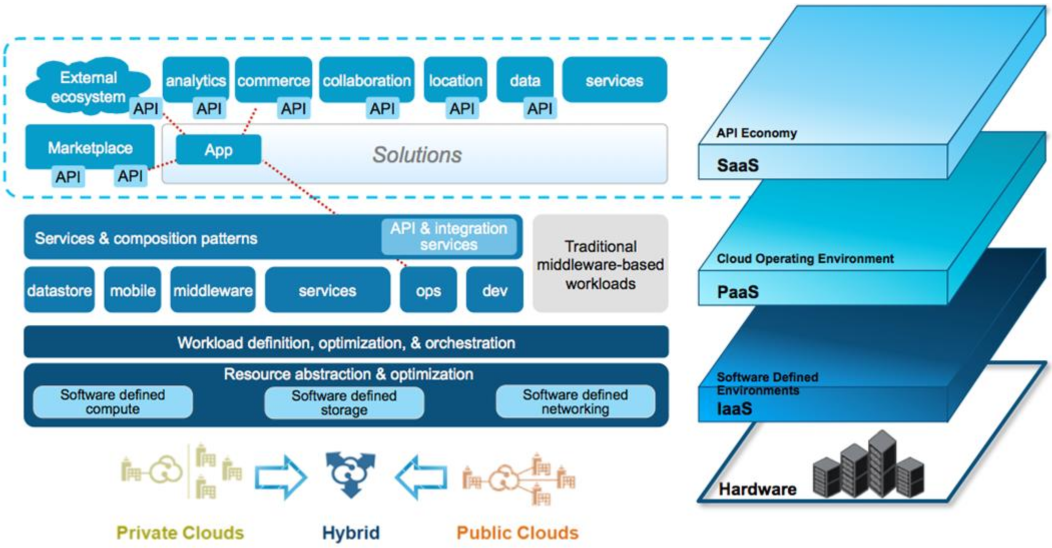

Figure 2.2 Cloud computing services (Diaz & Ferris 2013).

2.3.1 INFRASTRUCTURE AS A SERVICE

Infrastructure as a service offers on-demand virtual resources, such as processing, storage, or network infrastructure and other fundamental computing resources. A cloud infrastructure enables the on-demand provisioning of servers running several choices of operating systems and customised software (Ramezani 2016). Infrastructure services are considered to be the bottom layer of a cloud computing system (Buyya, Broberg & Goscinski 2011), providing customers with a choice of servers, operating systems, and a customized software stack. Although customers do not manage or control the underlying infrastructure, they do control the operating systems, storage, and deployed applications. They may also have limited control of selected networking components (e.g., host firewalls). In short, IaaS focuses on operations. Amazon EC2 is a good example of an infrastructure as a service (Ramezani 2016).

2.3.2 PLATFORM AS A SERVICE

Beyond infrastructure as a service, the next category of cloud services offers a higher level of abstraction for developing cloud-based applications–i.e., an environment where developers can create and deploy applications using programming languages, libraries, services, and tools. These types of services are known as platform as a service. Here, it is not necessary for a platform-as-a-service customer to know how many processors or how much memory an application might be using. Customers do not manage or control the underlying cloud infrastructure such as storage, operating systems, servers and network, but they do control the deployed applications and possibly the configuration settings of the hosting environment (Platform as a Service Magazine 2015). Moreover, multiple programming models and services such as authentication, data access and payments are offered as building blocks to new applications (Buyya, Broberg & Goscinski 2011). Platform as a service is designed for developers. Examples include dotCloud, CloudBees, or AppFog (Ramezani 2016).

2.3.3. SOFTWARE AS A SERVICE

On-premises software, often abbreviated as "on-prem software", is installed and executed on a personal computer rather than at a remote facility, such as a server farm or cloud. On-premises software is sometimes referred to as "shrink wrap" software, while off-premises software is commonly called SaaS or "computing in the cloud" (Mangaiyarkarasi, Sureshkumar & Elango 2013).

In infrastructure as a service, the applications reside at the top of the cloud stack and are accessed through a web browser. Given the benefits of software as a service, consumers are increasingly shifting from traditional desktop applications, such as word processing, spreadsheets, and email clients, to online software offered as a service. For customers, this option reduces the burden of software maintenance. For CSPs, this option simplifies development and testing (Buyya, Broberg & Goscinski 2011). SaaS consumers do not manage or control the underlying cloud infrastructure or the applications' capabilities, with the possible exception of limited configuration settings. Software as a service focuses on end users. Examples include Gmail, Microsoft Office 365, and Salesforce (Ramezani 2016).

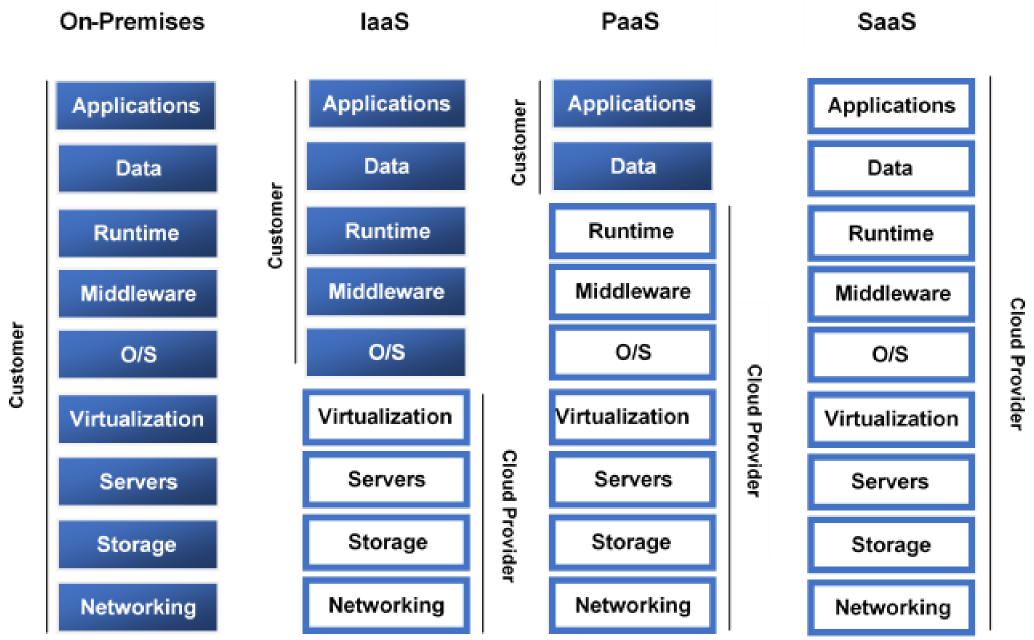

Figure 2.3 The distribution of responsibilities between customers and cloud providers given different service types (Ramezani 2016).

2.4 SERVICE-LEVEL AGREEMENT AND QUALITY OF SERVICE

According to the definition, a service-level agreement (SLA) is a part of a service contract where the level of service is formally defined. SLAs are offered by CSPs to express their commitment to deliver a certain QoS and determine the level(s) of service being sold in plain language terms (Buyya, Broberg & Goscinski 2011). QoS is a description or measurement of the overall performance of a service, such as a computer network or telephony or a cloud computing service, especially the performance seen by the users of the network. To quantitatively measure quality of service, several aspects related to network service are often considered, such as bit rate, transmission delay, packet loss, throughput, jitter, latency, availability etc. (Abdelmaboud et al. 2015). QoS is the ability to provide different priorities to different applications, users, or data flows, or to guarantee a certain level of performance to a data flow (Abdelmaboud et al. 2015; Heidari & Buyya 2019). In the rest of this section, we deal with the main QoS criteria related to this study.

- **CPU (Central Processing Unit):** CPU is the electronic circuitry within a computer that performs the instructions of a computer program by performing the basic arithmetic, logic, controlling, and input/output operations specified by the optimized instructions (Maalej et al. 2020; Tan & Demirel 2015). As Ramezani (2016) discussed, CPU utilization refers to a computer's usage of processing resources, or the amount of work handled by a CPU. Actual CPU utilization in cloud varies depending on the amount and type of the managed computing jobs/tasks. Certain jobs/tasks require heavy CPU time, while others require less because of non-CPU resource requirements. Appropriate CPU usage makes it easy to consume massive amounts of computation power for batch processing, data analysis, and high performance computing needs (Manvi & Shyam 2014).

- **Memory:** Cloud computer architecture is defined as a clustered structure of the memory resources in the form of virtual entities. Gone are the days when memory management was undertaken using static methods. As the cloud environment is dynamic and volatile, it is necessary to inculcate the dynamic memory allocation trends in cloud-based systems. The increased number of cores in cloud servers combined with the rapid adoption of virtualization technologies also creates a huge demand for memory (Manvi & Shyam 2014; Ramezani 2016).
- **Data Storage:** Data storage refers to saving data to a remote storage system maintained by a third party. The Internet provides the connection between the computer and the database. Cloud storage systems usually depend on hundreds of data servers. Because computers occasionally require maintenance or repair, it is important to store the same information on multiple machines, which is called redundancy. Without redundancy, a cloud storage system is unable to guarantee that clients will have access to their information at any given time. Most systems store the same data on servers that use different power supplies. As a result, users can access their data even if one power supply fails (Manvi & Shyam 2014; Ramezani 2016).
- **Bandwidth:** Bandwidth refers to the amount of data that can be sent from one point to another in a certain period of time. It is measured as a bit rate expressed in bits per second (bits/s) or multiples of it (kbit/s Mbit/s etc.) (de Oliveira & Silva 2020; Thangappan et al. 2020; Wu et al. 2020).
- **Latency:** Latency in cloud computing environments refers to a time interval between a customer request and a cloud computing service provider’s response or the time delay between the cause and

the effect of some physical changes in the system (Aburukba et al. 2019; Borylo et al. 2020; Caiza et al. 2020; Minnear 2011; Mulinka & Kencl 2015).

- **Cloud computing security:** Cloud computing security is a wide-ranging set of policies, applications, technologies, and controls utilised for protecting virtualized IP, applications, data, services, and the associated infrastructure of cloud computing. It is a sub-domain of computer security, network security, and, more generally, information security (Mthunzi et al. 2020; Subramanian & Jeyaraj 2018; Sun 2020; Wang et al. 2020).
- **Service Availability:** Service availability refers to the probability of receiving the proper service at any given time. It is usually expressed as SLA downtime in minutes per year or as the percentage of time the service will be up throughout the year. Thus, CSPs need to perform an availability analysis for quantifying the expected downtime that the service may experience over a period of time (Ghosh et al. 2014).
- **Price:** Price is considered as a quantity metric which plays an important role in performance measurement of CSPs. Thus, it is desirable for expressing price with respect to the features related to cloud service providers (Somu, Kirthivasan & VS 2017).

2.5 CLOUD SERVICE PROVIDERS (CSPs)

A CSP is a third party which provides their customers with cloud computing services such as infrastructure as a service, platform as a service and software as a service. The cloud services are hosted in a data centre that can be accessed and used by individuals or companies through network connectivity. Figure 2.4 and Table 2.1 shows some of the major CSPs and their approaches respectively.

Figure 2.4 Cloud service providers

Table 2.1 Services provided by the major cloud service providers

Cloud Service Provider	IAAS	PAAS	SAAS
Amazon	X	X	
Century Link	X	X	
Google	X	X	X
IBM	X	X	X
Microsoft	X	X	X
Rackspace	X	X	
Salesforce.com		X	X
SAP	X	X	X
Verizon Terremark	X	X	

2.6 PERFORMANCE MEASUREMENT OF CLOUD SERVICE PROVIDERS

Given the agility and flexibility cloud services offer, many businesses are opting to transfer all or part of their information systems to the cloud. Yet the growing number of CSPs is making it increasingly difficult to decide which CSPs are able to meet a customer’s requirements. To tackle this problem, a number of methods for selecting and measuring the performance of CSPs have been developed.

In this section, we review approaches which have been proposed and used for CSPs evaluation and selection in the past.

Chan & Chieu (2010) proposed a system for evaluating and selecting CSPs to match application requirements by using Singular Vector Decomposition (SVD). The SVD approach presents a compact and efficient knowledge representation mechanism to represent QoS requirements and the CSP relationship using the matrix and dimension reduction technique for extracting significant relationship among QoS requirements and the corresponding CSPs (Chan & Chieu 2010). In the proposed approach to evaluate and select CSPs according to the cloud consumer's requirements, the providers' attributes from the repository are transformed and processed in the form of Provider - QoS matrix for SVD transformation. The model proposed by Chan & Chieu (2010) enables the evaluation and selection of CSPs without a precise match of the required QoS attributes. Chauhan et al. (2011) presented a ranking method which ranks CSPs by matching the SLA attributes of the cloud service customer's requirements with the CSPs' provisions. They proposed a Dynamic Static Service Ranking System (SRS) with defined priorities which identifies the specific user's priorities for ranking CSPs. Each factor in the priority is given a weightage accordingly and the CSPs rankings are calculated based on it. Ghosh, Ghosh & Das (2014) and Ghosh et al. (2014) proposed a framework called SelCSP (Select CSPs) to evaluate and select trustworthy and competent CSPs. The proposed model estimates trustworthiness with respect to context-specific, dynamic trust and reputation feedback. The proposed model also computes the competence of a CSP based on the transparency of SLAs. They used a case study consisting of six CSPs to demonstrate the application of their proposed model. Rajarajeswari & Aramudhan (2015) proposed the Pointcare Plot method (PPM)-based mathematical model to find the rank of the CSPs in a federated cloud management system. The proposed model was used to find the most appropriate CSP for the incoming request in an efficient manner. Aruna & Aramudhan (2016) proposed a provider discovery algorithm and fuzzy sets ranking approach in the modified federated architecture and evaluated the performance of CSPs. The discovery approach shortlists CSPs according to the QoS indicators presented by the Service Measurement Index (SMI) with the SLA that provides improved performance. Furthermore, the cost is also included which represents fulfilment at the level of the end user. The ranking framework is based on a fuzzy set method comprising three general phases including problem decomposition, judgment of priorities and an aggregation of these priorities. Using a number of simple rules, the fuzzy set is combined with the QoS criteria. The Weighted Tuned Queuing Scheduling (WTOS) algorithm was proposed to resolve the issue of starvation in the existing architecture and managing the requests effectively. The results obtained demonstrate that the proposed framework has a better successful selection rate, average response time and less overhead with those of the existing framework that had supported the cloud computing environment. Somu, Kirthivasan & VS (2017) proposed a cloud service selection framework based on the Hypergraph-based Computational Model (HGCM) and the Minimum Distance-Helly Property (MDHP) algorithm to evaluate and select CSPs. The Helly feature of the hypergraph was used for assigning weights to the characteristics and decreasing the complexity of the ranking and selection model, while arithmetic residue and Expectation–Maximization (EM) algorithms were used for imputing missing values. The results obtained by MDHP under different case scenarios (dataset used by various research communities and synthetic dataset) show the ranking and selection algorithm is scalable and computationally appealing. Singh & Sidhu (2017) addressed the problem of evaluating and selecting trusted CSPs. They proposed a compliance-based multi-dimensional trust evaluation system (CMTES) that enables cloud service customers to determine trusted CSPs from different perspectives. The proposed method can help cloud service customers who want to select a CSP from a pool of CSPs, based on quality of service requirements. Chauhan et al. (2011) presented a ranking and selection method which evaluates and selects CSPs by matching the SLA factors of the given cloud customers' requirements with the CSPs' provision. Qu, Wang & Orgun (2013) presented a model to evaluate and select CSPs by aggregating the information from both the feedback from cloud customers and objective performance analysis from a trustworthy third party. According to the model, they first presented a structure that supports the cloud service provider's evaluation and selection method. After categorizing the subjective evaluation and objective evaluation, they proposed a CSP evaluation and selection method for aggregating all subjective assessments and objective assessments using a fuzzy simple additive weighting technique. Furthermore, to lessen the bias caused by unreasonable feedback from an amateur or malicious cloud service customer, an approach was presented to filter the feedback from such customers. After processing, the aggregated results can quantitatively indicate the performance of CSPs.

Huang, Hsu & Tzeng (2012) proposed a hybrid multi-criteria decision analysis (MCDM) model combining the Decision Making, Trial and Evaluation Laboratory (DEMATEL) and Analytic Network

Process (ANP) based framework and Grey Relational Analysis (GRA) to evaluate and select CSPs. They used DEMATEL, ANP and GRA methods to reduce the service quality gap, meet users' satisfaction and maximize profits in interdependence and feedback problems among a number of criteria. Kumar & Agarwal (2014) presented a framework for cloud service evaluation and a selection engine which acts as a tool to enable the consumers to select the most suitable CSPs from the Web Repository. The framework proposed by Kumar & Agarwal (2014) utilizes the Analytical Hierarchy Process (AHP) approach for the multi-criteria quality of service decision making which can accelerate the evaluation and selection process. Sahri et al. (2014) utilized AHP to help cloud customers to evaluate and select the optimal database as a service (DBaaS) cloud provider. AHP hierarchy consists of some key qualities of service attributes, distributed on three levels. Relative importance weights and rates are selected on a scale of 1–9. Supriya, Sangeeta & Patra (2016) compared various trust estimation methods using the MCDM process to evaluate and rank CSPs offering IaaS. The trust estimation of service providers uses the Cloud Service Measurement Initiative Consortium (CSMIC) parameters prioritized based on Finance, Security and Performance criteria. Singh & Sidhu (2017) addressed the problem of evaluating and selecting trusted CSPs They proposed a Compliance-based Multi-dimensional Trust Evaluation System (CMTES) that enables cloud service customers to determine trusted CSPs from different perspectives. The proposed method can help cloud service customers who want to select a CSP from a pool of CSPs based on QoS requirements. Garg, Versteeg & Buyya (2013) proposed a framework called SMICloud that can compare CSPs according to cloud customer requirements. They designed performance metrics to measure the QoS of an infrastructure as a service. In addition, they designed AHP-based ranking mechanism to compare different CSPs.

Table 2.2 Literature review of cloud service providers' performance measurement

Approaches	Advantage of the approach	Disadvantages of the approach	Study (References)
ANP, GRA, DEMATEL, Dynamic Static SRS;	• The consistency of the judgments. • Facilitating the process of assigning weights. • Handling both incomplete information and unclear problems more precise. • Analyzing the mutual influences among different factors and understands the complicated cause and effect relationships in the decision-making problem.	• Time-consuming. • Inconsistent results. • Complicated calculations. • Subjective weights. • Inappropriate for criteria utility. • Incomplete ranking. • Inefficient for problems with large scale.	Huang, Hsu & Tzeng (2012); Chauhan et al. (2011)
AHP	• Easy to use. • Being scalable and flexible. • Ability to check inconsistencies. • Capture both subjective and objective evaluation measures.	• Time-consuming. • Inappropriate for problems that contain complex interactions among the evaluation criteria and dimensions. • Inconsistencies between judgment and ranking criteria. • Ranking irregularities. • Subjective weights.	Kumar & Agarwal (2014); Sahri et al. (2014); Supriya, Sangeeta & Patra (2016)
SVD	• Easy to use. • Clear stages for performance measurement. • Flexibility.	• Inconsistent results. • Complicated calculations. • Subjective weights. • Incomplete ranking. • Inefficient for problems with large scale	Chan & Chieu (2010)
SelCSP	• Appropriate for uncertain environments.	• Subjective judgments • Incomplete ranking. • Time-consuming	Ghosh, Ghosh & Das (2014) and Ghosh et al. (2014)

PPM	• Easy to use. • Flexibility.	• Inappropriate for problems that contain complex interaction among the evaluation criteria and dimensions. • Inconsistencies between judgment and ranking criteria. • Subjective weights. • Limitation of the use of the scales.	Rajarajeswari & Aramudhan (2015)
Fuzzy set	• Appropriate to complex problems and uncertain environments. • Flexibility	• Ranking irregularities. • Subjective judgments. • Inefficient for problems with large scale.	Aruna & Aramudhan (2016); Qu, Wang & Orgun (2013)
HGCM/MDHP	• Scalable • Capture both subjective and objective evaluation measures. • Flexibility	• Subjective weights. • Ranking irregularities. • Inefficient for problems with large scale.	Somu, Kirthivasan & VS (2017)
CMTES	• Being scalable and flexible. • Clear stages for performance measurement. • Capture both subjective and objective evaluation measures	• Subjective weights. • Complicated calculations. • Ranking irregularities. • Inefficient for problems with large scale. • Inefficient for problems with large scale.	Singh & Sidhu (2017)

SMICloud	• Being scalable and flexible. • Capture both subjective and objective evaluation measures. • Easy to use.	• Time-consuming. • Inappropriate to problems that contain complex interaction among the evaluation criteria and dimensions. • Ranking irregularities. • Subjective weights.	Garg, Versteeg & Buyya (2013)

The performance measurement and selection problems of CSPs based on the QoS have been studied by many researchers. As outlined and discussed in Table 2.2, although some works have been done on performance measurement problems and the selection of CSPs, the existing approaches suffer from the following limitations and drawbacks:

1. The existing approaches for the performance measurement and selection of CSPs are unable to find a slight difference between a large number of CSPs in a highly intense competitive cloud marketplace.
2. Moreover, these approaches are unable to evaluate the efficiency of CSPs requiring complex calculations, being effort-intensive and time-consuming and having subjective weights and ranking irregularities.
3. Existing approaches are unable to evaluate the performance of CSPs in the presence of different types of data such as undesirable data, integer-valued data and stochastic data.

2.7 DATA ENVELOPMENT ANALYSIS

One of the most important techniques for performance evaluation is data envelopment analysis (DEA). DEA is a non-parametric method for measuring the efficiency of a set of decision-making units (DMUs) that convert multiple inputs into multiple outputs (Azadi & Saen 2011; Matin, Amin & Emrouznejad 2014). For the first time, DEA was proposed by Charnes, Cooper & Rhodes (1978). According to (DeToro 1995; Golany 1988; Post & Spronk 1999; Shafiee, Lotfi & Saleh 2014; Sheridan 1993; Zhu 2004), the main features of DEA are as follows:

1. Ability to process multiple elements, easy to use, and it can easily be incorporated with statistical methods.
2. There is no need to specify the relationships among the performance measures.

3. The concept of an efficient frontier which is used in DEA serves appropriately as an empirical standard of excellence.
4. DEA can analyse qualitative measures as well as quantitative measures simultaneously.
5. In the approach of utilizing DEA, there is no need to assume priority estimates. This feature increases the acceptability of its results.
6. DEA provides information about inefficient DMUs as well as efficient DMUs.
7. DEA is highly flexible and can be combined easily with other analytical methods such as statistical analysis and other multi-criteria decision-making techniques.

Because of the uniqueness of DEA, it has been widely developed and used to measure performance in different domains since 1978 (Emrouznejad & Yang 2018; Mirhedayatian, Azadi & Saen 2014). DEA for the first time was proposed by Charnes, Cooper & Rhodes (1978). Then Banker, Charnes & Cooper (1984) developed the CCR model and proposed the BCC model. To deal with nondiscretionary factors (factors which are beyond control of managers/decision makers), Banker & Morey (1986) proposed the DEA model with nondiscretionary factors. They measured the efficiency of DMUs using mathematical programming formulations when some inputs and outputs are nondiscretionary. Cooper, Park & Pastor (1999) proposed the Range Adjusted Measure (RAM) model based on the additive model of DEA. The RAM model can maximize outputs and minimize inputs simultaneously in the DEA context. Tone (2001) proposed a slacks-based measure (SBM) of the efficiency of DEA. The SBM model can deal directly with both the input excesses and the output shortfalls of the DMU under observation. Thanassoulis (2000) used DEA in the context of the regulator of water companies in the United Kingdom in 1994 for setting price limits. Manandhar & Tang (2002) proposed a framework to incorporate the intangible aspects into a DEA framework in the form of internal service quality. They suggested the simultaneous benchmarking of the performance of bank branches along multiple dimensions using a modified DEA formulation. Luo (2003) proposed a DEA model to evaluate the profitability and marketability efficiency of large banks. Ross & Droge (2004) applied the DEA methodology to evaluate the efficiency of units within a large-scale network of petroleum distribution facilities in the USA. Traditional DEA models are based on the complete homogeneity of DMUs while in many real applications, there are non-homogeneous DMUs. Saen, Memariani & Lotfi (2005) proposed an approach to determine the relative efficiency of slightly non-homogeneous DMUs via DEA. Amirteimoori & Kordrostami (2005) proposed a DEA-based approach to allocate fixed costs to DMUs and to allocate fixed inputs and set a fixed target. In order to deal with the dual-role factors (factors which can be considered as both input and output in the DEA context), Cook, Green & Zhu (2006) proposed a dual-factor DEA model. Chen et al. (2006) proposed a DEA non-linear programming model to evaluate the impact of information technology (IT) on multiple stages along with information on how to distribute IT-related resources so that efficiency is maximized. In order to achieve other targets in terms of efficiency in the DEA approach, Cooper et al. (2007) proposed the enhanced DEA Russell graph model as a non-radial model. Köksal & Aksu (2007) evaluated the comparative operating efficiency of 24 A-Group Travel Agencies which are operated internationally in Turkey using a DEA. To evaluate operating efficiency, they grouped A-Group Travel Agencies into "independently operating" and "operating under a chain brand". Lee (2008) used the multiple linear regression approach and DEA to examine the effectiveness of energy management. He utilized the regression method using environmental factors to calculate the predicted energy usage intensity of each evaluated building and then used DEA to calculate the overall energy efficiency, using the predicted energy usage intensity as output and the observed energy usage intensity as input. Xu, Li & Wu (2009) studied the supply chain performance evaluation of a furniture manufacturing industry in the southwest of China. They identified the main uncertainty factors affecting the evaluation process and then modelled and analysed them using a rough DEA model. Stewart (2010) developed the traditional DEA models to include long-term top management goals. Yu & Wen (2010) evaluated the urban environmental sustainability of 46 typical Chinese cities. They used the Malmquist Productivity Index (MPI) in DEA to identify changes between 2006–2007. Mousavi-Avval et al. (2011) used DEA to estimate the energy efficiencies of soybean producers based on eight energy inputs namely diesel fuel, machinery, chemicals, fertilizers, water for irrigation human labour, electricity and seed energy and a single output of grain yield. The study also helps to rank efficient and inefficient farmers and to identify optimal energy requirements and wasteful uses of energy. Kuah, Wong & Wong (2012) devised a genuine Knowledge Management (KM) performance measurement model in a stochastic setting based on DEA, Genetic Algorithm (GA) and Monte Carlo simulation. The proposed model assesses KM using a set of proxy measures related to the major KM processes. Khoshnevisan et al. (2013) used DEA to analyse the energy efficiency of wheat farms to separate efficient and inefficient growers and to calculate wasteful uses of energy. Azadi & Saen (2013) proposed a combination of quality function deployment (QFD) and imprecise DEA with the enhanced Russell graph measure for performance evaluation in healthcare. Azadi et al. (2014) proposed a two-stage target-setting DEA approach for the performance measurement of the green supply chain management of public transportation service providers. Hu & Liu (2015) proposed the preliminary analysis of undesirable output reduction targets and emission schedules in temporal–spatial comparisons according to DEA. They investigated the reduction targets of undesirable outputs such as the maximum, input, technical and ideal reduction targets. Yang, Lee & Hu (2016) adopted an extended urban metabolism framework for evaluation urban sustainability and utilized a DEA model with undesirable outputs that takes pollution into account to measure the aggregate urban input–output efficiency of Taiwan′s 22 administrative regions. Rebolledo-Leiva et al. (2017) proposed a four-step method for the joint use of the carbon footprint (CF) evaluation and DEA. They used an output-oriented DEA model to maximize production and reduce the CF, taking into account the economic and ecological perspectives simultaneously. Ang, Chen & Yang (2018) proposed group efficiency and group cross-efficiency models based on DEA to evaluate Taiwan hotel chains and subsidiary hotels with data from 2011 to 2015. Nahangi, Chen & McCabe (2019) proposed an approach based on DEA to identifying the efficiency of construction sites, also known as DMUs. Khezrimotlagh et al. (2019) proposed a DEA model to deal with big data. They utilized the proposed model to evaluate the efficiency of 30,099 electric power plants in the United States from 1996 to 2016.

Table 2.3 A summary of the reviewed literature related to the DEA approach in performance evaluation

References	Key Development
Charnes, Cooper & Rhodes (1978)	Proposed CCR model for the first time
Banker, Charnes & Cooper (1984)	Proposed BCC model for the first time
Banker & Morey (1986)	Proposed DEA model with nondiscretionary factors for the first time
Cooper, Park & Pastor (1999)	Proposed RAM model for the first time
Thanassoulis (2000)	Used DEA in the regulation of UK water utilities
Tone (2001)	Proposed SBM model for the first time
Manandhar & Tang (2002)	Proposed a framework to incorporate intangible aspects into a DEA framework
Luo (2003)	Used a combination of CRS DEA model and VRS DEA model
Ross & Droge (2004)	Used DEA within a large-scale network
Saen, Memariani & Lotfi (2005)	Proposed a non-homogeneous DEA model
Amirteimoori & Kordrostami (2005)	Proposed a DEA model to allocate fixed cost
Cook, Green & Zhu (2006)	Proposed dual-factor DEA model
Chen et al. (2006)	Proposed DEA non-linear programming model
Cooper et al. (2007)	Proposed enhanced DEA Russell graph model for the first time
Köksal & Aksu (2007)	Used DEA for performance evaluation of travel agencies
Lee (2008)	Proposed a combination of multiple linear regression approach and DEA approach
Xu, Li & Wu (2009)	Proposed a DEA model considering uncertainty factors
Stewart (2010)	Proposed goal-directed benchmarking DEA model
Yu & Wen (2010)	Used the MPI in DEA
Mousavi-Avval et al. (2011)	Used DEA for performance evaluation in energy sector
Kuah, Wong & Wong (2012)	Proposed a combination of Knowledge Management and DEA
Khoshnevisan et al. (2013)	Used DEA for analysing the energy efficiency in agriculture sector.

Azadi & Saen (2013)	Proposed a combination of QFD and imprecise DEA
Azadi et al. (2014)	Proposed a two-stage target-setting DEA approach
Hu & Liu (2015)	Proposed a DEA model in the presence of undesirable outputs.
Yang, Lee & Hu (2016)	Proposed a DEA model with undesirable outputs
Rebolledo-Leiva et al. (2017)	Proposed an output-oriented DEA model for maximizing production and reducing CF
Ang, Chen & Yang (2018)	Proposed group efficiency and group cross-efficiency models based on DEA
Nahangi, Chen & McCabe (2019)	Proposed a DEA approach for safety-based efficiency evaluation of construction sites
Khezrimotlagh et al., (2019)	Proposed DEA models to deal with big data

DEA uses linear programming to determine the relative efficiencies of a set of homogenous DMUs that employ multiple inputs to produce multiple outputs without requiring any assumptions about the functional form relating inputs to outputs. DEA forms an efficient combination of input and output variables by analyzing the historical data to make the efficiency boundary. A DMU is termed efficient if it lies on the boundary, otherwise it is inefficient. The three basic DEA models are the CCR model, the BCC model and the SBM model.

As Emrouznejad, Cabanda & Gholami (2010) stated, “DEA can be either input- or output-orientated. In the first case, the DEA method defines the frontier by seeking the maximum possible proportional reduction in input usage, with output levels held constant, for each firm. However, for the output-orientated case, the DEA method seeks the maximum proportional increase in output production, with input levels held fixed”.

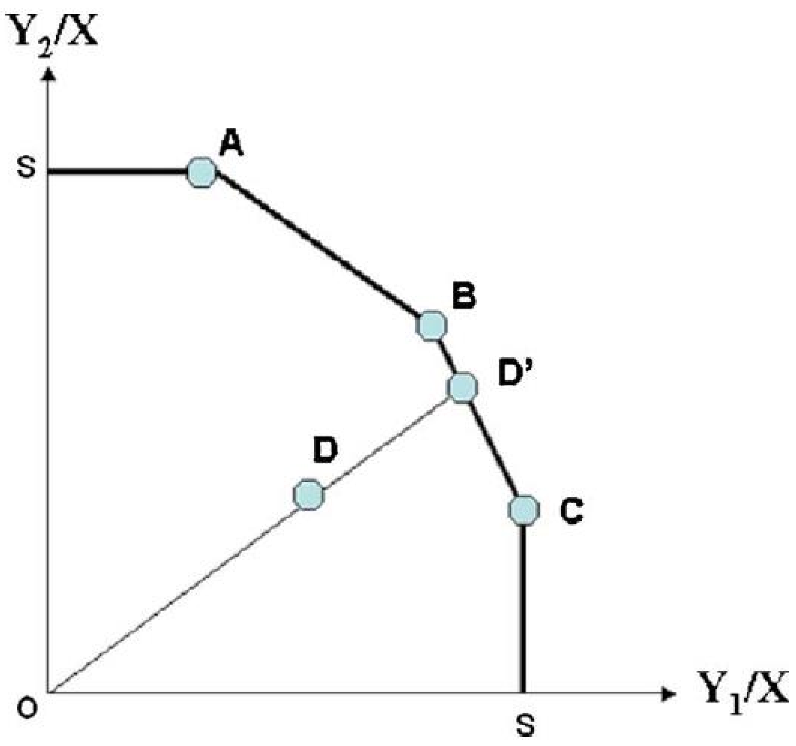

Figure 2.5 Output-oriented piecewise linear convex isoquant (Emrouznejad et al., 2010).

An output-oriented DEA model with input variables $(X_1, \dots, X_m)$ and output variables $(Y_1, \dots, X_s)$ with n DMUs $(j = 1, \dots, n)$ is shown in Model 2.1.

Figure 2.5 illustrates a two-output (Y1 and Y2) one-input X output-oriented piecewise convex linear hull under the assumptions of variable returns to scale (VRS). SS represents the full technical efficiency isoquant. Points A, B and C represent technically efficient DMUs on the frontier. If a given DMU uses one unit of input and produces outputs defined by point D, the technical inefficiency of DMU is presented as the distance D'D; this is the amount by which all outputs can be proportionally increased without increasing input. In percentage terms, it is expressed by the ratio OD/OD', which is the ratio by which all the outputs can be increased (Emrouznejad, Cabanda & Gholami 2010).

Model 2.1, which is an output-oriented DEA-VRS model, constructs the additive combinations of outputs and inputs to achieve a single virtual output and virtual input in the calculation of an efficiency score. Another DEA model, termed a 'multiplicative model', was introduced into the DEA literature by Charnes, Cooper & Rhodes (1978). An important property of this model is that it uses the concept of the geometric mean with non-dimensional (unit invariance) properties (Banker et al. 2004); thus it is more suitable for ratio data than the standard DEA.

Model 2.1

$$Max\ a$$

$$\sum_{j=1}^{n} \lambda_j X_{ij} \leq X_{io} \qquad ; i = 1, \dots, m$$

$$\sum_{j=1}^{n} \lambda_j Y_{rj} \geq a\, X_{ro} \qquad ; r = 1, \dots, s$$

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

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

Banker, Charnes & Cooper (1984), developed the CCR model to variable returns to scale which is called the BCC model. In other words, if the constraint $\sum_{j=1}^{n} \lambda_j = 1$ is adjoined in model CCR, the BCC model is obtained. This new constraint introduces an additional variable, μ_o, into the (dual) multiplier problems. This extra variable makes it possible to effect returns-to-scale assessments (decreasing, constant and increasing). Hence, the BCC model is also referred to as the VRS model and is different from the CCR model, which is referred to as the CRS (Constant Returns to Scale) model (Cooper, Seiford & Zhu 2011).

Tone (2001) proposed the SBM model for the first time. It has three variations, i.e., input, output, and non-oriented. The non-oriented model denotes both input and output-oriented models.

Model 2.2 is the input-oriented SBM.

Model 2.2

$$\rho_I^* = \min_{\lambda, s^-, s^+} 1 - \frac{1}{m} \sum_{i=1}^{m} \frac{s_i^-}{x_{io}},$$

s.t.

$$x_{io} = \sum_{j=1}^{n} x_{ij} \lambda_j + s_i^- \qquad (i = 1, \dots, m),$$

$$y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_i^+ \qquad (r = 1, \dots, s),$$

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

$$\lambda_j \geq 0 \ (\forall j), \quad s_i^- \geq 0 \ (\forall j), \ s_r^+ \geq 0 \ (\forall j),$$

Definition 2.1 (SBM-Input-Efficient). A $DMUo = (\boldsymbol{x}_o, \boldsymbol{y}_o)$ is SBM-input-efficient if $\rho_I^* = 1$ holds.

This means $\boldsymbol{s}^{-*} = 0$, i.e., all output slacks are zero. Nonetheless, input slacks may be nonzero.

The model 2.3 is the output-oriented SBM.

Model 2.3

$$\frac{1}{\rho_o^*} = \underset{\lambda,\ s^-,\ s^+}{\mathrm{m3ax}} \ 1 + \frac{1}{s} \sum_{r=1}^{s} \frac{s_r^+}{x_{io}},$$

s.t.

$$x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- \qquad (i = 1, \dots, m),$$

$$y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_i^+ \qquad (r = 1, \dots, s),$$

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

$$\lambda_j \geq 0 \ (\forall j), \quad s_i^- \geq 0 \ (\forall j), \ s_r^+ \geq 0 \ (\forall j),$$

Definition 2.2 (SBM-Input-Efficient). A $DMUo = (\boldsymbol{x}_o, \boldsymbol{y}_o)$ is SBM-out-efficient if $\rho_O^* = 1$ holds.

This means $\boldsymbol{s}^{+*} = 0$, i.e., all output slacks are zero. Nonetheless, input slacks may be nonzero.

The following model (model 2.4) measures the efficency of DMUs using the non-oriented SBM model.

Model 2.4

$$\rho_{Io}^{*} = \min_{\lambda, s^{-}, s^{+}} \frac{1-\frac{1}{m}\sum_{i=1}^{m}\frac{s_i^-}{x_{io}}}{1+\frac{1}{s}\sum_{r=1}^{s}\frac{s_r^+}{x_{io}}},$$

s.t.

$$x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- \qquad (i = 1, \dots, m),$$

$$y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_i^+ \qquad (r = 1, \dots, s),$$

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

$$\lambda_j \geq 0 \ (\forall j), \quad s_i^- \geq 0 \ (\forall j), \ \ s_r^+ \geq 0 \ (\forall j),$$

Definition 2.2 (SBM-Input-Efficient). A $DMUo = (\boldsymbol{x}_o, \boldsymbol{y}_o)$ is SBM-out-efficient if $\rho_{Io}^{*} = 1$ holds. This means $\boldsymbol{s}^{-*} = 0 \ \boldsymbol{s}^{+*} = 0$, i.e., all output slacks are zero. Nonetheless, input slacks may be nonzero. According to Tone (2001), Model 2.4 can be transformed into a linear programming model using the Charnes-Cooper transformation proposed by Charnes & Cooper (1962) as follows:

<u>Model 2.5</u>

$$\tau^{*} = \min_{t, \Lambda, \ \boldsymbol{S}^{-}, \boldsymbol{S}^{+}} t - \frac{1}{m}\sum_{i=1}^{m}\frac{s_i^-}{x_{io}}$$

Subject to:

$$1 = t + \ 1 + \frac{1}{s}\sum_{r=1}^{s}\frac{s_r^+}{y_{ro}}$$

$$tx_{io} = \sum_{j=1}^{n} x_{ij}\Lambda_j + S_i^- \qquad (i = 1, \dots, m),$$

$$ty_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - S_r^+ \qquad (r = 1, \dots, s),$$

$$\Lambda_j \geq 0 \ (\forall_j), S_i^- \geq 0 \ (\forall_i), S_i^+ \geq 0 \ (\forall_j), \ t > 0.$$

2.8 THE IMPLEMENTATION OF DEA IN THE CLOUD ENVIRONMENT

There is scant research on the application of DEA in the cloud environment. Kumar (2014) proposed a method for the performance evaluation of cloud services based on DEA, AHP and the technique for order of preference by similarity to ideal solution (TOPSIS). Xu, Ma & Wang (2015) presented a non-parametric DEA to evaluate the relative efficiency of IaaS services. In this approach, cloud services are described based on functional requirements such as storage, memory and CPU (Filiopoulou et al. 2018). Filiopoulou et al. (2018) proposed a DEA input-oriented model for the performance measurement of cloud services based on both functional and non-functional parameters such as reliability, security and cloud management features. They posited that both functional and non-functional parameters play key roles in enhancing cloud services and need to be taken into account in the performance evaluation of CSPs.

2.9 DISCRIMINATION POWER IN DATA ENVELOPMENT ANALYSIS

Traditional DEA models cannot differentiate between observations (DMUs) and many of them may possibly be efficient when evaluating efficiency scores. There are a number of studies in the literature to improve the discrimination power in DEA. In the first attempt to increase discrimination power, Sexton, Silkman & Hogan (1986) ranked DMUs based on a cross-efficiency model, which eliminates unrealistic weighting schemes in performance measurement through DEA. Doyle & Green (1994) developed the cross-efficiency model proposed by Sexton, Silkman & Hogan (1986) in a number of directions. They grounded an intuitive understanding of cross-efficiency in the concept of peer-assessment, contrasted with self-appraisal implied by simple efficiency, and discussed the relative merits of each. Furthermore, they proposed aggressive and benevolent concepts in the cross-efficiency model. Wang & Chin (2010) proposed some secondary goals for cross-efficiency assessment and discussed that the ideal points in the model of Liang et al. (2008) cannot be possible for inefficient DMUs. Jahanshahloo, Lotfi, et al. (2011) presented a symmetric weight assignment technique (SWAT) for cross-efficiency assessment that does not influence feasibility and rewards DMUs that make a symmetric selection of weights. Lim (2012) proposed aggressive and benevolent formulations of cross-efficiency assessment by maximizing (or minimizing) the cross-efficiency of the worst (or best) peer DMU. Khodabakhshi & Aryavash (2017) applied an optimistic–pessimistic method to present a cross-efficiency approach. In the proposed approach, the estimated scores are obtained using both the weaknesses and strengths of DMUs. Moreover, in this method, the optimal weights are uniquely determined and there is no need to set any secondary goals.

Liu, Wang & Lv (2017) proposed an iterative algorithm to achieve an aggressive game cross-efficiency assessment combining the aggressive and the game cross-efficiency assessments presented by Liang et al. (2008). Liu (2018) considered cross-efficiency intervals which take the aggressive and benevolent formulations into account concurrently, and their variances for ranking DMUs. Liu, Wang & Lv (2017) proposed an aggressive game cross-efficiency concept and an aggressive secondary model to minimize the cross-efficiencies of other DMUs under the constraints. They developed an iterative algorithm to obtain aggressive game cross-efficiency. Örkcü et al. (2019) proposed a neutral cross-assessment to measure the efficiency of the basic two-stage systems in DEA and presented a neutral cross-efficiency model.

Super-efficiency is another stream to increase discrimination power in DEA (Farzipoor Saen 2008). For the first time, Andersen & Petersen (1993) (AP) proposed the super-efficiency model. The proposed model modifies the envelopment linear programming formulation so that the corresponding column of the DMUs being scored is removed from the coefficient matrix (Farzipoor Saen 2008). However, Thrall (1996) showed that the AP model might result in instability when some inputs are close to zero. To address this issue, the MAJ (Mehrabian, Alirezaee & Jahanshahloo 1999) model and SBM were presented. Lovell & Rouse (2003) presented a super-efficiency model that is able to generate the super-efficiency scores for both feasible solutions and infeasible solutions using a scalar. Chen (2005) discussed that to fully differentiate super-efficiency, both output-oriented and input-oriented super-efficiency DEA models are required. Farzipoor Saen (2008) proposed a super-efficiency measure to rank suppliers in the presence of volume discount offers. The model proposed by Farzipoor Saen (2008) can solve the infeasibility problem with a super-efficiency measure which is superior in comparison with the MAJ model. Cook et al. (2009) developed a VRS super-efficiency DEA model to address the infeasibility issue in super-efficiency. They defined the proposed model using both an input and output efficiency score. Lee, Chu & Zhu (2011) presented a two-step process to develop the works of Chen (2005) and Cook et al. (2009). They calculated the super-efficiency scores irrespective of the feasible and infeasible solutions of the standard VRS super-efficiency model. Chen & Liang (2011) modified the super-efficiency DEA model according to simultaneous input–output projection as a way to systematically characterize super-efficiency in both inputs and outputs. The model proposed by Chen & Liang (2011) addressed the infeasibility problem in the super-efficiency DEA measure. Chen & Liang (2011) discussed that the super-efficiency two-stage model proposed by Lee, Chu & Zhu (2011) can be solved in a single DEA-based model. Chen, Du & Huo (2013) addressed the infeasibility problem of the radial super-efficiency DEA measures under VRS using a Nerlove–Luenberger (N–L) measure of super-efficiency and according to the directional distance function. Lee & Zhu (2012) discussed that the super-efficiency models proposed by Lee, Chu & Zhu (2011) and Chen & Liang (2011) can be feasible when input data are positive whereas it can be infeasible when some of the inputs are zero. They modified the models of Lee, Chu & Zhu (2011) and Chen & Liang (2011) and proposed a new version of the super-efficiency model which is always feasible when data are non-negative.

Another approach to increase discrimination power in DEA is based on the optimistic and the pessimistic viewpoints. Entani, Maeda & Tanaka (2002) considered the efficiency measure in DEA from both the optimistic and the pessimistic perspectives. In the model proposed by Entani, Maeda & Tanaka (2002), the worst and best possible relative efficiencies are used to constitute an interval. Wang, Chin & Yang (2007) presented a model to integrate the two different efficiencies into a geometric average efficiency, which estimates the overall performance of each DMU. They showed that the geometric average efficiency has more discrimination power than either of the two efficiencies. Wang & Luo (2006) proposed two virtual DMUs named ideal DMU (IDMU) and anti-ideal DMU (AIDMU) into the DEA approach. IDMU AND AIDMU estimate DMUs performance from both the perspectives of the best possible relative efficiency and of the worst possible relative efficiency, respectively. Azizi & Ajirlu (2010) proposed DEA models based on an IDMU named bounded DEA models. They integrated both efficiencies in the form of an interval to estimate the overall performance of a DMU. Wang, Chin & Luo (2011) presented a couple of DEA models for cross-efficiency estimation using a virtual IDMU and a virtual AIDMU. The proposed models determine the input and output weights from the viewpoint of distance from either IDMU or AIDMU without considering aggressive or benevolent concepts in the DEA context. Sun, Wu & Guo (2013) proposed a DEA model to rank efficient DMUs using IDMU, AIDMU and common weights in the performance evaluation of flexible manufacturing systems (FMSs). Shi, Wang & Chen (2019) proposed a neutral cross-efficiency evaluation model to rank efficient DMU using an ideal frontier and anti-ideal aspects. However, the methods proposed by Sun, Wu & Guo (2013) and Shi, Wang & Chen (2019) do not consider mixed ideal and anti-ideal aspects nor do they consider decision-makers' preferences in the performance evaluation process.

Table 2.4 summarises the reviewed literature related to discrimination power in DEA.

Table 2.4 A summary of the reviewed literature related to the discrimination power in DEA

References	Key Development
Sexton, Silkman & Hogan (1986)	Proposed cross-efficiency DEA model for the first time.
Doyle & Green (1994)	Proposed aggressive and benevolent concepts for the first time in cross-efficiency model.
Wang & Chin (2010)	Proposed several secondary goals in the cross-efficiency context.
Jahanshahloo, Lotfi, et al. (2011)	Proposed SWAT in cross-efficiency assessment.
Lim (2012)	Developed aggressive and benevolent formulations in cross-efficiency.
Khodabakhshi & Aryavash (2017)	Applied an optimistic–pessimistic approach to present a cross-efficiency approach.
Liu, Wang & Lv (2017)	Proposed an iterative algorithm in DEA.
Liu (2018)	Considered cross-efficiency intervals in the aggressive and benevolent formulations.
Örkcü et al. (2019)	Proposed neutral cross-efficiency model.
Andersen & Petersen (1993)	Proposed the super-efficiency model for the first time.
Mehrabian, Alirezaee & Jahanshahloo (1999)	Proposed a combination of the MAJ model and SBM model.
Lovell & Rouse (2003)	Presented a super-efficiency model for both feasible solutions and infeasible solutions.

Farzipoor Saen (2008)	Solved the infeasibility problem with a super-efficiency measure.
Cook et al. (2009)	Developed a VRS super-efficiency DEA model.
Lee, Chu & Zhu (2011)	Presented a VRS super-efficiency DEA model.
(Chen & Liang 2011)	Addressed the infeasibility problem in the super-efficiency DEA measure.
Chen, Du & Huo (2013)	Addressed a Nerlove–Luenberger (N–L) measure of super-efficiency based on DDF.
Entani, Maeda & Tanaka (2002)	Proposed optimistic and the pessimistic viewpoints in DEA.
Wang & Luo (2006)	Proposed ideal and anti-ideal DMU models for the first time.
Azizi & Ajirlu (2010)	Proposed DEA bounded DEA models based on ideal DMU (IDMU).
Wang, Chin & Luo (2011)	Proposed a cross-efficiency model and virtual ideal DMU.
Sun, Wu & Guo (2013)	Proposed a DEA model based on DMUs ideal IDMU, AIDMU and common weights.
Shi, Wang & Chen (2019)	Proposed a neutral cross-efficiency based on IDMU and AIDMU

2.10 UNDESIRABLE OUTPUTS

DEA assumes that producing more outputs relative to less inputs is one criterion of efficiency. However, some outputs may be undesirable, such as pollution or noise (Cooper et al. 2007). Therefore, the results of an efficiency evaluation are likely to be less than optimal if undesirable outputs are not taken into consideration in the model. The first attempts to consider undesirable outputs in DEA are found in Pittman (1983). Färe et al. (1993) proposed a performance index called the hyperbolic efficiency measure which treats desirable outputs and undesirable outputs in a variety of ways. They viewed the proposed measure as an alternative to the "enhanced" multilateral productivity index proposed by Pittman (1983). Scheel (2001) proposed various methods to treat undesirable outputs in DEA. He introduced radial measures, which suppose that any change of the output level involves both desirable and undesirable outputs. In order to provide a better representation of the production technology, Hailu & Veeman (2001) developed the Chavas-Cox method for non-parametric analysis by incorporating undesirable outputs. They also constructed inner and outer non-parametric technology constraints to deal with undesirable outputs. Seiford & Zhu (2002) subsequently presented a DEA model to improve model performance by increasing desirable outputs and decreasing undesirable outputs. Färe & Grosskopf (2004) proposed an alternative method for Seiford and Zhu's method, which allows decision makers to explicitly model a joint environmental technology and measure performance in terms of increased desirable output and decreased undesirable output. This is achieved by adopting a directional distance function which may be estimated using the usual linear programming techniques employed in DEA. Korhonen & Luptacik (2004) proposed two different approaches to deal with desirable and undesirable in DEA. They first measured technical efficiency by relating the desirable outputs to the inputs. Then they measured ecological efficiency by relating the desirable outputs to the undesirable outputs. Jahanshahloo et al. (2005) presented a non-radial DEA model that simultaneously considers both undesirable inputs and outputs. Pathomsiri et al. (2008) evaluated the performance of 56 US airports between 2000 and 2003. The models proposed by Pathomsiri et al. (2008) joined the production of both good and bad outputs using the non-parametric directional output distance function. Yang & Pollitt (2009) proposed several DEA-based performance assessment models according to a research sample of the Chinese coal-fired power plants in order to deal with undesirable outputs jointly produced with the desirable outputs and to deal with uncontrollable variables. Saen (2010a) presented a model for supplier evaluation and selection in the existence of both imprecise data and undesirable outputs. Tone & Tsutsui (2011) proposed a hybrid non-parametric DEA model to measure efficiency in the presence of radial and non-radial inputs/outputs. They extended the proposed model to cope with non-separable desirable and undesirable outputs in DEA. He et al. (2013) evaluated the energy efficiency and productivity change of 50 companies in China's steel and iron industry in the presence of undesirable outputs. They discussed that omitting undesirable outputs in performance evaluation would result in biased technical change efficiency. Huang et al. (2014) proposed a DEA model based on the slacks-based measure, super efficiency and undesirable output for the performance evaluation of regional eco-efficiency in China. Chen et al. (2015) presented an enhanced directional distance measure model to deal with bad and good outputs while allowing some inputs and outputs to be zero using the evaluation of

CO_2 emissions in 111 countries. Rashidi & Saen (2015) evaluated eco-efficiency according to green criteria and the potential for undesirable output abatement and energy saving. Li et al. (2017) presented a DEA model to measure the performance of ecological systems in various regions of China. The proposed model considers both good outputs and undesirable variables in performance evaluation. Cecchini et al. (2018) evaluated environmental efficiency analysis and estimated CO_2 abatement costs in dairy cattle farms in Italy using an SBM model in the presence of undesirable outputs. Tamaki et al. (2016) evaluated the performance of public transport systems using three approaches: DEA, order-m and order-α, and they analysed the transport systems and environmental load of various cities by calculating the shadow prices of undesirable outputs (CO_2) emitted from public transport. Wang et al. (2019) combined an SBM model with environmental impacts as undesirable outputs with spatial analysis techniques to measure the environmental efficiency of 21 cities in China.

2.11 INTEGER-VALUED DATA

Conventional DEA models also assume all inputs and outputs have real values. Howeve, in many real-world applications, some inputs and outputs only have integer values. As an example, analyzing the efficiency of hospitals requires inputs such as the number of doctors and nurses, and outputs such as the number of surgeries. These attributes are integer-valued data (Du et al. 2012). Integer-valued data was first incorporated into DEA by Lozano & Villa (2006). They proposed a mixed integer linear programming (MILP) DEA model to guarantee the required integrality of the computed targets. Matin & Kuosmanen (2009) improved Lozano and Villa's model by composing a new axiomatic foundation, which resulted in a novel MILP DEA model that is consistent with the minimum extrapolation principle in the Banker-Charnes-Cooper model (Wu & Zhou 2015). Chen et al. (2012) incorporated undesirable factors into an integer-valued DEA to evaluate the operational efficiencies of city bus systems considering safety records. They also developed the proposed model using the super-efficiency approach to increase the discrimination power and the performance of efficient DMUs. Chen et al. (2013) modified the directional distance function (DDF) in order to integrate integer-valued data under the super-efficiency concept. Wu & Zhou (2015) proposed a mixed-objective integer DEA model to deal with input excesses and output shortfalls concurrently. Karimi, Khorram & Moeini

(2016) proposed a procedure to evaluate integer congestion (PEIC) in a set of DMUs. Khoveyni et al. (2019) presented various MILP-DEA models using the slack based congestion method. The proposed model first identifies DMUs with all the possible candidates to exhibit integer congestion. Then it detects the integer congestion status to the left and right of the candidate DMUs.

2.12 CHANCE-CONSTRAINED DEA

Chance-constrained programming (CCP) is a type of stochastic optimization method to deal with optimization problems with random variables included in constraints and objective function (Charnes & Cooper 1959). Olson & Swenseth (1987) discussed that CCP is a means to describe constraints in the form of probability levels of achievement. The consideration of chance constraints allows the decision-maker to consider objectives in terms of their attainment probability. Significant contributions to chance-constrained DEA (CCDEA) were made by (Sengupta 1982, 1987, 1990, 1997, 1998, 2000). A significant feature of the studies conducted by Sengupta is that stochastic variables can be integrated into DEA after whcih the DEA model is reformulated into a deterministic equivalent (Azadi & Saen 2011). Land, Lovell & Thore (1993) used the CCP concept to develop efficient frontiers in DEA. Olesen & Petersen (1995) presented a CCDEA model that uses a piecewise linear envelopment of confidence regions for observed stochastic multiple inputs and multiple outputs (Azadi & Saen 2011). Cooper, Huang & Li (1996) integrated satisficing concepts into CCDEA. Morita & Seiford (1999) discussed the efficiency analysis of DMUs in the presence of stochastic variations in inputs and outputs data. The obtained efficiency results of the CCDEA model proposed by Morita & Seiford (1999) were more reliable and robust in comparison with traditional DEA models. Sueyoshi (2000) proposed DEA future analysis to integrate stochastic outputs into an analytical framework. He used both the estimation technique of the Program Evaluation and Review Technique/Critical Path Method and CCP to reformulate the stochastic DEA model. Huang & Li (2001) developed two traditional DEA models based on random disturbances. Cooper et al. (2002) presented a series of CCP models to substitute traditional DEA formulations with stochastic counterparts. Azadi & Saen (2011) presented a CCDEA to deal with dual-role factors to evaluate and select third-party reverse logistics (3PL). Chen & Zhu (2019) relaxed the uncorrelation assumption in the CCDEA models. They developed CCDEA from the Gaussian model to a distributional robust model to deal with chance distributions of inputs and outputs.

2.13 NETWORK DEA

Despite the considerable advantages of conventional DEA models, they are unable to measure the efficiency of DMUs with network structures (Lewis & Sexton 2004). Färe & Grosskopf (1996) were among the first to take into account the internal structure of the DMUs in performance measurement and proposed network activity analysis models (Despotis, Koronakos & Sotiros 2016). Considering the internal structure of DMUs, Lewis & Sexton (2004) proposed a network DEA model that could be either input-oriented or output-oriented and allows for any of the four standard assumptions regarding returns to scale in any Sub-DMU, and makes adjustments for site characteristics in each Sub-DMU. Kao & Hwang (2008) considered the series relationship of the two divisions and showed that overall efficiency is a product of the efficiencies of these two divisions. Liang, Cook & Zhu (2008) examined and extended the two-stage processes where all outputs from stage one are the inputs to stage two using game theory concepts. Tone & Tsutsui (2009) presented a weighted SBM network DEA model which accounts for the importance of each division. The proposed model is able to gauge the overall efficiency and multi-divisional efficiencies in a unified framework. Li et al. (2012) extended the model proposed by Liang, Cook & Zhu (2008). They considered a two-stage DEA model in which the outputs of stage one and additional inputs to stage two are assumed as inputs for stage two. Moreno & Lozano (2014) proposed a network DEA model to measure the performance of NBA basketball teams. They also compared their obtained results with the single-stage DEA approach. Mirhedayatian, Azadi & Saen (2014) proposed a network DEA for the performance evaluation of green supply chain management in the presence of undesirable outputs and dual-role factor. Avkiran (2015) proposed a dynamic network DEA in commercial banking with an emphasis on testing robustness. Galagedera et al. (2018) presented a network DEA model to evaluate mutual fund (MF) performance in a multi-dimensional framework. They conceptualized MF management process as a serially linked three-stage process comprising resource management, operational management and portfolio management processes. Kalantary & Saen (2019) proposed a dynamic inverse network DEA based on the range adjusted measure (RAM) model. The proposed model changes both inputs and outputs of DMUs while efficiency scores remain unchanged. Chen, Cook &

Zhu (2020) introduced a conic relaxation model that searches for the global optimum to the general multiplier-based network DEA model. They reformulated the general network DEA models and relaxed the existing models into second-order cone programming (SOCP) problems.

2.14 FREE DISPOSAL HULL (FDH)

Classical DEA models evaluate the efficiency of DMUs using virtual DMUs in order to undertake a more accurate performance evaluation. Deprins & Simar (1984) proposed the free disposal hull (FDH) model which computes the efficiency of DMUs with an actual observed performance of DMUs (Ray 2004). Kerstens & Eeckaut (1999) proposed a more general way to determine CRS for both efficient and inefficient DMUs. Its efficacy was shown by considering variations on an FDH model. Cherchye, Kuosmanen & Post (2001) generalised the directional distance function (DDF) framework towards non-convex FDH efficiency analysis. They reformulated the profit interpretation of DDF for non-convex FDH efficiency scores. They also derived a general enumeration formula to compute DDF relative to the non-convex technology. Briec & Kerstens (2006) developed a series of deterministic non-convex technologies, nonparametric incorporating conventional CRS assumptions into the non-convex FDH technology to calculate output and graph measures of technical efficiency and indicate the general advantage of such a solution strategy by enumeration. Soleimani-Damaneh & Mostafaee (2009) discussed the estimation of CRS in FDH models and provided some stability intervals to preserve the CRS classification. They showed that the proposed approach can be obtained using a polynomial-time algorithm according to the computation of certain ratios of inputs and outputs without solving any mathematical programming problem. Soleimani-Damaneh & Reshadi (2007) proposed a polynomial-time algorithm to estimate CRS in FDH models, greatly reducing computational time. Diewert & Fox (2014) proposed a DEA model to decompose productivity growth for a panel of production units into explanatory factors according to the FDH approaches pioneered by Tulkens & Eeckaut (2006). Shiraz et al. (2015) presented a MAJ-FDH model which is oriented input-output and always feasible. Fukuyama & Shiraz (2015) proposed a number of cost-effectiveness measures on convex and non-convex FDH technologies. Fukuyama et al. (2016) investigated the basic monotonicity properties of least-distance (in) performance assessment on the FDH technologies. They showed that any known FDH least-distance measure violates strong monotonicity over the strongly Pareto-Koopmans efficient frontier. Tavakoli & Mostafaee (2019) developed FDH technology into a two-stage network DEA. In addition, they proposed an approach to compute the efficiency scores of DMUs.

2.15 QUALITY OF SERVICE

Quality of service (QoS) describes or measures the overall performance of a service such as a cloud computing service or computer network (Abdelmaboud et al. 2015; Hayyolalam & Kazem 2018). All the cloud services including Software as a Service (SaaS), Infrastructure as a Service (IaaS) and Platform as a Service (PaaS) have some QoS indicators for service evaluation. Some essential QoS indicators that are mostly used in the literature are discussed in this subsection. Service availability refers to the probability of receiving the proper service at any given time. It is usually expressed as service level agreements (SLAs) downtime in minutes per year or as the percentage of time the service will be up throughout the year. Thus, cloud service providers need to perform an availability analysis to quantify the expected downtime that the service may experience over a period of time (Jula, Sundararajan & Othman 2014). Reputation is another significant indicator for the performance assessment of cloud services. Ludwig et al. (2003) defined the reputation of a service as the reputation value resulting from the customers' feedback.

Wang, Yang & Mi (2015) and Chen et al. (2016) defined latency in cloud computing environments as a time interval between a customer request and a cloud computing service provider’s response or the time delay between the cause and the effect of some physical changes in the system. Throughput is another significant metric for the performance evaluation of cloud services. Wang et al. (2013) defined throughput as the total invocations of the service within a specified period of time. Levitin, Xing & Dai (2018) described cloud computing security as a set of technologies, applications, policies, and controls applied to protect virtualized IP, applications, data, services, and the associated infrastructure of cloud computing. It is a sub-domain of computer security, network security, and, more generally, information security. Price is also considered as a quantity metric which plays a significant role in the performance assessment of CSPs. Therefore, it is desirable to consider price with respect to the features related to CSPs (Filiopoulou et al. 2018). Reliability in cloud computing is the ratio of error messages to total messages (Bao & Dou 2012). Several scholars have addressed reliability in the literature (Younes, Essaaidi & El Moussaoui 2014; Zhang, Liu & Liu 2015; Zhou et al. 2014).

2.16 SUPPLY CHAIN

Kozlenkova et al. (2015) stated that “in business and finance, a supply chain is a system of organizations, people, activities, information, and resources involved in moving a product or service from supplier to customer. Supply chain activities involve the transformation of natural resources, raw materials, and components into a finished product that is delivered to the end customer. Cloud supply chain activities include providing computing infrastructure, software development platforms, and software to the end customer. In a cloud supply chain, IaaS is often provided to PaaS suppliers; PaaS suppliers deliver their services to SaaS suppliers; and all services can be delivered to cloud service customers. In terms of DEA, the three cloud services IaaS, PaaS, and SaaS are considered as three stages in the chain, while the providers are the decision-making units. Lindner et al. (2010) stated that a cloud supply chain is two or more parties linked by the provision of cloud services, related information and funds. Leukel, Kirn & Schlegel (2011) provided a completely different definition of a cloud supply chain as supply chain operations such as transportation, warehousing etc. as software-based services.

2.17. RESEARCH GAPS

In summary, although a great deal of work has been undertaken in evaluating and selecting CSPs, these methods and frameworks have various limitations and gaps. The existing approaches for performance measurement and the selection of CSPs are unable to find slight differences between a large number of CSPs in a highly intense competitive cloud marketplace. Moreover, these approaches are unable to evaluate the efficiency of CSPs requiring complex calculations, being effort-intensive, time-consuming, using subjective weights, and being inefficient for problems with large-scale and ranking irregularities. In addition, the previous research has developed and used different methods to evaluate the efficiency of CSPs. However, none are able to evaluate the CSPs in a supply chain as a unified system. Moreover, no existing model is able to provide customers with an optimal CSP composition given their QoS priorities, such as cost or latency. The techniques for evaluating and selecting CSPs have ranged from simple weighted scoring methods to advanced mathematical programming methods. However, despite the importance of undesirable outputs, integer-valued data, and stochastic data as part of an efficiency evaluation, these factors have not received attention and no studies address these conditions in terms of CSPs or DEA.

To address these gaps in the literature, this study presents a series of new DEA models. The proposed models benefit from a high discrimination power in the performance assessment of CSPs and are able to provide a complete ranking of all CSPs. Moreover, the proposed models can easily be computerized, enabling them to serve as a decision-making tool to assist decision makers. Furthermore, the proposed models are more objective and equitable than the existing methods for performance measurement and the selection of CSPs. The proposed models also consider both undesirable outputs and weight restrictions in a DEA model simultaneously. Moreover, the proposed models are able to evaluate the efficiency of CSPs in a cloud supply chain both in separate stages of the chain and overall using an integrated model and based on QoS indicators. In addition to this, the proposed models can evaluate the efficiency of CSPs in a cloud supply chain in the presence of undesirable outputs, integer-valued data and stochastic data.

2.18. SUMMARY

This chapter reviews the literature related to this study. First, the chapter provides a preliminary introduction to cloud computing, cloud layer architecture and services, infrastructure as a service, platform as a service and software as a service. Then service-level agreements, quality of service and cloud service providers were discussed. After this, the performance measurement of cloud service providers, DEA, the implementation of DEA in the cloud environment, the discrimination power in DEA, undesirable outputs, integer-valued data, chance-constrained DEA, network DEA, free disposable hull and supply chains were discussed. Finally, the research gaps in the literature were identified.

Chapter 3 PERFORMANCE MEASUREMENT OF CLOUD SERVICE SUPPLIERS AND CLOUD SUPPLY CHAINS

3.1 INTRODUCTION

Cloud computing is a paradigm in the IT area which provides a wide range of services to users such as on-demand computing resource access, dynamic and elastic scaling, virtualized resources and metered resource usage. Although cloud computing provides customers with numerous benefits, the evaluation and selection of a suitable CSP for a given task is a complicated process (Sun et al., 2019). This is because a number of factors need to be taken into account, such as quality of service, the nature of the data available for cloud service providers, the importance of the service to the customer, resource constraints and the relationship between criteria. Each aspect plays a key role in the evaluation and selection of cloud service providers. Due to the complexity of the task, there is no generally accepted? process for the evaluation and selection of cloud service providers; nevertheless, developing and proposing appropriate approaches or frameworks to assist the cloud computing user are still required (Devi & Shanmugalakshmi 2020; Duan 2017; Gireesha et al. 2020; Lang, Wiesche & Krcmar 2018; Ramachandran & Chang 2016).

Therefore, in this chapter, we propose a framework that addresses some key challenges in the evaluation and selection of cloud service providers. The framework comprises several models developed in this research for the performance evaluation. These evaluation models estimate the efficiency of cloud service providers and identify the optimal CSPs in terms of QoS. Five out of seven models evaluate the performance of a cloud service provider in one-stage structures while the other two evaluate them in network structures. It should be noted that in one-stage structures, we evaluate the performance of cloud service providers for one of the cloud services, such as IaaS, PaaS or SaaS. This is in the network structure, where the performance of the providers is evaluated in the context of an entire supply chain, where multiple services (IaaS, PaaS and SaaS) interact to achieve a business objective or goal. In the proposed models, we remove some barriers to the performance evaluation and selection process of cloud service providers such as complex calculations, being effort-intensive, being time-consuming, and having subjective weight and ranking irregularities. This framework supports cloud computing customers' decisions to select the optimal CSPs in a highly competitive market based on QoS.

3.2 THE PERFORMANCE MEASUREMENT FRAMEWORK

With respect to organizations' need for cloud computing service, a performance measurement framework for evaluating and selecting potential cloud service providers is required. The performance measurement framework should be able to assist cloud customers and decision makers to select optimal CSPs. Such a framework should improve the performance of cloud service providers by identifying inefficient resources in providing cloud computing services. Furthermore, some existing major barriers to performance measurement and the selection of cloud service providers in such a framework need to be taken into account and obviated, such as subjective weight and ranking irregularities.

The performance measurement framework proposed in this study addresses these requirements by proposing and applying a number of new models. The framework is designed to increase discrimination power and consider the decision-makers' viewpoints in terms of the performance evaluation of cloud service providers. The models proposed in this framework are more objective and equitable than the existing methods for performance measurement and the selection of CSPs. In addition, the network structure of cloud services is taken into account in the proposed framework, which can deal with undesirable outputs, integer-valued data and stochastic data.

In order to achieve these objectives, the following indicators are taken into consideration.

1. QoS indicators: Appropriate cloud service providers can be selected using a number of indicaors such as price, latency, avaiabilty, security, memory, storage, data transfer and CPU among others.
2. Performance measurement modeling of cloud service providers: According to Sarkis (2000), Peng Wong & Yew Wong (2008) and Park, Ok & Ha (2018), the best approach for evaluating the performance of a set of DMUs is DEA and network DEA. With respect to the objectives defined in

this study and the existing constraints in the performance measurement of cloud service providers, we develop models based on DEA and network DEA methods.

3. Considering the decision makers' viewpoints: The subjective judgements of decision makers and managers play a key role in the performance evalution of cloud service providers and need to be taken into account in the process.
4. Data type: In the performance evalution of cloud supply chains provided by cloud service providers, there are different types of data, such as undesirable data, integer-valued data, and stochastic data. The obtained results can be inaccurate or changed if these data are not considered in the performance measurement process. Based on these indicators and parameters, the following objectives are defined in the performance measurement framework:
    a) Increasing discrimination power in evaluating and selecting cloud service providers.
    b) Minimizing the impact of decision makers' subjective judgments in the performance measurement process.
    c) Considering different types of data in the performance evaluation of cloud service providers.
    d) Proposing an optimal composition of cloud service providers.

The performance measurement framework is designed according to the above objectives to address the significant issues in the performance evaluation of cloud service providers.

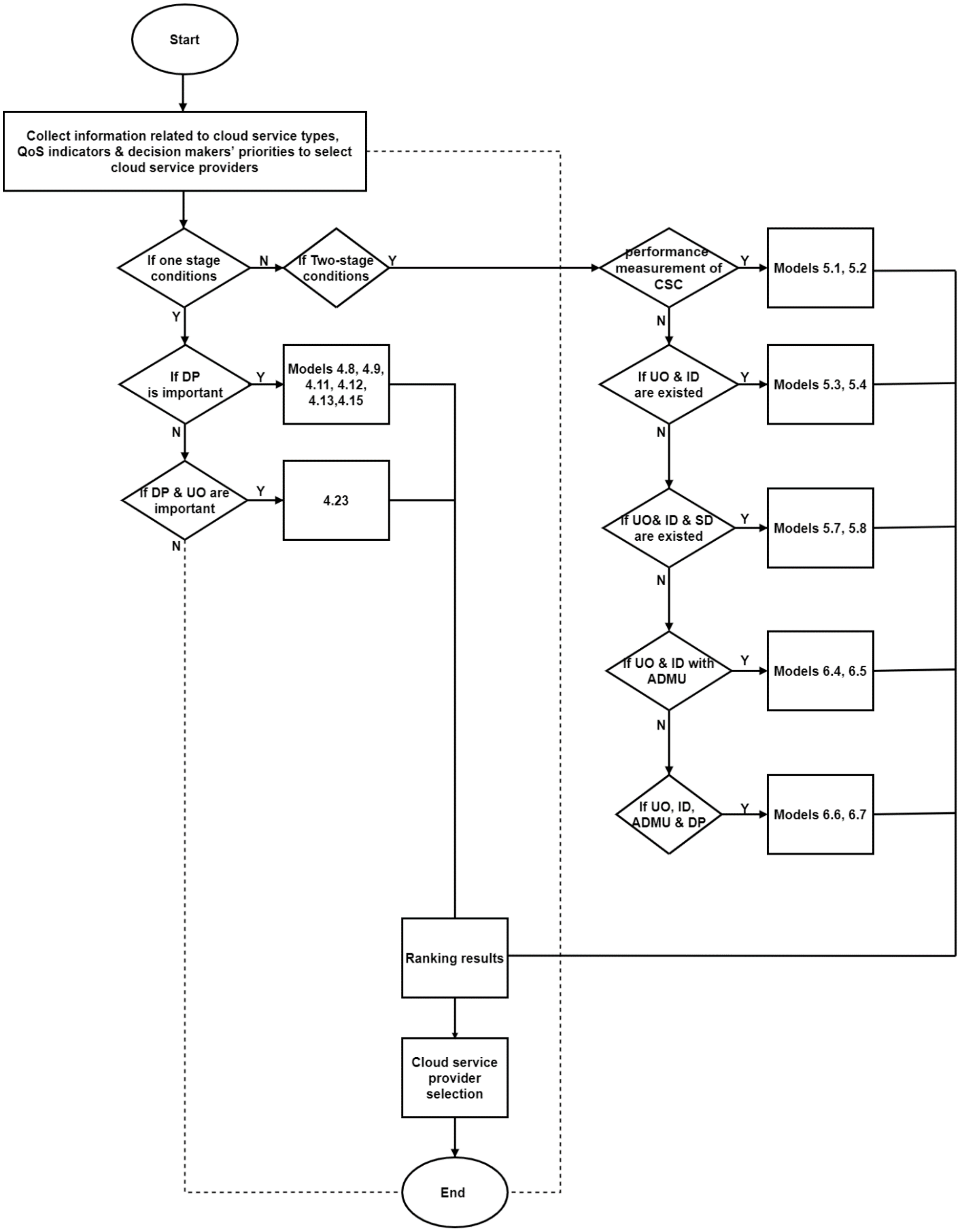

Figure 3.1 Performance evaluation framework of Cloud Service Providers

As shown in Figure 3.1, the performance evaluation framework of cloud service providers begins with collecting information related to cloud service types, QoS indicators and decision makers' priorities. Then, if decision makers deal with one-stage conditions and the discrimination power (DP) of cloud service providers is important, one of Models 4.8, 4.9, 4.11, 4.12, 4.13, 4.15 (as elaborated in Chapter 4) can be selected and applied. Furthermore, if both DP and undesirable outputs (UO) need to be taken into consideration in the performance evaluation of cloud service providers, Model 4.23 (as discussed in Chapter 4) can be applied. If decision makers deal with the network structures of cloud services in the performance evaluation process, Models 5.1 and 5.2 proposed in Chapter 5 are employed. Furthermore, in the presence of UO and integer data (ID) Models 5.3 and 5.4 are suitable in this situation. Models 5.7 and 5.8 are suitable for the performance measurement of cloud supply chains in the existence of undesirable outputs (UO), integer data (ID) and stochastic data (SD). In addition, while Models 6.4, 6.5 proposed in Chapter 6 are more objective than Models 5.7 and 5.8, Models 6.6 and 6.7 benefit from more discrimination power in comparison with Models 6.4 and 6.5. The performance evaluation framework proposed in this study has been designed based on real situations and different scenarios in cloud computing environments.

3.2.1 ONE-STAGE PERFORMANCE EVALUATION MODELS WITH A SIGNIFICANT ABILITY TO DISTINGUISH BETWEEN CLOUD SERVICE SUPPLIERS

As discussed, the majority of performance evaluation problems of cloud services are solved using MCDM and MCDA methods such as AHP, GRA and ANP. However, these approaches suffer from complex calculations, being time-consuming, being effort-intensive, and having subjective weight and ranking irregularities (Sun et al. 2019). Thus, advanced approaches and frameworks to evaluate and select CSPs need to be developed and applied to decrease the risk of an inappropriate CSP selection. On the other hand, DEA is one of the most important techniques to evaluate the performance of a set of decision-making units (DMUs). Because of the numerous advantages of DEA compared to other decision-making approaches, it has been developed and used for performance measurement and selection problems in many decision-making settings (Emrouznejad & Yang 2018; Mirhedayatian, Azadi & Saen 2014). Nevertheless, its applications in the cloud computing environment are scarce and in its infancy. In addition, the rapid development of cloud computing and the sharp increase in the number of CSPs have resulted in the majority of CSPs providing cloud services with highly competitive performance and prices. As a result, the current methods are unable to differentiate between CSPs which have very close or almost the same ranks. Moreover, all weights given to QoS criteria by decision makers for the evaluation and selection of CSPs are subjective measures, which results in an unfair assessment. Consequently, more objective methods with high discrimination power are needed to solve or address performance evaluation and selection problems.

In this study, to increase discrimination power in the performance evaluation of cloud service providers, mixed ideal and anti-ideal DEA models are developed. The proposed models are designed based on the distances between two special DMUs, namely the ideal DMU and the anti-ideal DMU. There are two advantages of the proposed ranking methods. First, they consider both pessimistic and optimistic scenarios of DEA, so they are more equitable than methods which are based on only one of these scenarios. The second strength of this approach is its discrimination power, enabling it to provide a complete ranking for all CSPs. Our proposed method can help customers to choose the most appropriate CSP while at the same time, it helps end users to identify inefficient CSPs to improve their performance in the marketplace. Furthermore, we propose a network DEA to measure the differences between the performance of CSPs. When network dimensions are taken into consideration, a more comprehensive analysis is enabled where divisional efficiency is reflected in overall efficiency estimates. This helps managers and decision makers in organizations to make accurate decisions in evaluating and selecting cloud service providers. We use a non-oriented network SBM model to measure the performance of cloud service providers. Finally, to increase the discrimination power in the performance evaluation of cloud service providers, a super-efficiency DEA model considering undesirable outputs and weight restrictions is proposed. The proposed model is not only able to rank all CSPs completely in the presence of undesirable outputs, it can also deal with decision makers' viewpoints in performance measurement in a more objective way.

3.2.2 EVALUATING EFFICIENCY IN CLOUD SUPPLY CHAINS

As discussed, considerable research has been undertaken to solve the problem of evaluating the efficiency of CSPs. Nevertheless, no study addresses the efficiency of providers in the context of an entire supply chain, where multiple services interact to achieve a business objective or goal. In this study, a cloud supply chain consists of two stages. Infrastructure as a Service (IaaS) is stage 1 and Platform as a Service (PaaS) is stage 2. Figure 5.3 in Chapter 5 illustrates a cloud supply chain. The inputs for the IaaS stage are price (stochastic), latency (undesirable/stochastic), memory, CPU (integer-valued), and data transfer which are used as the intermediate inputs/outputs. Price is also used as an additional input for the second stage (PaaS). The outputs for the PaaS stage are availability (stochastic), the number of security certifications (integer-valued), and service time delays (undesirable/stochastic). The current models for performance evaluation and the selection problem ignore the undesirable factors, integer-valued data, and stochastic data, such as latency, security certifications, and service prices, which can lead to inaccurate results. Furthermore, no existing approach to evaluate and select CSPs is able to provide customers with an optimal CSP composition given their QoS priorities, such as cost or availability in a cloud supply chain.

In this study, to addresses the efficiency of providers in the context of an entire supply chain, where multiple services interact to achieve a business objective or goal, a novel two-stage network DEA is proposed. The proposed model can consider undesirable outputs, integer-valued data, and stochastic data in the performance measurement of cloud supply chains provided by cloud service providers. Moreover, the model is developed further to be more objective and benefits from higher discrimination power. The proposed model is also able to provide the optimal composition of cloud service providers to suit a customer's priorities and requirements. In addition, the proposed model is able to evaluate the individual efficiency of providers in each stage of the supply chain and across the entire chain.

3.3 SUMMARY

In this chapter, we presented a framework for the performance measurement of cloud service providers. The framework can evaluate the performance of a cloud service provider in both one-stage structures and network structures under different conditions. Furthermore, it allows cloud service providers to identify their inefficient resources and improve their performance. The framework for the performance measurement of cloud service providers is introduced briefly in this chapter and will be elaborated in the next chapters.

Chapter 4 ONE-STAGE PERFORMANCE EVALUATION MODELS WITH A SIGNIFICANT ABILITY TO DISTINGUISH BETWEEN CLOUD SERVICE SUPPLIERS

4.1 INTRODUCTION

To select optimal CSPs which satisfy consumers' needs, cloud computing service customers need effective and powerful methods to evaluate the performance of CSPs. This chapter proposes DEA models to estimate the performance of CSPs. These models provide an opportunity for cloud computing service customers to evaluate and select the optimal CSPs based on QoS criteria in a highly competitive market. Three out of five proposed models in this chapter are based on the enhanced Russell model (ERM) (An et al. 2015; Halická & Trnovská 2018) to increase the discrimination power in the evaluation and selection of CSPs. The proposed models are designed based on the distances to two special decision-making units (DMUs), namely the ideal DMU and the anti-ideal DMU. The fourth model is a network DEA model, which increases discrimination power in the performance measurement of cloud service providers. The fifth model is a novel super-efficiency DEA model for evaluating and benchmarking CSPs in the presence of undesirable outputs and weight restrictions. The proposed model not only increases discrimination power in the performance evaluation of cloud service providers, it can also deal with undesirable outputs and decision makers' subjective judgments in the performance evaluation process. Also, in this chapter, each of the proposed methods is evaluated using a real data set and is compared with the methods in the literature to show the advantages of the proposed models.

4.2 A MIXED IDEAL AND ANTI-IDEAL DEA MODEL TO INCREASE DISCRIMINATION POWER IN THE PERFORMANCE EVALUATION OF CLOUD SERVICE PROVIDERS

In this section, we propose three novel models based on the enhanced Russell model (ERM) to increase the discrimination power in the evaluation and selection of CSPs. The proposed models are designed based on the distances to two special decision-making units (DMUs), namely the ideal DMU and the anti-ideal DMU. There are two advantages of the proposed ranking methods. First, they consider both pessimistic and optimistic scenarios of data envelopment analysis (DEA), so they are more equitable than methods that are based on only one of these scenarios. The second strength of this approach is its discrimination power, enabling it to provide a complete ranking for all CSPs. The proposed method can help customers to choose the most appropriate CSP while at the same time, it helps end users or cloud services consumers to identify inefficient CSPs to improve their performance in the marketplace.

4.2.1 PRELIMINARIES

The primary concepts that are applied in developing the new models are presented in this sub-section. Assume that there are n DMUs where each $DMU_j \, (j = 1,...,n)$ uses m inputs, $x_{ij} (i = 1, \dots, m)$ to produce s outputs, $y_{rj} (r = 1, \dots, s)$. Also, assume that the data set is positive and deterministic. The nomenclatures used in this paper are listed as follows:

Table 4.1 The nomenclatures

DMUp	:	DMU under evaluation
DMUj	:	*j*th DMU
m	:	Number of inputs
s	:	Number of outputs
x_{ij}	:	*i*th input of DMUj
y_{rj}	:	*r*th output of DMUj
λ_j	:	Intensity
v_i	:	Weight of *i*th input
u_r	:	Weight of *r*th output
ρ_p^*	:	ERM efficiency score for DMU*p*
θ_i	:	Input contraction
φ_r	:	Output extension
ψ_p^*	:	ERM efficiency score for DMUp from dual model
μ_i	:	Dual variable
ω_r	:	Dual variable
α	:	Dual variable
β	:	Dual variable

The non-radial ERM model is considered for measuring the relative efficiency of the DMU under evaluation, DMU_p, as follows (Pastor, Ruiz & Sirvent 1999):

$$\rho_p^* = \min \frac{\frac{1}{m}\sum_{i=1}^{m}\theta_i}{\frac{1}{s}\sum_{r=1}^{s}\varphi_r} \qquad)4.1($$

$$
\begin{aligned}
& s.t. \\
& \sum_{j=1}^{n}\lambda_j x_{ij} \le \theta_i x_{ip}, \qquad i = 1,\dots,m, \\
& \sum_{j=1}^{n}\lambda_j y_{rj} \ge \varphi_r y_{rp}, \qquad r = 1,\dots,s, \\
& \theta_i \le 1, \quad \varphi_r \ge 1, \qquad \forall i, r, \\
& \lambda_j \ge 0, \qquad j = 1,\dots,n.
\end{aligned}
$$

Definition 4.1. ERM-efficiency: Optimal ρ_p^* of the model (4.1) is called the ERM efficiency score of DMU_p. DMU_p is ERM efficient, if and only if $\rho_p^* = 1$. This condition is equivalent to $\theta_i^* = 1$ and $\varphi_r^* = 1$ for each i and r in any optimal solution (Pastor, Ruiz & Sirvent 1999).

4.2.2 DUAL LINEAR FORM OF THE ERM MODEL

Our proposed ideas are based on the multiplier form. The dual linear form of the ERM model is as follows:

$$
\begin{aligned}
& \psi_p^* = \max \ \alpha - \beta \\
& s.t. \\
& \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \qquad j = 1,\dots,n, && (4.2a) \\
& v_i x_{ip} - \mu_i \le \frac{1}{m}, \qquad i = 1,\dots,m, && (4.2b) \quad (4.2) \\
& \frac{\alpha}{s} - u_r y_{rp} + \omega_r \le 0, \qquad r = 1,\dots,s, && (4.2c) \\
& \sum_{i=1}^{m}\mu_i - \sum_{r=1}^{s}\omega_r - \beta \le 0, && (4.2d) \\
& v_i, u_r, \mu_i, \omega_r \ge 0, \qquad \forall i, r,
\end{aligned}
$$

$\alpha, \beta \geq 0.$

where '*' stands for the optimal solution and ψ_p^* is the optimal value of the objective function. This value is also known as the ERM efficiency of DMUp. In model 4.2, $v_i, u_r, \mu_i, \omega_r$ and also α and β are dual variables corresponding to the constraints of the linear version of model 4.1.

4.2.3 MODEL DEVELOPMENT

Usually, many DMUs are recognized as being efficient so it is important to be able to discriminate between them. Model 4.2 is a linear programming [3]model and has some useful properties. The following straightforward theorem shows this model is always feasible.

Theorem 4.1: Model 4.2 is always feasible.

Proof:

It is evident to see that the vector $(\alpha = 0, \beta = 0, v_i, u_r, \mu_i, \omega_r = 0;\ \forall i, r)$ is a feasible solution for Model 4.2.

The following theorem guarantees that the efficiency score obtained by Model 4.2 always lies between zero and one.

Theorem 4.2: In each optimal solution of Model 4.2, we always have $0 \leq \psi_p^* \leq 1$.

Proof:

As mentioned in Theorem 1, the vector $(\alpha = 0, \beta = 0, v_i, u_r, \mu_i, \omega_r = 0;\ \forall i, r)$ is a feasible solution for Model 4.2 and the objective value for this solution is $\psi_p = 0$. Since the model has a maximization form, then clearly, we conclude that $0 = \psi_p \leq \psi_p^*$.

[3] Linear programming or linear optimization is an approach for achieving the best outcome (such as maximum profit or lowest cost) in a mathematical model whose requirements are represented by linear relationships. It is a type of mathematical programming (Borgwardt & Viss 2020; Strandmark, Qu & Curtois 2020).

On the other hand, we sum all the relations (4.2b) and (4.2c) up over the i and r, respectively. The results are as follows:

$$(4.2b) \Rightarrow \sum_{i=1}^{m} v_i x_{ip} - \sum_{i=1}^{m} \mu_i \leq 1, \tag{4.3}$$

$$(4.2c) \Rightarrow \alpha - \sum_{r=1}^{s} u_r y_{rp} + \sum_{r=1}^{s} \omega_r \leq 0, \tag{4.4}$$

Considering (4.3), (4.4) and (4.2d), we have $\alpha - \beta + \sum_{i=1}^{m} v_i x_{ip} - \sum_{r=1}^{s} u_r y_{rp} \leq 1$, and according to (4.2a) we conclude that, always $\alpha - \beta \leq 1$, and thus $\psi_p^* \leq 1$. This fact completes the proof.

In addition to the above properties, this model fails to present a complete ranking of DMUs. In order to remove this difficulty and increase the power of discrimination among efficient DMUs, we further extend the ERM model and propose three new models. For this purpose, we define two special artificial DMUs, namely IDMU and AIDMU. We assume that the IDMU is an artificial DMU which uses the lowest number of inputs to produce the largest number of outputs. We denote it by $IDMU = (\overline{x}, \overline{y})$, where $\overline{x}_i = \min_{j=1,\dots,n} \{x_{ij}\}$ and $\overline{y}_r = \max_{j=1,\dots,n} \{y_{rj}\}$. Clearly, this DMU has the optimal performance. We assume that the AIDMU is an artificial DMU which uses the largest number of inputs to produce the lowest number of outputs. We denote it by $AIDMU = (\underline{x}, \underline{y})$, where $\underline{x}_i = \max_{j=1,\dots,n} \{x_{ij}\}$ and $\underline{y}_r = \min_{j=1,\dots,n} \{y_{rj}\}$. Clearly, this DMU has the worst performance. To increase the discrimination power of the efficient DMUs, we add another concept to the ERM model to present a complete ranking. For this purpose, we consider the distances of each DMU to IDMU and AIDMU. First of all, we define the distances.

The distance between DMU_p and the ideal DMU, i.e. DMU is denoted by d_p^+ and is defined as follows:

$$d_p^+ = \sum_{r=1}^{s} u_r (\overline{y}_r - y_{rp}) + \sum_{r=1}^{s} v_i (x_{ip} - \overline{x}_i), \tag{4.5}$$

It is obvious that the smaller the value of d_p^+, the closer the distance to the ideal point and therefore this translates to better performance. The distance between DMU_p and the anti-ideal DMU (AIDMU) is denoted by d_p^- and is defined as follows:

$$d_p^- = \sum_{r=1}^{s} u_r (y_{rp} - \underline{y}_r) + \sum_{r=1}^{s} v_i (\underline{x}_i - x_{ip}), \tag{4.6}$$

It is obvious that the larger the value of d_p^-, the farther the distance to the anti-ideal point which indicates a worthy performance. To use the two distances simultaneously, we introduce the concept of relative closeness. Therefore, for DMU_p, the relative closeness is defined as follows:

$$RC_p = \frac{d_p^-}{d_p^- + d_p^+} = \frac{\sum_{r=1}^{s} u_r (y_{rp} - \underline{y}_r) + \sum_{r=1}^{s} v_i (\underline{x}_i - x_{ip})}{\left(\sum_{r=1}^{s} u_r (y_{rp} - \underline{y}_r) + \sum_{r=1}^{s} v_i (\underline{x}_i - x_{ip})\right) + \left(\sum_{r=1}^{s} u_r (\bar{y}_r - y_{rp}) + \sum_{r=1}^{s} v_i (x_{ip} - \bar{x}_i)\right)}$$

$$\Rightarrow RC_p = \frac{\sum_{r=1}^{s} u_r (y_{rp} - \underline{y}_r) + \sum_{r=1}^{s} v_i (\underline{x}_i - x_{ip})}{\left(\sum_{r=1}^{s} u_r (\bar{y}_r - \underline{y}_r) + \sum_{r=1}^{s} v_i (\underline{x}_i - \bar{x}_i)\right)} \tag{4.7}$$

It is obvious that the larger the distance d_p^- and the smaller the distance d_p^+ means a larger RC_p index than before and it indicates the performance of DMU_p is close to the performance of the IDMU.

Based on these distances, we define three new modifications and develop the ERM model to increase its discrimination power.

4.2.4 THE ERM MODEL BASED ON DISTANCE TO IDEAL ASPECT (ERM-IDMU/METHOD 1)

The first modification in the ERM model is based on the closeness to ideal point 5 which can be stated as the following Model 4.8:

$$
\begin{aligned}
&\max \ \ (\alpha-\beta)-\varepsilon\left(\sum_{r=1}^{s} u_r(\overline{y}_r-y_{rp})+\sum_{r=1}^{s} v_i(x_{ip}-\overline{x}_i)\right)\\
&s.t.\\
&\sum_{r=1}^{s} u_r y_{rj}-\sum_{i=1}^{m} v_i x_{ij}\le 0, \qquad j=1,\dots,n,\\
&v_i x_{ip}-\mu_i\le \frac{1}{m}, \qquad i=1,\dots,m,\\
&\frac{\alpha}{s}-u_r y_{rp}+\omega_r\le 0, \qquad r=1,\dots,s,\\
&\sum_{i=1}^{m}\mu_i - \sum_{r=1}^{s}\omega_r-\beta\le 0,\\
&v_i,u_r,\mu_i,\omega_r\ge 0, \qquad \forall i,r,\\
&\alpha,\beta\ge 0.
\end{aligned}
\tag{4.8}
$$

As previously mentioned, the use of ε only has a theoretical justification. In practice, the calculations must be performed in two steps, and two models must be solved to obtain the optimal solution and identify the corresponding score for ranking. In the first stage, we solve Model 4.2. Assume that $\psi_p^*=(\alpha^*-\beta^*)$ is the optimal solution which shows the efficiency status of the DMU. Then, in the second stage, we solve the following Model 4.9:

$$
\begin{aligned}
&\xi_p^*=\min \ \sum_{r=1}^{s} u_r(\overline{y}_r-y_{rp})+\sum_{r=1}^{s} v_i(x_{ip}-\overline{x}_i)\\
&s.t.\\
&\sum_{r=1}^{s} u_r y_{rj}-\sum_{i=1}^{m} v_i x_{ij}\le 0, \qquad j=1,\dots,n,\\
&v_i x_{ip}-\mu_i\le \frac{1}{m}, \qquad i=1,\dots,m,\\
&\frac{\alpha^*}{s}-u_r y_{rp}+\omega_r\le 0, \qquad r=1,\dots,s,\\
&\sum_{i=1}^{m}\mu_i - \sum_{r=1}^{s}\omega_r-\beta^*\le 0,\\
&v_i,u_r,\mu_i,\omega_r\ge 0, \qquad \forall i,r
\end{aligned}
\tag{4.9}
$$

Assume that ξ_p^* is the optimal value of the objective function. We know that the less distance there is to IDMU, the more important the DMU_p is. So, we define the following index for ranking.

$$
\theta_p=(\psi_p^*,\delta_p^*) \tag{4.10}
$$

where $\delta_p^* = 1/\xi_p^*$. Index 10 has a lexicographic property. That is, to compare DMU_p and DMU_q there are two cases:

i. If $\psi_p^* > \psi_q^*$, then DMU_p has a better ranking than DMU_q

ii. If $\psi_p^* = \psi_q^*$ and if $\delta_p^* > \delta_q^*$, then DMU_p has a better ranking than DMU_q

4.2.5 THE ERM MODEL BASED ON DISTANCE TO ANTI-IDEAL ASPECT (ERM-AIDMU/METHOD 2)

The second modification in the ERM model is based on the distance from anti-ideal point 4 which can be stated as the following Model 4.11:

$$
\begin{aligned}
&\max\ (\alpha-\beta)-\varepsilon\left(\sum\nolimits_{r=1}^{s} u_r(y_{rp}-\underline{y}_r)+\sum\nolimits_{r=1}^{s} v_i(\underline{x}_i - x_{ip})\right)\\
&s.t.\\
&\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \qquad j=1,\ldots,n,\\
&v_i x_{ip} - \mu_i \le \frac{1}{m}, \qquad i=1,\ldots,m,\\
&\frac{\alpha}{s} - u_r y_{rp} + \omega_r \le 0, \qquad r=1,\ldots,s,\\
&\sum_{i=1}^{m} \mu_i - \sum_{r=1}^{s} \omega_r - \beta \le 0,\\
&v_i, u_r, \mu_i, \omega_r \ge 0, \qquad \forall i,r,\\
&\alpha,\beta \ge 0.
\end{aligned}
\tag{4.11}
$$

This model must be performed in two steps, and two models must be solved to obtain the optimal solution and identify the corresponding score for ranking. In the first stage, we solve Model 4.2. Assume that $\psi_p^* = (\alpha^* - \beta^*)$ is the optimal solution. Then, in the second stage, we solve the following Model 4.12:

$$
\begin{aligned}
&\sigma_p^* = \min \ \sum\nolimits_{r=1}^{s} u_r (y_{rp} - \underline{y}_r) + \sum\nolimits_{r=1}^{s} v_i (\underline{x}_i - x_{ip}) \\
&\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \qquad j = 1,\dots,n, \\
&v_i x_{ip} - \mu_i \le \frac{1}{m}, \qquad i = 1,\dots,m, \\
&\frac{\alpha^*}{s} - u_r y_{rp} + \omega_r \le 0, \qquad r = 1,\dots,s, \\
&\sum_{i=1}^{m} \mu_i - \sum_{r=1}^{s} \omega_r - \beta^* \le 0, \\
&v_i, u_r, \mu_i, \omega_r \ge 0, \qquad \forall i, r
\end{aligned} \tag{4.12}
$$

Assume that σ_p^* is the optimal value of the objective function. We know the greater the distance to AIDMU, the more important DMU_p is. So, we use the ranking index $\theta_p = (\psi_p^*, \delta_p^*)$, where $\delta_p^* = \sigma_p^*$.

4.2.6 THE ERM MODEL BASED ON MIXED ASPECTS (ERM-MAIDMU/ METHOD 3)

The third modification in the ERM model is based on the closeness to the ideal point and the distance from the anti-ideal point and can be stated as the following Model 4.13:

$$
\begin{aligned}
&\max \ (\alpha - \beta) - \varepsilon \left(\frac{w_1 \left(\sum\nolimits_{r=1}^{s} u_r (y_{rp} - \underline{y}_r) + \sum\nolimits_{r=1}^{s} v_i (\underline{x}_i - x_{ip}) \right)}{w_1 \left(\sum\nolimits_{r=1}^{s} u_r (y_{rp} - \underline{y}_r) + \sum\nolimits_{r=1}^{s} v_i (\underline{x}_i - x_{ip}) \right) + w_2 \left(\sum\nolimits_{r=1}^{s} u_r (\overline{y}_r - y_{rp}) + \sum\nolimits_{r=1}^{s} v_i (x_{ip} - \overline{x}_i) \right)} \right) \\
&s.t. \\
&\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \qquad j = 1,\dots,n, \\
&v_i x_{ip} - \mu_i \le \frac{1}{m}, \qquad i = 1,\dots,m, \\
&\frac{\alpha}{s} - u_r y_{rp} + \omega_r \le 0, \qquad r = 1,\dots,s, \\
&\sum_{i=1}^{m} \mu_i - \sum_{r=1}^{s} \omega_r - \beta \le 0, \\
&v_i, u_r, \mu_i, \omega_r \ge 0, \qquad \forall i, r, \\
&\alpha, \beta \ge 0.
\end{aligned} \tag{4.13}
$$

where w1 and w2 are the weights that show the importance of the distances to the ideal DMU and to the anti-ideal DMU, respectively. These values are determined by the decision maker (DM) and are based on their preferences over the aspects in relation to the three proposed models. Hence, the DM's opinion can be involved in the solution. This model must be performed in two steps, and two models must be solved in order to obtain the optimal solution and identify the corresponding score for ranking. In the first stage, we solve Model 4.2. Assume that $\psi_p^* = (\alpha^* - \beta^*)$ is the optimal solution. Then, in the second stage, we solve the following Model 4.14:

$$
\begin{aligned}
&\min \ \frac{w_1\left(\sum_{r=1}^{s} u_r(y_{rp} - \underline{y}_r) + \sum_{r=1}^{s} v_i(\underline{x}_i - x_{ip})\right)}{w_1\left(\sum_{r=1}^{s} u_r(y_{rp} - \underline{y}_r) + \sum_{r=1}^{s} v_i(\underline{x}_i - x_{ip})\right) + w_2\left(\sum_{r=1}^{s} u_r(\overline{y}_r - y_{rp}) + \sum_{r=1}^{s} v_i(x_{ip} - \overline{x}_i)\right)} \\
&s.t. \\
&\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \qquad j = 1,\dots,n, \\
&v_i x_{ip} - \mu_i \le \frac{1}{m}, \qquad i = 1,\dots,m, \\
&\frac{\alpha^*}{s} - u_r y_{rp} + \omega_r \le 0, \qquad r = 1,\dots,s, \\
&\sum_{i=1}^{m} \mu_i - \sum_{r=1}^{s} \omega_r - \beta^* \le 0, \\
&v_i, u_r, \mu_i, \omega_r \ge 0, \qquad \forall i, r
\end{aligned}
\tag{4.14}
$$

Based on the Charnes-Cooper translation and by introducing the following changes:

$$u_r = tu_r, \quad r = 1,\dots,s, \ ; \quad v_i = tv_i, \quad i = 1,\dots,m$$

$$\mu_i = t\mu_i, \ i = 1,\dots,m, \ ; \ \omega_r = t\omega_r, \ r = 1,\dots,s$$

then Model 4.15 can be stated as follows:

$$
\begin{aligned}
&\gamma_p^* = \min \ w_1 \left(\sum_{r=1}^{s} u_r (y_{rp} - \underline{y}_r) + \sum_{r=1}^{s} v_i (\underline{x}_i - x_{ip}) \right) \\
&s.t. \\
&w_1 \left(\sum_{r=1}^{s} u_r (y_{rp} - \underline{y}_r) + \sum_{r=1}^{s} v_i (\underline{x}_i - x_{ip}) \right) + w_2 \left(\sum_{r=1}^{s} u_r (\overline{y}_r - y_{rp}) + \sum_{r=1}^{s} v_i (x_{ip} - \overline{x}_i) \right) = 1 \\
&\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \qquad j = 1,\dots,n, \\
&v_i x_{ip} - \mu_i - \frac{t}{m} \le 0, \qquad i = 1,\dots,m, \\
&\frac{\alpha^*}{s} t - u_r y_{rp} + \omega_r \le 0, \qquad r = 1,\dots,s, \\
&\sum_{i=1}^{m} \mu_i - \sum_{r=1}^{s} \omega_r - t\beta^* \le 0, \\
&v_i, u_r, \mu_i, \omega_r, t \ge 0, \qquad \forall i, r
\end{aligned}
\tag{4.15}
$$

Assume that γ_p^* is the optimal value of the objective function. We define the ranking index $\theta_p = (\psi_p^*, \delta_p^*)$, where $\delta_p^* = \gamma_p^*$.

The results of the different aspects of the proposed model are not necessarily similar because different methods follow different viewpoints. In fact, efficient DMUs are always thought to perform better than inefficient DMUs. However, in some cases, when both efficient and inefficient DMUs are evaluated based on the worst possible relative efficiency, the efficient DMU will be scored as having poorer relative efficiency. Hence, in some situations, efficient DMUs are not considered to perform better than inefficient DMUs so the conclusion is obviously uncertain. Therefore, there is a clear need to combine the optimistic and the pessimistic viewpoints to arrive at a fair assessment of each DMU (Wang & Luo 2006).

4.2.7 ILLUSTRATION

Consider seven DMUs that use two inputs to produce one output of unity, as depicted in Table 4.2.

Table 4.2 Data set for the numerical example and the results obtained using different models.

DMUs	A	B	C	D	E	F	G	H	I
Input 1	1	2	5	10	1	2/3	4	8	12
Input 2	8	5	2	0.5	10	6	3	1.1	0.5
Output	1	1	1	1	1	1	1	1	1
CCR Efficiency	1	1	1	1	1	1	1	1	1
ERM Efficiency	1	1	1	1	0.9	1	1	1	0.917
AP Super Efficiency	1.1429	1.0755	1.1200	1.1500	1	1	1	1	1

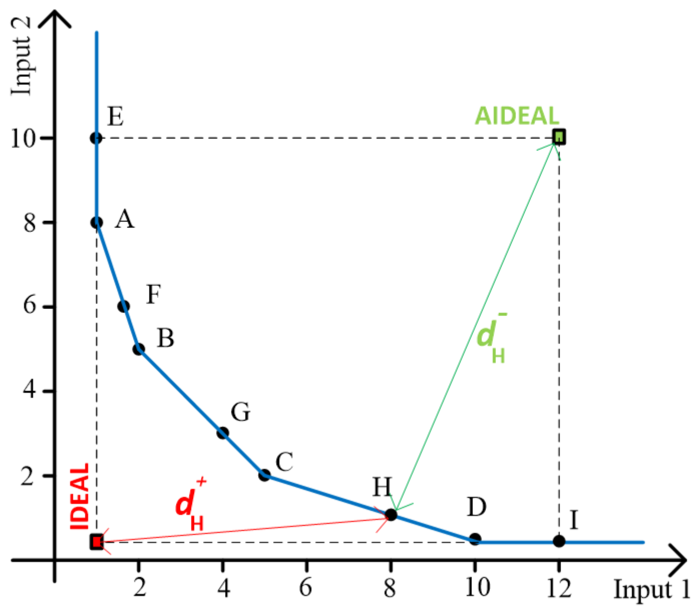

Figure 4.1 The efficiency status of the DMUs

Table 4.2 shows the data set for the illustrative example and the efficiency scores using CCR, ERM and AP models. Figure 4.1 illustrates the efficiency status of the DMUs. As can be seen in Table 4.2 and Figure 4.1, all the DMUs are efficient using the CCR model. Clearly, DMUs "E, I" are inefficient and the CCR model cannot report these inefficiencies. However, the ERM model is able to recognize these inefficiencies, but as shown in row 6 of Table 4.2, the ERM model fails to present a complete ranking of the efficient DMUs. Based on the results in the last row of Table 4.2, the AP method is not able to rank the non-extreme CCR efficient DMUs "E, F, G, H, I" and it fails to present a complete ranking. To increase the discrimination power of the DMUs, these DMUs are evaluated by applying the three proposed methods. The efficiency ranking results of the DMUs when applying the three proposed models are presented in Table 4.3.

Table 4.3 The results of our methods

Method		A	B	C	D	E	F	G	H	I
Ideal	Index	(1,0.990)	(1,3.560)	(1,1.823)	(1,0.016)	(0.9,0.297)	(1,2.310)	(1,2.027)	(1,1.799)	(0.917,3.25)
	Rank	6	1	4	7	9	2	3	5	8
Anti-Ideal	Index	(1,7.85)	(1,0.96)	(1, 7.06)	(1, 304.3)	(0.9, 4.41)	(1, 1.13)	(1, 0.95)	(1, 6.90)	(0.917, 43.04)
	Rank	1	7	2	4	9	5	6	3	8
Mixed	Index	(1,3.28)	(1,3.33)	(1, 1.50)	(1, 5.00)	(0.9, 2.71)	(1, 3.85)	(1, 1.53)	(1, 2.14)	(0.917, 1.79)
	Rank	4	3	7	1	9	2	6	5	8

As seen from Table 4.2, although in the mixed model we have $\delta_E^* = 2.71 > \delta_I^* = 1.79$, since we have $\psi_E^* = 0.9 < \psi_I^* = 0.917$, then we should have a better ranking for $\mathrm{DMU_I}$. This example shows that the proposed modified models give a complete ranking of the efficient DMUs. Furthermore, unlike other IDMU and AIDMU models, the preferences of decision-maker(s) have been taken into account in the models proposed in this study.

4.2.8 DATA AND EVALUATION OF THE PROPOSED MODELS

There are several research studies in the existing literature that are based on the concepts of ideal or anti-ideal DMUs. Table 4.3 presents a comparison of the structures and models in this area. From Table 4.3, we can see the proposed methodology has several advantages over other models. From Table 4.3, it is clear that our proposed methodology is more comprehensive than the other methods and considers more cases. In Table 4.3 ✓and × denote whether the different assessment aspects were considered or not considered, respectively.

Table 4.4 Comparison of the DEA models based on ideal and anti-ideal concepts

Studies	**Radial / Non-Radial**	**Complete Ranking**	**Using Ideal Point**	**Using Anti-Ideal Point**	**Optimistic efficiency**	**Pessimistic efficiency**	**Mixed aspect Model**	**Using Decision Maker's Opinion**	**Extra Description**
(Wang & Luo 2006)	Radial	×	√	×	√	×	×	×	-
(Rezaie et al. 2009)	Non-Radial	×	√	×	×	×	×	×	Non-Parametric
(Azizi & Ajirlu 2010)	Radial	×	√	√	√	√	×	×	Interval Efficiency
(Jahanshahloo et al. 2010)	Non-Radial	√	√	×	×	×	×	×	Common weights
(Jahanshahloo, Hosseinzadeh Lotfi, et al. 2011)	Non-Radial	√	√	×	×	×	×	×	Interval Efficiency
(Wang, Chin & Luo 2011)	Radial	√	√	√	√	√	×	×	Cross Efficiency
(Sun, Wu & Guo 2013)	Non-Radial	√	√	×	×	×	×	×	Common weights

(Barzegarinegad, Jahanshahloo & Rostamy-Malkhalifeh 2014)	Radial	×	√	×	×	×	×	×	Multi-Objective
(He et al. 2016)	Radial	×	√	×	×	×	×	×	Interval Efficiency
(Kritikos 2017)	Non-Radial	√	√	√	×	×	×	×	Common weights
(Shi, Wang & Chen 2019)	Non-Radial	√	√	√	×	×	×	×	Neutral cross-efficiency
The proposed models in this study	Non-Radial	√	√	√	√	√	√	√	ERM

The proposed models (IDMU, AIDMU and MAIDMU) are evaluated in this section. The selected companies are considered as DMUs. The research sample includes the top 82 IaaS providers. Of the 18 CSPs, Amazon Web Service (AWS), Microsoft Azure, IBM SoftLayer, etc[4] had the most data based on the QoS indicators. Hence, we removed the CSPs which did not have complete QoS data as the final research sample for all indicators after excluding missing values and incomplete indicators. The data are gathered from different resources such as websites, telephone calls and chats, and personal contact with sales employees who offered IaaS in 2017 and 2018. QoS signifies a number of non-functional attributes of services including availability, latency, price and security in the cloud domain. The data are described in Table 4.4. Two inputs are considered in the study, price and latency. The case study outputs consist of six items: memory, storage, data transfer, CPU, availability and the number of security certifications. It should be noted that due to limitations in collecting further information, in this study we selected two inputs. However, the models proposed in this research are able to provide accurate efficiency scores irrespective of the number of inputs and outputs. The values are presented in Table 4.5.

[4] Several CSPs asked the authors during the data collection process not to reveal their private information, hence we have removed the name of the CSPs from the tables in this study.

Table 4.5 Characteristics of 18 cloud service providers (CSPs) used for evaluation.

CSPs (DMUs)	Inputs		Outputs					
	Price (monthly/$)	Latency (ms)	Memory (GB)	Storage (GB)	Data transfer (TB)	CPU	Availability (Monthly)	The number of security certifications
1	80	433	8	80	5	2	100%	5
2	140.79	49	7	100	3.2	2	99.98%	3
3	80	46	8	80	5	4	100%	4
4	80	39	8	200	8	6	99.94%	1
5	158	45	2	500	0.5	4	100%	3
6	110	41	4	100	3	2	99.99%	4
7	150	68	16	384	8	6	99.994%	4
8	160	32	16.384	170	2	8	99.99%	1
9	156.24	40	2	40	10	2	100%	4
10	87.88	46	2.048	90	3	3	99.99%	2
11	16.65	152	0.5	20	0.5	1	99.89%	1
12	15	40	0.5	10	3	1	99.93%	1
13	79	71	8	80	5	2	100%	2
14	83.00	62	7	100	3	1	100%	4
15	64.95	62	4	250	3	2	100%	1
16	5	45	1	20	1	1	99.98%	1
17	219	46	8	300	10	8	99.74%	2
18	82.60	32	2	100	18	2	99.99%	1
Average	98.23	74.94	5.80	145.8	5.07	3.17	99.97	2.44
Std. Dev.	57.035	93.34	4.77	134.2	4.38	2.33	0.06	1.42
Min	5	32	0.5	10	0.5	1	99.74	1
Max	219	433	16.38	500	18	8	100	5
Ideal Point	5	32	16.384	500	18	8	100	5
Anti-Ideal Point	219	433	0.5	10	0.5	1	99.747	1

ms = milliseconds; TB = Terabyte; GB = Gigabyte

4.2.9 RESULTS AND DISCUSSION

The results of the ERM efficiency scores for the CSPs are presented in Table 4.5. Based on the ERM model, twelve *CSPs (i.e. 67%)* are recognized as being relatively efficient with a technical efficiency score of one, and the remaining six *CSPs (i.e. 33%)* are inefficient. The average *CSP* technical efficiency score is 0.802. Technical efficiency varies from 0.219 to 1 with a standard deviation of 0.29. The wide variation in technical efficiency indicates a substantial inefficiency between the *CSPs* in the studied area.

Table 4.6 ERM efficiency scores for the 18 CSPs

CSPs (DMUs)	**ERM Efficiency Score (ψ_j^*)**
1	0.219
2	0.478
3	1.000
4	1.000
5	1.000
6	1.000
7	1.000
8	1.000
9	1.000
10	0.454
11	0.224
12	1.000
13	0.537
14	0.528

15	1.000
16	1.000
17	1.000
18	1.000
Average	0.802
Std. Dev	0.299
Min	0.219
Max	1.000

DMU 1 had the lowest ERM efficiency score, $\psi_1^* = 0.219$. The main reason for this may be due to the fact that DMU 1 had the highest level of the second input 'latency', consuming nearly six times more than the average value. The second worst ERM efficiency score is assigned to DMU 11 at $\psi_{11}^* = 0.224$. From Table 4.5, it can be seen that DMU 11 has the lowest score for four outputs, which is the reason for its low efficiency score. In fact, the output value of DMU 11 is less than the average value.

The ERM results in a high number of efficient DMUs. We evaluate the CSPs by applying the proposed ranking models to increase the level of discrimination. DMUs can be ranked as an optimistic, pessimistic or a mixed viewpoint. In the case of an optimistic viewpoint, we use Model 4.8 and Model 4.9 to measure the distance to the ideal point.

The results of these models are shown in Table 4.6. DMU 7 was ranked in first position and is recognized as the optimal CSP in terms of distance to the ideal aspect. Also, of the efficient DMUs, DMU 12 has the furthest distance to the ideal point and so is placed in 12th position.

Table 4.7 Ranking based on the ideal aspect

CSPs (DMUs)	ξ_j^*	δ_j^*	$\theta_j = (\psi_j^*, \delta_j^*)$	Rank
1	1.3670	0.7315	(0.219, 0.7315)	18
2	2.1472	0.4657	(0.478, 0.4657)	15
3	2.6686	0.3747	(1, 0.3747)	3
4	2.1060	0.4748	(1, 0.4748)	2
5	11.266	0.0887	(1, 0.0887)	10
6	5.0840	0.1966	(1, 0.1966)	7
7	1.7914	0.5581	(1, 0.5581)	1
8	2.8079	0.3561	(1, 0.3561)	4
9	6.4389	0.1553	(1, 0.1553)	8
10	2.4259	0.4122	(0.454, 0.4122)	16
11	4.5372	0.2203	(0.224, 0.2203)	17
12	19.318	0.0517	(1, 0.051764)	12
13	4.2481	0.2353	(0.537, 0.2353)	13
14	3.5150	0.2844	(0.528, 0. 2844)	14
15	7.4699	0.1338	(1, 0. 1338)	9
16	11.375	0.0879	(1, 0. 0879)	11
17	2.9914	0.3342	(1, 0.3342)	5
18	3.5017	0.2855	(1, 0.2855)	6

In the case of the pessimistic viewpoint, Model 4.11 and Model 4.12, which measure the distance to the anti-ideal point, can be applied. The results of these models are shown in Table 4.7. DMU 12 was ranked in first position and is recognized as the optimal CSPs in terms of distance to the anti-ideal aspect.

Table 4.8 Ranking based on the anti-ideal aspect

CSPs (DMUs)	σ_p^*	δ_j^*	$\theta_j = (\psi_j^*, \delta_j^*)$	Rank
1	1.0156	1.0156	(0.219, 1.0156)	18
2	7.5524	7.5524	(0.478, 7.5524)	15
3	7.0524	7.0524	(1, 7.0524)	11
4	7.4191	7.4191	(1, 7.4191)	10
5	51.175	51.175	(1, 51.175)	2
6	43.283	43.283	(1, 43.28)	4
7	5.9633	5.9633	(1, 5.9633)	12
8	18.794	18.794	(1, 18.794)	7
9	43.331	43.331	(1, 43.331)	3
10	5.4805	5.4805	(0.454, 5.4805)	16
11	7.0196	7.0196	(0.224, 7.0196)	17
12	1.0202	1.0202	(1, 1.0202)	1
13	6.7804	6.7804	(0.537, 6.7804)	13
14	5.8937	5.8937	(0.528, 5.8937)	14
15	12.673	12.673	(1, 12.6739)	8
16	25.961	25.961	(1, 25.961)	5
17	23.203	23.203	(1, 23.203)	6
18	11.685	11.685	(1, 11.685)	9

Table 4.9 Ranking based on mixed aspects

CSPs (DMUs)	γ_p^*	δ_j^*	$\theta_j = (\psi_j^*, \delta_j^*)$	Rank
1	0.4855	0.4855	(0.219,0.4855)	18
2	0.8645	0.8645	(0.478,0.8645)	15
3	0.7584	0.7584	(1, 0. 7584)	6
4	0.7349	0.7349	(1, 0. 7349)	2
5	0.7583	0.7583	(1, 0. 7583)	7
6	0.9316	0.9316	(1, 0. 9316)	3
7	0.7410	0.7410	(1, 0. 7410)	8
8	0.7764	0.7764	(1, 0. 7764)	5
9	0.9362	0.9362	(1, 0. 9362)	2
10	0.7948	0.7948	(0.454, 0. 7948)	16
11	0.7273	0.7273	(0.224, 0. 7273)	17
12	0.8161	0.8161	(1, 0. 8161)	4
13	0.7420	0.7420	(0.537, 0. 7420)	13
14	0.7442	0.7442	(0.528, 0. 7442)	14
15	0.7349	0.7349	(1, 0. 7349)	10
16	0.7349	0.7349	(1, 0. 7349)	11
17	0.9376	0.9376	(1, 0. 9376)	1
18	0.7349	0.7349	(1, 0. 7349)	9

In the case of the mixed viewpoint, Model 4.13 and Model 4.15, which simultaneously consider the distance to the ideal point and the distance to the anti-ideal point, are applied. Here, the importance of the distance to the ideal point and the distance to the anti-ideal point are considered equally, i.e., w1=w2=0.5. It should be noted that by considering equal weights (w1=w2=0.5) for the distance to the ideal point and the distance to the anti-ideal point, cloud service providers are rated with better discrimination power. The results of these models are presented in Table 4.8. The ranking positions are shown in the last column of Table 4.8.

DMU 17 was ranked in the first position and is recognized as the optimal CSP in terms of distance to both the ideal and anti-ideal points. This DMU was ranked in 5th and 6th place, respectively for the ideal and anti-ideal aspects. Second place was assigned to DMU 7. Also, DMU 3 was ranked in the 12^{th} position of the efficient DMUs.

Table 4.9 confirms that our proposed model can rank DMUs based on the different aspects. Here, we compare the proposed method with the modified ERM method by Izadikhah, Farzipoor Saen & Ahmadi (2017), the well-known cross efficiency method and the AP method. Table 4.9 shows the results.

Table 4.10 Results of other ranking methods

	Efficiency Evaluation	**Ranking Methods**		
CSPs (DMUs)	ERM score	AP ranking	Cross-efficiency ranking	Modified ERM ranking
1	0.219	17	17	18
2	0.478	16	16	15
3	1.000	6	1	6
4	1.000	5	5	3
5	1.000	4	2	5
6	1.000	11	4	11
7	1.000	9	3	7
8	1.000	3	8	2
9	1.000	8	7	8
10	0.454	15	15	16
11	0.224	18	18	17
12	1.000	12	11	12
13	0.537	14	14	13
14	0.528	13	13	14
15	1.000	7	12	10
16	1.000	1	6	1

17	1.000	10	9	9
18	1.000	2	10	4

In addition to these results, our proposed model can use the decision-maker's opinion whereas none of the other existing methods has this ability. The Spearman's rank correlation coefficient between the mixed viewpoint of our method and the cross-efficiency ranking method is 0.800 with a p-value of 0.001765 which means there is a significant correlation coefficient between the results.

4.3 EFFICIENCY MEASUREMENT OF CLOUD SERVICE PROVIDERS USING NETWORK DEA

In this section, we propose a network DEA model to increase discrimination power in the performance of cloud service providers.

Tone (2001) proposed a slack-based measure (SBM) of efficiency in DEA. It can deal directly with the input excesses and the output shortfalls of the DMU. As Tone (2001) discussed, the SBM model is monotone and units are invariant decreasing with respect to input excess and output shortfall. Moreover, this measure is determined by consulting the reference set of the DMU and is not affected by statistics over the whole data set. Furthermore, the SBM model has a close connection with CCR, BCC and the Russell Measure (RM) models.

The input-oriented and output-oriented SBM models are given as follows.

4.3.1 SBM AND NETWORK SBM

Model 4.16

$$\rho_I^* = \min_{\lambda,s^-,s^+} 1 - \frac{1}{m}\sum_{i=1}^{m}\frac{s_i^-}{x_{io}},$$

s.t.

$$x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- \qquad (i = 1,\dots,m),$$

$$y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_i^+ \qquad (r = 1,\dots,s),$$

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

$\lambda_j \geq 0 \ (\forall j), \quad s_i^- \geq 0 \ (\forall j), \ s_r^+ \geq 0 \ (\forall j),$

Model 4.17

$$\frac{1}{\rho_o^*} = \max_{\lambda,\ s^-,\ s^+} 1 + \frac{1}{s}\sum_{r=1}^{s}\frac{s_r^+}{x_{io}},$$

s.t.

$$x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- \qquad (i = 1, \dots, m),$$

$$y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_i^+ \qquad (r = 1, \dots, s),$$

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

$\lambda_j \geq 0 \ (\forall j), \quad s_i^- \geq 0 \ (\forall j), \ s_r^+ \geq 0 \ (\forall j),$

The following model measures the efficency of DMUs using the non-oriented SBM model.

Model 4.18

$$\rho_{Io}^* = \min_{\lambda, s^-, s^+} \frac{1 - \frac{1}{m}\sum_{i=1}^{m}\frac{s_i^-}{x_{io}}}{1 + \frac{1}{s}\sum_{r=1}^{s}\frac{s_r^+}{x_{io}}},$$

s.t.

$$x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- \qquad (i = 1, \dots, m),$$

$$y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_i^+ \qquad (r = 1, \dots, s),$$

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

$\lambda_j \geq 0 \ (\forall j), \quad s_i^- \geq 0 \ (\forall j), \ s_r^+ \geq 0 \ (\forall j),$

Tone & Tsutsui (2009) proposed the SBM network model. According to the formulations in Tone & Tsutsui (2009), the SBM network model is not only able to estimate the overall DMU efficiency, it is also able to estimate divisional efficiency, which can increase discrimination power in the performance evaluation process. The relative efficiency technique that is utilized in this study is the non-oriented network slacks-based measure (SBM) for overall and divisional efficiency. The estimated efficiency for a CSP (DMU) is based on both input and output slacks (inefficiencies).

The non-oriented network SBM model is defined as follows:

Model 4.19

$$\rho_o^* = \min_{\lambda^k, s^{k-}, s^{k+}} \frac{\left[1 - \frac{1}{m_k}\left(\sum_{i=1}^{m_k} \frac{S_i^{k-}}{x_{io}^k}\right)\right]}{\left[1 + \frac{1}{r_k}\left(\sum_{r=1}^{r_k} \frac{S_r^{k+}}{y_{ro}^k}\right)\right]}$$

$x_o^k = x_{ij}{}^k\lambda^k + S^{k-} \qquad (k = 1, \dots, K),$

$y_o^k = y_{ij}{}^k\lambda^k - S^{k+} \qquad (k = 1, \dots, K),$

$z_o^{(k,h)} = z_{ij}{}^{k,h}\lambda^h \qquad (\forall(k,h)),$

$z_o^{(k,h)} = z_{ij}{}^{k,h}\lambda^k \qquad (\forall(k,h)),$

$e\lambda^k = 1 \qquad (k = 1, \dots, K),$

$\lambda^k \geq 0, \quad S^{k-} \geq 0, \quad S^{k+} \geq 0, \quad \forall k,$

We deal with n DMUs $j = (1, \dots, n)$ consisting of K divisions $(k = 1, \dots, K)$. Let m_k and r_k be the numbers of inputs and outputs to Division k, respectively. We denote the link leading from Division k to Division h by (k, h) and the set of links by L. The observed data are $\{x_j^k \in R_+^{m_k}\}$ $(j = 1, \dots, n; k = 1, \dots, K)$ (input resources to DMU_j at Division k), $\{y_j^k \in R_+^{r_k}\}$ $(j = 1, \dots, n; k = 1, \dots, K)$ (output products from DMU_j at Division k) and $\{z_j^{(k,h)} \in R_+^{t_{(k,h)}}\}$ $(j = 1, \dots, n;\ (k, h))$ (linking intermediate products from Division k to Division h) where $t_{(k,h)}$ is the number of items in Link (k, h).

Also the non-oriented network SBM model's divisional efficiency is defined as follows:

Model 4.20

$$\rho_k = \frac{1 - \frac{1}{m_k}\left(\sum_{i=1}^{m_k} \frac{S_i^{k-*}}{x_{io}^k}\right)}{1 + \frac{1}{r_k}\left(\sum_{r=1}^{r_k} \frac{S_i^{k-*}}{x_{io}^k}\right)} \quad (k = 1, \dots, K)$$

The network SBM model is a composite formulation of the input-oriented and output-oriented SBM models proposed by Tone (2001). Generally, the SBM network model can be formulated under input, output and non-oriented forms, and can be designated as constant returns to scale (CRS) or variable returns to scale (VRS) as demonstrated later in this study.

SBM is unit invariant and can accept variables measured in various dimensions, i.e. the optimal solution is not affected by variables measured in dissimilar units. Nonetheless, the SBM model is not translation invariant, denoting that the optimal solution will be impacted by data transformation that may be undertaken by researchers during the data collection. Lastly, SBM can accept all types of data including negative, zero or positive numbers for output variables; however it accepts only semi-positive data such as zero or positive numbers for input variables (Avkiran 2015; Cooper, Seiford & Tone 2006).

In the next sub-section, we use the network SBM model to increase discrimination power in the performance of cloud service providers.

4.3.2 EVALUATION OF THE PROPOSED METHODOLOGY

This section details our network DEA approach for selecting CSPs. In this study, the selected companies are considered to be DMUs. The condition of homogeneity has also been met in order to ensure a fair and comparable evaluation. We use the data set utilized in the previous section to evaluate the proposed network DEA approach for selecting CSPs. Descriptive statistics on data for the crucial input and output variables adjusted for computing services in IaaS are shown in Table 4.10. As above, two inputs are considered in this study, price and latency. The case study outputs consist of six items, memory, storage, data transfer, CPU, availability and the number of security certifications.

Table 4.11 Attributes of the 18 cloud service providers (CSPs) and their values

CSPs	Inputs		Intermediate inputs/ outputs			Outputs		
	Price (monthly/$)	Latency (ms)	Memory (GB)	Storage (GB)	CPU	Availability %	The number of security certifications	Data transfer (TB)
1	80	433	8	80	2	100	5	5
2	140.79	49	7	100	2	99.9898	3	3.2
3	80	46	8	80	4	100	4	5
4	80	39	8	200	6	99.9453	1	8
5	158	45	2	500	4	100	3	0.5
6	110	41	4	100	2	99.9987	4	3
7	150	68	16	384	6	99.994	4	8
8	160	32	16.384	170	8	99.9993	1	2
9	156.24	40	2	40	2	100	4	10
10	87.88	46	2.048	90	3	99.9968	2	3
11	16.65	152	0.5	20	1	99.8938	1	0.5
13	15	40	0.5	10	1	99.9303	1	3
14	79	71	8	80	2	100	2	5
14	83.00	62	7	100	1	100	4	3
15	64.95	62	4	250	2	100	1	3
16	5	45	1	20	1	99.9876	1	1
17	219	46	8	300	8	99.7473	2	10
18	82.60	32	2	100	2	99.999	1	18

Sarkis (2007) stated that the DEA method and its appropriate applications are greatly dependent on the data set that is used as an input to the productivity model. Although there are numerous models based on DEA, some data have certain characteristics that may not be acceptable for the execution of DEA models. One of them is normalized data. To be more precise, one of the best ways to make sure there is not much imbalance in the data sets is to ensure that they are at the same or similar magnitude. One way to address this is to mean normalize the data. There are two steps in the mean normalizing process: first, finding the mean of the data set for each input and output; and second to divide each input or output by the mean for that specific factor. The normalized data are given in Table 4.11

Table 4.12 The normalized data for the attributes of the 18 CSPs

CSPs	**Inputs**		**Intermediate inputs/ outputs**			**Outputs**		
	Price (monthly/$)	Latency (ms)	Memory (GB)	Storage (GB)	CPU	Availability	The number of security certifications	Data transfer (TB)
1	0.814	5.777	1.378	0.548	0.631	1.0002	2.045	0.986
2	1.433	0.653	1.206	0.685	0.631	1.0001	1.227	0.631
3	0.814	0.613	1.378	0.548	1.263	1.0002	1.636	0.986
4	0.814	0.520	1.378	1.371	1.894	0.9997	0.409	1.578
5	1.608	0.600	0.344	3.429	1.263	1.0002	1.227	0.098
6	1.119	0.547	0.689	0.685	0.631	1.0002	1.636	0.592
7	1.527	0.907	2.757	2.634	1.894	1.0002	1.636	1.578
8	1.628	0.426	2.823	1.166	2.526	1.0002	0.409	0.394
9	1.590	0.533	0.344	0.274	0.631	1.0002	1.636	1.973
10	0.894	0.613	0.352	0.617	0.947	1.0002	0.818	0.592
11	0.169	2.028	0.086	0.137	0.315	0.9992	0.409	0.098
13	0.152	0.533	0.086	0.068	0.315	0.9995	0.409	0.592
14	0.804	0.947	1.378	0.548	0.631	1.0002	0.818	0.986
14	0.844	0.827	1.206	0.685	0.315	1.0002	1.636	0.592
15	0.661	0.827	0.689	1.714	0.631	1.0002	0.409	0.5921
16	0.050	0.600	0.172	0.137	0.315	1.0001	0.409	0.197
17	2.229	0.613	1.378	2.057	2.526	0.9977	0.818	1.973
18	0.840	0.426	0.344	0.685	0.631	1.0002	0.409	3.552

4.3.3 RESULTS AND DISCUSSION

First, we calculate the efficiency scores for 18 CSPs using model 4.18 with the assumptions of CRS and VRS with two inputs and six outputs avoiding the links between divisions. This is shown in the last two columns in Table 4.12 and in red and yellow in Figure 4.2. Then, we calculate the efficiency scores of the CSPs using the network SBM model, which are shown in columns 2, 3, and 4 in Table 4.12. In addition, the overall efficiencies of the CSPs (column 4) obtained using network SBM are shown in Figure 4.2 in blue.

Table 4.13 The results

CSPs	***Stage 1 efficiency***	***Stage 2 efficiency***	***Network SBM (Overall efficiency)***	***SBM (CRS[5])***	***SBM (VRS[6])***
1	*0.517*	*1*	*0.758*	*0.1271652*	*1*
2	*0.642*	*0.503*	*0.572*	*0.3300747*	*1*
3	*0.896*	*0.750*	*0.823*	*1*	*1*
4	*1*	*0.468*	*0.734*	*1*	*1*
5	*1*	*0.134*	*0.567*	*1*	*1*
6	*0.690*	*0.562*	*0.626*	*1*	*1*
7	*1*	*0.923*	*0.961*	*1*	*1*
8	*1*	*0.435*	*0.717*	*1*	*1*
9	*0.587*	*1*	*0.793*	*1*	*1*
10	*0.662*	*0.421*	*0.541*	*0.2949357*	*1*
11	*0.298*	*0.375*	*0.336*	*0.1259881*	*1*
12	*1*	*1*	*1*	*1*	*1*
13	*0.753*	*0.600*	*0.676*	*0.3853003*	*1*
14	*0.7213567*	*1*	*0.860*	*0.3615719*	*1*
15	*1*	*0.360*	*0.680*	*1*	*1*
16	*1*	*1*	*1*	*1*	*1*
17	*1*	*0.749*	*0.874*	*1*	*1*
18	*1*	*1*	*1*	*1*	*1*

[5] C*onstant returns to scale. For more details, please see Cooper et al. (2011).*

[6] Variable returns to scale. *For more details, please see Avkiran, (2015).*

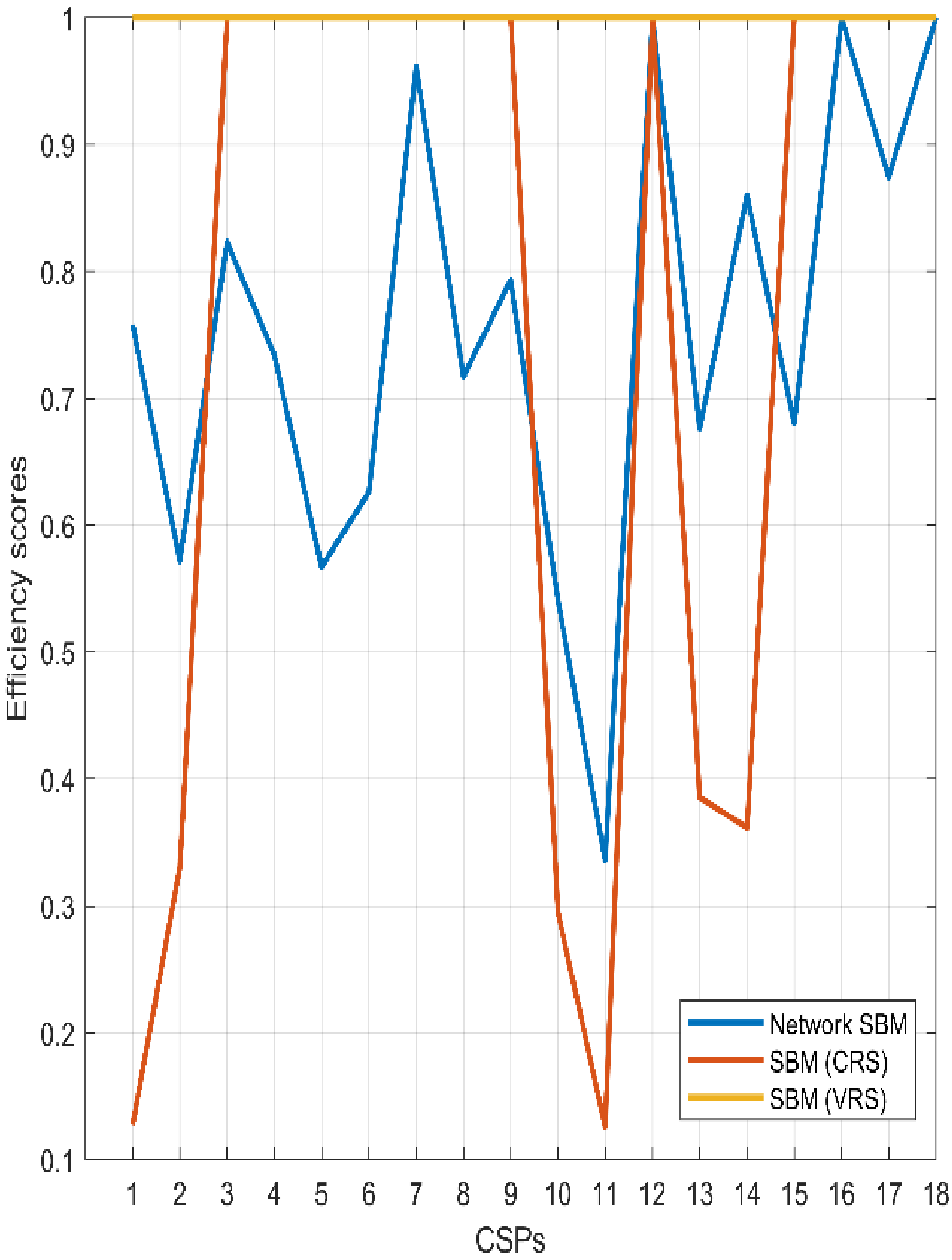

Figure 4.2 The obtained results using network DEA, SBM (CRS) and SBM (VRS).

The comparison results presented in Table 4.12 show that all the CSPs are efficient with the assumption of SBM (VRS). In addition, Figure 4.2 in yellow demonstrates that all the CSPs are on the efficiency frontier, which means the decision maker is unable to discern differences between CSPs. Furthermore, the SBM (VRS) results in Table 4.12 and Figure 4.2 show that while CSP 11 has the lowest efficiency, a large number of CPSs (CSPs 3, 4, 5, 6, 7, 8, 9, 12, 15, 16, 17 and 18) are efficient and are on the efficiency frontier. As seen, the SBM model with VRS and CRS assumptions cannot differentiate between most of the CSPs. In such conditions, many researchers implement outlier detection analysis, such as super efficiency, to rank the DMUs.

In addition, the results provided in Table 4.12 and Figure 4.2 show that 3 of the 18 CSPs (CSPs 12, 16 and 18) are efficient using the network SBM model, and that CSP 11 has the lowest efficiency among CSPs. In general, the efficiency scores using conventional DEA models tend to be higher than those of the network DEA models. This is mainly due to differences between the number of inputs and outputs in the two models. As previously mentioned, the traditional DEA models do not consider the links and the gap between the conventional and network models which implies the ‘‘networking effects”. As Tone and Tsutsui (2009) discussed, the VRS models have at least one efficient DMU within each division. In this case study, there are 11 efficient CSPs out of 18. However, CSP 12, 16 and 18 are the only CSPs that are efficient in both stages one and two. This means the other CSPs need to improve their efficiency at least in one stage to be efficient. Column 5 in Table 4.12 shows that CSPs 1 and 11 have the lowest efficiency value compared to the other CSPs. These inefficiencies are mainly related to their high latency. However, in stage two, CSP 5 had the lowest efficiency value mainly due to having a lower amount of data transfer compared to the other CSPs.

4.4 A NEW RUSSELL DIRECTIONAL DISTANCE FUNCTION (DDF) MODEL FOR PERFORMANCE EVALUATION OF CLOUD SERVICE PROVIDERS

In this section, to increase discrimination power and deal with undesirable outputs and decision makers' subjective judgments in the performance evaluation of cloud service providers, we propose a new super-efficiency Russell directional distance function (SRDDF) DEA model. The model has several salient features as follows:

1. The suggested approach has the ability to differentiate between CSPs even if they have been given the same rank or are rated the same using current methods.
2. The proposed approach is more flexible to select a similar directional vector or specific vector for all DMUs or each DMU respectively;
3. It considers undesirable outputs and weight restrictions simultaneously to measure the performance of CSPs;
4. It provides a more objective performance benchmark assessment; and
5. It can easily be computerized to serve as a decision-making tool.

4.4.1 PROPOSED MODEL

Table 4.13 presents the nomenclatures used in this paper.

Table 4.14 The nomenclatures

$j = 1, \dots, n$	Collection of CSPs (DMUs)
$r = r, \dots, s$	The set of outputs
$i = 1, \dots, m$	The set of inputs
x_{ij}	The ith input of DMU_j
y_{rj}	The rth output of DMU_j
x_{io}	The ith input of DMU_o

y_{ro}	The rth output of DMU_o
$g_i^x = \theta_i$	The ith variable for input reduction
$g_r^y = \varphi_r$	The rth variable for output increase
μ_j	Intensifier variable for the nonactive part of inputs
λ_j	Intensifier variable for the active part of inputs
$\alpha_{t_1}, \alpha_{t_2}$	Intensifier variable for trade-offs
$m_{rt_2}^1, m_{rt_2}^2$	Matrices of trade-offs
t	Positive variable used for linearization
d_i^x	The ith element of input reduction vector.
d_r^y	The rth element of output increment vector.
ρ	Objective function
$s_1+s_2=s$	Two subsets of outputs whose sum is equal to s
$m_1+m_2=m$	Two subsets of inputs whose sum is equal to m

Assume that there are *n* DMUs, which consume *m* inputs to produce *s* outputs. Here, a model for dealing with the undesirable outputs, the homogeneous weight restriction among outputs, and ratio data is introduced. In the proposed model, the undesirable outputs are considered by the weak disposability principle. Also, to take into account the manager's viewpoint on the weights of outputs, homogeneous weight restrictions for the outputs is used. The homogeneous weight restrictions are shown by the trade-off principle in the envelopment form. *M* shows the trade-off matrix. To force the benchmarks to be less than or equal to 100%, an extra constraint is added for the ratio data in model (4.22). However, since the inputs are reduced, the extra constraint is redundant and can be ignored. The objective function of model (4.22) is the sum of the changes in the inputs and outputs to project the DMU under evaluation on the Pareto efficiency frontier. Otherwise, the DMU under evaluation is Pareto efficient. Thus, if the objective function of model (4.22) is equal to zero, then the DMU under evaluation is Pareto efficient. If the objective function is positive, the DMU under evaluation is inefficient.

To introduce the efficiency measurement function, the directional distance function (DDF) is used. The DDF helps to project an inefficient DMU on the efficiency frontier. By selecting the direction $\begin{pmatrix} -d_i^x \\ d_r^y \end{pmatrix} = \begin{pmatrix} -x_{io} \\ y_{ro} \end{pmatrix}$, the inefficient point $\begin{pmatrix} x_{io} \\ y_{ro} \end{pmatrix}$ is projected on the efficiency frontier. In model (1), ρ is function of the variables g_i^x, g_r^y.

$$Max \quad \rho(g_i^x, g_r^y)$$

$$s.t. \quad (x_{io} - g_i^x d_i^x, y_{ro} + g_r^y d_r^y) \in PPS, \; i = 1, \dots, m; \; r = 1, \dots, s$$

(4.21)

where d is the direction and g is a variable. Also, $g_o^y = \varphi_r$, $d_r^y = y_{ro}$, $g_o^x = \theta_i$, and $d_i^x = x_{io}$. As is seen, there are two types of inputs, including real inputs and ratio inputs (m_1+m_2=m). Also, there are two types of outputs, including the desirable and undesirable outputs. The undesirable outputs are based on weak feasibility and their associated constraints have an equal sign. Note that s_1+s_2=s.

$$\begin{aligned}
\min \quad & \rho = \frac{1 - \frac{1}{m}\sum_{i=1}^{m} \theta_i}{1 + \frac{1}{s}\sum_{r=1}^{s} \varphi_r} \\
s.t. \quad & \sum_{j=1}^{n} (\lambda_j + \mu_j) x_{ij} = x_{io} - \theta_i x_{io}, && i = 1, \dots, m_1 \\
& \sum_{j=1}^{n} (\lambda_j + \mu_j) x_{ij} = x_{io} - \theta_i x_{io}, && i = 1, \dots, m_2, \\
& \sum_{j=1}^{n} \lambda_j y_{rj} + \sum_{t_1=1}^{g_1} \alpha_{t_1} m_{rt_2}^1 = y_{ro} + \varphi_r y_{ro}, && r = 1, \dots, s_1, \qquad (4.22) \\
& \sum_{j=1}^{n} \lambda_j y_{rj} + \sum_{t_2=1}^{g_2} \alpha_{t_2} m_{rt_2}^2 = y_{ro}, && r = 1, \dots, s_2, \\
& \theta_i \geq 0, \varphi_r \geq 0, && for\ all\ i, r, \\
& \sum_{j=1}^{n} (\lambda_j + \mu_j) = 1, \qquad \lambda_j \geq 0, \mu_j \geq 0, && for\ all\ j.
\end{aligned}$$

Model (4.22) is nonlinear. To linearize model (4.22), the following change of variables is used. The α_{t_1} and α_{t_2} are the intensity coefficients of trade-off matrices m^1 and m^2.

$t\theta_i = \theta_i, t\varphi_r = \varphi_r, t\lambda_j = \lambda_j, t\mu_j = \lambda_j, t\alpha_{t_1} = \alpha_{t_1,} t\alpha_{t_2} = \alpha_{t_2}\ for\ all\ i, j, r, t_1, t_2$
(4.23)

Model (4) is the linear version of model (2).

$$
\begin{aligned}
\min\quad & \rho = t - \frac{1}{m}\sum_{i=1}^{m} \theta_i \\
s.t.\quad & \sum_{j=1}^{n} (\lambda_j + \mu_j) x_{ij} = t x_{io} - \theta_i x_{io}, && i = 1, \dots, m_1 \\
& \sum_{j=1}^{n} (\lambda_j + \mu_j) x_{ij} = t x_{io} - \theta_i x_{io}, && i = 1, \dots, m_2, \\
& \sum_{j=1}^{n} \lambda_j y_{rj} + \sum_{t_1=1}^{g_1} \alpha_{t_1} m_{rt_2}^{1} = t y_{ro} + \varphi_r y_{ro}, && r = 1, \dots, s_1, \qquad (4.24) \\
& \sum_{j=1}^{n} \lambda_j y_{rj} + \sum_{t_2=1}^{g_2} \alpha_{t_2} m_{rt_2}^{2} = t y_{ro}, && r = 1, \dots, s_2, \\
& 1 + \frac{1}{s}\sum_{r=1}^{s} \varphi_r = t, \theta_i \geq 0, \varphi_r \geq 0, \mathrm{t} > 0, && for\ all\ i, r, \\
& \sum_{j=1}^{n} (\lambda_j + \mu_j) = t, \quad \lambda_j \geq 0, \mu_j \geq 0, && for\ all\ \ j.
\end{aligned}
$$

Theorem 1: The DMU$_o$ is Pareto efficient if and only if $\rho^* = 0$.

Proof: If DMU$_o$ is Pareto efficient, then there is no way to improve the inputs and outputs of DMU$_o$. Thus, for each i and *r*, we have $\theta_i^* = 0$ and $\varphi_r^* = 0$. If $\rho^* = 0$, then $\sum_{i=1}^{m} \theta_i^* + \sum_{r=1}^{s} \varphi_r^* = 0$. Since $\theta_i \geq 0$ and $\varphi_r \geq 0$, we have $\theta_i^* = 0$. $\varphi_r^* = 0$ implies that DMU$_o$ is Pareto efficient. □

To rank the DMUs, the following super-efficiency approach is proposed. In the super-efficiency approach, the DMU$_o$ is removed and the least distance of DMU$_o$ until the new PPS is measured. If the objective function of Model (4.25) is bigger than 1, then the associated DMU$_o$ is super-efficient.

$$
\begin{aligned}
\min \quad & \rho = t + \frac{1}{m}\sum_{i=1}^{m}\theta_i \\
s.t. \quad & \sum_{\substack{j=1 \\ j\neq o}}^{n} (\lambda_j + \mu_j)x_{ij} = tx_{io} + \theta_i x_{io}, && i = 1,\dots,m_1 \\
& \sum_{\substack{j=1 \\ j\neq o}}^{n} (\lambda_j + \mu_j)x_{ij} = tx_{io} + \theta_i x_{io}, && i = 1,\dots,m_2, \\
& \sum_{\substack{j=1 \\ j\neq o}}^{n} \lambda_j y_{rj} + \sum_{t_1=1}^{g_1} \alpha_{t_1} m_{rt_2}^{1} = ty_{ro} - \varphi_r y_{ro}, && r = 1,\dots,s_1, \\
& \sum_{\substack{j=1 \\ j\neq o}}^{n} \lambda_j y_{rj} + \sum_{t_2=1}^{g_2} \alpha_{t_2} m_{rt_2}^{2} = ty_{ro}, && r = 1,\dots,s_2, \\
& 1 - \frac{1}{s}\sum_{r=1}^{s}\varphi_r = t, \theta_i \geq 0, \varphi_r \geq 0, \mathrm{t} > 0, && for\ all\ i, r, \\
& \sum_{\substack{j=1 \\ j\neq o}}^{n} (\lambda_j + \mu_j) = t, \quad \lambda_j \geq 0, \mu_j \geq 0, && for\ all\ \ j.
\end{aligned}
\tag{4.25}
$$

4.4.2 EXPERIMENTATION AND RESULTS

To evaluate our model (Model 4.25), we assessed its efficacy on a dataset of CSPs (i.e., DMUs). The data set in the previous sections is used for evaluation as listed in Table 4.14. Two inputs, price and availability, and six outputs, memory, storage, data transfer, CPU, availability, and the number of security certifications were considered. Latency was the undesirable output. According to the decision maker, the importance of some indicators such as price is expressed by giving it a weight.

Table 4.15 Characteristics of 18 CSPs

CSPs	Inputs		Outputs					
	Price (monthly/$)	Availability (Monthly	Memory (GB)	Storage (GB)	Data transfer (TB)	CPU	Latency (ms)	The number of security certifications
1	80	100%	8	80	5	2	433	5
2	140.79	99.9898%	7	100	3.2	2	49	3
3	80	100%	8	80	5	4	46	4
4	80	99.9453%	8	200	8	6	39	1
5	158	100%	2	500	0.5	4	45	3
6	110	99.9987%	4	100	3	2	41	4
7	150	99.994%	16	384	8	6	68	4
8	160	99.9993%	16.384	170	2	8	32	1
9	156.24	100%	2	40	10	2	40	4
10	87.88	99.9968%	2.048	90	3	3	46	2
11	16.65	99.8938%	0.5	20	0.5	1	152	1
12	15	99.9303%	0.5	10	3	1	40	1
13	79	100%	8	80	5	2	71	2
14	83.00	100%	7	100	3	1	62	4
15	64.95	100%	4	250	3	2	62	1

16	5	99.9876%	1	20	1	1	45	1
17	219	99.7473%	8	300	10	8	46	2
18	82.60	99.999%	2	100	18	2	32	1

Using these parameters and data, the results of our analysis are presented in Table 4.15.

Table 4.16 The efficiency results

CSPs	Efficiency scores	Targets							Efficiency scores	Rank
		x_1	x_2	y_1	y_2	y_3	y_4	y_5		
1	1.000	80.000	100.000	8.000	80.000	5.000	2.000	5.000	5.625	1
2	0.055	137.143	99.990	7.000	100.000	3.374	2.020	16.777	0.055	18
3	0.155	80.000	99.998	8.000	89.320	5.233	4.000	20.241	0.155	13
4	1.000	80.000	99.950	8.000	200.000	8.000	6.000	1.000	1.1452	5
5	1.000	158.000	100.000	2.000	500.000	0.500	4.000	3.000	1.0438	9
6	0.074	107.761	100.000	4.000	100.000	3.398	2.008	21.750	0.074	15
7	1.000	150.000	99.990	16.000	384.000	8.000	6.000	4.000	1.2019	4
8	1.000	160.000	100.000	16.384	170.000	2.000	8.000	1.000	1.0571	7
9	0.064	150.029	100.000	2.000	44.296	10.000	2.012	21.861	0.064	16
10	0.520	87.880	100.000	2.403	90.583	3.105	3.000	11.471	0.520	12
11	1.000	16.650	99.890	0.500	0.500	0.500	1.000	1.000	1.2724	3
12	1.000	15.000	99.930	0.500	0.500	3.000	1.000	1.000	1.0002	11
13	0.056	79.000	99.999	8.000	81.063	5.000	2.018	11.395	0.056	17
14	0.089	83.000	100.000	7.000	7.000	5.320	2.083	198.881	0.089	14
15	1.000	64.950	100.000	4.000	4.000	3.000	2.000	1.000	1.0563	8
16	1.000	5.000	99.990	1.000	1.000	1.000	1.000	1.000	2.3774	2
17	1.000	219.000	99.750	8.000	8.000	10.000	8.000	2.000	1.0321	10
18	1.000	82.600	100.000	2.000	2.000	18.000	2.000	1.000	1.0607	6

Table 4.16 provides the obtained results and targets. As shown in column 2 using Model 4.24, eleven of the eighteen CSPs are efficient. Now, we run Model 4.25 for making differences between the performance of efficient CSPs. Column 10 in Table 4.16 demonstrates the obtained

results using Model 4.25. As we can see, all the CSPs are ranked completely so that the efficiency score of each CSP is different from the others.

4.4 SUMMARY

The rapid development of cloud computing and the considerable increase in the number of cloud service providers (CSPs) have resulted in many challenges in the suitability and selection of the optimal CSPs according to quality of service (QoS) requirements. One of the biggest challenges in the performance evaluation of cloud services is low discrimination power in the existing approaches. Low discrimination power refers to a weakness in distinguishing between the performance of CSPs when they are rated the same. In this chapter, we proposed and applied several distinctive DEA models to increase discrimination power in the performance evaluation of cloud services. The first three models proposed in this chapter are based on the enhanced Russell model (ERM) ideal DMU, anti-ideal DMU and a mixed ideal and anti-ideal DMU aspects in the performance evaluation process. The proposed models provide cloud service customers with an inclusive rank. Method 4, which is a network SBM model, increases discrimination power in the performance of cloud service providers whereas method 5 not only increases discrimination power in the performance evaluation of cloud service providers, it can also deal with undesirable outputs and decision makers' subjective judgments in the performance evaluation process.

Chapter 5 EVALUATING EFFICIENCY IN CLOUD SUPPLY CHAINS

5.1 INTRODUCTION

To date, considerable research has been undertaken to solve the problem of evaluating the efficiency of cloud service providers. However, no study addresses the efficiency of providers in the context of an entire supply chain, where multiple services interact to achieve a business objective or goal. DEA is, arguably, one of the most powerful methods for solving performance evaluation problems (Aparicio et al. 2020; Avilés-Sacoto et al. 2020; Kaffash et al. 2020). However, the current models ignore undesirable outputs, integer-valued data, and stochastic data, such as latency, the number of security certificates[7] and product prices, which can lead to inaccurate results. To address these shortcomings, the primary objective of this paper is to design a decision support system that accurately evaluates the efficiency of multiple cloud service providers in a supply chain. The system comprises a suite of two-stage network data envelopment analysis models that consider undesirable outputs, integer-valued data, and stochastic data. This study is the first attempt to address this problem. The evaluation results from a case study involving 24 cloud service providers show that the proposed system is able to evaluate the individual efficiency of providers in each stage of the supply chain and across the entire chain properly. The system is also able to provide the optimal composition of cloud service providers to suit a customer's priorities and requirements.

[7] The number of licenses that CSPs have for showing their services security level.

5.2 THE CLOUD ENVIRONMENT AND ITS SUPPLY CHAIN

As shown in Figure 2.2, cloud computing services can be divided into three categories according to the abstraction level of the service provided and the provider's business model. These categories are IaaS, PaaS, and SaaS (Buyya, Broberg & Goscinski 2011).

A supply chain is the system of organizations, people, activities, information, and resources involved in moving a product or service from a supplier to a customer (Reefke & Sundaram 2018). Cloud supply chain activities include providing computing infrastructure, software development platforms, and software to the end customer. In a cloud supply chain, IaaS is often provided to PaaS suppliers; PaaS suppliers deliver their services to SaaS suppliers; and all services can be delivered to cloud service customers. Figure 5.1 illustrates the cloud supply chain. In terms of DEA, the three cloud services, IaaS, PaaS, and SaaS, are considered as three stages in the chain, while the providers are the decision-making units.

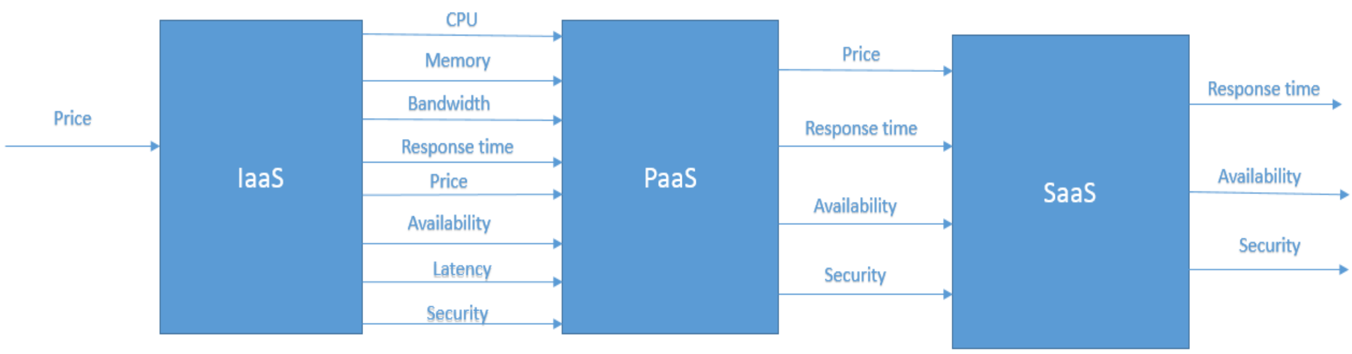

Figure 5.1 The cloud supply chain

5.3 A DECISION SUPPORT SYSTEM FOR EVALUATING THE PERFORMANCE OF CLOUD SERVICE PROVIDERS (CSPs)

In this section we present a decision-making system for evaluating CSPs across a cloud supply chain as follows.

Begin.

Step1: Construct the cloud supply chain

Step 2: Determine the decision-making variables (i.e., the inputs, the intermediate, and the outputs variables).

Step 3: Build three separate DEA ranking models to consider the undesirable, integer, and stochastic variables.

Step 4: Integrate the three models from Step 3 into one model that considers all three variables.

Step 5: Identify the three different variable types (undesirable outputs, integer data and stochastic data.

Step 6: Determine the scope of the problem, i.e., the number of stages (services) in the supply chain that need to be considered given the customer's priorities.

Step 7: Select the relevant DEA ranking models based on the number of stages and the types of decision-making variables.

Step 8: Analyse the results of the evaluation.

Step 9: Recommend the highest-ranking CSPs.

End.

Figure 5.2 illustrates the decision support system based on this algorithm, hereafter referred to as the CSP performance evaluation system (CSP-PE). The input variables are:

1. The decision-making variables in the constructed cloud supply chain.
2. The types of decision-making variables (integer, undesirable, stochastic).
3. The customer's priorities. For example, low response times or cost reduction (low costs for purchasing cloud computing service).
4. The number of stages, i.e., how many different types of cloud services are included in the supply chain (IaaS, PaaS, and SaaS). Depending on the scope of the evaluation problem, there can be between one and three stages.

The system's engine comprises four main components.

(a) Component 1: This component considers two-stage SBM network DEA. It comprises two models, Models (5.1) and (5.2) which are detailed in the remainder of this chapter.

(b) Component 2: This component considers two of the three variables: undesirable outputs and integer-valued data. Again, this component comprises two models, Models (5.3) and (5.4) proposed in the rest of this chapter.

(c) Component 3: This component considers all three variables: undesirable outputs, integer-valued data, and stochastic data and comprises Models (5.5) and (5.6) proposed in the rest of this chapter.

(d) Component 4 is a unified, deterministic equivalent for Models (5.5) and (5.6). This component also considers each of the three types of decision-making variables with two models, Models (5.7) and (5.8) proposed in the rest of this chapter.

Each of these components and their models are discussed in more detail in the next sections.

CSP-PE's structure is based on the cloud supply chain. Its output is a ranked list of CSPs and the optimal composition of those CSPs, given the customer's requirements. For example, a customer may benefit more from choosing IaaS_1, PaaS_5, and SaaS_3 rather than choosing one provider that offers all three services.

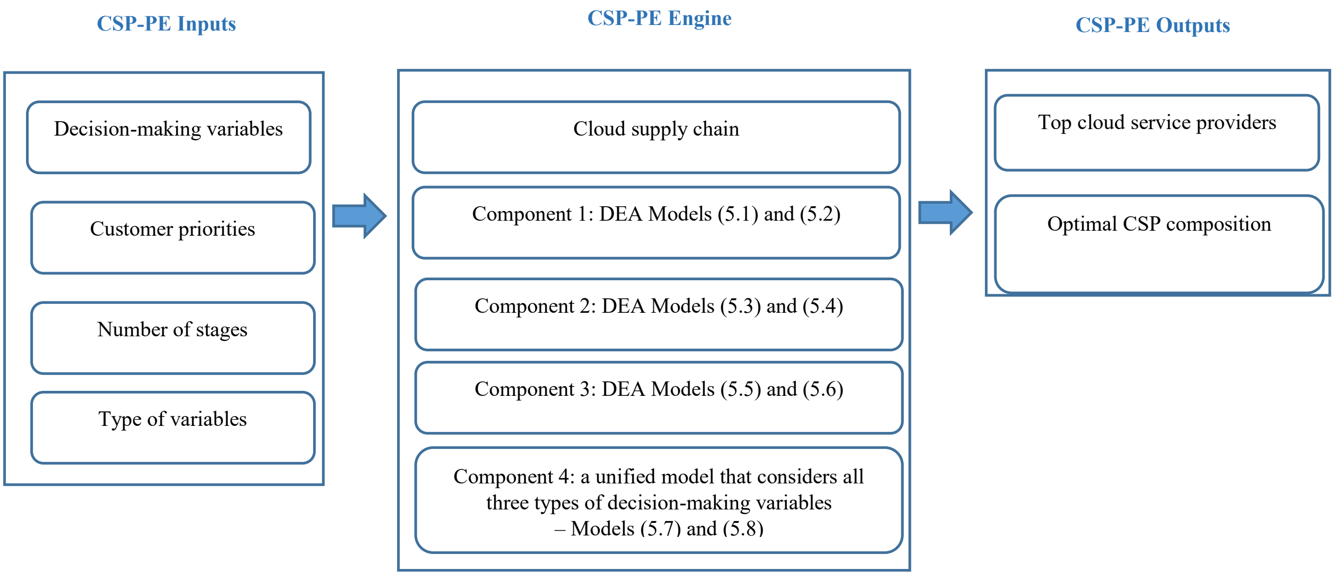

Figure 5.2: CSP-PE, a decision support system for evaluating the performance of cloud service providers

Figure 5.1 CSP-PE, a decision support system

5.4 THE PROPOSED DEA MODELS

5.4.1. TWO-STAGE SBM NETWORK DEA MODELS WITH UNDESIRABLE OUTPUTS AND INTEGER-VALUED DATA

The relative efficiency technique used in this study is based on the SBM network DEA approach proposed by Tone & Tsutsui (2009). SBM network DEA is a powerful approach for evaluating the efficiency of each stage as well as the overall efficiency of the CSP. We begin by introducing the notations used in this study. Suppose that n decision-making units (DMUs) ($j = 1, \dots, n$) somehow need to be evaluated over a set of inputs $I(= \{1, \dots, |I|\})$ and a set $R(= \{1, \dots, |R|\})$ of outputs $R(= \{1, \dots, |R|\})$. Each observation of DMU_k is characterized by the magnitudes of the inputs to be consumed $\boldsymbol{x}_k = \left(x_{1k}, \dots, x_{|I|k}\right)$ and the outputs to be produced $\boldsymbol{y}_k = (y_{1k}, \dots, y_{|R|k})$. Moreover, it is assumed that each DMU is divided into two sub-DMUs (Stage 1 and Stage 2), where Stage 1 of DMU$_k$ uses $\boldsymbol{x}_k$ and Stage 2 of DMU$_j$ produces $\boldsymbol{y}_k$. There is also a set of intermediate measures $L(=\{1, \dots, |L|\})$, where each $\boldsymbol{z}_k = (z_{1k}, \dots, z_{|L|k})$ simultaneously plays the role of the outputs and inputs for Stage 1 and Stage 2, respectively.

In these two-stage network DEA models, the SBM efficiency of Stage 1 is formulated as follows:

$$
\begin{aligned}
&\boldsymbol{\rho_1^* = \min \tau - \frac{1}{|I|}\left(\sum\nolimits_{i \in I} \frac{s_i^-}{x_{io}}\right)} && \\
&\boldsymbol{\tau + \frac{1}{|L|}\left(\sum\nolimits_{l \in L} \frac{S_r^+}{z_{lo}}\right) = 1} && \\
&\boldsymbol{\tau x_{io} - s_i^- = \sum\nolimits_{j=1}^{n} x_{ij} \lambda_j} && \boldsymbol{i \in I} \\
&\boldsymbol{\tau z_{lo} + s_l^+ = \sum\nolimits_{j=1}^{n} z_{lj} \lambda_j} && \boldsymbol{l \in L} \\
&\boldsymbol{\sum\nolimits_{j \in J} \lambda_j = \tau} && \\
&\boldsymbol{\lambda_j \geq 0} && \boldsymbol{j \in J} \\
&\boldsymbol{s_i^- \geq 0} && \boldsymbol{i \in I} \\
&\boldsymbol{s_l^+ \geq 0} && \boldsymbol{l \in L}
\end{aligned}
\tag{5.1}
$$

where $\boldsymbol{s}^- = (s_1^-, \dots, s_{|I|}^-)$ are the input excesses and $\boldsymbol{s}^+ = (s_1^+, \dots, s_{|L|}^+)$ are the output (intermediate) shortfalls, also known as slacks. Let an optimal solution for Model (5.1) be $(\boldsymbol{\lambda}^*, \boldsymbol{s}^{-*}, \boldsymbol{s}^{+*}) \in \mathbb{R}^{n+|I|+|L|}$.

Similarly, the SBM efficiency of Stage 2 can be measured by

$$
\begin{aligned}
&\boldsymbol{\rho_2^* = \min \tau - \frac{1}{|L|}\left(\sum\nolimits_{l \in L} \frac{S_r^+}{z_{lo}}\right)} && \\
&\boldsymbol{\tau + \frac{1}{|R|}\left(\sum\nolimits_{r \in R} \frac{S_r^+}{y_{ro}}\right) = 1} && \\
&\boldsymbol{\tau z_{lo} - s_l^- = \sum\nolimits_{j=1}^{n} z_{lj} \lambda_j} && \boldsymbol{l \in L} \\
&\boldsymbol{\tau y_{ro} + s_r^+ = \sum\nolimits_{j=1}^{n} y_{rj} \lambda_j} && \boldsymbol{r \in R} \\
&\boldsymbol{\sum\nolimits_{j \in J} \lambda_j = \tau} && \\
&\boldsymbol{\lambda_j \geq 0} && \boldsymbol{j \in J} \\
&\boldsymbol{s_l^- \geq 0} && \boldsymbol{l \in L} \\
&\boldsymbol{s_r^+ \geq 0} && \boldsymbol{r \in R}
\end{aligned}
\tag{5.2}
$$

Here, $\boldsymbol{s}^- = \left(s_1^-, \ldots, s_{|L|}^-\right)$ are the input (intermediate) excesses, and $\boldsymbol{s}^+ = \left(s_1^+, \ldots, s_{|R|}^+\right)$ are the output shortfalls, again, known as slacks. Also λ is the intensity vector. Let an optimal solution for Model $\left(\boldsymbol{\lambda}^*, \boldsymbol{s}^{-*}, \boldsymbol{s}^{+*}\right) \in \mathbb{R}^{n+|L|+|R|}$.

Definition 5.1. The optimal objective values ρ_1^* and ρ_2^* are the SBM efficiency of Stages 1 and 2, respectively, for DMU_o.

Definition 2. The overall SBM efficiency of DMU_o is $\frac{\rho_1^*+\rho_2^*}{2}$. If we have $\rho_1^* = \rho_2^* = 1$, then DMU_o shows the SBM efficiency overall.

Two types of integer and non-integer measures are considered. It is assumed that I^{IN} and I^{NI} are two mutually exclusive and collectively exhaustive input subsets for integer and non-integer-valued inputs. Mathematically, $I^{IN} \cup I^{NI} = I$ and $I^{IN} \cap I^{NI} = \phi$. Similarly, we let the integer- and non-integer-valued outputs and intermediate measures be respectively R^{IN}, R^{NI} and L^{IN}, L^{NI}. In addition, all the outputs and intermediate measures are partitioned into four subsets ($R^{INU}, R^{NIU}, R^{IND}, R^{NID}$ and $L^{INU}, L^{NIU}, L^{IND}, L^{NID}$) so as to consider the undesirable outputs, where the superscript *INU* represents the integer-valued undesirable variables, NIU denotes the non-integer-valued undesirable variables, INU, IND, NID, integer-valued desirable, and non-integer-valued desirable measures. We suggest the following mixed-integer linear programming to measure the SBM efficiency of Stage 1 in the presence of integer-valued inputs and (un)desirable intermediate variables:

$$
\begin{array}{lll}
\boldsymbol{\rho_1^* = \min \tau - \frac{1}{|I^{NI}|+|I^{IN}|}\left(\sum_{i\in I^{NI}}\frac{s_i^-}{x_{io}} + \sum_{i\in I^{IN}}\frac{(s_i^-+t_i^-)}{x_{io}}\right)} & & \\
\boldsymbol{1 = \tau + \frac{\left(\sum_{l\in L^{NID}}\frac{s_l^{+D}}{z_{lo}}+\sum_{l\in L^{IND}}\frac{(s_l^{+D}+t_l^{+D})}{z_{lo}}+\sum_{l\in L^{NIU}}\frac{s_l^{-U}}{z_{lo}}+\sum_{l\in L^{INU}}\frac{(s_l^{-U}+t_l^{-U})}{z_{lo}}\right)}{|L^{NID}|+|L^{IND}|+|L^{NIU}|+|L^{INU}|}} & & \\
\boldsymbol{\tau x_{io} - s_i^- = \sum_{j=1}^{n} x_{ij}\lambda_j} & \boldsymbol{i \in I^{NI}} & \\
\boldsymbol{\bar{x}_i - s_i^- = \sum_{j=1}^{n} x_{ij}\lambda_j} & \boldsymbol{i \in I^{IN}} & \\
\boldsymbol{\tau x_{io} - t_i^- = \bar{x}_i} & \boldsymbol{i \in I^{IN}} & \\
\boldsymbol{\tau z_{lo} + s_l^{+D} = \sum_{j=1}^{n} z_{lj}\lambda_j} & \boldsymbol{l \in L^{NID}} & \\
\boldsymbol{\bar{z}_l + s_l^{+D} = \sum_{j=1}^{n} z_{lj}\lambda_j} & \boldsymbol{l \in L^{IND}} & \\
\boldsymbol{\tau z_{lo} + t_l^{+D} = \bar{z}_l} & \boldsymbol{l \in L^{IND}} & \\
\boldsymbol{\tau z_{lo} - s_l^{-U} = \sum_{j=1}^{n} z_{lj}\lambda_j} & \boldsymbol{l \in L^{NIU}} & \\
\boldsymbol{\bar{z}_l - s_l^{-U} = \sum_{j=1}^{n} z_{lj}\lambda_j} & \boldsymbol{l \in L^{INU}} & (5.3) \\
\boldsymbol{\tau z_{lo} - t_l^{-U} = \bar{z}_l} & \boldsymbol{l \in L^{INU}} & \\
\boldsymbol{\sum_{j\in J}\lambda_j = \tau} & & \\
\boldsymbol{\lambda_j \geq 0} & \boldsymbol{j \in J} & \\
\boldsymbol{s_i^- \geq 0} & \boldsymbol{i \in I} & \\
\boldsymbol{t_i^- \geq 0} & \boldsymbol{i \in I^{IN}} & \\
\boldsymbol{s_l^{+D} \geq 0} & \boldsymbol{l \in L^{D}} & \\
\boldsymbol{t_l^{+D} \geq 0} & \boldsymbol{l \in L^{IND}} & \\
\boldsymbol{s_l^{-U} \geq 0} & \boldsymbol{l \in L^{U}} & \\
\boldsymbol{t_l^{-U} \geq 0} & \boldsymbol{l \in L^{INU}} & \\
\boldsymbol{\bar{x}_i \in \mathbb{Z}_+} & \boldsymbol{i \in I^{IN}} & \\
\boldsymbol{\bar{z}_l \in \mathbb{Z}_+} & \boldsymbol{l \in L^{IN}} &
\end{array}
$$

where $L^D = L^{IND} \cup L^{NID}$, $L^U = L^{INU} \cup L^{NIU}$, $\bar{x}_i$ and $\bar{z}_l$ are integer decision variables that indicate integer-valued reference point for input $i \in I^{IN}$ and intermediate $l \in L^{IN}$. It should be noted that, here, there are two types of slacks[8]: one for the integer-valued inputs, the other for the intermediate variables. The first type of slack, $i \in I^{IN}$, i.e. s_i^-, is the difference between the combination $\sum_{j=1}^{n} x_{ij}\lambda_j$ and the integer-valued $\bar{x}_i$. The second type of slack t_i^- is the difference between the integer-valued $\bar{x}_i$ and the projection τx_{io}. As a result, $s_i^- + t_i^-$ is the total slack for an integer-valued x_i. Similarly, the total slack for the integer-valued desirable measures and the undesirable intermediate measures are $s_l^{+D} + t_l^{+D}$ and $s_l^{-U} + t_l^{-U}$, respectively. These values are

[8] Slacks are related to inputs surpass and outputs shortage in the DEA context.

considered in the objective function of Model (5.3) along with a set of normalization constraints.

Similarly, Model (5.4) evaluates efficiency in Stage 2 according to the integer-valued intermediate variables and the (un)desirable outputs. Model (5.2) is formulated as follows:

$$
\begin{aligned}
&\boldsymbol{\rho_2^* = \min \tau - \frac{1}{|L^{NI}|+|L^{IN}|}\left(\sum_{l\in L^{NI}} \frac{s_l^-}{z_{lo}} + \sum_{l\in L^{IN}} \frac{(s_l^- + t_l^-)}{z_{lo}}\right)} && \\
&\boldsymbol{1 = \tau + \frac{\left(\sum_{r\in R^{NID}} \frac{s_r^{+D}}{y_{ro}} + \sum_{r\in R^{IND}} \frac{(s_r^{+D}+t_r^{+D})}{y_{ro}} + \sum_{r\in R^{NIU}} \frac{t_r^{-U}}{y_{ro}} + \sum_{r\in R^{INU}} \frac{(s_r^{-U}+t_r^{-U})}{y_{ro}}\right)}{|R^{NID}|+|R^{IND}|+|R^{NIU}|+|R^{INU}|}} && \\
&\boldsymbol{\tau z_{lo} - s_l^- = \textstyle\sum_{j=1}^{n} z_{lj}\lambda_j} && \boldsymbol{l \in L^{NI}} \\
&\boldsymbol{\bar{z}_l - s_l^- = \textstyle\sum_{j=1}^{n} z_{lj}\lambda_j} && \boldsymbol{l \in L^{IN}} \\
&\boldsymbol{\tau z_{lo} - t_l^- = \bar{z}_l} && \boldsymbol{l \in L^{IN}} \\
&\boldsymbol{\tau y_{ro} + s_r^{+D} = \textstyle\sum_{j=1}^{n} y_{rj}\lambda_j} && \boldsymbol{r \in R^{NID}} \\
&\boldsymbol{\bar{y}_r + s_r^{+D} = \textstyle\sum_{j=1}^{n} y_{rj}\lambda_j} && \boldsymbol{r \in R^{IND}} \\
&\boldsymbol{\tau y_{ro} + t_r^{+D} = \bar{y}_r} && \boldsymbol{r \in R^{IND}} \\
&\boldsymbol{\tau y_{ro} - s_r^{+U} = \textstyle\sum_{j=1}^{n} y_{rj}\lambda_j} && \boldsymbol{r \in R^{NIU}} \\
&\boldsymbol{\bar{y}_r - s_r^{+U} = \textstyle\sum_{j=1}^{n} y_{rj}\lambda_j} && \boldsymbol{r \in R^{INU}} \qquad (5.4) \\
&\boldsymbol{\tau y_{ro} - t_r^{+U} = \bar{y}_r} && \boldsymbol{r \in R^{INU}} \\
&\boldsymbol{\textstyle\sum_{j\in J} \lambda_j = \tau} && \\
&\boldsymbol{\lambda_j \geq 0} && \boldsymbol{j \in J} \\
&\boldsymbol{s_l^- \geq 0} && \boldsymbol{l \in L} \\
&\boldsymbol{t_l^- \geq 0} && \boldsymbol{l \in L^{IN}} \\
&\boldsymbol{s_r^{+D} \geq 0} && \boldsymbol{r \in R^{D}} \\
&\boldsymbol{t_r^{+D} \geq 0} && \boldsymbol{r \in R^{IND}} \\
&\boldsymbol{s_r^{+U} \geq 0} && \boldsymbol{r \in R^{U}} \\
&\boldsymbol{t_r^{+U} \geq 0} && \boldsymbol{r \in R^{INU}} \\
&\boldsymbol{z_l \in \mathbb{Z}_+} && \boldsymbol{l \in L^{IN}} \\
&\boldsymbol{y_r \in \mathbb{Z}_+} && \boldsymbol{r \in R^{IN}}
\end{aligned}
$$

where $R^D = R^{IND} \cup R^{NID}$ and $R^U = R^{INU} \cup R^{NIU}$.

Models (5.3) and (5.4) evaluate the individual efficiency of each DMU_o in Stage 1 and Stage 2, respectively, given integer-valued and undesirable data. The overall SBM efficiency of a DMU_o is derived from an average of the efficiency in Stages 1 and 2 (see Definition 2).

5.4.2 TWO-STAGE SBM NETWORK DEA MODELS WITH UNDESIRABLE OUTPUTS, INTEGER-

Most DEA and network DEA models treat data as being deterministic. Subsequently, the relative efficiencies of the DMUs are also deterministic. However, measuring the efficiency of CSPs in practical applications often involves random variables and uncertainty. Hence, Models (5.5) and (5.6) rely on a chance-constrained programming approach that allows for random variations in the data. As discussed by Zha, Zhao & Bian (2016), chance-constrained programming can robustly deal with data uncertainty when that uncertainty is caused by random errors in the data set. By incorporating different levels of random errors into the model, chance-constrained programming can show the influence the "randomness" has had on the evaluation results. Moreover, this approach focuses on real units and the uncertainty inherent in individual inputs, intermediate variables, and outputs.

To this end, we use $\tilde{\boldsymbol{x}}_j = (\tilde{x}_{1j}, \dots, \tilde{x}_{|I|j})$, $\tilde{\boldsymbol{y}}_j = (\tilde{y}_{1j}, \dots, \tilde{y}_{|R|j})$, and $\tilde{\boldsymbol{z}}_j = (z_{1j}, \dots, \tilde{z}_{|L|j})$ to represent random input, output, and intermediate vectors, respectively. Let $\boldsymbol{x}_j = (x_{1j}, \dots, x_{|I|j})$, $\boldsymbol{y}_j = (y_{1j}, \dots, y_{|R|j})$, and $\boldsymbol{z}_j = (z_{1j}, \dots, z_{|L|j})$ represent the expected input, output, and intermediate vector values for each $\text{DMU}_j; j = 1, \dots, n$.

The following pair of stochastic SBM models is used to evaluate the SBM efficiency of Stages 1 and 2[9]:

[9] For more details about the chance-constrained programming approach used in this paper please see (Cooper et al. 2002, 2004).

$$\boldsymbol{\rho_1^* = \min \rho_1}$$

$$\begin{aligned}
&P\left\{\rho_1 - \tau + \frac{\left(\sum_{i\in I^{NI}}\frac{s_i^-}{x_{io}} + \sum_{i\in I^{IN}}\frac{(s_i^- + t_i^-)}{x_{io}}\right)}{|I^{NI}|+|I^{IN}|} \le 0\right\} \ge 1-\alpha \\
&P\left\{\tau + \frac{\left(\sum_{l\in L^{NID}}\frac{S_l^{+D}}{z_{lo}} + \sum_{l\in L^{IND}}\frac{\left(S_l^{+D}+t_l^{+D}\right)}{z_{lo}} + \sum_{l\in L^{NIU}}\frac{s_l^{-U}}{z_{lo}} + \sum_{l\in L^{INU}}\frac{\left(s_l^{-U}+t_l^{-U}\right)}{z_{lo}}\right)}{|L^{NID}|+|L^{IND}|+|L^{NIU}|+|L^{INU}|} = 1\right\} \ge 1-\alpha \\
&P\left\{\textstyle\sum_{j=1}^{n} \tilde{x}_{ij}\lambda_j + s_i^- \le \tau\tilde{x}_{io}\right\} \ge 1-\alpha && i \in I^{NI} \\
&P\left\{\textstyle\sum_{j=1}^{n} \tilde{x}_{ij}\lambda_j + s_i^- \le \bar{x}_i\right\} \ge 1-\alpha && i \in I^{IN} \\
&P\{\bar{x}_i + t_i^- \le \tau\tilde{x}_{io}\} \ge 1-\alpha && i \in I^{IN} \\
&P\left\{\textstyle\sum_{j=1}^{n} \tilde{z}_{lj}\lambda_j - \tau\tilde{z}_{lo} \ge s_l^{+D}\right\} \ge 1-\alpha && l \in L^{NID} \\
&P\left\{\textstyle\sum_{j=1}^{n} \tilde{z}_{lj}\lambda_j - \bar{z}_l \ge s_l^{+D}\right\} \ge 1-\alpha && l \in L^{IND} \\
&P\left\{\bar{z}_l - \tau\tilde{z}_{lo} \ge t_l^{+D}\right\} \ge 1-\alpha && l \in L^{IND} \\
&P\left\{\textstyle\sum_{j=1}^{n} \tilde{z}_{lj}\lambda_j + s_l^{+U} \ge \tau\tilde{z}_{lo}\right\} \ge 1-\alpha && l \in L^{NIU} \\
&P\left\{\textstyle\sum_{j=1}^{n} \tilde{z}_{lj}\lambda_j + s_l^{+U} \ge \bar{z}_l\right\} \ge 1-\alpha && l \in L^{INU} \\
&P\left\{\tau\tilde{z}_{lo} - t_l^{+U} \le \bar{z}_l\right\} \ge 1-\alpha && l \in L^{INU} \\
&\textstyle\sum_{j\in J} \lambda_j = \tau \\
&\lambda_j \ge 0 && j \in J \\
&s_i^- \ge 0 && i \in I \\
&t_i^- \ge 0 && i \in I^{IN} \\
&s_l^{+D} \ge 0 && l \in L^{D} \\
&t_l^{+D} \ge 0 && l \in L^{IND} \\
&s_l^{-U} \ge 0 && l \in L^{U} \\
&t_l^{-U} \ge 0 && l \in L^{INU} \\
&\bar{x}_i \in \mathbb{Z}_+ && i \in I^{IN} \\
&\bar{z}_l \in \mathbb{Z}_+ && l \in L^{IN}
\end{aligned} \tag{5.5}$$

$$
\begin{aligned}
&\boldsymbol{\rho_2^* = \min \rho_2} && \\
&P\left\{\rho_2 - \tau \frac{1}{|L^{NI}|+|L^{IN}|}\left(\sum_{l\in L^{NI}} \frac{s_l^-}{z_{lo}} + \sum_{l\in L^{IN}} \frac{(s_l^- + t_l^-)}{z_{lo}}\right) \le 0\right\} \ge 1-\alpha && \\
&P\left\{\tau + \frac{\left(\sum_{r\in R^{NID}} \frac{s_r^{+D}}{y_{ro}} + \sum_{r\in R^{IND}} \frac{(s_r^{+D}+t_r^{+D})}{y_{ro}} + \sum_{r\in R^{NIU}} \frac{t_r^{-U}}{y_{ro}} + \sum_{r\in R^{INU}} \frac{(s_r^{-U}+t_r^{-U})}{y_{ro}}\right)}{|R^{NID}|+|R^{IND}|+|R^{NIU}|+|R^{INU}|} = 1\right\} \ge 1-\alpha && \\
&P\left\{\sum_{j=1}^n \tilde{z}_{lj}\lambda_j + s_l^- \le \tau\tilde{z}_{lo}\right\} \ge 1-\alpha && l\in L^{NI} \\
&P\left\{\sum_{j=1}^n \tilde{z}_{lj}\lambda_j + s_l^- \le \bar{z}_l\right\} \ge 1-\alpha && l\in L^{IN} \\
&P\{\bar{z}_l + t_l^- \le \tau\tilde{z}_{lo}\} \ge 1-\alpha && l\in L^{IN} \\
&P\left\{\sum_{j=1}^n \tilde{y}_{rj}\lambda_j - \tau\tilde{y}_{ro} \ge s_r^{+D}\right\} \ge 1-\alpha && r\in R^{NID} \\
&P\left\{\sum_{j=1}^n \tilde{y}_{rj}\lambda_j - \bar{y}_r \ge s_r^{+D}\right\} \ge 1-\alpha && r\in R^{IND} \\
&P\{\bar{y}_r - \tau\tilde{y}_{ro} \ge t_r^{+D}\} \ge 1-\alpha && r\in R^{IND} \\
&P\left\{\sum_{j=1}^n \tilde{y}_{rj}\lambda_j + s_r^{+U} \ge \tau\tilde{y}_{ro}\right\} \ge 1-\alpha && r\in R^{NIU} \qquad (5.6) \\
&P\left\{\sum_{j=1}^n \tilde{y}_{rj}\lambda_j + s_r^{+U} \ge \bar{y}_r\right\} \ge 1-\alpha && r\in R^{INU} \\
&P\{\tau\tilde{y}_{ro} - t_r^{+U} \le \bar{y}_r\} \ge 1-\alpha && r\in R^{INU} \\
&\sum_{j\in J} \lambda_j = \tau && \\
&\lambda_j \ge 0 && j\in J \\
&s_l^- \ge 0 && l\in L \\
&t_l^- \ge 0 && l\in L^{IN} \\
&s_r^{+D} \ge 0 && r\in R^D \\
&t_r^{+D} \ge 0 && r\in R^{IND} \\
&s_r^{+U} \ge 0 && r\in R^U \\
&t_r^{+U} \ge 0 && r\in R^{INU} \\
&\bar{z}_l \in \mathbb{Z}_+ && l\in L^{IN} \\
&\bar{y}_r \in \mathbb{Z}_+ && r\in R^{IN}
\end{aligned}
$$

where P means probability, and $\alpha \in (0,1)$ is a predetermined number. Chance constraints prevent violations with a probability of at most α (Zhou et al. 2017).

Models (5.7) and (5.8) are the deterministic equivalents of Models (5.5) and (5.6), which demonstrate how the optimal objective values ρ_1^* and ρ_2^* required for Definition 1 are determined. Readers interested in the more detailed development of these models can consult Cooper et al. (2002).

$$
\begin{aligned}
&\rho_1^* = \min \tau - \frac{1}{|I^{NI}|+|I^{IN}|}\left(\sum_{i\in I^{NI}} \frac{s_i^-}{x'_{io}} + \sum_{i\in I^{IN}} \frac{(s_i^- + t_i^-)}{x'_{io}}\right) && \\
&1 = \tau + \frac{\left(\sum_{l\in L^{NID}} \frac{s_l^{+D}}{z'_{lo}} + \sum_{l\in L^{IND}} \frac{\left(s_l^{+D}+t_l^{+D}\right)}{z'_{lo}} + \sum_{l\in L^{NIU}} \frac{s_l^{-U}}{z'_{lo}} + \sum_{l\in L^{INU}} \frac{\left(s_l^{-U}+t_l^{-U}\right)}{z'_{lo}}\right)}{|L^{NID}|+|L^{IND}|+|L^{NIU}|+|L^{INU}|} && \\
&\tau x'_{io} - s_i^- = \sum_{j=1}^{n} x'_{ij}\lambda_j && i \in I^{NI} \\
&\bar{x}_i - s_i^- = \sum_{j=1}^{n} x'_{ij}\lambda_j && i \in I^{IN} \\
&\tau x'_{io} - t_i^- = \bar{x}_i && i \in I^{IN} \\
&\tau z'_{lo} + s_l^{+D} = \sum_{j=1}^{n} z'_{lj}\lambda_j && l \in L^{NID} \\
&\bar{z}_l + s_l^{+D} = \sum_{j=1}^{n} z'_{lj}\lambda_j && l \in L^{IND} \\
&\tau z'_{lo} + t_l^{+D} = \bar{z}_l && l \in L^{IND} \\
&\tau z'_{lo} - s_l^{-U} = \sum_{j=1}^{n} z'_{lj}\lambda_j && l \in L^{NIU} \\
&\bar{z}_l - s_l^{-U} = \sum_{j=1}^{n} z'_{lj}\lambda_j && l \in L^{INU} \qquad (5.7) \\
&\tau z'_{lo} - t_l^{-U} = \bar{z}_l && l \in L^{INU} \\
&\sum_{j\in J} \lambda_j = \tau && \\
&\lambda_j \geq 0 && j \in J \\
&s_i^- \geq 0 && i \in I \\
&t_i^- \geq 0 && i \in I^{IN} \\
&s_l^{+D} \geq 0 && l \in L^{D} \\
&t_l^{+D} \geq 0 && l \in L^{IND} \\
&s_l^{-U} \geq 0 && l \in L^{U} \\
&t_l^{-U} \geq 0 && l \in L^{INU} \\
&\bar{x}_i \in \mathbb{Z}_+ && i \in I^{IN} \\
&\bar{z}_l \in \mathbb{Z}_+ && l \in L^{IN}
\end{aligned}
$$

where

$$
x'_{ij} = \begin{cases} x_{io} + \sigma_{io}^{I}\Phi^{-1}(\alpha), & i \in I^{NI}, j = o \\ x_{io} + \left\lceil \sigma_{io}^{I}\Phi^{-1}(\alpha) \right\rceil, & i \in I^{IN}, j = o \\ x_{ij} - \sigma_{ij}^{I}\Phi^{-1}(\alpha), & i \in I^{NI}, j \neq o \\ x_{ij} - \left\lceil \sigma_{ij}^{I}\Phi^{-1}(\alpha) \right\rceil, & i \in I^{IN}, j \neq o \end{cases}
$$

$$
z'_{lo} = \begin{cases} z_{lo} - \sigma_{lo}^{o}\,\Phi^{-1}(\alpha), & l \in L^{NID} \\ z_{lo} - \left\lceil \sigma_{lo}^{o}\,\Phi^{-1}(\alpha) \right\rceil, & l \in L^{INU} \\ z_{lo} + \sigma_{lo}^{o}\,\Phi^{-1}(\alpha), & l \in L^{NIU} \\ z_{lo} + \left\lceil \sigma_{lo}^{o}\,\Phi^{-1}(\alpha) \right\rceil, & l \in L^{INU} \end{cases}
$$

$$z'_{lj} = \begin{cases} z_{lj} + \sigma^{o}_{lj}\,\Phi^{-1}(\alpha), & l \in L^{NID} \\ z_{lj} + \lceil \sigma^{o}_{lj}\,\Phi^{-1}(\alpha) \rceil, & l \in L^{INU} \\ z_{lj} - \sigma^{o}_{lj}\,\Phi^{-1}(\alpha), & l \in L^{NIU} \\ z_{lj} - \lceil \sigma^{o}_{lj}\,\Phi^{-1}(\alpha) \rceil, & l \in L^{INU} \end{cases}$$

$$\begin{array}{ll}
\rho_2^* = \min \tau - \frac{1}{|L^{NI}|+|L^{IN}|}\left(\sum_{l\in L^{NI}} \frac{s_l^-}{z''_{lo}} + \sum_{l\in L^{IN}} \frac{(s_l^- + t_l^-)}{z''_{lo}}\right) & \\
1 = \tau + \frac{\left(\sum_{r\in R^{NID}} \frac{s_r^{+D}}{y''_{ro}} + \sum_{r\in R^{IND}} \frac{(s_r^{+D}+t_r^{+D})}{y''_{ro}} + \sum_{r\in R^{NIU}} \frac{t_r^{-U}}{y''_{ro}} + \sum_{r\in R^{INU}} \frac{(s_r^{-U}+t_r^{-U})}{y''_{ro}}\right)}{|R^{NID}|+|R^{IND}|+|R^{NIU}|+|R^{INU}|} & \\
\tau z''_{lo} - s_l^- = \sum_{j=1}^{n} z''_{lj}\lambda_j & l \in L^{NI} \\
\bar{z}_l - s_l^- = \sum_{j=1}^{n} z''_{lj}\lambda_j & l \in L^{IN} \\
\tau z''_{lo} - t_l^- = \bar{z}_l & l \in L^{IN} \\
\tau y''_{ro} + s_r^{+D} = \sum_{j=1}^{n} y''_{rj}\lambda_j & r \in R^{NID} \\
\bar{y}_r + s_r^{+D} = \sum_{j=1}^{n} y''_{rj}\lambda_j & r \in R^{IND} \\
\tau y''_{ro} + t_r^{+D} = \bar{y}_r & r \in R^{IND} \\
\tau y''_{ro} - s_r^{+U} = \sum_{j=1}^{n} y''_{rj}\lambda_j & r \in R^{NIU} \\
\bar{y}_r - s_r^{+U} = \sum_{j=1}^{n} y''_{rj}\lambda_j & r \in R^{INU} \\
\tau y''_{ro} - t_r^{+U} = \bar{y}_r & r \in R^{INU} \\
\sum_{j\in J} \lambda_j = \tau & \\
\lambda_j \geq 0 & j \in J \\
s_l^- \geq 0 & l \in L \\
t_l^- \geq 0 & l \in L^{IN} \\
s_r^{+D} \geq 0 & r \in R^{D} \\
t_r^{+D} \geq 0 & r \in R^{IND} \\
s_r^{+U} \geq 0 & r \in R^{U} \\
t_r^{+U} \geq 0 & r \in R^{INU} \\
z_l \in \mathbb{Z}_+ & l \in L^{IN} \\
y_r \in \mathbb{Z}_+ & r \in R^{IN}
\end{array} \tag{5.8}$$

where

$$z''_{lj} = \begin{cases} z_{lo} + \sigma^{I}_{lo}\Phi^{-1}(\alpha), & l \in L^{NI}, j = o \\ z_{lo} + \lceil \sigma^{I}_{lo}\Phi^{-1}(\alpha) \rceil, & l \in L^{IN}, j = o \\ z_{lj} - \sigma^{I}_{lj}\Phi^{-1}(\alpha), & l \in L^{NI}, j \neq o \\ z_{lj} - \lceil \sigma^{I}_{lj}\Phi^{-1}(\alpha) \rceil, & l \in L^{IN}, j \neq o \end{cases}$$

$$
\boldsymbol{y_{ro}''} = \begin{cases} \boldsymbol{y_{ro} - \sigma_{ro}^{o}\,\Phi^{-1}(\alpha)}, & \boldsymbol{r \in R^{NID}} \\ \boldsymbol{y_{ro} - \lceil \sigma_{ro}^{o}\,\Phi^{-1}(\alpha) \rceil}, & \boldsymbol{r \in R^{INU}} \\ \boldsymbol{y_{ro} + \sigma_{ro}^{o}\,\Phi^{-1}(\alpha)}, & \boldsymbol{r \in R^{NIU}} \\ \boldsymbol{y_{ro} + \lceil \sigma_{ro}^{o}\,\Phi^{-1}(\alpha) \rceil}, & \boldsymbol{r \in R^{INU}} \end{cases}
$$

$$
\boldsymbol{y_{rj}''} = \begin{cases} \boldsymbol{y_{rj} + \sigma_{lj}^{o}\,\Phi^{-1}(\alpha)}, & \boldsymbol{r \in R^{NID}} \\ \boldsymbol{y_{rj} + \lceil \sigma_{lj}^{o}\,\Phi^{-1}(\alpha) \rceil}, & \boldsymbol{r \in R^{INU}} \\ \boldsymbol{y_{rj} - \sigma_{lj}^{o}\,\Phi^{-1}(\alpha)}, & \boldsymbol{r \in R^{NIU}} \\ \boldsymbol{y_{rj} - \lceil \sigma_{lj}^{o}\,\Phi^{-1}(\alpha) \rceil}, & \boldsymbol{r \in R^{INU}} \end{cases}
$$

Here Φ^{-1} is the inverse of the standard normal distribution function and σ is the standard deviation.

It is worth noting that, in this study, we presume the inputs, intermediate variables, and random variables have normal distributions and known parameters because normal distributions are less restrictive and can be used to transform other types of distributions into approximately normal form (Cooper, Huang & Li 1996; Zhou et al. 2017).

5.5 CASE STUDY

To evaluate the proposed models, we prepared a data set on a sample of the top IaaS and PaaS providers. Each company in this study was considered to be a decision-making unit within a two-stage cloud supply chain comprising IaaS as Stage 1 and PaaS as Stage 2. Figure 5.3 shows the two stages of the cloud supply chain structure. The inputs for the IaaS stage were price (stochastic), latency (undesirable/stochastic), memory and CPU (integer-valued), and data transfer was used as the intermediate inputs/outputs. Price was also used as an additional input for the second stage (PaaS). The outputs for the PaaS stage were availability (stochastic), the number of security certifications (integer-valued), and service time delays (undesirable/stochastic). Details of the sample data set are provided in Table 5.1.

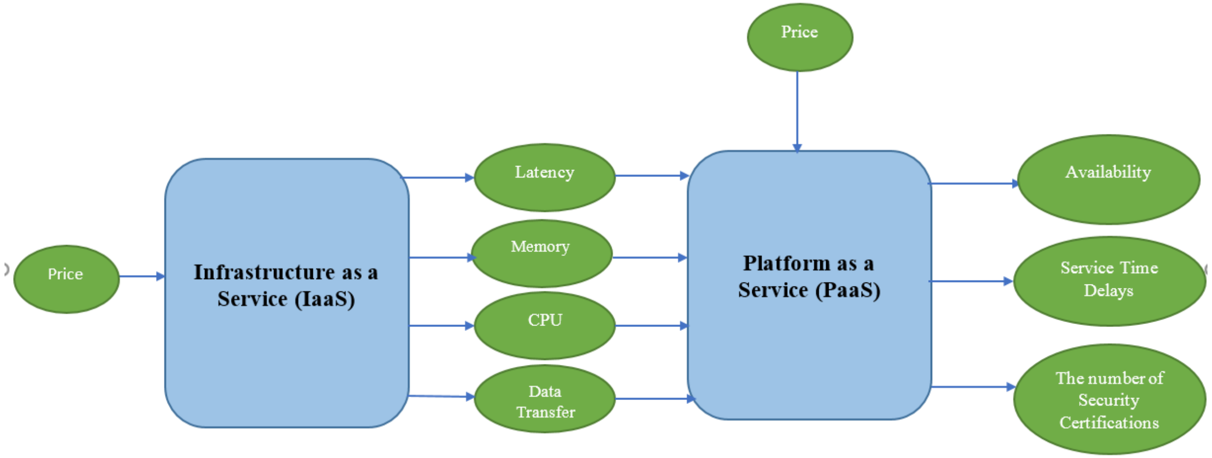

Figure 5.2 The two-stage cloud supply chain structure

Table 5.1 The data set of 24 CSPs

CSPs (DMUs)	Stage 1 Input	Stage 2 Input	Outputs			Intermediate variables			
	Price 1 (Monthly/$)	Price 2 (1000 Devices/$)	Availability (Percentage)	Security (Number)	Delayed service time (second)	Latency (Millisecond)	Memory	CPU	Data transfer
1	80	35	100%	5	78	433	8	2	5
2	140.79	50	99.99%	3	101	49	7	2	3.2
3	80	47	100%	4	52	46	8	4	5
4	80	59	99. 95%	1	163	39	8	6	8
5	158	50	100%	3	41	45	2	4	0.5
6	110	45	100%	4	33	41	4	2	3
7	150	42	99.99%	4	139	68	16	6	8
8	156.24	49	100%	4	64	40	2	2	10
9	87.88	37	100%	2	149	46	2.048	3	3
10	16.65	49	99.92%	1	176	152	0.5	1	0.5
11	15	31	99.93%	1	180	40	0.5	1	3
12	79	40	100%	2	59	71	8	2	5
13	83	31	100%	4	115	62	7	1	3
14	64.95	43	100%	1	152	62	4	2	3

15	219	37	99.97%	2	119	46	8	8	10
16	150	42	99.99%	4	26	68	16	6	8
17	140	42	99.99%	4	176	70	16	6	6
18	110	45	100%	4	143	41	4	2	3
19	80	47	100%	4	154	46	8	4	4
20	83	31	100%	4	179	62	7	1	3
21	15	34	99.94%	1	165	40	0.5	1	3
22	80	62	99.96%	1	134	40	8	6	8
23	15	31	99.99%	1	126	40	0.5	1	3
24	221	38	99.93%	2	177	48	8	8	10

Table 5.2 The results obtained using Models (1)-(8)

CSPs (DMUs)	Models										
	Model (5.1)	**Model** Error! Reference source not found.		**Model (5.3)**	**Model (5.4)**		**Model (5.5)**	**Model (5.6)**		**Model (5.7)**	**Model (5.8)** $\alpha = 0.8$
	Stage 1	**Stage 2**	**Overall**	**Stage 1**	**Stage 2**	**Overall**	**Stage 1**	**Stage 2**	**Overall**	**Stage 1**	**Stage 2**	**Overall**
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.529	1.000	0.764
2	0.453	0.556	0.505	0.240	0.751	0.496	0.456	0.751	0.604	0.287	0.559	0.423
3	1.000	0.527	0.763	0.493	0.467	0.480	1.000	0.623	0.811	0.601	0.492	0.547
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.778	0.147	0.463
5	0.149	1.000	0.575	0.142	1.000	0.571	0.149	1.000	0.575	0.227	1.000	0.614
6	0.499	1.000	0.750	0.256	0.500	0.378	0.499	1.000	0.750	0.293	1.000	0.647
7	1.000	0.283	0.641	1.000	0.390	0.695	1.000	0.293	0.646	0.750	0.236	0.493
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.234	1.000	0.617
9	0.555	0.574	0.564	0.317	0.741	0.529	0.558	0.637	0.597	0.256	0.468	0.362
10	0.314	1.000	0.657	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

11	1.000	0.870	0.935	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
12	0.753	0.389	0.571	0.599	0.261	0.430	0.753	1.000	0.877	0.492	1.000	0.746
13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.404	1.000	0.702
14	0.596	0.317	0.457	0.411	0.398	0.405	0.596	0.372	0.484	0.339	0.258	0.299
15	1.000	0.261	0.630	1.000	0.303	0.652	1.000	0.261	0.630	0.829	0.197	0.513
16	1.000	1.000	1.000	1.000	0.195	0.598	1.000	1.000	1.000	0.718	1.000	0.859
17	1.000	0.297	0.648	1.000	0.491	0.745	1.000	0.313	0.657	0.643	0.238	0.440
18	0.499	0.722	0.611	0.256	1.000	0.628	0.499	1.000	0.750	0.281	0.931	0.606
19	0.941	0.474	0.708	0.488	1.000	0.744	1.000	0.495	0.748	0.518	0.362	0.440
20	1.000	0.848	0.924	1.000	1.000	1.000	1.000	1.000	1.000	0.329	1.000	0.665
21	0.956	0.894	0.925	0.956	1.000	0.978	0.956	1.000	0.978	1.000	0.520	0.760
22	0.970	0.274	0.622	1.000	0.330	0.665	1.000	0.287	0.643	0.760	0.141	0.451
23	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
24	0.972	0.250	0.611	1.000	0.423	0.712	1.000	0.250	0.625	0.823	0.184	0.504

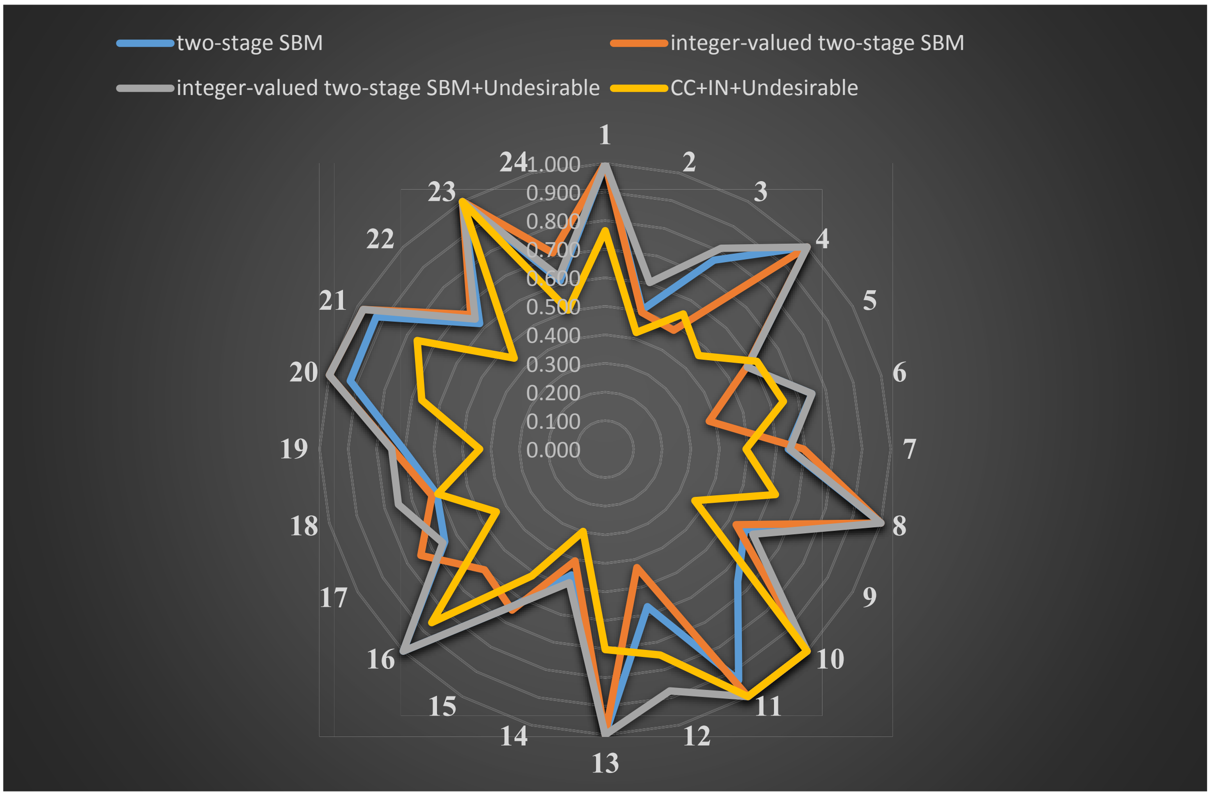

Figure 5.3 Ranking scores with different models

5.6 RESULTS AND DISCUSSION

To illustrate the rationality of the proposed models, we compared the results of the efficiency evaluations for each of the 24 CSPs obtained using Models (5.1) to (5.8) as shown in Table 5.2. We first calculated the efficiency scores for Stage 1 (IaaS) using Model (5.1) and for Stage 2 (PaaS) using Model (5.2). The average of these two scores represents the overall efficiency [column 4] and show that six of the 24 CSPs (CSPs 1, 4, 6, 13, 16, and 23) would operate efficiently in this supply chain. This is because these CSPs satisfy Definition 2 and, according to this definition, they are efficient in both stages. CSP 14 had the worst overall efficiency mainly due to its low performance in Stage 2.

The overall efficiency scores when considering undesirable outputs are shown in column 7. Compared to the overall efficiency scores without this condition [column 4], eight of the 24 CSPs (CSPs 1, 4, 8, 10, 11, 13, 20 and 23) were efficient. CSPs 6 and 16 were only efficient when undesirable outputs were not considered, and CSPs 10, 11, and 20 changed from inefficient to efficient once the undesirable outputs were included in the evaluation.

The efficiency scores that consider both undesirable outputs and integer-valued variables are shown under Model (5.5) – Stage 1, Model (5.6) – Stage 2 and Overall (5.5) & (5.6). In comparison to the overall efficiency scores with undesirable outputs (the comparison of column 7 and column 10), only the status of CSP 16 changed from inefficient to efficient.

The last column demonstrates the overall efficiency score using Model (5.7) and Model (5.8) when taking undesirable outputs, integer-valued, and stochastic data into account. Corresponding to $\alpha = 80\%$, only CSPs 10, 11, and 23 have overall efficiency. This means that considering the new condition of stochastic data means many CSPs are no longer efficient. Notably, only CSP 23 showed efficiency under all conditions. Therefore, this CSP may serve as a benchmark for other CSPs wishing to improve their performance in different conditions. The worst CSP overall was CSP 14. This CSP showed poor performance in both Stages 1 and 2 (see columns 11 and 12), which resulted in an overall efficiency score of 0.299.

However, although CSP 23 (which provides IaaS and PaaS) was efficient in all conditions, this does not necessarily mean that cloud customers should purchase both IaaS and PaaS services from this provider. CSP-PE also provides a range of different CSP compositions for designing a cloud supply chain. For example, the IaaS option (Stage 1) provided by CSP 20 may result in a more optimal composition when combined with the PaaS option (Stage 2) provided by CSP 21.

Figure 5.4 illustrates a radar chart of the salient features of the evaluation results produced by each of the different models. Each CSP (DMU) forms an individual axis, organized radially around a point. The nodes on each axis represent the overall efficiency scores produced by the each of the different models. As shown, CSP 23 was positioned on the edge of the graph by all models and, therefore, is considered to be the most efficient CSP. However, although CSP 4 sits on the edge when undesirable outputs and integer-valued data are considered both separately and together, CSP 4 would be a risky choice for customers who need to deal with uncertainty as evidenced by the drop in efficiency scores from 1 to 0.463.

5.7 SUMMARY

The main contribution of this chapter is to present a reliable method for evaluating the performance of multiple CSPs from a supply chain perspective. Using this method, each CSP is individually assessed with an efficiency measure as part of a logical and sequential process using a series of two-stage SBM network DEA models. Additionally, the supply chain's inputs, intermediate variables, and outputs are concurrently considered, along with undesirable factors, integer-valued data, and stochastic data, to result in an overall performance measure in the context of the chain.

Our findings demonstrate the advantages of network DEA as a tool for determining performance efficiency at each stage of a cloud supply chain as well as the chain's overall efficiency. This technique offers rigor to studies on efficiency assessment in cloud supply chains but can also be used as a basic reference for researchers and practitioners when developing and applying DEA models to evaluate cloud network performance. Furthermore, this technique also addresses other significant issues in the cloud computing domain. For managers, the proposed models are able to identify inefficient CSPs or aspects of a CSP's service that need to be improved. Such insights provide valuable information to help cloud customers optimize the composition of their cloud services.

Chapter 6 NON-CONVEX TECHNOLOGY IN CLOUD SUPPLY CHAINS

INTRODUCTION

In most of the DEA models including the models proposed in the previous chapter for the performance measurement of cloud supply chains, DMU performance is often assessed against virtual benchmarks, not real-life observations. Moreover, the models proposed in the previous chapter do not benefit from a high discrimination power in the performance measurement of the cloud supply chain. To this end, in this chapter we develop the proposed models presented in Chapter 5 using non-convex technology which allows each DMU to be evaluated against actual DMUs instead of artificial ones. Furthermore, we incorporate the super-efficiency concept in the proposed model to increase discrimination power in relation to the performance of cloud service providers. Validation studies on a real-world data set of 24 cloud service providers demonstrate how greatly the efficiency and rankings of CSPs can change under different scenarios in the real world.

6.2 PROPOSED MODEL

The relative efficiency technique used in this study is based on models 3 and 4 proposed in Chapter 5.

As discussed in Chapter 5, the models consider two types of measures, integers and non-integers. It is assumed that I^{IN} and I^{NI} are two mutually exclusive but collectively exhaustive input subsets of integer- and non-integer-valued inputs. Mathematically, $I^{IN} \cup I^{NI} = I$ and $I^{IN} \cap I^{NI} = \phi$. Similarly, let the integer- and non-integer-valued outputs and intermediate measures be denoted as R^{IN}, R^{NI} and L^{IN}, L^{NI}, respectively. In addition, all the outputs and intermediate measures are partitioned into four subsets ($R^{INU}, R^{NIU}, R^{IND}, R^{NID}$ and $L^{INU}, L^{NIU}, L^{IND}, L^{NID}$) so as to consider the undesirable outputs.

The superscript *INU* represents the undesirable integer-valued variables; NIU denotes the undesirable non-integer-valued variables; INU, IND, and NID are the desirable integer-valued and non-integer-valued measures, respectively. We suggest the following MILP to measure the SBM efficiency of Stage 1 in the presence of integer-valued inputs and (un) desirable intermediate variables:

Model 6.1

$$\rho_1^* = \min \tau - \frac{1}{|I^{NI}| + |I^{IN}|}\left(\sum\nolimits_{i\in I^{NI}} \frac{s_i^-}{x_{io}} + \sum\nolimits_{i\in I^{IN}} \frac{(s_i^- + t_i^-)}{x_{io}}\right)$$

Subject to

$$1 = \tau + \frac{\left(\sum_{l\in L^{NID}} \frac{S_l^{+D}}{z_{lo}} + \sum_{l\in L^{IND}} \frac{\left(S_l^{+D} + t_l^{+D}\right)}{z_{lo}} + \sum_{l\in L^{NIU}} \frac{s_l^{-U}}{z_{lo}} + \sum_{l\in L^{INU}} \frac{\left(s_l^{-U} + t_l^{-U}\right)}{z_{lo}}\right)}{|L^{NID}| + |L^{IND}| + |L^{NIU}| + |L^{INU}|}$$

$$\tau x_{io} - s_i^- = \sum_{j=1}^{n} x_{ij}\lambda_j \qquad i \in I^{NI}$$

$$\bar{x}_i - s_i^- = \sum_{j=1}^{n} x_{ij}\lambda_j \qquad i \in I^{IN}$$

$$\tau x_{io} - t_i^- = \bar{x}_i \qquad i \in I^{IN}$$

$$\tau z_{lo} + s_l^{+D} = \sum_{j=1}^{n} z_{lj}\lambda_j \qquad l \in L^{NID}$$

$$\bar{z}_l + s_l^{+D} = \sum_{j=1}^{n} z_{lj}\lambda_j \qquad l \in L^{IND}$$

$$\tau z_{lo} + t_l^{+D} = \bar{z}_l \qquad l \in L^{IND}$$

$$\tau z_{lo} - s_l^{-U} = \sum\nolimits_{j=1}^{n} z_{lj}\lambda_j \qquad l \in L^{NIU}$$

$$\bar{z}_l - s_l^{-U} = \sum_{j=1}^{n} z_{lj}\lambda_j \qquad l \in L^{INU}$$

$$\tau z_{lo} - t_l^{-U} = \bar{z}_l \qquad l \in L^{INU}$$

$$\sum_{j\in J} \lambda_j = \tau$$

$$\lambda_j \geq 0 \qquad j \in J$$

$$s_i^- \geq 0 \qquad i \in I$$

$$t_i^- \geq 0 \qquad i \in I^{IN}$$

$$s_l^{+D} \geq 0 \qquad l \in L^D$$

$$t_l^{+D} \geq 0 \qquad l \in L^{IND}$$

$$s_l^{-U} \geq 0 \qquad l \in L^U$$

$$t_l^{-U} \geq 0 \qquad l \in L^{INU}$$

$$\bar{x}_i \in \mathbb{Z}_+ \qquad i \in I^{IN}$$

$$\bar{z}_l \in \mathbb{Z}_+ \qquad l \in L^{IN}$$

where $L^D = L^{IND} \cup L^{NID}$, $L^U = L^{INU} \cup L^{NIU}$, $\bar{x}_i$, and $\bar{z}_l$ are integer decision variables that indicate the integer-valued reference point for the input $i \in I^{IN}$ and the intermediate output $l \in L^{IN}$. Note that there are two types of slacks here: one for the integer-valued inputs, the other for the intermediate variables. The first type of slack, $i \in I^{IN}$, i.e., s_i^-, is the difference between the $\sum_{j=1}^{n} x_{ij}\lambda_j$ and the integer-valued $\bar{x}_i$. The second type of slack t_i^- is the difference between the integer-valued $\bar{x}_i$ and the projection τx_{io}. As a result, $s_i^- + t_i^-$ is the total slack for an integer-valued x_i. Similarly, the total slack for the integer-valued desirable measure is $s_l^{+D} + t_l^{+D}$ and $s_l^{-U} + t_l^{-U}$ for the undesirable intermediate measures. These values are considered in the objective function of Model (3) along with a set of normalization constraints.

Similarly, Model (6.2) evaluates the efficiency in Stage 2 based on the integer-valued intermediate variables and the undesirable outputs. Model (6.2) is formulated as follows:

Model 6.2

$$\rho_1^* = \min \tau - \frac{1}{|L^{NI}| + |L^{IN}|}\left(\sum\nolimits_{l\in L^{NI}} \frac{s_i^-}{z_{lo}} + \sum\nolimits_{l\in L^{IN}} \frac{(s_i^- + t_i^-)}{z_{lo}}\right)$$

Subject to

$$1 = \tau + \frac{\left(\sum_{r\in R^{NID}} \frac{s_r^{+D}}{y_{ro}} + \sum_{r\in R^{IND}} \frac{(S_r^{+D} + t_r^{+D})}{y_{ro}} + \sum_{r\in R^{NIU}} \frac{t_r^{-U}}{y_{ro}} + \sum_{r\in R^{INU}} \frac{(s_r^{-U} + t_r^{-U})}{z_{lo}}\right)}{|R^{NID}| + |R^{IND}| + |R^{NIU}| + |R^{INU}|}$$

$$\tau Z_{lo} - s_i^- = \sum_{j=1}^{n} z_{lj}\lambda_j \qquad i \in I^{NI}$$

$$\bar{z}_l - s_l^- = \sum_{j=1}^{n} z_{lj}\lambda_j \qquad i \in I^{IN}$$

$$\tau z_{lo} - t_l^- = \bar{z}_l \qquad i \in I^{IN}$$

$$\tau y_{ro} + s_r^{+D} = \sum_{j=1}^{n} y_{rj}\lambda_j \qquad l \in L^{NID}$$

$$\bar{y}_r + s_r^{+D} = \sum_{j=1}^{n} y_{rj}\lambda_j \qquad l \in L^{IND}$$

$$\tau y_{ro} + t_r^{+D} = \bar{y}_r \qquad l \in L^{IND}$$

$$\tau y_{ro} - s_r^{+U} = \sum\nolimits_{j=1}^{n} y_{rj}\lambda_j \qquad l \in L^{NIU}$$

$$\bar{y}_r - s_l^{+U} = \sum_{j=1}^{n} y_{rj}\lambda_j \qquad l \in L^{INU}$$

$\tau y_{ro} - t_r^{+U} = \bar{y}_r \qquad l \in L^{INU}$

$\sum_{j \in J} \lambda_j = \tau$

$\lambda_j \geq 0 \qquad j \in J$

$s_l^- \geq 0 \qquad i \in I$

$t_l^- \geq 0 \qquad i \in I^{IN}$

$s_r^{+D} \geq 0 \qquad l \in L^D$

$t_r^{+D} \geq 0 \qquad l \in L^{IND}$

$s_r^{+U} \geq 0 \qquad l \in L^U$

$t_r^{+U} \geq 0 \qquad l \in L^{INU}$

$z_l \in \mathbb{Z}_+ \qquad i \in I^{IN}$

$y_r \in \mathbb{Z}_+ \qquad l \in L^{IN}$

where $R^D = R^{IND} \cup R^{NID}$ and $R^U = R^{INU} \cup R^{NIU}$.

Definition 6.1. The optimal objective values $\boldsymbol{\rho_1^*}$ and $\boldsymbol{\rho_2^*}$ represent the SBM efficiency of Stages 1 and 2 for $\boldsymbol{DMU_o}$, respectively.

Definition 6.2. The overall SBM efficiency of DMU_o is $\frac{\rho_1^* + \rho_2^*}{2}$. If $\rho_1^* = \rho_2^* = 1$, the overall SBM-efficiency of DMU_o is good.

These two models evaluate the individual efficiency of each DMU_o in Stage 1 (Model (6.1) and Stage 2 (Model (6.2). The overall SBM efficiency of DMU_o is the average of Stages 1 and 2 (see Definition 6.2). However, Models (6.1) and (6.2) are based on convexity assumptions, plus the efficiency scores are derived from artificial observations of DMU performance. Deprins et al.'s (1984) FDH model computes efficiency based on actual observations of a DMU's performance by comparing inefficient DMUs with efficient DMUs to represent the efficiency frontier more precisely and in more practical terms.

Formally, the FDH technology is represented by its production possibility set (PPS) as follows:

$$P_{FDH} = \left\{(x,y) \middle| x \geq \sum_{j=1}^{J} \lambda_j x_j; y \leq \sum_{j=1}^{J} \lambda_j y_j; \right. \qquad (6.3)$$

$$\left. \sum_{j=1}^{J} \lambda_j = 1\ x, y \geq 0; \lambda_j \in |\{0,1\}; j = 1, \dots, J \right\}$$

where $x_j \geq 0$, $y_j \geq 0$ are observed input and output quantities for all DMUs, respectively. In effect, a point is in FDH's PPS set if all of its input coordinates are at least as large as the corresponding observed value vector for x_j for any $j = 1, \dots, J$ and the output coordinates are no greater than the corresponding observed value vector for y_j for this same j (Shabani, Torabipour & Saen 2015).

At this juncture, the FDH model and Models 6.1 and 6.2 are integrated for a more accurate efficiency evaluation. Model 6.4 is the integration of Models 6.1 and 6.3 for stage 1.

<u>Model 6.4</u>

$$\rho_1^* = \min \tau - \frac{1}{|I^{NI}| + |I^{IN}|}\left(\sum\nolimits_{i \in I^{NI}} \frac{s_i^-}{x_{io}} + \sum\nolimits_{i \in I^{IN}} \frac{(s_i^- + t_i^-)}{x_{io}}\right)$$

Subject to

$$1 = \tau + \frac{\left(\sum_{l \in L^{NID}} \frac{S_l^{+D}}{z_{lo}} + \sum_{l \in L^{IND}} \frac{\left(S_l^{+D} + t_l^{+D}\right)}{z_{lo}} + \sum_{l \in L^{NIU}} \frac{s_l^{-U}}{z_{lo}} + \sum_{l \in L^{INU}} \frac{\left(s_l^{-U} + t_l^{-U}\right)}{z_{lo}}\right)}{|L^{NID}| + |L^{IND}| + |L^{NIU}| + |L^{INU}|}$$

$$\tau x_{io} - s_i^- = \sum_{j=1}^{n} x_{ij} \lambda_j \qquad i \in I^{NI}$$

$$\bar{x}_i - s_i^- = \sum_{j=1}^{n} x_{ij}\lambda_j \qquad i \in I^{IN}$$

$$\tau x_{io} - t_i^- = \bar{x}_i \qquad i \in I^{IN}$$

$$\tau z_{lo} + s_l^{+D} = \sum_{j=1}^{n} z_{lj}\lambda_j \qquad l \in L^{NID}$$

$$\bar{z}_l + s_l^{+D} = \sum_{j=1}^{n} z_{lj}\lambda_j \qquad l \in L^{IND}$$

$$z_{lo} + t_l^{+D} = \bar{z}_l \qquad l \in L^{IND}$$

$$\tau z_{lo} - s_l^{-U} = \textstyle\sum_{j=1}^{n} z_{lj}\lambda_j \qquad l \in L^{NIU}$$

$$\bar{z}_l - s_l^{-U} = \sum_{j=1}^{n} z_{lj}\lambda_j \qquad l \in L^{INU}$$

$$\tau z_{lo} - t_l^{-U} = \bar{z}_l \qquad l \in L^{INU}$$

$$\sum_{j \in J} \lambda_j = \tau$$

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

$$\lambda_j \in \{0, 1\} \qquad j = 1, \dots, n$$

$$s_i^- \geq 0 \qquad i \in I$$

$$t_i^- \geq 0 \qquad i \in I^{IN}$$

$$s_l^{+D} \geq 0 \qquad l \in L^D$$

$$t_l^{+D} \geq 0 \qquad l \in L^{IND}$$

$$s_l^{-U} \geq 0 \qquad l \in L^U$$

$$t_l^{-U} \geq 0 \qquad l \in L^{INU}$$

$\bar{x}_i \in \mathbb{Z}_+$ $\qquad i \in I^{IN}$

$\bar{z}_l \in \mathbb{Z}_+$ $\qquad l \in L^{IN}$

Model 6.5 is the integration of Models 6.2 and 6.3 for stage 2.

<u>Model 6.5</u>

$$\rho_1^* = \min \tau - \frac{1}{|L^{NI}| + |L^{IN}|}\left(\sum\nolimits_{l \in L^{NI}} \frac{s_i^-}{z_{lo}} + \sum\nolimits_{l \in L^{IN}} \frac{(s_i^- + t_i^-)}{z_{lo}}\right)$$

s.t.

$$1 = \tau + \frac{\left(\sum_{r \in R^{NID}} \frac{s_r^{+D}}{y_{ro}} + \sum_{r \in R^{IND}} \frac{(S_r^{+D} + t_r^{+D})}{y_{ro}} + \sum_{r \in R^{NIU}} \frac{t_r^{-U}}{y_{ro}} + \sum_{r \in R^{INU}} \frac{(s_r^{-U} + t_r^{-U})}{z_{lo}}\right)}{|R^{NID}| + |R^{IND}| + |R^{NIU}| + |R^{INU}|}$$

$$\tau Z_{lo} - s_i^- = \sum_{j=1}^{n} z_{lj}\lambda_j \qquad i \in I^{NI}$$

$$\bar{z}_l - s_l^- = \sum_{j=1}^{n} z_{lj}\lambda_j \qquad i \in I^{IN}$$

$$\tau z_{lo} - t_l^- = \bar{z}_l \qquad i \in I^{IN}$$

$$\tau y_{ro} + s_r^{+D} = \sum_{j=1}^{n} y_{rj}\lambda_j \qquad l \in L^{NID}$$

$$\bar{y}_r + s_r^{+D} = \sum_{j=1}^{n} y_{rj}\lambda_j \qquad l \in L^{IND}$$

$$\tau y_{ro} + t_r^{+D} = \bar{y}_r \qquad l \in L^{IND}$$

$$\tau y_{ro} - s_r^{+U} = \sum_{j=1}^{n} y_{rj}\lambda_j \qquad l \in L^{NIU}$$

$$\bar{y}_r - s_l^{+U} = \sum_{j=1}^{n} y_{rj}\lambda_j \qquad l \in L^{INU}$$

$$\tau y_{ro} - t_r^{+U} = \bar{y}_r \qquad l \in L^{INU}$$

$$\sum_{j \in J} \lambda_j = \tau$$

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

$$\lambda_j \in \{0, 1\} \qquad j = 1, \dots, n$$

$$s_l^- \geq 0 \qquad i \in I$$

$$t_l^- \geq 0 \qquad i \in I^{IN}$$

$$s_r^{+D} \geq 0 \qquad l \in L^D$$

$$t_r^{+D} \geq 0 \qquad l \in L^{IND}$$

$$s_r^{+U} \geq 0 \qquad l \in L^U$$

$$t_r^{+U} \geq 0 \qquad l \in L^{INU}$$

$$z_l \in \mathbb{Z}_+ \qquad i \in I^{IN}$$

$$y_r \in \mathbb{Z}_+ \qquad l \in L^{IN}$$

Although Models (6.4) and (6.5) consider undesirable outputs, integer-valued data, and compare inefficient DMUs with actual efficient DMUs, they cannot differentiate among DMUs and most of them become efficient. To date, few researchers have explored the discriminatory power of the FDH model.

Of these few, Andersen & Petersen (1993) developed a modified FDH model to discriminate efficient DMUs. Van Puyenbroeck (1998) further modified Andersen and Petersen’s model. Mehrabian, Alirezaee & Jahanshahloo (1999) (MAJ) proposed the MAJ-FDH model. The drawback of all these models is that they sometimes produce such as infeasible solution(s). However, Shiraz et al. (2015) have since modified the MAJ-FDH model to ensure feasible and optimal solution(s). Models (6.6) and (6.7) are based on Shiraz et al. (2015) to differentiate between efficient DMUs and produce optimal and feasible solution(s).

Model 6.6

$$\tau^* = \min \tau = \frac{1}{|I^{NI}|+|I^{IN}|}\left(\sum_{i\in I^{NI}}\frac{\tilde{x}_i}{x_{io}} + \sum_{i\in I^{IN}}\frac{\tilde{x}_i}{x_{io}}\right)$$

s.t.

$$1 = \frac{1}{|L^{NID}|+|L^{IND}|+|L^{NIU}|+|L^{INU}|}\Big(\sum_{l\in L^{NID}}\frac{\tilde{z}_l}{z_{lo}} + \sum_{l\in L^{IND}}\frac{\tilde{z}_l}{z_{lo}} + \sum_{l\in L^{NIU}}\frac{\tilde{z}_l}{z_{lo}} + \sum_{l\in L^{INU}}\frac{\tilde{z}_l}{z_{lo}}\Big)$$

$$\tilde{x}_i \geq \sum_{\substack{j=1\\ j\neq o}}^{n} x_{ij}\lambda_j \qquad i\in I^{NI}$$

$$\tilde{\tilde{x}}_i \geq \sum_{\substack{j=1\\ j\neq o}}^{n} x_{ij}\lambda_j \qquad i\in I^{IN}$$

$$\tilde{\tilde{x}}_i \geq x_{io} \qquad i\in I^{IN}$$

$$\tilde{z}_l \leq \sum_{\substack{j=1\\ j\neq o}}^{n} z_{lj}\lambda_j \qquad l\in L^{NID}$$

$$\tilde{\tilde{z}}_l \leq \sum_{\substack{j=1\\ j\neq o}}^{n} z_{lj}\lambda_j \qquad l\in L^{NID}$$

$$\tilde{z}_l \le z_{lo} \qquad l \in L^{NID}$$

$$\tilde{z}_l \le \sum_{\substack{j=1 \\ j \neq o}}^{n} z_{lj}\lambda_j \qquad l \in L^{NIU}$$

$$\tilde{\tilde{z}}_l \le \sum_{\substack{j=1 \\ j \neq o}}^{n} z_{lj}\lambda_j \qquad l \in L^{INU}$$

$$\tilde{\tilde{z}}_l \le z_{lo} \qquad l \in L^{INU}$$

$\tilde{x}_i \ge t x_{io}$ and $\tilde{z}_l \le t z_{lo}$

$$\sum_{j \in J} \lambda_j = \tau$$

$$\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j = 1$$

$$\lambda_j \in \{0, 1\} \qquad j = 1, \dots, n, \quad j \neq o$$

$$s_i^- \ge 0 \qquad i \in I$$

$$t_i^- \ge 0 \qquad i \in I^{IN}$$

$$s_l^{+D} \ge 0 \qquad l \in L^D$$

$$t_l^{+D} \ge 0 \qquad l \in L^{IND}$$

$$s_l^{-U} \ge 0 \qquad l \in L^U$$

$$t_l^{-U} \ge 0 \qquad l \in L^{INU}$$

$$\bar{x}_i \in \mathbb{Z}_+ \qquad i \in I^{IN}$$

$$\bar{z}_l \in \mathbb{Z}_+ \qquad l \in L^{IN}$$

$$\tilde{z} \ge 0,$$

$$t \ge 0.$$

For stage 2 of the cloud supply chain, Model (6.7) not only makes difference between the performance of cloud service providers, it also produces optimal and feasible solution(s).

<u>Model 6.7</u>

$$\rho_1^* = \min \tau - \frac{1}{|L^{NI}| + |L^{IN}|}\left(\sum\nolimits_{l \in L^{NI}} \frac{\widetilde{z}_l}{z_{lo}} + \sum\nolimits_{l \in L^{IN}} \frac{\widetilde{z}_l}{z_{lo}}\right)$$

s.t.

$$1 = \frac{1}{|R^{NID}| + |R^{IND}| + |R^{NIU}| + |R^{INU}|}\left(\sum\nolimits_{r \in R^{NID}} \frac{\widetilde{y}_l}{y_{ro}} + \sum\nolimits_{r \in R^{IND}} \frac{\widetilde{y}_l}{y_{ro}} + \sum\nolimits_{r \in R^{NIU}} \frac{\widetilde{y}_l}{y_{ro}} + \sum\nolimits_{r \in R^{INU}} \frac{\widetilde{y}_l}{y_{ro}}\right)$$

$$\tilde{z}_l \geq \sum_{\substack{j=1 \\ j \neq o}}^{n} x_{lj}\lambda_j \qquad i \in I^{NI}$$

$$\tilde{\tilde{z}}_l \geq \sum_{\substack{j=1 \\ j \neq o}}^{n} z_{lj}\lambda_j \qquad i \in I^{IN}$$

$$\tilde{\tilde{z}}_l \geq z_{lo} \qquad i \in I^{IN}$$

$$\tilde{y}_r \leq \sum_{\substack{j=1 \\ j \neq o}}^{n} y_{rj}\lambda_j \qquad l \in L^{NID}$$

$$\tilde{\tilde{y}}_l \leq \sum_{\substack{j=1 \\ j \neq o}}^{n} y_{rj}\lambda_j \qquad l \in L^{IND}$$

$$\tilde{y}_r \leq y_{ro} \qquad l \in L^{IND}$$

$$\tilde{y}_r \leq \sum_{\substack{j=1 \\ j \neq o}}^{n} y_{rj}\lambda_j \qquad l \in L^{NIU}$$

$$\tilde{\tilde{y}}_r \leq \sum_{\substack{j=1 \\ j \neq o}}^{n} y_{rj}\lambda_j \qquad l \in L^{INU}$$

$$\tilde{\tilde{y}}_r \leq y_{ro} \qquad l \in L^{INU}$$

$\tilde{z}_l \geq tz_{lo}$ and $\tilde{y}_r \leq ty_{ro}$

$$\sum_{j \in J} \lambda_j = \tau$$

$$\sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j = 1$$

$$\lambda_j \in \{0, 1\} \qquad j = 1, \dots, n$$

$$s_l^- \geq 0 \qquad i \in I$$

$$t_l^- \geq 0 \qquad i \in I^{IN}$$

$$s_r^{+D} \geq 0 \qquad l \in L^D$$

$$t_r^{+D} \geq 0 \qquad l \in L^{IND}$$

$$s_r^{+U} \geq 0 \qquad l \in L^U$$

$$t_r^{+U} \geq 0 \qquad l \in L^{INU}$$

$$z_l \in \mathbb{Z}_+ \qquad i \in I^{IN}$$

$$y_r \in \mathbb{Z}_+ \qquad l \in L^{IN}$$

Definition 3. The overall SBM efficiency of DMU_o is $\frac{\rho_1^* + \rho_2^*}{2}$.

The next section presents a practical DEA evaluation with the above integrated models through an empirical study on 24 CSPs.

6.3 DATA AND EVALUATION OF THE PROPOSED MODELS

To evaluate the proposed models, we use the data set in Chapter 5. Each company was considered to be a DMU within a two-stage cloud supply chain: IaaS as Stage 1 and PaaS as Stage 2, as shown in Figure 6.1 The input for the IaaS stage was price. The intermediate inputs/outputs were latency (undesirable), memory and CPU (both integer-valued), and data transfer. Price was also used as an additional input for the second stage (PaaS). The outputs for the PaaS stage were availability, the number of security certifications (integer-valued), and service time delays (undesirable). Details of the sample data set are provided in Table 6.1.

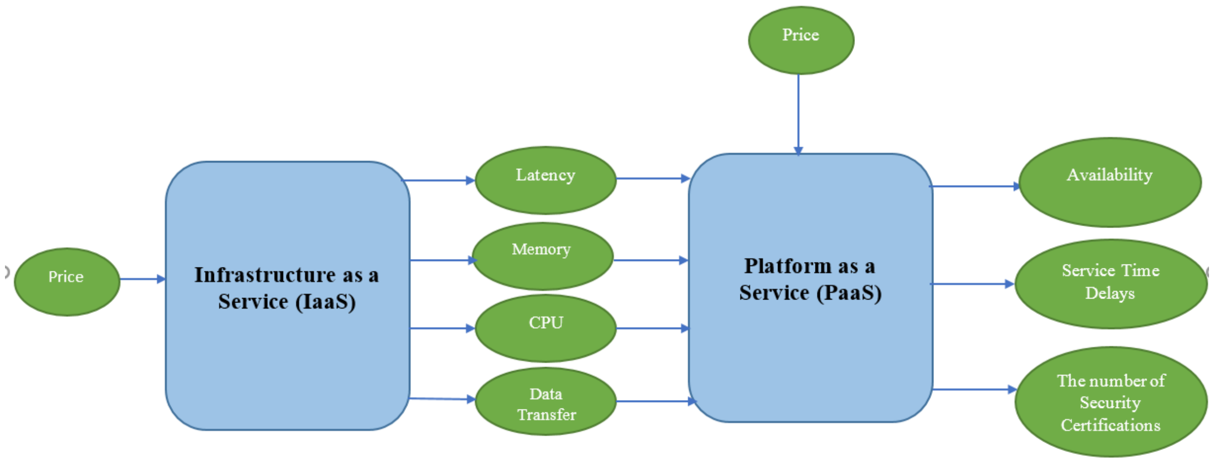

Figure 6.1 The two-stage cloud supply chain

Table 6.1 The data set of 24 CSPs

CSPs (DMUs)	Stage 1 Input	Stage 2 Input	Outputs			Intermediate variables			
	Price 1 (monthly/$)	Price 2 (1000 devices/$)	Availability (percentage)	Security (number)	Delayed service time (second)	Latency (millisecond)	Memory	CPU	Data transfer
1	80	35	100%	5	78	433	8	2	5
2	140.79	50	99.99%	3	101	49	7	2	3.2
3	80	47	100%	4	52	46	8	4	5
4	80	59	99. 95%	1	163	39	8	6	8
5	158	50	100%	3	41	45	2	4	0.5
6	110	45	100%	4	33	41	4	2	3
7	150	42	99.99%	4	139	68	16	6	8
8	156.24	49	100%	4	64	40	2	2	10
9	87.88	37	100%	2	149	46	2.048	3	3
10	16.65	49	99.92%	1	176	152	0.5	1	0.5
11	15	31	99.93%	1	180	40	0.5	1	3
12	79	40	100%	2	59	71	8	2	5
13	83	31	100%	4	115	62	7	1	3
14	64.95	43	100%	1	152	62	4	2	3
15	219	37	99.97%	2	119	46	8	8	10

16	150	42	99.99%	4	26	68	16	6	8
17	140	42	99.99%	4	176	70	16	6	6
18	110	45	100%	4	143	41	4	2	3
19	80	47	100%	4	154	46	8	4	4
20	83	31	100%	4	179	62	7	1	3
21	15	34	99.94%	1	165	40	0.5	1	3
22	80	62	99.96%	1	134	40	8	6	8
23	15	31	99.99%	1	126	40	0.5	1	3
24	221	38	99.93%	2	177	48	8	8	10

Table 6.2 The results from Models (1)-(9)

CSPs (DMUs)	**Model (5.1)**			**Model (5.3)**	**Model (5.4)**		**Model (5.5)**	**Model (5.6)**		**Model (5.7)**	**Model (5.8)**		**Rank**
	Stage 1	**Stage 2**	**Overall**	**Stage 1**	**Stage 2**	**Overall**	**Stage 1**	**Stage 2**	**Overall**	**Stage 1**	**Stage 2**	**Overall**	
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	2.159	2.03	1.93	3
2	0.453	0.556	0.505	0.240	0.751	0.496	1.000	1.000	1.000	1.09	1.13	1.11	17
3	1.000	0.527	0.763	0.493	0.467	0.480	1.000	1.000	1.000	1.25	1.17	1.21	15
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.86	1.69	1.775	5
5	0.149	1.000	0.575	0.142	1.000	0.571	0.643	1.000	0.8215	0.643	1.000	0.8215	24
6	0.499	1.000	0.750	0.256	0.500	0.378	1.000	1.000	1.000	1.05	1.03	1.04	19
7	1.000	0.283	0.641	1.000	0.390	0.695	1.000	0.791	0.8955	1.000	0.791	0.8955	23

8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	2.159	2.01	2.084	2
9	0.555	0.574	0.564	0.317	0.741	0.529	1.000	1.000	1.000	1.07	1.10	1.085	18
10	0.314	1.000	0.657	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.63	1.45	11
11	1.000	0.870	0.935	1.000	1.000	1.000	1.000	1.000	1.000	1.68	1.52	1.6	6
12	0.753	0.389	0.571	0.599	0.261	0.430	1.000	1.000	1.000	1.05	1.000	1.025	20
13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.92	1.79	1.855	4
14	0.596	0.317	0.457	0.411	0.398	0.405	1.000	1.000	1.000	1.19	1.13	1.16	16
15	1.000	0.261	0.630	1.000	0.303	0.652	1.000	1.000	1.000	1.51	1.211	1.36	12
16	1.000	1.000	1.000	1.000	0.195	0.598	1.000	1.000	1.000	1.29	1.22	1.24	14
17	1.000	0.297	0.648	1.000	0.491	0.745	1.000	1.000	1.000	1.58	1.527	1.475	9
18	0.499	0.722	0.611	0.256	1.000	0.628	1.000	1.000	1.000	1.29	1.35	1.32	13
19	0.941	0.474	0.708	0.488	1.000	0.744	1.000	1.000	1.000	1.32	1.65	1.485	10
20	1.000	0.848	0.924	1.000	1.000	1.000	1.000	1.000	1.000	1.75	1.39	1.57	7
21	0.956	0.894	0.925	0.956	1.000	0.978	1.000	1.000	1.000	1.49	1.61	1.55	8
22	0.970	0.274	0.622	1.000	0.330	0.665	1.000	0.894	0.947	1.000	0.894	0.947	22
23	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	2.412	2.161	2.286	1
24	0.972	0.250	0.611	1.000	0.423	0.712	1.000	0.994	0.997	1.000	0.994	0.997	21

To assess the rationality of the proposed models, we compared the results of the efficiency evaluations for each of the 24 CSPs obtained using the proposed models as shown in Table 6.1. We first calculated the efficiency scores for Stage 1 (IaaS) using Model (5.1) and for Stage 2 (PaaS) using Model (5.22). The average of these two scores represents the overall efficiency and shows that six of the 24 CSPs (CSPs 1, 4, 6, 13, 16, and 23) operate efficiently in this supply chain. This is because these CSPs satisfy Definition 6.2 and, according to this definition, they are efficient in both stages. CSP 14 had the worst overall efficiency mainly due to its low performance in Stage 2.

When considering both undesirable outputs and integer-valued data, the overall efficiency scores change, as shown in Table 6.2 column 7. Compared to the overall efficiency scores without this condition [column 4], eight of the 24 CSPs (CSPs 1, 4, 8, 10, 11, 13, 20 and 23) were efficient. CSPs 6 and 16 were only efficient when undesirable outputs were not considered, and CSPs 10, 11, and 20 changed from inefficient to efficient once the undesirable outputs were included in the evaluation.

Column 10 in Table 6.2 shows the overall efficiency scores with undesirable outputs, integer-valued data, and introduces the comparisons to actual efficiency. These scores are expected to be more precise and practical. On the surface, the results indicate that many CSPs are now more efficient than before due to the convex assumptions in Models (6.4) and (6.5). However, their discrimination power has decreased, and they are unable to inform decision makers to select the optimal among a large number of CSP.

Columns 11 and 12 in Table 6.2 show the overall efficiency scores obtained from Models 6.6 and 6.7, with the overall rankings in the last column. Unlike the overall efficiency scores in column

8, CSP 23 has moved from its place among the pack to the most efficient CSP, while CSP 5 remained the least efficient.

6.4 SUMMARY

The current weaknesses with many DEA models to provide a holistic performance evaluation of DMU efficiency that reflects real-world needs limits their potential applications. In this chapter, we attempted to build an integrated DEA that includes all the benefits of existing DEA models and none of the shortcomings. The model combines a two-stage SBM network model, non-convex technology, and super efficiency for performance evaluation of cloud supply chain. The SBM model simultaneously considers both undesirable outputs and integer-valued data. The non-convex technology measures DMU efficiency against actual performance metrics, and a super efficiency technique ranks each DMU according to its performance, offering managers a high level of discriminatory power. The results of an empirical DEA with 24 CSPs show how much efficiency varies when none, some, or all of these factors are considered in the model.

Chapter 7 CONCLUSIONS AND FUTURE WORKS

7.1 CONCLUSIONS

Cloud services have become a significant paradigm for meeting organizations' diverse information technology needs. In the cloud marketplace, many CSPs offer a wide range of services with highly competitive prices and performance. In such an intensely competitive marketplace, finding the optimal CSPs that can satisfy the QoS indicators is demanding (Fan, Yang & Pei 2014). One of the major issues with current approaches to evaluate and select the optimal CSPs is that none of them can distinguish the differences between the performance of CSPs in a highly competitive marketplace. Furthermore, a majority of the current approaches are effort-intensive and time-consuming and provide ranking irregularities in the performance measurement of cloud service providers. Moreover, the existing approaches for performance evaluation and the selection of CSPs suffer from subjective judgments, which results in a biased evaluation. In addition, no study addresses the performance of CSPs in the context of an entire supply chain, where multiple services interact to achieve a business objective or goal. Furthermore, the existing approaches do not consider integer-valued data, undesirable outputs, and stochastic data which can result in inaccurate results.

Considering these issues and that the initial investment in cloud computing services can be both costly and time-consuming, performance measurement techniques such as DEA and network DEA can serve as appropriate decision support system (DSS) tools. This study investigates and develops a number of advanced DEA and network DEA models for the evaluation and selection of CSPs. The models proposed and applied in this study have been designed to help managers and decision-makers to select appropriate CSPs in a highly competitive market.

In summary, the main contributions of this research are as follows:

1. **Developing a mixed ideal and anti-ideal DEA model:** This study developed three novel models using the ideal concept in one of the most rigorous DEAs, the enhanced Russell model, to increase the ability to discriminate between CSPs. The proposed models can rank and select DMUs with higher sensitivity. Due to the importance of selecting the optimal CSPs and the advantages of the proposed models, we apply them to evaluate and select the optimal CSPs in a real case study. The results and analysis show that there are six main advantages of the proposed ranking methods.

2. **Proposing network SBM model for performance evaluation of cloud service providers:** This study illustrates how the network SBM DEA model, which is a benchmark frontier technique, can be applied in evaluating the performance of CSPs. The findings of this model demonstrate the advantages of network DEA in determining the efficiency of each stage as well as the overall efficiency of CSPs. In addition, the model is able to identify the inefficient aspects of the CSPs which need to be improved. This is a key advantage of the network DEA approach over all the other existing approaches. The current study proposes a rigorous technique to assess the efficiency of CSPs for the first time and can be used as a reference for researchers and practitioners seeking to develop other DEA and network DEA models to evaluate the performance of CSPs and, most importantly, to address other significant issues in the cloud domain.

3. **Developing a new super-efficiency DEA model in the presence of undesirable outputs and weight restrictions:** In this study, we developed a novel DEA which considers both undesirable outputs and weight restrictions simultaneously to evaluate and select CSPs. In

addition, the proposed model provides a more objective performance assessment in comparison with other approaches; the model also benefits from high discrimination power in the performance measurement of cloud service providers.

4. **Developing a decision support system for evaluating efficiency in cloud supply chains**: In this study, we designed a decision support system that accurately evaluates the efficiency of multiple cloud service providers in a supply chain. The system comprises a suite of two-stage network DEAs that consider undesirable outputs, integer-valued data, and stochastic data. This study is the first attempt to address this problem. The models proposed in this study provide cloud computing customers with an optimal CSP composition given their individual priorities, such as cost or latency.

5. **Developing a novel network SBM DEA model for performance evaluation of cloud supply chains:** In this study, we proposed a non-convex technology to evaluate the efficiency of a cloud supply chain. The proposed model compares DMUs with actual DMUs instead of artificial ones. Since the developed model was not able to rank DMUs fully, the super efficiency technique was developed to address this issue in the performance evaluation process. The obtained results using the proposed model show how much efficiency scores can change under different conditions. Moreover, not only are the obtained efficiency scores of CSPs based on actual CSPs, there is also a discrimination between a large number of efficient CSPs which provide the cloud supply chain.

6. **Evaluation of the Proposed Models:** The models proposed in this study are evaluated using a real data set. Moreover, each of the models proposed in this study is compared

against the models in the literature to show the advantages of the proposed models. The results obtained from the mixed ideal and anti-ideal DEA model show the discrimination power and ability of the model to rank non-extreme efficient DMUs which results in the complete ranking of all DMUs. The results obtained using the super-efficiency model demonstrates the advantage of the proposed model to rank extremely efficient CSPs and the ability to deal with undesirable factors and decision maker's subjective judgments in the performance evaluation process. The results obtained in Chapters 5 and 6 also show how well the performance of cloud supply chains can be evaluated using the proposed model under different conditions.

Table 7.1 Contributions of the current research

The proposed models	Key contributions
Developing a mixed ideal and anti-ideal DEA model	• Increasing the ability to differentiate between the performance of CSPs. • Considering both pessimistic and optimistic scenarios of data envelopment analysis in the proposed models.
New super-efficiency DEA model in the presence of undesirable outputs and weight restrictions	• The ability to differentiate between CSPs even if they have been given the same rank or are rated the same using the current methods. • More flexibility to select a similar directional vector or specific vector for all DMUs or each DMU respectively. • The ability to consider undesirable outputs and weight restrictions simultaneously to measure the performance of CSPs. .
New network SBM DEA models for the performance evaluation of cloud supply chains	• For the first time, we proposed models to evaluate the overall efficiency of providers within a cloud supply chain. • The proposed network DEA models concurrently consider undesirable outputs, integer-valued data, and stochastic data simultaneously. • The proposed models provide cloud computing customers with an optimal CSP composition given their individual priorities, such as cost or latency. • The applicability and capability of the proposed models were evaluated through a case study involving 24 cloud service providers.

New super network SBM FDH models	• For the first time, both integer-valued data and undesirable outputs are considered simultaneously in super DEA and network DEA models. • For the first time, a combination of SBM FDH models is developed. • The proposed models are more objective compared to conventional SBM DEA models. • The proposed models not only rank all CSPs, they also compute a single efficient CSP.

In this research, we encountered two major limitations. Firstly, due to limitations in collecting information from CSPs, we had to select 8 metrics including price, latency, availability and the number of security certifications. However, there are other metrics such as reliability, usability and portability which can be used to measure the performance of CSPs. Furthermore, a few CSPs provided three types of cloud computing services, IaaS, PaaS and SaaS. Hence, we had to measure the performance of cloud supply chains in two-stage structures.

7.2 FUTURE WORKS

Further research can be undertaken based on the results of this study in the domain of cloud computing as follows:

- In the proposed mixed ideal and anti-ideal DEA model, we assumed that there was no undesirable data or integer-valued data in the performance evaluation of CSPs, however, we showed that some data can be undesirable and integer-valued in the performance evaluation process of CSPs. To obtain more accurate results, developing a mixed ideal and anti-ideal DEA model in the presence of both undesirable and integer-valued data could be an interesting topic for researchers. Furthermore, the results of the new model can be compared with the proposed mixed ideal and anti-ideal DEA model in this study.

- In this study, we focussed on the efficiency of IaaS and PaaS using a two-stage network DEA. Our next study will explore a new dynamic network DEA for the performance measurement of cloud supply chains consisting of IaaS, PaaS, and SaaS. To carry out this research as well as develop the new model, new data spanning different time periods of the cloud supply chain is required. The newly proposed dynamic network DEA model can give a more accurate analysis of the performance of cloud supply chains for different time periods.

- Another future research direction can be developing new DEA models for the performance evaluation of trust-based CSPs in the presence of dual-factors. To be more precise, in many practical applications, there are certain factors such as service-quality credence (CRE) and service-quality experience (EXP) which can play the role of both inputs and outputs named dual-role factors (Azadi and Saen, 2011). The new model can be developed based on the models proposed in this study and incorporating statistical methods in the performance measurement of trust-based CSPs.

7.3 APPENDIX

Table 7.2 Abbreviations

AIDMU	Anti-Ideal Decision-Making Units
AHP	Analytical Hierarchy Process
ANP	Analytic Network Process
AP	Andersen and Petersen
AWS	Amazon Web Services
BCC	Banker, Charnes, and Cooper
CCDEA	Chance-constrained Data Envelopment Analysis
CCR	Charnes, Cooper and Rhodes
CCP	Chance-Constrained Programming
CRS	Constant Return To Scale
DBaaS	Database As A Service
DEA	Data Envelopment Analysis
DM	Decision maker
DEMATEL	Decision Making, Trial and Evaluation Laboratory
DSS	Decision Support System
CSMIC	The Cloud Service Measurement Initiative Consortium

CMTES	Compliance-Based Multi-Dimensional Trust Evaluation System
DDF	Directional distance function
FMSs	Flexible manufacturing systems
EM	Expectation–Maximization
ERM	Enhanced Russell Model
FDH	Free Disposal Hull
IBM	International Business Machines
IS	Information Systems
IaaS	Instructure as a System
IT	Information Technology
GA	Genetic Algorithm
GRA	Grey Relational Analysis
KM	Knowledge Management
MAJ	Mehrabian, Alirezaee, and Jahanshahloo
MCDM	Multiple-Criteria Decision-Making
MCDA	Multiple-Criteria Decision Analysis
MDHP	Minimum Distance-Helly Property
MILP	Mixed integer linear programming
NIST	The National Institute of Standards and Technology
PaaS	Platform as a Service
PEIC	Procedure to evaluate Integer Congestion

PPM	Poincare Plot Method
PPS	Production Possibility Set
RAM	Range Adjusted Measure
QFD	Quality function deployment
QOS	Quality of Service
SaaS	Software as a Service
SBM	Slacks-Based Measure
SOCP	Second order cone programming
SelCSP	Select Cloud Service Provider
SLA	Service Level Agreement
SMI	Service Measurement Index
SVD	Singular Vector Decomposition
SRS	Service Ranking System
SWAT	Symmetric weight assignment technique
TOPSIS	Technique for order of preference by similarity to ideal solution
VPN	Virtual private network
VRS	Variable returns to scale
WTQS	The Weighted Tuned Queuing Scheduling

Table 7.3 The data set for CSPs

CSPs (DMUs)	Inputs		Outputs					
	Price (monthly/$)	Latency (ms)	Memory (GB)	Storage (GB)	Data transfer (TB)	CPU	Availability (Monthly)	The number of security certifications
1	80	433	8	80	5	2	100%	5
2	140.79	49	7	100	3.2	2	99.98%	3
3	80	46	8	80	5	4	100%	4
4	80	39	8	200	8	6	99.94%	1
5	158	45	2	500	0.5	4	100%	3
6	110	41	4	100	3	2	99.99%	4
7	150	68	16	384	8	6	99.994%	4
8	160	32	16.384	170	2	8	99.99%	1
9	156.24	40	2	40	10	2	100%	4
10	87.88	46	2.048	90	3	3	99.99%	2
11	16.65	152	0.5	20	0.5	1	99.89%	1
12	15	40	0.5	10	3	1	99.93%	1
13	79	71	8	80	5	2	100%	2
14	83.00	62	7	100	3	1	100%	4
15	64.95	62	4	250	3	2	100%	1
16	5	45	1	20	1	1	99.98%	1
17	219	46	8	300	10	8	99.74%	2
18	82.60	32	2	100	18	2	99.99%	1

Table 7.4 The data set for cloud supply chains

CSPs (DMUs)	Stage 1 Input	Stage 2 Input	Outputs			Intermediate variables			
	Price 1 (Monthly/$)	Price 2 (1000 Devices/$)	Availability (Percentage)	Security (Number)	Delayed service time (second)	Latency (Millisecond)	Memory	CPU	Data transfer
1	80	35	100%	5	78	433	8	2	5
2	140.79	50	99.99%	3	101	49	7	2	3.2
3	80	47	100%	4	52	46	8	4	5
4	80	59	99. 95%	1	163	39	8	6	8
5	158	50	100%	3	41	45	2	4	0.5
6	110	45	100%	4	33	41	4	2	3
7	150	42	99.99%	4	139	68	16	6	8
8	156.24	49	100%	4	64	40	2	2	10
9	87.88	37	100%	2	149	46	2.048	3	3
10	16.65	49	99.92%	1	176	152	0.5	1	0.5
11	15	31	99.93%	1	180	40	0.5	1	3
12	79	40	100%	2	59	71	8	2	5
13	83	31	100%	4	115	62	7	1	3
14	64.95	43	100%	1	152	62	4	2	3
15	219	37	99.97%	2	119	46	8	8	10
16	150	42	99.99%	4	26	68	16	6	8
17	140	42	99.99%	4	176	70	16	6	6
18	110	45	100%	4	143	41	4	2	3
19	80	47	100%	4	154	46	8	4	4
20	83	31	100%	4	179	62	7	1	3
21	15	34	99.94%	1	165	40	0.5	1	3
22	80	62	99.96%	1	134	40	8	6	8
23	15	31	99.99%	1	126	40	0.5	1	3
24	221	38	99.93%	2	177	48	8	8	10

7.3.1 Implementation for the proposed models

Mixed ideal and anti-ideal DEA model

γ_1^* = min 0.5*(u1*(8-5.80) + v1*(98.23-80);

0.5*(u1*(8-5.80) + v1*(98.23-80) + 0.5*(u1*(5.80-8) + v1*(80-98.23) = 1

u1*8 + u2*7 + u3*8 + u4*8 + u5*2 + u6*4 + u7*16 + u8*16.384 + u9*2 + u10*2.048 + u11*0.5 + u12*0.5 + u13*8 + u14*7 + u15*4 + u16*1 + u17*8 + u18*2 + u1*80 + u2*100 + u3*80 + u4*200 + u5*500 + u6*100 + u7*384 + u8*170+ u9*40 + u10*90 + u11*20+ u12*10+ u13*80 + u14*100 + u15*250 + u16*20 + u17*300 + u18* 100 + u1*5 + u2*3.2 + u3*5 + u4*8 + u5*0.5 + u6*3 + u7*8 + u8*2 + u9*10 + u10*3+ u11*0.5 + u12*3 + u13*5 + u14*3 + u15*3 + u16*1 + u17*10 + u18*18 + u1*2 + u2*2 + u3*4 + u4*6 + u5*4 + u6*2 + u7*6 + u8*8 + u9*2 + u10*3+ u11*1 + u12*1 + u13*2 + u14*1 + u15*2 + u16*1 + u17*8 + u18*2 + u1*100 + u2*99.98 + u3*100 + u4*99.94 + u5*100 + u6*99.99 + u7*99.994 + u8*99.99 + u9*100 + u10*99.99+ u11*99.89 + u12*99.93 + u13*100 + u14*100 + u15*100 + u16*99.98 + u17*99.74 + u18*99.99 + u1*5 + u2*3 + u3*4 + u4*1 + u5*3 + u6*4 + u7*4 + u8*1 + u9*4 + u10*2+ u11*1 + u12*1 + u13*2 + u14*4 + u15*1 + u16*1 + u17*2+ u18*1 – v1*80 + v2*140.79 + v3*80 + v4*80 + v5*158 + v6*110 + v7*150 + v8*160 + v9*156.24 + v10*87.88 + v11*16.65 + v12*15 + v13*79 + v14*83 + v15*64.95 + v16*5 + v17*219 + v18* 82.60 + v1*433 + v2*49 + v3*46 + v4*39 + v5*45 + v6*41 + v7*68 + v8*32 + v9*40 + v10*46 + v11*152 + v12*40 + v13*71 + v14*62 + v15*62 + v16*45 + v17*46 + v18* 32<= 0;

v1*80 + v1*433 –M1-t/m <= 0;

a*t – u1*8 + u1*80 + u1*5 + u1*2 + u1*100 + u1*5 + w1<= 0;

M1–w1– t*B*<= 0;

Network SBM model

Min = 0.4*(1-0.5*(S1/80 + S2/433));

80 = 80*L11 + 140.79*L21 + 80*L31 + 80*L41 + 158*L51 + 110*L61 + 150*L71 + 160*L81 + 156.24*L91 + 87.88*L101 + 16.65*L111 + 15*L121 + 79*L131 + 83*L141 + 64.95*L151 + 5*L161 + 219*L171 + 82.60*L181 + S1;

433 = 433*L11 + 49*L21 + 46*L31 + 39*L41 + 45*L51 + 41*L61 + 68*L71 + 32*L81 + 40*L91 + 46*L101 + 152*L111 + 40*L121 + 71*L131 + 62*L14 + 62*L15 + 45*L16 + 46*L17 + 32*L18 + S2;

100 = 100*L12 + 99.9898*L22 + 100*L32 + 99.9453*L42 + 100*L52 + 99.9987*L62 + 99.994*L72 + 99.9993*L82 + 100*L92 + 99.9968*L102 + 99.8938*L112 + 99.9303*L122 + 100*L132 + 100*L142 + 100*L152 + 99.9876*L162 + 99.7473*L172 + 99.999*L182 - S3;

5 = 5*L12 + 3*L22 + 4*L32 + 1*L42 + 3*L52 + 4*L62 + 4*L72 + 1*L82 + 4*L92 + 2*L102 + 1*L112 + 1*L122 + 2*L132 + 4*L142 + 1*L152 + 1*L162 + 2*L172 + 1*L182 - S4;

5 = 5*L12 + 3.2*L22 + 5*L32 + 8*L42 + 0.5*L52 + 3*L62 + 8*L72 + 2*L82 + 10*L92 + 3*L102 + 0.5*L112 + 3*L122 + 5*L132 + 3*L142 + 3*L152 + 1*L162 + 10*L172 + 18*L182 - S5;

L11 + L21 + L31 + L41 + L51 + L61+ L71 + L81 + L91 + L101 + L111 + L121 + L131 + L141 + L151 + L161+ L171 + L181 = 1;

L12 + L22 + L32 + L42 + L52 + L62+ L77 + L82 + L92 + L102 + L112 + L122 + L132 + L142 + L152 + L162+ L172 + L182 = 1;

8 = 8*L11 + 7*L21 + 8*L31 + 8*L41 + 2*L51 + 4*L61 + 16*L71 + 16.384*L81 + 2*L91 + 2.048*L101 + 0.5*L111 + 0.5*L121 + 8*L131 + 7*L141 + 4*L151 + 1*L161 + 8*L171 + 2*L181;

8 = 8*L12 + 7*L22 + 8*L32 + 8*L42 + 2*L52 + 4*L62 + 16*L72 + 16.384*L82 + 2*L92 + 2.048*L102 + 0.5*L112 + 0.5*L122 + 8*L132 + 7*L142 + 4*L152 + 1*L162 + 8*L172 + 2*L182;

80 = 80*L11 + 100*L21 + 80*L31 + 200*L41 + 500*L51 + 100*L61 + 384*L71 + 170*L81 + 40*L91 + 90*L101 + 20*L111 + 10*L121 + 80*L131 + 100*L141 + 250*L151 + 20*L161 + 300*L171 + 100*L181;

80= 80*L12 + 100*L22 + 80*L32 + 200*L42 + 500*L52 + 100*L62 + 384*L72+ 170*L82 + 40*L92 + 90*L102 + 20*L112 + 10*L122 + 80*L132 + 100*L142 + 250*L152 + 20*L162 + 300*L172 + 100*L182;
2 = 2*L11 + 2*L21 + 4*L31 + 6*L41 + 4*L51 + 2*L61 + 6*L71 + 8*L81 + 2*L91 + 3*L101 + 1*L111 + 1*L121 + 2*L131 + 1*L141 + 2*L151 + 1*L161 + 8*L171 + 2*L181;
2 = 2*L12 + 2*L22 + 4*L32 + 6*L42 + 4*L52 + 2*L62 + 6*L72 + 8*L82 + 2*L92 + 3*L102 + 1*L112 + 1*L122 + 2*L132 + 1*L142 + 2*L152 + 1*L162 + 8*L172 + 2*L182;

Super-efficiency Russell directional distance function (SRDDF) DEA model

```
Max= (ti (1)+ti (2)   + ( po1 (1) + po1 (2) + po1 (3) +  po1 (5) +  po1 (7)  ) ) ;


@SUM (DMUS (J) : X (j,1)*(la (j)+m(j))) <= t*X (w,1)-(ti (1)*X (w,1)) ;
@SUM (DMUS (J) : X (j,2)*(la (j)+m(j)))  <= t*X (w,2)-(ti (2)*X (w,2));

(@SUM (DMUS (J) : Y (j,1)*(la (j) ))+  @SUM (DMUSS (jJ): q (jj,1)*lan (jj))  ) >= t*Y
(W,1)+(po1(1)*Y (W,1)) ;
(@SUM (DMUS (J) : Y (j,2)*(la (j) ))+  @SUM (DMUSS (jJ): q (jj,2)*lan (jj))  ) >= t*Y
(w,2)+(po1(2)*Y (W,2)) ;
(@SUM (DMUS (J) : Y (j,3)*(la (j) ))+  @SUM (DMUSS (jJ): q (jj,3)*lan (jj))  ) >= t*Y
(w,3)+(po1(3)*Y (W,3));

(@SUM (DMUS (J) : Y (j,4)*(la (j) ))+  @SUM (DMUSS (jJ): q (jj,4)*lan (jj))   )=t* Y
(w,4);

@GIN(h);
(@SUM (DMUS (J) : Y (j,5)*(la (j) ))+  @SUM (DMUSS (jJ): q (jj,5)*lan (jj))   )>= h;
h=t*Y (w, 5)+(po1(5)*Y (w, 5));


(@SUM (DMUS (J) : Y (j,6)*(la (j) ))+  @SUM (DMUSS (jJ): q (jj,6)*lan (jj))   )= t*Y
(w,6);


@GIN(k);
(@SUM (DMUS (J) : Y (j,7)*(la (j) ))+  @SUM (DMUSS (jJ): q (jj,7)*lan (jj))   )>= k;
k=t*Y (w, 7)+(po1(7)*Y (w, 7));


@SUM (DMUS (J) : (la (j)+m(j))) =t;

a=(1/(1+(ti (1)+ti (2)   + ( po1 (1) + po1 (2) + po1 (3) +  po1 (5) +  po1 (7)  ) ) ));
```

a=(1/(1+(ti (1)/X (w,1)+ti (2)/X (w,2)   + (po1 (1)/y(w,1) + po1 (2)/y (w,2) + po1 (3)/y (w,3) + po1 (5)/y (w,5) + po1 (7)/y (w,7)))));
b=X (w,1)-(ti (1)*X (w,1)) ;
c=X (w,2)-(ti (2)*X (w,2)) ;
e=Y (W,1)+(po1(1)*Y (W,1)) ;
f=Y (W,2)+(po1(2)*Y (W,2)) ;
g=Y (W,3)+(po1(3)*Y (W,3)) ;
l= Y (w,4) ;
r=Y (w, 5)+(po1(5)*Y (w, 5));
n= Y (w,6) ;
o=Y (w, 7)+(po1(7)*Y (w, 7));

Network FDH SBM model with integer-valued data and undesirable outputs

P2* = min t-1/LN+LI*(lELNI*sl/433+8+2+5 + lELIN*sl+tl/433+8+2+5);

1=t+NID*srd/100+5+78+433+IND*srd+trd/100+5+78+NIU*tru/100+5+78+433+inu*sru+tru/100+5+78+433/RNID+RIND+R+NIU+RINU;

t*433-sl= 433*L11 + 49*L21 + 46*L31 + 39*L41 + 45*L51 + 41*L61 + 68*L71 + 40*L81 + 46*L91 + 152*L101 + 40*L111 + 71*L121 + 62*L131 + 62*L141 + 46*L151 + 68*L161 + 70*L171 + 41*L181 + 46*L191 + 62*L201 + 40*L211 + 40*L221 + 48*L231+ 48*L241;

zl-sl= 433*L11 + 49*L21 + 46*L31 + 39*L41 + 45*L51 + 41*L61 + 68*L71 + 40*L81 + 46*L91 + 152*L101 + 40*L111 + 71*L121 + 62*L131 + 62*L141 + 46*L151 + 68*L161 + 70*L171 + 41*L181 + 46*L191 + 62*L201 + 40*L211 + 40*L221 + 48*L231+ 48*L241;

t*433-tl-zl;

t*100+srd = 100*L12 + 99.9898*L22 + 100*L32 + 99.9453*L42 + 100*L52 + 99.9987*L62 + 99.994*L72 + 99.9993*L82 + 100*L92 + 99.9968*L102 + 99.8938*L112 + 99.9303*L122 + 100*L132 + 100*L142 + 100*L152 + 99.9876*L162 + 99.7473*L172 + 100*L182 +100*L192 + 100*L202 + 99.94*L212 + 99.96*L222 + 99.99*L232 + 99.93*L242;

100+srd = 100 = 100*L12 + 99.9898*L22 + 100*L32 + 99.9453*L42 + 100*L52 + 99.9987*L62 + 99.994*L72 + 99.9993*L82 + 100*L92 + 99.9968*L102 + 99.8938*L112 + 99.9303*L122 + 100*L132 + 100*L142 + 100*L152 + 99.9876*L162 + 99.7473*L172 + 100*L182 +100*L192 + 100*L202 + 99.94*L212 + 99.96*L222 + 99.99*L232 + 99.93*L242;
t*8+trd = yr;

t*8-sru = 8*L11 + 7*L21 + 8*L31 + 8*L41 + 2*L51 + 4*L61 + 16*L71 + 16.384*L81 + 2*L91 + 2.048*L101 + 0.5*L111 + 0.5*L121 + 8*L131 + 7*L141 + 4*L151 + 1*L161 + 8*L171 + 2*L18 + 8*L191 + 7*L201 + 0.5*L211 + 8*L221 + 0.5*L231 + 8*L21;

t*78-sru = 78*L11 + 101*L21 + 52*L31 +163*L41 + 41*L51 + 4*L33 + 139*L71 + 64*L81 + 149*L91 + 176*L101 + 180*L111 + 59*L121 + 115*L131 + 152*L141 + 119*L151 + 26*L161 + 8*L176 + 143*L18 + 154*L191 + 179*L201 + 165*L211 + 134*L221 + 126*L231 + 177*L241;

t*78-tru = 78;

7.4 REFERENCES

Abdelmaboud, A., Jawawi, D.N., Ghani, I., Elsafi, A. & Kitchenham, B. 2015, 'Quality of service approaches in cloud computing: A systematic mapping study', *Journal of Systems and Software*, vol. 101, pp. 159-179.

Aburukba, R.O., AliKarrar, M., Landolsi, T. & El-Fakih, K. 2019, 'Scheduling Internet of Things requests to minimize latency in hybrid Fog–Cloud computing', *Future Generation Computer Systems*, vol. 111, pp. 539-551.

Alabool, H.M. & Mahmood, A.K.B. 2016, 'A novel evaluation framework for improving trust level of Infrastructure as a Service', *Cluster Computing*, vol. 19, no. 1, pp. 389-410.

Amirteimoori, A. & Kordrostami, S. 2005, 'Allocating fixed costs and target setting: A DEA-based approach', *Applied Mathematics and Computation*, vol. 171, no. 1, pp. 136-151.

An, Q., Pang, Z., Chen, H. & Liang, L. 2015, 'Closest targets in environmental efficiency evaluation based on enhanced Russell measure', *Ecological Indicators*, vol. 51, pp. 59-66.

Andersen, P. & Petersen, N.C. 1993, 'A procedure for ranking efficient units in data envelopment analysis', *Management science*, vol. 39, no. 10, pp. 1261-1264.

Ang, S., Chen, M. & Yang, F. 2018, 'Group cross-efficiency evaluation in data envelopment analysis: An application to Taiwan hotels', *Computers & Industrial Engineering*, vol. 125, pp. 190-199.

Angel Diaz & Chris Ferris 2013, *IBM's open cloud architecture*, IBM, <http://www.ibm.com/>.

Aparicio, J., Ortiz, L., Pastor, J.T. & Zabala-Iturriagagoitia, J.M. 2020, 'Introducing cross-productivity: a new approach for ranking productive units over time in Data Envelopment Analysis', *Computers & Industrial Engineering*, p. 106456.

Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R.H., Konwinski, A., Lee, G., Patterson, D.A., Rabkin, A. & Stoica, I. 2009, *Above the clouds: A berkeley view of cloud computing*, Technical Report UCB/EECS-2009-28, EECS Department, University of California, Berkeley.

Aruna, L. & Aramudhan, M. 2016, 'Framework for ranking service providers of federated cloud architecture using fuzzy sets', *International Journal of Technology*, vol. 7, no. 4, pp. 643-653.

Ataş, G. & Gungor, V.C. 2014, 'Performance evaluation of cloud computing platforms using statistical methods', *Computers & Electrical Engineering*, vol. 40, no. 5, pp. 1636-1649.

Atmaca, T., Begin, T., Brandwajn, A. & Castel-Taleb, H. 2015, 'Performance evaluation of cloud computing centers with general arrivals and service', *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 8, pp. 2341-238.

Avilés-Sacoto, S.V., Cook, W.D., Güemes-Castorena, D. & Zhu, J. 2020, 'Modelling Efficiency in Regional Innovation Systems: A Two-Stage Data Envelopment Analysis Problem with Shared Outputs within Groups of Decision-Making Units', *European Journal of Operational Research*.

Avkiran, N.K. 2015, 'An illustration of dynamic network DEA in commercial banking including robustness tests', *Omega*, vol. 55, pp. 141-150.

Azadi, M., Emrouznejad, A., Ramezani, F. & Hussain, F.K. 2019, 'Efficiency measurement of cloud service providers using network data envelopment analysis', *IEEE Transactions on Cloud Computing*.

Azadi, M. & Saen, R.F. 2011, 'A new chance-constrained data envelopment analysis for selecting third-party reverse logistics providers in the existence of dual-role factors', *Expert Systems with applications*, vol. 38, no. 10, pp. 12231-12236.

Azadi, M. & Saen, R.F. 2013, 'A combination of QFD and imprecise DEA with enhanced Russell graph measure: A case study in healthcare', *Socio-Economic Planning Sciences*, vol. 47, no. 4, pp. 281-291.

Azadi, M., Shabani, A., Khodakarami, M. & 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*, vol. 70, pp. 324-338.

Azizi, H. & Ajirlu, S.F. 2010, 'Measurement of overall performances of decision-making units using ideal and anti-ideal decision-making units', *Computers & Industrial Engineering*, vol. 59, no. 3, pp. 411-418.

Banker, R.D., Charnes, A. & Cooper, W.W. 1984, 'Some models for estimating technical and scale inefficiencies in data envelopment analysis', *Management science*, vol. 30, no. 9, pp. 1078-92.

Banker, R.D., Cooper, W.W., Seiford, L.M., Thrall, R.M. & Zhu, J. 2004, 'Returns to scale in different DEA models', *European Journal of Operational Research*, vol. 154, no. 2, pp. 345-362.

Banker, R.D. & Morey, R.C. 1986, 'Efficiency analysis for exogenously fixed inputs and outputs', *Operations Research*, vol. 34, no. 4, pp. 513-521.

Barzegarinegad, A., Jahanshahloo, G.R. & Rostamy-Malkhalifeh, M. 2014, 'A Full Ranking for Decision Making Units Using Ideal and Anti-Ideal Points in DEA', *The Scientific World Journal*, vol. 2014, pp. 1-8.

Borgwardt, S. & Viss, C. 2020, 'An implementation of steepest-descent augmentation for linear programs', *Operations Research Letters*.

Borylo, P., Tornatore, M., Jaglarz, P., Shahriar, N., Chołda, P. & Boutaba, R. 2020, 'Latency and energy-aware provisioning of network slices in cloud networks', *Computer Communications*.

Briec, W. & Kerstens, K. 2006, 'Input, output and graph technical efficiency measures on non-convex FDH models with various scaling laws: An integrated approach based upon implicit enumeration algorithms', *Top*, vol. 14, no. 1, pp. 135-166.

Buyya, R., Broberg, J. & Goscinski, A. (eds) 2011, *Cloud Computing, Principles and Paradigms*, vol. 81 John Wiley & Sons, Hoboken, NJ.

Buyya, R., Pandey, S. & Vecchiola, C. 2009, 'Cloudbus toolkit for market-oriented cloud computing', Springer, pp. 24-44.

Caiza, G., Saeteros, M., Oñate, W. & Garcia, M.V. 2020, 'Fog computing at industrial level, architecture, latency, energy, and security: A review', *Heliyon*, vol. 6, no. 4, p. 3706.

Cecchini, L., Venanzi, S., Pierri, A. & Chiorri, M. 2018, 'Environmental efficiency analysis and estimation of CO2 abatement costs in dairy cattle farms in Umbria (Italy): A SBM-DEA model with undesirable output', *Journal of cleaner production*, vol. 197, pp. 895-907.

Chan, H. & Chieu, T. 2010, 'Ranking and mapping of applications to cloud computing services by SVD', IEEE, pp. 362-369.

Charnes, A. & Cooper, W.W. 1959, 'Chance-constrained programming', *Management science*, vol. 6, no. 1, pp. 73-79.

Charnes, A. & Cooper, W.W. 1962, 'Programming with linear fractional functionals', *Naval Research logistics quarterly*, vol. 9, no. 3-4, pp. 181-186.

Charnes, A., Cooper, W.W. & Rhodes, E. 1978, 'Measuring the efficiency of decision making units', *European Journal of Operational Research*, vol. 2, no. 6, pp. 429-944.

Chauhan, T., Chaudhary, S., Kumar, V. & Bhise, M. 2011, 'Service level agreement parameter matching in cloud computing', IEEE, pp. 564-570.

Chen, C.-M., Du, J., Huo, J. & Zhu, J. 2012, 'Undesirable factors in integer-valued DEA: Evaluating the operational efficiencies of city bus systems considering safety records', *Decision Support Systems*, vol. 54, no. 1, pp. 330-335.

Chen, K., Cook, W.D. & Zhu, J. 2020, 'A conic relaxation model for searching for the global optimum of network data envelopment analysis', *European Journal of Operational Research*, vol. 280, no. 1, pp. 242-253.

Chen, Y., Huang, J., Lin, C. & Shen, X. 2016, 'Multi-objective service composition with QoS dependencies', *IEEE Transactions on Cloud Computing*.

Chen, K. & Zhu, J. 2019, 'Computational tractability of chance constrained data envelopment analysis', *European Journal of Operational Research*, vol. 274, no. 3, pp. 1037-1046.

Chen, P.-C., Yu, M.-M., Chang, C.-C., Hsu, S.-H. & Managi, S. 2015, 'The enhanced Russell-based directional distance measure with undesirable outputs: numerical example considering CO2 emissions', *Omega*, vol. 53, pp. 30-40.

Chen, Y. 2005, 'Measuring super-efficiency in DEA in the presence of infeasibility', *European Journal of Operational Research*, vol. 161, no. 2, pp. 545-551.

Chen, Y., Djamasbi, S., Du, J. & Lim, S. 2013, 'Integer-valued DEA super-efficiency based on directional distance function with an application of evaluating mood and its impact on performance', *International Journal of Production Economics*, vol. 146, no. 2, pp. 550-556.

Chen, Y., Du, J. & Huo, J. 2013, 'Super-efficiency based on a modified directional distance function', *Omega*, vol. 41, no. 3, pp. 621-625.

Chen, Y. & Liang, L. 2011, 'Super-efficiency DEA in the presence of infeasibility: One model approach', *European Journal of Operational Research*, vol. 213, no. 1, pp. 359-360.

Chen, Y., Liang, L., Yang, F. & Zhu, J. 2006, 'Evaluation of information technology investment: a data envelopment analysis approach', *Computers & Operations Research*, vol. 33, no. 5, pp. 1368-1379.

Cherchye, L., Kuosmanen, T. & Post, T. 2001, 'FDH directional distance functions with an application to European commercial banks', *Journal of Productivity Analysis*, vol. 15, no. 3, pp. 201-15.

Choudhury, P., Sharma, M., Vikas, K., Pranshu, T. & Satyanarayana, V. 2012, 'Service ranking systems for cloud vendors', vol. 433, Trans Tech Publ, pp. 3949-3953.

Cook, W.D., Green, R.H. & Zhu, J. 2006, 'Dual-role factors in data envelopment analysis', *IIE Transactions*, vol. 38, no. 2, pp. 105-115.

Cook, W.D., Liang, L., Zha, Y. & Zhu, J. 2009, 'A modified super-efficiency DEA model for infeasibility', *Journal of the Operational Research Society*, vol. 60, no. 2, pp. 276-281.

Cooper, W., Huang, Z. & Li, S.X. 1996, 'Satisficing DEA models under chance constraints', *Annals of operations research*, vol. 66, no. 4, pp. 279-295.

Cooper, W., Huang, Z., Li, S.X., Parker, B.R. & Pastor, J.T. 2007, 'Efficiency aggregation with enhanced Russell measures in data envelopment analysis', *Socio-Economic Planning Sciences*, vol. 41, no. 1, pp. 1-21.

Cooper, W.W., Deng, H., Huang, Z. & Li, S.X. 2002, 'Chance constrained programming approaches to technical efficiencies and inefficiencies in stochastic data envelopment analysis', *Journal of the Operational Research Society*, vol. 53, no. 12, pp. 1347-1356.

Cooper, W.W., Deng, H., Huang, Z. & Li, S.X. 2004, 'Chance constrained programming approaches to congestion in stochastic data envelopment analysis', *European Journal of Operational Research*, vol. 155, no. 2, pp. 487-501.

Cooper, W.W., Park, K.S. & Pastor, J.T. 1999, 'RAM: a range adjusted measure of inefficiency for use with additive models, and relations to other models and measures in DEA', *Journal of Productivity analysis*, vol. 11, no. 1, pp. 5-42.

Cooper, W.W., Seiford, L.M. & Tone, K. 2006, *Introduction to data envelopment analysis and its uses: with DEA-solver software and references*, Springer Science & Business Media.

Cooper, W.W., Seiford, L.M. & Zhu, J. 2011, *Handbook on data envelopment analysis*, vol. 164, Springer Science & Business Media.

de Oliveira, S.G. & Silva, L. 2020, 'An ant colony hyperheuristic approach for matrix bandwidth reduction', *Applied Soft Computing*, p. 106434.

Deprins, D. & Simar, L. 1984, 'Measuring Labor Effi ciency in Post Of (fices. In Marchand M., P. Pestieau and H. Tulkens (eds.), The Performance of Public Enterprises: Concepts and Measurements', Series Measuring Labor Effi ciency in Post Of (fices. In Marchand M., P. Pestieau and H. Tulkens (eds.), The Performance of Public Enterprises: Concepts and Measurements Amsterdam, North (Holland. pp. 243.

Despotis, D.K., Koronakos, G. & Sotiros, D. 2016, 'The "weak-link" approach to network DEA for two-stage processes', *European Journal of Operational Research*, vol. 254, no. 2, pp. 481-492.

DeToro, I. 1995, 'The 10 pitfalls of benchmarking', *Quality progress*, vol. 28, no. 1, p. 61.

Devi, P., Gupta, S. & Choudhary, S. 2014, 'Parallel Efficiency for Graph Search'', *International Journal of Enhanced Research in Science Technology & Engineering*, vol. 3, no. 4, pp. 278-82.

Devi, R. & Shanmugalakshmi, R. 2020, 'Cloud providers ranking and selection using quantitative and qualitative approach', *Computer Communications*, vol. 154, pp. 370-379.

Diewert, W.E. & Fox, K.J. 2014, 'Reference technology sets, Free Disposal Hulls and productivity decompositions', *Economics Letters*, vol. 122, no. 2, pp. 238-242.

Doyle, J. & Green, R. 1994, 'Efficiency and cross-efficiency in DEA: Derivations, meanings and uses', *Journal of the operational research society*, vol. 45, no. 5, pp. 567-578.

Du, J., Chen, C.-M., Chen, Y., Cook, W.D. & Zhu, J. 2012, 'Additive super-efficiency in integer-valued data envelopment analysis', *European Journal of Operational Research*, vol. 218, no. 1, pp. 186-192.

Duan, Q. 2017, 'Cloud service performance evaluation: status, challenges, and opportunities–a survey from the system modeling perspective', *Digital Communications and Networks*, vol. 3, no. 2, pp. 101-11.

Elgendy, I., Zhang, W., Liu, C. & Hsu, C.-H. 2018, 'An efficient and secured framework for mobile cloud computing', *IEEE Transactions on Cloud Computing*.

Emrouznejad, A., Cabanda, E. & Gholami, R. 2010, 'An alternative measure of the ICT-Opportunity Index', *Information & Management*, vol. 47, no. 4, pp. 246-54.

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

Entani, T., Maeda, Y. & Tanaka, H. 2002, 'Dual models of interval DEA and its extension to interval data', *European Journal of Operational Research*, vol. 136, no. 1, pp. 32-45.

Erl, T. 2005, 'Service-Oriented Architecture: Concepts, Technology, and Design, Prentice Hall PTR, NJ, USA, 2005'.

Fan, W., Yang, S. & Pei, J. 2014, 'A novel two-stage model for cloud service trustworthiness evaluation', *Expert Systems*, vol. 31, no. 2, pp. 136-53.

Färe, R. & Grosskopf, S. 1996, 'Productivity and intermediate products: A frontier approach', *Economics letters*, vol. 50, no. 1, pp. 65-70.

Färe, R. & Grosskopf, S. 2004, 'Modeling undesirable factors in efficiency evaluation: comment', *European Journal of Operational Research*, vol. 157, no. 1, pp. 242-245.

Färe, R., Grosskopf, S., Lovell, C.K. & Yaisawarng, S. 1993, 'Derivation of shadow prices for undesirable outputs: a distance function approach', *The review of economics and statistics*, pp. 374-80.

Farzipoor Saen, R. 2008, 'Using super-efficiency analysis for ranking suppliers in the presence of volume discount offers', *International Journal of Physical Distribution & Logistics Management*, vol. 38, no. 8, pp. 637-651.

Filiopoulou, E., Mitropoulou, P., Lionis, N. & Michalakelis, C. 2018, 'On the efficiency of cloud providers: A DEA approach incorporating categorical variables', *IEEE Transactions on Cloud Computing*.

Fukuyama, H., Leth Hougaard, J., Sekitani, K. & Shi, J. 2016, 'Efficiency measurement with a non-convex free disposal hull technology', *Journal of the Operational Research Society*, vol. 67, no. 1, pp. 9-19.

Fukuyama, H. & Shiraz, R.K. 2015, 'Cost-effectiveness measures on convex and nonconvex technologies', *European Journal of Operational Research*, vol. 246, no. 1, pp. 307-319.

Galagedera, D.U., Roshdi, I., Fukuyama, H. & Zhu, J. 2018, 'A new network DEA model for mutual fund performance appraisal: An application to US equity mutual funds', *Omega*, vol. 77, pp. 168-179.

Gallupe, R.B. 2007, 'The tyranny of methodologies in information systems research 1', *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, vol. 38, no. 3, pp. 20-28.

Garg, S.K., Versteeg, S. & Buyya, R. 2013, 'A framework for ranking of cloud computing services', *Future Generation Computer Systems*, vol. 29, no. 4, pp. 1012-1023.

Ghanbari, H., Simmons, B., Litoiu, M. & Iszlai, G. 2012, 'Feedback-based optimization of a private cloud', *Future Generation Computer Systems*, vol. 28, no. 1, pp. 104-111.

Ghosh, N., Ghosh, S.K. & Das, S.K. 2014, 'SelCSP: A framework to facilitate selection of cloud service providers', *IEEE transactions on cloud computing*, vol. 3, no. 1, pp. 66-79.

Ghosh, R., Longo, F., Frattini, F., Russo, S. & Trivedi, K.S. 2014, 'Scalable analytics for IaaS cloud availability', *IEEE Transactions on Cloud Computing*, vol. 2, no. 1, pp. 57-70.

Gireesha, O., Somu, N., Krithivasan, K. & VS, S.S. 2020, 'IIVIFS-WASPAS: An integrated Multi-Criteria Decision-Making perspective for cloud service provider selection', *Future Generation Computer Systems*, vol. 103, pp. 91-110.

Golany, B. 1988, 'An interactive MOLP procedure for the extension of DEA to effectiveness analysis', *Journal of the Operational Research Society*, vol. 39, no. 8, pp. 725-734.

Haff, G. 2009, 'Just don't call them private clouds', *CNET News*, vol. 27.

Hayyolalam, V. & Kazem, A.A.P. 2018, 'A systematic literature review on QoS-aware service composition and selection in cloud environment', *Journal of Network and Computer Applications*, vol. 110, pp. 52-74.

Hailu, A. & Veeman, T.S. 2001, 'Non-parametric productivity analysis with undesirable outputs: an application to the Canadian pulp and paper industry', *American Journal of Agricultural Economics*, vol. 83, no. 3, pp. 605-616.

Halabi, T. & Bellaiche, M. 2017, 'Towards quantification and evaluation of security of Cloud Service Providers', *Journal of Information Security and Applications*, vol. 33, pp. 55-65.

Halická, M. & Trnovská, M. 2018, 'The Russell measure model: Computational aspects, duality, and profit efficiency', *European journal of operational research*, vol. 268, no. 1, pp. 386-397.

He, F., Xu, H., Chen, R. & Zhu, L. 2016, 'Interval efficiency improvement in DEA by using ideal points', *Measurement*, vol. 87, pp. 138-145.

He, F., Zhang, Q., Lei, J., Fu, W. & Xu, X. 2013, 'Energy efficiency and productivity change of China's iron and steel industry: Accounting for undesirable outputs', *Energy Policy*, vol. 54, pp. 204-213.

Heidari, S. & Buyya, R. 2019, 'Quality of Service (QoS)-driven resource provisioning for large-scale graph processing in cloud computing environments: Graph Processing-as-a-Service (GPaaS)', *Future Generation Computer Systems*, vol. 96, pp. 490-501.

Hu, X. & Liu, C. 2015, 'Managing undesirable outputs in the Australian construction industry using Data Envelopment Analysis models', *Journal of cleaner production*, vol. 101, pp. 148-157.

Huang, C.-Y., Hsu, P.-C. & Tzeng, G.-H. 2012, 'Evaluating cloud computing based telecommunications service quality enhancement by using a new hybrid MCDM model', *Intelligent Decision Technologies*, Springer, pp. 519-36.

Huang, J., Yang, X., Cheng, G. & Wang, S. 2014, 'A comprehensive eco-efficiency model and dynamics of regional eco-efficiency in China', *Journal of Cleaner Production*, vol. 67, pp. 228-38.

Huang, Z. & Li, S.X. 2001, 'Stochastic DEA models with different types of input-output disturbances', *Journal of Productivity Analysis*, vol. 15, no. 2, pp. 95-113.

Hussain, A., Chun, J. & Khan, M. 2020, 'A novel customer-centric Methodology for Optimal Service Selection (MOSS) in a cloud environment', *Future Generation Computer Systems*, vol. 105, pp. 562-580.

Izadikhah, M., Farzipoor Saen, R. & Ahmadi, K. 2017, 'How to Assess Sustainability of Suppliers in the Presence of Dual-Role Factor and Volume Discounts? A Data Envelopment Analysis Approach', *Asia-Pacific Journal of Operational Research*, vol. 34, no. 3, pp. 1-25.

Jadeja, Y. & Modi, K. 2012, 'Cloud computing-concepts, architecture and challenges', IEEE, pp. 877-880.

Jahanshahloo, G., Lotfi, F.H., Jafari, Y. & Maddahi, R. 2011, 'Selecting symmetric weights as a secondary goal in DEA cross-efficiency evaluation', *Applied Mathematical Modelling*, vol. 35, no. 1, pp. 544-549.

Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., Khanmohammadi, M., Kazemimanesh, M. & Rezaie, V. 2010, 'Ranking of units by positive ideal DMU with common weights', *Expert Systems with Applications*, vol. 37, no. 12, pp. 7483-7488.

Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., Rezaie, V. & Khanmohammadi, M. 2011, 'Ranking DMUs by ideal points with interval data in DEA', *Applied Mathematical Modelling*, vol. 35, no. 1, pp. 218-229.

Jahanshahloo, G.R., Lotfi, F.H., Shoja, N., Tohidi, G. & Razavyan, S. 2005, 'Undesirable inputs and outputs in DEA models', *Applied Mathematics and Computation*, vol. 169, no. 2, pp. 917-925.

Jula, A., Sundararajan, E. & Othman, Z. 2014, 'Cloud computing service composition: A systematic literature review', *Expert systems with applications*, vol. 41, no. 8, pp. 3809-24.

Kaffash, S., Azizi, R., Huang, Y. & Zhu, J. 2020, 'A survey of data envelopment analysis applications in the insurance industry 1993–2018', *European Journal of Operational Research*, vol. 284, no. 3, pp. 801-13.

Kalantary, M. & Saen, R.F. 2019, 'Assessing sustainability of supply chains: An inverse network dynamic DEA model', *Computers & Industrial Engineering*, vol. 135, pp. 1224-1238.

Kao, C. 2016, 'Efficiency decomposition and aggregation in network data envelopment analysis', *European Journal of Operational Research*, vol. 255, no. 3, pp. 778-786.

Kao, C. & 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*, vol. 185, no. 1, pp. 418-429.

Kao, C. & Liu, S.-T. 2009, 'Stochastic data envelopment analysis in measuring the efficiency of Taiwan commercial banks', *European Journal of Operational Research*, vol. 196, no. 1, pp. 312-322.

Karimi, B., Khorram, E. & Moeini, M. 2016, 'Identification of congestion by means of integer-valued data envelopment analysis', *Computers & Industrial Engineering*, vol. 98, pp. 513-521.

Kerstens, K. & Eeckaut, P.V. 1999, 'Estimating returns to scale using non-parametric deterministic technologies: A new method based on goodness-of-fit', *European Journal of Operational Research*, vol. 113, no. 1, pp. 206-214.

Khezrimotlagh, D., Zhu, J., Cook, W.D. & Toloo, M. 2019, 'Data envelopment analysis and big data', *European Journal of Operational Research*, vol. 274, no. 3, pp. 1047-1054.

Khodabakhshi, M. & Aryavash, K. 2017, 'The cross-efficiency in the optimistic–pessimistic framework', *Operational Research*, vol. 17, no. 2, pp. 619-632.

Khoshnevisan, B., Rafiee, S., Omid, M. & Mousazadeh, H. 2013, 'Applying data envelopment analysis approach to improve energy efficiency and reduce GHG (greenhouse gas) emission of wheat production', *Energy*, vol. 58, pp. 588-593.

Khoveyni, M., Eslami, R., Fukuyama, H., Yang, G.-l. & Sahoo, B.K. 2019, 'Integer data in DEA: Illustrating the drawbacks and recognizing congestion', *Computers & Industrial Engineering*.

Köksal, C.D. & Aksu, A.A. 2007, 'Efficiency evaluation of A-group travel agencies with data envelopment analysis (DEA): A case study in the Antalya region, Turkey', *Tourism Management*, vol. 28, no. 3, pp. 830-834.

Korhonen, P.J. & Luptacik, M. 2004, 'Eco-efficiency analysis of power plants: An extension of data envelopment analysis', *European journal of operational research*, vol. 154, no. 2, pp. 437-446.

Kozlenkova, I.V., Hult, G.T.M., Lund, D.J., Mena, J.A. & Kekec, P. 2015, 'The role of marketing channels in supply chain management', *Journal of Retailing*, vol. 91, no. 4, pp. 586-609.

Kritikos, M.N. 2017, 'A full ranking methodology in data envelopment analysis based on a set of dummy decision making units', *Expert Systems with Applications*, vol. 77, pp. 211-225.

Kuah, C.T., Wong, K.Y. & Wong, W.P. 2012, 'Monte Carlo data envelopment analysis with genetic algorithm for knowledge management performance measurement', *Expert Systems with Applications*, vol. 39, no. 10, pp. 9348-9358.

Kumar, S. 2014, 'Evaluating Cloud Services Using DEA, AHP and TOPSIS model', Indian Institute of Technology, Kanpur.

Land, K.C., Lovell, C. & Thore, S. 1993, 'Chance-constrained data envelopment analysis', *Managerial and decision economics*, vol. 14, no. 6, pp. 541-554.

Lang, M., Wiesche, M. & Krcmar, H. 2018, 'Criteria for selecting cloud service providers: a delphi study of quality-of-service attributes', *Information & Management*, vol. 55, no. 6, pp. 746-58.

Lee, H.-S., Chu, C.-W. & Zhu, J. 2011, 'Super-efficiency DEA in the presence of infeasibility', *European Journal of Operational Research*, vol. 212, no. 1, pp. 141-147.

Lee, H.-S. & Zhu, J. 2012, 'Super-efficiency infeasibility and zero data in DEA', *European Journal of Operational Research*, vol. 216, no. 2, pp. 429-33.

Lee, W.-S. 2008, 'Benchmarking the energy efficiency of government buildings with data envelopment analysis', *Energy and Buildings*, vol. 40, no. 5, pp. 891-5.

Leukel, J., Kirn, S. & Schlegel, T. 2011, 'Supply chain as a service: a cloud perspective on supply chain systems', *IEEE Systems Journal*, vol. 5, no. 1, pp. 16-27.

Levitin, G., Xing, L. & Dai, Y. 2018, 'Co-residence based data vulnerability vs. security in cloud computing system with random server assignment', *European Journal of Operational Research*, vol. 267, no. 2, pp. 676-686.

Lewis, H.F. & Sexton, T.R. 2004, 'Network DEA: efficiency analysis of organizations with complex internal structure', *Computers & Operations Research*, vol. 31, no. 9, pp. 1365-1410.

Li, W., Li, Z., Liang, L. & Cook, W.D. 2017, 'Evaluation of ecological systems and the recycling of undesirable outputs: an efficiency study of regions in China', *Socio-Economic Planning Sciences*, vol. 60, pp. 77-86.

Li, Y., Chen, Y., Liang, L. & Xie, J. 2012, 'DEA models for extended two-stage network structures', *Omega*, vol. 40, no. 5, pp. 611-618.

Liang, L., Cook, W.D. & Zhu, J. 2008, 'DEA models for two-stage processes: Game approach and efficiency decomposition', *Naval Research Logistics (NRL)*, vol. 55, no. 7, pp. 643-653.

Liang, L., Wu, J., Cook, W.D. & Zhu, J. 2008, 'Alternative secondary goals in DEA cross-efficiency evaluation', *International Journal of Production Economics*, vol. 113, no. 2, pp. 1025-1030.

Lindner, M., Galán, F., Chapman, C., Clayman, S., Henriksson, D. & Elmroth, E. 2010, 'The cloud supply chain: A framework for information, monitoring, accounting and billing', *2nd International ICST Conference on Cloud Computing (CloudComp 2010)*.

Lim, S. 2012, 'Minimax and maximin formulations of cross-efficiency in DEA', *Computers & Industrial Engineering*, vol. 62, no. 3, pp. 726-731.

Liu, S.-T. 2018, 'A DEA ranking method based on cross-efficiency intervals and signal-to-noise ratio', *Annals of Operations Research*, vol. 261, no. 1-2, pp. 207-232.

Liu, W., Wang, Y.-M. & Lv, S. 2017, 'An aggressive game cross-efficiency evaluation in data envelopment analysis', *Annals of Operations Research*, vol. 259, no. 1-2, pp. 241-258.

Lovell, C.K. & Rouse, A. 2003, 'Equivalent standard DEA models to provide super-efficiency scores', *Journal of the Operational Research Society*, vol. 54, no. 1, pp. 101-108.

Lozano, S. & Villa, G. 2006, 'Data envelopment analysis of integer-valued inputs and outputs', *Computers & Operations Research*, vol. 33, no. 10, pp. 3004-14.

Luo, X. 2003, 'Evaluating the profitability and marketability efficiency of large banks: An application of data envelopment analysis', *Journal of Business research*, vol. 56, no. 8, pp. 627-635.

Ludwig, H., Keller, A., Dan, A., King, R.P. & Franck, R. 2003, 'Web service level agreement (WSLA) language specification', *Ibm corporation*, pp. 815-824.

Maalej, S., Zayoud, A., Abdelaziz, I., Saad, I. & Zaghdoudi, M. 2020, 'Thermal performance of finned heat pipe system for Central Processing Unit cooling', *Energy Conversion and Management*, vol. 218, p. 112977.

Mahmoud, M.S. & Xia, Y. 2019, *Networked Control Systems: Cloud Control and Secure Control*, Butterworth-Heinemann.

Manandhar, R. & Tang, J.C. 2002, 'The evaluation of bank branch performance using data envelopment analysis: A framework', *The Journal of High Technology Management Research*, vol. 13, no. 1, pp. 1-17.

Mangaiyarkarasi, K., Sureshkumar, K. & Elango, N. 2013, 'Comparative study on performance test methodologies-traditional and cloud', *International Journal of Advanced Research in Computer Science and Applications*, vol. 1, no. 2, pp. 1-11.

Manvi, S.S. & Shyam, G.K. 2014, 'Resource management for Infrastructure as a Service (IaaS) in cloud computing: A survey', *Journal of network and computer applications*, vol. 41, pp. 424-440.

Matin, R.K., Amin, G.R. & Emrouznejad, A. 2014, 'A modified semi-oriented radial measure for target setting with negative data', *Measurement*, vol. 54, pp. 152-8.

Matin, R.K. & Kuosmanen, T. 2009, 'Theory of integer-valued data envelopment analysis under alternative returns to scale axioms', *Omega*, vol. 37, no. 5, pp. 988-95.

Mehrabian, S., Alirezaee, M.R. & Jahanshahloo, G.R. 1999, 'A complete efficiency ranking of decision making units in data envelopment analysis', *Computational optimization and applications*, vol. 14, no. 2, pp. 261-266.

Mell, P. & Grance, T. 2011, 'The NIST definition of cloud computing'.

Menzel, M., Ranjan, R., Wang, L., Khan, S.U. & Chen, J. 2015, 'CloudGenius: a hybrid decision support method for automating the migration of web application clusters to public clouds', *IEEE Transactions on Computers*, vol. 64, no. 5, pp. 1336-1348.

Minnear, R. 2011, 'Latency: The Achilles heel of cloud computing', *Cloud Computing Journal*.

Mirhedayatian, S.M., Azadi, M. & Saen, R.F. 2014, 'A novel network data envelopment analysis model for evaluating green supply chain management', *International Journal of Production Economics*, vol. 147, pp. 544-554.

Moltó, G., Caballer, M. & De Alfonso, C. 2016, 'Automatic memory-based vertical elasticity and oversubscription on cloud platforms', *Future Generation Computer Systems*, vol. 56, pp. 1-10.

Moreno, P. & Lozano, S. 2014, 'A network DEA assessment of team efficiency in the NBA', *Annals of Operations Research*, vol. 214, no. 1, pp. 99-124.

Morita, H. & Seiford, L.M. 1999, 'Characteristics on stochastic DEA efficiency: Reliability and probability being efficient', *Journal of the Operations Research Society of Japan*, vol. 42, no. 4, pp. 389-404.

Mousavi-Avval, S.H., Rafiee, S., Jafari, A. & Mohammadi, A. 2011, 'Optimization of energy consumption for soybean production using Data Envelopment Analysis (DEA) approach', *Applied Energy*, vol. 88, no. 11, pp. 3765-3772.

Mthunzi, S.N., Benkhelifa, E., Bosakowski, T., Guegan, C.G. & Barhamgi, M. 2020, 'Cloud computing security taxonomy: From an atomistic to a holistic view', *Future Generation Computer Systems*, vol. 107, pp. 620-644.

Mulinka, P. & Kencl, L. 2015, 'Learning from cloud latency measurements', IEEE, pp. 1895-901.

Nahangi, M., Chen, Y. & McCabe, B. 2019, 'Safety-based efficiency evaluation of construction sites using data envelopment analysis (DEA)', *Safety science*, vol. 113, pp. 382-8.

Niu, L., Lu, J. & Zhang, G. 2009, 'Cognition-driven decision support for business intelligence (pp. 3–18)', *Berlin, Germany: Springer-Verlag*, vol. 10, pp. 978-973.

Olesen, O.B. & Petersen, N. 1995, 'Chance constrained efficiency evaluation', *Management Science*, vol. 41, no. 3, pp. 442-457.

Olson, D.L. & Swenseth, S.R. 1987, 'A linear approximation for chance-constrained programming', *Journal of the Operational Research Society*, vol. 38, no. 3, pp. 261-267.

Örkcü, H.H., Özsoy, V.S., Örkcü, M. & Bal, H. 2019, 'A neutral cross efficiency approach for basic two stage production systems', *Expert Systems with Applications*, vol. 125, pp. 333-344.

Park, S., Ok, C. & Ha, C. 2018, 'A stochastic simulation-based holistic evaluation approach with DEA for vendor selection', *Computers & Operations Research*, vol. 100, pp. 368-78.

Pastor, J.T., Ruiz, J.L. & Sirvent, I. 1999, 'An enhanced DEA Russell graph efficiency measure', *European Journal of Operational Research*, vol. 115, no. 1, pp. 596–607.

Pathomsiri, S., Haghani, A., Dresner, M. & Windle, R.J. 2008, 'Impact of undesirable outputs on the productivity of US airports', *Transportation Research Part E: Logistics and Transportation Review*, vol. 44, no. 2, pp. 235-259.

Peng Wong, W. & Yew Wong, K. 2008, 'A review on benchmarking of supply chain performance measures', *Benchmarking: An international journal*, vol. 15, no. 1, pp. 25-51.

Pittman, R.W. 1983, 'Multilateral productivity comparisons with undesirable outputs', *The Economic Journal*, vol. 93, no. 372, pp. 883-91.

Post, T. & Spronk, J. 1999, 'Performance benchmarking using interactive data envelopment analysis', *European journal of operational research*, vol. 115, no. 3, pp. 472-87.

Qu, L., Wang, Y. & Orgun, M.A. 2013, 'Cloud service selection based on the aggregation of user feedback and quantitative performance assessment', IEEE, pp. 152-9.

Rajarajeswari, C. & Aramudhan, M. 2015, 'RANKING OF CLOUD SERVICE PROVIDERS IN CLOUD', *Journal of Theoretical and Applied Information Technology*, vol. 78, no. 2, p. 212.

Ramachandran, M. & Chang, V. 2016, 'Towards performance evaluation of cloud service providers for cloud data security', *International Journal of Information Management*, vol. 36, no. 4, pp. 618-625.

Ramezani, F. 2016, 'Autonomic system for optimal resource management in cloud environments'.

Ranjan, R., Benatallah, B. & Wang, M. 2011, 'A cloud resource orchestration framework for simplifying the management of web applications', Springer, pp. 248-249.

Rashidi, K. & Saen, R.F. 2015, 'Measuring eco-efficiency based on green indicators and potentials in energy saving and undesirable output abatement', *Energy Economics*, vol. 50, pp. 18-26.

Ray, S.C. 2004, *Data envelopment analysis: theory and techniques for economics and operations research*, Cambridge university press.

Rebolledo-Leiva, R., Angulo-Meza, L., Iriarte, A. & González-Araya, M.C. 2017, 'Joint carbon footprint assessment and data envelopment analysis for the reduction of greenhouse gas emissions in agriculture production', *Science of the Total Environment*, vol. 593, pp. 36-46.

Reefke, H. & Sundaram, D. 2018, 'Sustainable supply chain management: Decision models for transformation and maturity', *Decision Support Systems*.

Rezaie, V., Jahanshahloo, G.R., Hosseinzadeh Lotfi, F. & Khanmohammadi, M. 2009, 'Ranking DMUs by Ideal Points in DEA', *Int. Journal of Math. Analysis*, vol. 3, no. 19, pp. 919-929.

Ross, A.D. & Droge, C. 2004, 'An analysis of operations efficiency in large-scale distribution systems', *Journal of Operations Management*, vol. 21, no. 6, pp. 673-688.

Saen, R.F. 2009, 'A decision model for ranking suppliers in the presence of cardinal and ordinal data, weight restrictions, and nondiscretionary factors', *Annals of Operations Research*, vol. 172, no. 1, p. 177.

Saen, R.F. 2010a, 'Developing a new data envelopment analysis methodology for supplier selection in the presence of both undesirable outputs and imprecise data', *The International Journal of Advanced Manufacturing Technology*, vol. 51, no. 9-12, pp. 1243-1250.

Saen, R.F. 2010b, 'Restricting weights in supplier selection decisions in the presence of dual-role factors', *Applied Mathematical Modelling*, vol. 34, no. 10, pp. 2820-2830.

Saen, R.F., Memariani, A. & Lotfi, F.H. 2005, 'Determining relative efficiency of slightly non-homogeneous decision making units by data envelopment analysis: a case study in IROST', *Applied Mathematics and Computation*, vol. 165, no. 2, pp. 313-328.

Sahri, S., Moussa, R., Long, D.D. & Benbernou, S. 2014, 'DBaaS-Expert: A recommender for the selection of the right cloud database', Springer, pp. 315-324.

Sarkis, J. 2000, 'A comparative analysis of DEA as a discrete alternative multiple criteria decision tool', *European journal of operational research*, vol. 123, no. 3, pp. 543-557.

Sarkis, J. 2007, 'Preparing your data for DEA', *Modeling data irregularities and structural complexities in data envelopment analysis*, Springer, pp. 305-320.

Scheel, H. 2001, 'Undesirable outputs in efficiency valuations', *European journal of operational research*, vol. 132, no. 2, pp. 400-410.

Seiford, L.M. & Zhu, J. 2002, 'Modeling undesirable factors in efficiency evaluation', *European journal of operational research*, vol. 142, no. 1, pp. 16-20.

Sengupta, J.K. 1982, 'Efficiency measurement in stochastic input-output systems', *International Journal of Systems Science*, vol. 13, no. 3, pp. 273-287.

Sengupta, J.K. 1987, 'Data envelopment analysis for efficiency measurement in the stochastic case', *Computers & operations research*, vol. 14, no. 2, pp. 117-129.

Sengupta, J.K. 1990, 'Transformations in stochastic DEA models', *Journal of Econometrics*, vol. 46, no. 1-2, pp. 109-123.

Sengupta, J.K. 1997, 'Stochastic efficiency measurement: A new approach', *Applied Economics Letters*, vol. 4, no. 2, pp. 125-128.

Sengupta, J.K. 1998, 'Efficiency distribution and the cost frontier', *Applied stochastic models and data analysis*, vol. 14, no. 1, pp. 67-76.

Sengupta, J.K. 2000, 'Efficiency analysis by stochastic data envelopment analysis', *Applied Economics Letters*, vol. 7, no. 6, pp. 379-383.

Sexton, T.R., Silkman, R. & Hogan, A. 1986, 'Measuring efficiency: An assessment of data envelopment analysis', *New Directions For Program Evaluation*, no. 32.

Shabani, A., Faramarzi, G.R., Farzipoor Saen, R. & Khodakarami, M. 2017, 'Simultaneous evaluation of efficiency, input effectiveness, and output effectiveness: An application in after sales services agents', *Benchmarking: An International Journal*, vol. 24, no. 7, pp. 1854-1870.

Shabani, A., Torabipour, S.M.R. & Saen, R.F. 2015, 'A new super-efficiency dual-role FDH procedure: an application in dairy cold chain for vehicle selection', *International Journal of Shipping and Transport Logistics*, vol. 7, no. 4, pp. 426-456.

Shafiee, M., Lotfi, F.H. & Saleh, H. 2014, 'Supply chain performance evaluation with data envelopment analysis and balanced scorecard approach', *Applied Mathematical Modelling*, vol. 38, no. 21-22, pp. 5092-112.

Sheridan, J.H. 1993, 'Whe5re benchmarkers go wrong', *Industry Week*, vol. 242, no. 6, pp. 28-34.

Shi, H., Wang, Y. & Chen, L. 2019, 'Neutral cross-efficiency evaluation regarding an ideal frontier and anti-ideal frontier as evaluation criteria', *Computers & Industrial Engineering*, vol. 132, pp. 385-394.

Shiraz, R.K., Tavana, M., Di Caprio, D. & Vakili, J. 2015, 'An improved non-convex model for discriminating efficient units in free disposal hull', *Measurement*, vol. 69, pp. 222-235.

Singh, S. & Sidhu, J. 2017, 'Compliance-based multi-dimensional trust evaluation system for determining trustworthiness of cloud service providers', *Future Generation Computer Systems*, vol. 67, pp. 109-132.

Soleimani-Damaneh, M. & Mostafaee, A. 2009, 'Stability of the classification of returns to scale in FDH models', *European Journal of Operational Research*, vol. 196, no. 3, pp. 1223-8.

Soleimani-Damaneh, M. & Reshadi, M. 2007, 'A polynomial-time algorithm to estimate returns to scale in FDH models', *Computers & operations research*, vol. 34, no. 7, pp. 2168-2176.

Somu, N., Kirthivasan, K. & VS, S.S. 2017, 'A computational model for ranking cloud service providers using hypergraph based techniques', *Future Generation Computer Systems*, vol. 68, pp. 14-30.

Stewart, T.J. 2010, 'Goal directed benchmarking for organizational efficiency', *Omega*, vol. 38, no. 6, pp. 534-539.

Strandmark, P., Qu, Y. & Curtois, T. 2020, 'First-order Linear Programming in a Column Generation Based Heuristic Approach to the Nurse Rostering Problem', *Computers & Operations Research*, p. 104945.

Subramanian, N. & Jeyaraj, A. 2018, 'Recent security challenges in cloud computing', *Computers & Electrical Engineering*, vol. 71, pp. 28-42.

Sueyoshi, T. 2000, 'Stochastic DEA for restructure strategy: an application to a Japanese petroleum company', *Omega*, vol. 28, no. 4, pp. 385-398.

Sun, J., Wu, J. & Guo, D. 2013, 'Performance ranking of units considering ideal and anti-ideal DMU with common weights', *Applied Mathematical Modelling*, vol. 37, no. 9, pp. 6301-6310.

Sun, L., Dong, H., Hussain, O.K., Hussain, F.K. & Liu, A.X. 2019, 'A framework of cloud service selection with criteria interactions', *Future Generation Computer Systems*, vol. 94, pp. 749-764.

Sun, P.J. 2020, 'Security and privacy protection in cloud computing: Discussions and challenges', *Journal of Network and Computer Applications*, p. 102642.

Supriya, M., Sangeeta, K. & Patra, G. 2016, 'Trustworthy Cloud Service Provider Selection using Multi Criteria Decision Making Methods', *Engineering Letters*, vol. 24, no. 1.

Tamaki, T., Nakamura, H., Fujii, H. & Managi, S. 2016, 'Efficiency and emissions from urban transport: Application to world city-level public transportation', *Economic Analysis and Policy*.

Tan, S.O. & Demirel, H. 2015, 'Performance and cooling efficiency of thermoelectric modules on server central processing unit and Northbridge', *Computers & Electrical Engineering*, vol. 46, pp. 46-55.

Tavakoli, I.M. & Mostafaee, A. 2019, 'Free disposal hull efficiency scores of units with network structures', *European Journal of Operational Research*, vol. 277, no. 3, pp. 1027-1036.

Thanassoulis, E. 2000, 'The use of data envelopment analysis in the regulation of UK water utilities: water distribution', *European Journal of Operational Research*, vol. 126, no. 2, pp. 436-453.

Thangappan, T., Therese, B., Adi, S. & Gadda, S.S. 2020, 'Review on Dynamic Bandwidth Allocation of GPON and EPON', *Journal of Electronic Science and Technology*, p. 100044.

Thrall, R. 1996, 'Duality, classification and slacks in DEA. Annals of Operations Research 66: 109-138.

Tone, K. 2001, 'A slacks-based measure of efficiency in data envelopment analysis', *European journal of operational research*, vol. 130, no. 3, pp. 498-509.

Tone, K. & Tsutsui, M. 2009, 'Network DEA: A slacks-based measure approach', *European journal of operational research*, vol. 197, no. 1, pp. 243-252.

Tone, K. & Tsutsui, M. 2011, 'Applying an efficiency measure of desirable and undesirable outputs in DEA to US electric utilities', *Journal of CENTRUM Cathedra: The Business and Economics Research Journal*, vol. 4, no. 2, pp. 236-249.

Tulkens, H. & Eeckaut, P.V. 2006, 'Non-frontier measures of efficiency, progress and regress for time series data', *Public goods, environmental externalities and fiscal competition*, Springer, pp. 373-394.

Varia, J. 2008, 'Amazon white paper on cloud architectures', Series Amazon white paper on cloud architectures Sept.

Wang, J., Wang, S., Li, S., Cai, Q. & Gao, S. 2019, 'Evaluating the energy-environment efficiency and its determinants in Guangdong using a slack-based measure with environmental undesirable outputs and panel data model', *Science of the Total Environment*, vol. 663, pp. 878-888.

Wang, Y.-M. & Chin, K.-S. 2010, 'Some alternative models for DEA cross-efficiency evaluation', *International Journal of Production Economics*, vol. 128, no. 1, pp. 332-338.

Wang, Y.-M., Chin, K.-S. & Luo, Y. 2011, 'Cross-efficiency evaluation based on ideal and anti-ideal decision making units', *Expert systems with applications*, vol. 38, no. 8, pp. 10312-10319.

Wang, Y.-M. & Luo, Y. 2006, 'DEA efficiency assessment using ideal and anti-ideal decision making units', *Applied Mathematics and Computation*, vol. 173, no. 2, pp. 902-915.

Wang, Y., Chin, K.-S. & Yang, J.-B. 2007, 'Measuring the performances of decision-making units using geometric average efficiency', *Journal of the Operational Research Society*, vol. 58, no. 7, pp. 929-937.

Wang, D., Yang, Y. & Mi, Z. 2015, 'QoS-Based and Network-Aware Web Service Composition across Cloud Datacenters', *KSII Transactions on Internet & Information Systems*, vol. 9, no. 3.

Wang, S., Sun, Q., Zou, H. & Yang, F. 2013, 'Particle swarm optimization with skyline operator for fast cloud-based web service composition', *Mobile Networks and Applications*, vol. 18, no. 1, pp. 116-121.

Wang, Z., Wang, N., Su, X. & Ge, S. 2020, 'An empirical study on business analytics affordances enhancing the management of cloud computing data security', *International Journal of Information Management*, vol. 50, pp. 387-394.

Wu, J. & Zhou, Z. 2015, 'A mixed-objective integer DEA model', *Annals of Operations Research*, vol. 228, no. 1, pp. 81-95.

Wu, L., Gan, C., Qiao, H., Hui, J. & Xu, Z. 2020, 'Bandwidth Allocation Algorithm Based on Differential BRP Models in Ethernet PON', *Computer Networks*, p. 107291.

Xu, C., Ma, Y. & Wang, X. 2015, 'A non-parametric data envelopment analysis approach for cloud services evaluation', Springer, pp. 250-5.

Xu, J., Li, B. & Wu, D. 2009, 'Rough data envelopment analysis and its application to supply chain performance evaluation', *International Journal of Production Economics*, vol. 122, no. 2, pp. 628-638.

Yang, H. & Pollitt, M. 2009, 'Incorporating both undesirable outputs and uncontrollable variables into DEA: The performance of Chinese coal-fired power plants', *European journal of operational research*, vol. 197, no. 3, pp. 1095-105.

Yang, W.-C., Lee, Y.-M. & Hu, J.-L. 2016, 'Urban sustainability assessment of Taiwan based on data envelopment analysis', *Renewable and Sustainable Energy Reviews*, vol. 61, pp. 341-353.

Younes, A., Essaaidi, M. & El Moussaoui, A. 2014, 'SFL algorithm for QoS-based cloud service composition', *International Journal of Computer Applications*, vol. 97, no. 17, pp. 42-49.

Yu, Y. & Wen, Z. 2010, 'Evaluating China's urban environmental sustainability with Data Envelopment Analysis', *Ecological Economics*, vol. 69, no. 9, pp. 1748-1755.

Zha, Y., Zhao, L. & Bian, Y. 2016, 'Measuring regional efficiency of energy and carbon dioxide emissions in China: A chance constrained DEA approach', *Computers & Operations Research*, vol. 66, pp. 351-361.

Zhang, M., Liu, L. & Liu, S. 2015, 'Genetic algorithm based QoS-aware service composition in multi-cloud', *2015 IEEE Conference on Collaboration and Internet Computing (CIC)*, IEEE, pp. 113-118.

Zhou, Z., Lin, L., Xiao, H., Ma, C. & Wu, S. 2017, 'Stochastic network DEA models for two-stage systems under the centralized control organization mechanism', *Computers & Industrial Engineering*, vol. 110, pp. 404-412.

Zhou, A., Wang, S., Zheng, Z., Hsu, C.-H., Lyu, M.R. & Yang, F. 2014, 'On cloud service reliability enhancement with optimal resource usage', *IEEE Transactions on Cloud Computing*, vol. 4, no. 4, pp. 452-466.

Zhu, J. 2004, 'A buyer–seller game model for selection and negotiation of purchasing bids: extensions and new models', *European Journal of Operational Research*, vol. 154, no. 1, pp. 150-156.