



UNIVERSITY OF GOTHENBURG  
SCHOOL OF BUSINESS, ECONOMICS AND LAW

**Sustainable supplier selection in the logistics industry: A comparison of  
alternative approaches**

Kamran Rashidi<sup>a</sup>

*Kamran.rashidi@handels.gu.se*

<sup>a</sup> *Dept of Business Administration, School of Business, Economics and Law, University of  
Gothenburg, SE-405 30 Gothenburg, Sweden*

*To my lovely sister, Leila.*

## **Abstract**

---

Supplier selection has become one of the most crucial tasks in supply chain management, especially in the procurement function. In recent years, the importance of selecting the best possible suppliers has been enhanced due to the emergence of sustainability issues. Manufacturers have been obliged and/or encouraged by various stakeholders to embed environmental and social concerns into their supply chain activities. As a consequence of this evolution, procurement managers have started to not only evaluate the suppliers' economic abilities, but also their competencies in environmental and social aspects. In addition to traditional economic criteria, therefore, environmental criteria (e.g. energy consumption, greenhouse gas emissions) and social criteria (e.g. labour health and work safety, diverse education programs for employees) have been added to the process of evaluating suppliers. One group of suppliers with a crucial role in any chain is the group of logistics service providers that need to be evaluated and selected based on all three lines of sustainability, i.e. economic, environmental, and social competencies.

Researchers have continuously proposed a number of diverse methods for handling the problem of sustainable supplier selection efficiently. Three of the most widely applied methods in the literature in this field are the Technique for Order of Preference by Similarity to the Ideal Solution (TOPSIS), Data Envelopment Analysis (DEA) and the Analytical Hierarchy Process (AHP). Each method is applied to evaluate a set of suppliers, given a set of variables/criteria, to provide a ranking of the suppliers. As discussed in the literature, many strategic decisions can be made based on the outcomes of these methods, e.g. sourcing and benchmarking strategies. This study aims at comparing the outcomes of these three methods based on a common data set. The comparisons are illustrated with an empirical application for measuring the sustainability of a set of logistics service providers. In other words, this study sheds light on the aspect regarding the extent to which the outcomes of these methods are reliable for making strategic decisions in a supplier management system.

The results reveal that each method produces a unique ranking of the logistics service providers under evaluation. Despite positive correlation coefficients between the rankings yielded, it is not possible to find only one supplier as the best or worst in the list. More specifically, the supplier rankings are influenced by the nature of the algorithm underpinning the evaluation methods and/or type of data used (e.g. fuzzy or non-fuzzy data). Therefore, due to this inconsistency between the outcomes of AHP, DEA, and TOPSIS, it is challenging to make decisions regarding sourcing and benchmarking strategies. It is not possible to find the best supplier(s) for either single- or multiple-sourcing strategies. Of course, considering the outcomes of only one method is not reliable, but the buying company can create a network of top suppliers based on the outcomes of each method and then further analyse their performance for the final decision. Furthermore, the comparison between AHP and DEA shows that each method provides a different benchmarking strategy for suppliers that need to improve their performance.

**Keywords:** Sustainable supplier evaluation; Logistics industry; Sourcing; Benchmarking; Technique for Order of Preference by Similarity to the Ideal Solution (TOPSIS); Data Envelopment Analysis (DEA); and Analytical Hierarchy Process (AHP).

## Acknowledgements

---

I would like to thank all persons and institutions which contributed to this study with their support and encouragement throughout this dissertation. Please accept my sincere apologies if I forget to mention all the names herein.

First and foremost, I am deeply grateful to the *University of Gothenburg* for providing me this opportunity to obtain a Doctor of Philosophy degree at the School of Business, Economics, and Law. Undoubtedly, I would not have been able to do this study without all the support I received from the University of Gothenburg. I would also like to thank all my colleagues at the Industrial and Financial Management & Logistics section.

I would like to express my deepest appreciation to my main supervisor, *Professor Kevin Cullinane*, for his continuous support and invaluable guidance throughout this dissertation. I will never forget how supportive he has been during these four years. He has not only supervised me through this study but also provided indispensable lessons for my future career in academia. I would also like to thank my co-supervisor, *Associate Professor Jonas Flodén*, for his constructive comments on my 'Kappa' that helped me a lot.

I am grateful to *Professor Rickard Bergqvist*, who has contributed to the quality of this study through his invaluable comments on my dissertation in my final internal seminar. I am also in debt to *Professor Michael Browne*, *Senior Lecturer Elisabeth Karlsson*, and *Senior Lecturer Conny Overland*, who helped me a lot to strengthen my teaching competencies.

I also want to thank postgraduate studies officer, *Kajsa Lundh*, and the Industrial and Financial Economics & Logistics section administrator, *Wiviann Hall*, for all their administrative support during my studies at the Department of Business Administration.

All my friends have also been instrumental in paving the path of this study. Thank-you to all those friends who have encouraged me while doing this research. I especially thank *Mahmoud* and *Parisa* for all their advice and assistance during these four years.

Last but not least, I would like to express my deepest appreciation to my *family* for their continuous support. The most important thing, among many others, my parents have taught me was the crucial role of education in one's life. I would not have been able to do this research without their help, love, and passion.

*Kamran Rashidi*

October 2019,  
Gothenburg

## **List of appended papers<sup>1</sup>**

---

*Paper I*) Rashidi, K., & Cullinane, K.P.B. (2019), “Evaluating the sustainability of national logistics performance using Data Envelopment Analysis”, *Transport Policy*, 74, 35-46.

<https://doi.org/10.1016/j.tranpol.2018.11.014>

*Paper II*) Rashidi, K., & Cullinane, K.P.B. (Under review), “Techniques Applied for the Selection of Sustainable Suppliers: A Systematic Review of the Literature”, *Submitted to International Journal of Information Technology and Management*.

*Paper III*) Rashidi, K., Noorizadeh, A., and Cullinane, K.P.B. (Under review), “Applying the Triple Bottom Line in Supplier Selection: A Meta-Review of the State of the Art”, *Revised and Re-submitted to International journal of Cleaner Production*.

*Paper IV*) Rashidi, K., and Cullinane, K.P.B. (2019), “A comparison of fuzzy DEA and fuzzy TOPSIS in sustainable supplier selection: Implications for sourcing strategy”, *Expert Systems with Applications*, Vol. 121, 266-281. <https://doi.org/10.1016/j.eswa.2018.12.025>

*Paper V*) Rashidi, K. “AHP versus DEA for the gradual improvement of unsustainable suppliers: A comparative analysis”. *Presented at the eighth annual Swedish transportation research conference (Linköping, 22-23 October 2019)*.

---

### **<sup>1</sup> The author’s contribution to the papers:**

The original ideas in all papers were mine. Given the original ideas, the concept and design of each paper were discussed in several meetings with my supervisor. This process was ongoing until finalising the paper. All data for all papers were gathered by the author. In terms of data analysis, I have done all analyses performed in Papers I, II, IV, and V. In Paper III, due to the size of the data set, I have done almost 80 percent of the analysis conducted. I also implemented the major part of data interpretation for all papers. I drafted all the papers and then sent them to my supervisor for further critical review. Throughout this process, my supervisor critically reviewed the papers and asked me to revise them accordingly. In addition, I have done the major part of the revision process, including answering the comments and preparing a detailed response to the reviewers’ comments. This process was also critically reviewed by my supervisor before resubmitting the revised version of the papers.

## TABLE OF CONTENTS

---

<b>1. Introduction</b> .....	1
<b>1.1. Background</b> .....	1
1.1.1. <i>Traditional vs. non-traditional supplier evaluation/selection</i> .....	1
1.1.2. <i>Logistics service providers</i> .....	3
1.1.3. <i>Evaluation approaches</i> .....	4
<b>1.2. Problem description</b> .....	5
1.2.1. <i>Sourcing problem</i> .....	5
1.2.2. <i>Benchmarking problem</i> .....	6
<b>1.3. The overall aim, sub-objectives, and research questions</b> .....	6
<b>1.4. Summary of papers</b> .....	10
1.4.1. <i>Paper I</i> .....	11
1.4.2. <i>Paper II</i> .....	11
1.4.3. <i>Paper III</i> .....	12
1.4.4. <i>Paper IV</i> .....	12
1.4.5. <i>Paper V</i> .....	13
<b>1.5. Delimitations</b> .....	13
<b>1.6. Disposition of the dissertation</b> .....	14
<b>2. Theoretical views</b> .....	15
<b>2.1. Supply chain management</b> .....	15
<b>2.2. Logistics industry</b> .....	15
<b>2.3. Procurement</b> .....	16
2.3.1. <i>Procuring logistics services</i> .....	17
2.3.2. <i>Supplier management</i> .....	18
<b>2.4. Sustainability approach</b> .....	19
<b>2.5. Modelling approaches for supplier evaluation</b> .....	22
2.5.1. <i>Performance evaluation methods</i> .....	23
2.5.1.1. <i>Qualitative methods</i> .....	24
2.5.1.2. <i>Mathematical programming methods</i> .....	25
2.5.1.3. <i>Uncertain decision methods</i> .....	26
2.5.1.4. <i>Multi-criteria decision-making methods</i> .....	26
2.5.1.5. <i>Artificial intelligence methods</i> .....	27
<b>3. Methodology</b> .....	29
<b>3.1. Research design</b> .....	29

3.1.1. <i>Research method</i> .....	30
<b>3.2.   Data</b> .....	33
<b>3.3.   Measurement methods</b> .....	35
3.3.1.   TOPSIS .....	35
3.3.2.   AHP .....	36
3.3.3.   DEA .....	36
<b>3.4.   Validity and reliability</b> .....	37
3.4.1.   Validity .....	38
3.4.2.   Reliability .....	39
<b>4.   Findings</b> .....	41
<b>4.1.   Paper I</b> .....	41
“Evaluating the sustainability of national logistics performance using Data Envelopment Analysis” ..	41
<b>4.2.   Paper II and Paper III</b> .....	42
“Techniques Applied for the Selection of Sustainable Suppliers: A Systematic Review of the Literature”	42
.....	42
“Applying the Triple Bottom Line in Supplier Selection: A Meta-Review of the State of the Art” .....	42
<b>4.3.   Paper IV</b> .....	43
“A comparison of fuzzy DEA and fuzzy TOPSIS in sustainable supplier selection: Implications for	
sourcing strategy” .....	43
<b>4.4.   Paper V</b> .....	44
“AHP versus DEA for the gradual improvement of unsustainable suppliers: A comparative analysis”	
.....	44
<b>5.   Conclusion</b> .....	45
<b>5.1.   Research question 1</b> .....	45
RQ1: Why does sustainability performance need to be considered more in the logistics industry? .....	45
<b>5.2.   Research question 2</b> .....	46
RQ 2: How has the sustainable supplier selection process been studied in the literature?.....	46
<b>5.3.   Research question 3</b> .....	46
RQ 3: To what extent are the results achieved from different evaluation methods consistent?.....	46
<b>5.4.   Research question 4</b> .....	48
RQ 4: Why do sourcing and benchmarking decisions change based on the type of method selected? ..	48
<b>5.5.   Future research directions</b> .....	49
<b>6.   References</b> .....	51

## **LIST OF FIGURES**

---

<b>Figure 1.</b> Supplier evaluation (an illustrative example) .....	5
<b>Figure 2.</b> Interrelations between the main objective, the sub-objectives, the research questions, and the papers .....	9
<b>Figure 3.</b> Purchasing process.....	17
<b>Figure 4.</b> Issues within the three pillars of sustainability .....	20
<b>Figure 5.</b> Supplier selection approaches and methods.....	25
<b>Figure 6.</b> Comparing Qualitative and Quantitative Research.....	31
<b>Figure 7.</b> Data sources for Paper I.....	33
<b>Figure 8.</b> Data sources for Paper II and Paper III.....	34
<b>Figure 9.</b> All data sources.....	35
<b>Figure 10.</b> Conceptual framework of research design.....	37

# 1. Introduction

---

*This section sets out with the aim of providing a holistic view of this dissertation. The key concepts are explained, and the interrelationships between these concepts are clarified. Furthermore, the problem description, overall aim and research purposes, and delimitations of this research are discussed. In other words, this section helps you – as a reader – achieve the gist of this dissertation within a few pages.*

## **1.1. Background**

Suppliers play a crucial role in any supply chain. They can contribute to the performance of a supply chain in either a positive or negative way. A supplier provides materials, goods, or services to a buying company. The buying company converts the products purchased into final products or uses them for other purposes in its supply chain. Suppliers are able to decrease the total cost, lead time, delivery time, and/or increase the quality of the final products, service level, and customers' satisfaction. In a similar vein, Handfield et al. (2015) claim that it is beneficial for both suppliers and buyers to establish a strong partnership in which the buyers' needs can be better understood by the suppliers; the suppliers can easier adapt to changing requirements, and more importantly, the performance of both parties can be improved. To have such a partnership, the buying company first needs to evaluate a set of suppliers in the market and then select the best possible supplier(s) among the set.

When a buying company decides to buy a product and/or outsource all or part of an activity, a remarkable number of potential suppliers exists if the market is competitive and not a monopoly. In a competitive market, it is challenging for the buying company to evaluate and select the best available supplier(s) among all other alternatives. In doing so, some preferences – also known as variables, criteria, or indexes – need to be defined and selected by the buying company. Afterward, the potential suppliers are evaluated and ranked based on the preferences defined to choose the best possible alternative (supplier). Based on the outcomes yielded from this evaluation process, the buying company begins the contracting process to achieve an agreement with the selected supplier. Given the contract signed by both parties, orders are placed by the buying company and delivered to the customer (buying company) by the supplier based on an agreed delivery timetable.

### *1.1.1. Traditional vs. non-traditional supplier evaluation/selection*

Traditionally, suppliers have only been evaluated based on economic criteria such as cost, price, quality, delivery, performance history, technical capability, etc. (Weber et al., 1991; Bhutta, 2003). According to Huang and Keskar (2007), the literature profoundly emphasised cost during the 1970s and 1980s. Later, during the 1990s, customer responsiveness and cycle time were embedded into the supplier evaluation process. Supplier flexibility was added to the evaluation in the late 1990s. Over the past two to three decades, environmental and social concerns have also been recognized

as significant characteristics of suppliers (Igarashi et al., 2013). Since then, it has been accepted, at least by a large number of both practitioners and academics, that there is no conflict between sustainability and efficiency in terms of economic aspects (Gimenez & Tachizawa, 2012; Govindan et al., 2015). More specifically, by embedding the sustainability concept into its long-term vision, a company can take advantage of it, rather than being an inconvenience that imposes extra cost (Mahler, 2007; Pant, 2005). In other words, buying companies have recognized that not only economic criteria mentioned above but also environmental (e.g. energy consumption, waste, recycling, greenhouse gas emissions) and social criteria (e.g. information disclosure, social well-being, child labour, work safety) can contribute to and improve the overall performance of their supply chains (Markley & Davis, 2007). Hollos et al. (2012) state that sustainable supplier cooperation has a positive impact on buying companies' total performance. They further discuss that evaluating the sustainability performance of suppliers cannot be an approach against the traditional perspectives in companies. Buying companies, in most cases, traditionally created their business models based on cost minimisation or profit maximisation, continuous quality improvement, higher service levels. Shortly after the emergence of the sustainability paradigm, it was discovered that the sustainability approach enables buying companies to not only achieve the aims of their traditional business models but also help them simultaneously minimise/maximise the negative/positive effects on the environment and society through synergy. Klassen and McLaughlin (1996), in the conclusion of their research, claim that the marketplace rewards companies that are implementing redesigned products and processes to minimise the negative environmental impacts of their operations. Likewise, Capon et al. (1990) found a positive correlation between social responsibility and firms' performance; the better the social performance, the more the firm is productive.

Despite numerous advantages, however, there are some challenges for the implementation of the sustainability approach. Boström et al. (2015) identified six gaps as different challenges for the implementation of sustainability, including geographical gaps, information and knowledge gaps, communication gaps, compliance or implementation gaps, power gaps, and credibility or legitimacy gaps. Many sustainability practices revolve around decreasing geographical distances between diverse parties involved in a supply chain, e.g. between suppliers and purchasers. This in fact moves from a globalised supply chain to a localised supply chain, which is difficult to perform. As a consequence of outsourcing activities to increase sustainability performance, the need has emerged for reliable and verified information in different links within a supply chain. Accessibility to such information is problematic. To enhance a supply chain's sustainability performance, the parties involved are required to communicate and collaborate efficiently, which is difficult to execute. Helin and Babri (2015) analysed a code of ethics in a supplier audit process, demonstrating the difficulties between the buyer within the first and second tiers of suppliers in terms of communication for enhancing sustainability performance. For many companies and actors involved in a supply chain, it is also troublesome to guarantee or even investigate compliance or implementation steps of the sustainability programs due to the lack of monitoring and verification systems. Boström et al. (2015) also claimed that equal distribution of power between the actors in

a chain is another major challenge for implementation of sustainability. Credibility or legitimacy gap sheds light on this aspect that sometimes a little improvement in sustainability is enough for a company to survive in the market while continuous improvement is more favourable. Chkanikova and Mont (2015) also reported several barriers, including regulatory, resource, market, and social barriers. In terms of regulatory barriers, there is no unified set of regulations between countries of different parties involved in a supply chain. Definitely, limited access to financial resources challenges the implementation of sustainability practices. Furthermore, customers' low desire and interest in paying extra cost for sustainability considerations in the market is considered another barrier. Last but not least, there are some social barriers, for instance, customers' lack of commitment as well as considering sustainability as a socially constructed phenomenon rather than a necessary reality that needs to be handled efficiently.

One of procurement managers' significant tasks involves selecting the most appropriate supplier who is compatible with sustainability requirements because, as Krause et al. (2009) claimed, a firm is only as sustainable as its suppliers. While a business can flourish considering sustainability aspects in its supplier selection process, it can fade due to the poor performance of suppliers in its supply chain. A firm's high level of environmental performance could be ruined by its suppliers' poor environmental performance (Faruk et al., 2001). Shane and Spicer (1983) stated that firms' stock price decreased based on their poor pollution control performance. This implies that a firm's future profitability will be negatively changed if the firm violates environmental standards. The same situation may occur for a firm when unpleasant results of social concerns related to its suppliers influence its public image. For instance, Mattel and Nike are but two examples of the many companies that have paid high costs for their suppliers' poor environmental and social sustainability performance. According to *The New York Times*, Mattel recalled roughly a million children's toys after its suppliers used lead-contaminated paint in the production process (Story, 2007). Similarly, *The Guardian* reported that Nike has been widely criticised, as its sub-contractors were using child labour in sweatshops (Day, 2001).

### *1.1.2. Logistics service providers*

In any supply chain, there exists a wide range of suppliers with different roles and activities. Some are involved in primary activities, while others are selected to complement support activities. They are spread out over supply chains, from upstream to downstream. Some supply raw materials, some provide repair and maintenance services, some operate as consultancy companies providing guides, and some handle logistics operations. However, for practical purposes and conformity to real-world situations, one type of the above supplier groups, and then a sub-set of this respective group, that is homogenous in terms of the type of activity needs to be selected. This enables the researcher to show the applicability of her/his findings based on a real-world empirical application. In doing so, the last group of suppliers that handle logistics operations, called logistics service providers, is selected for achieving the purpose of this research. The logistics function encompasses diverse activities such as warehousing, packaging, transportation, materials handling (Jahre & Johan Hatteland, 2004; Roorda et al., 2010; Van Hoek, 2002). Another reason for choosing a set

of logistics service providers is that most companies nowadays buy the required logistics services for their operation from one or several logistics service providers. There are a remarkable number of reasons for outsourcing logistics operations. For instance, buying companies with outsourced logistics activities can focus more on their core activities and decreasing transportation costs through taking advantage of economies of scale provided by logistics service providers. The last reason, and maybe the utmost for choosing a set of logistics service providers as suppliers in this research, is the remarkable share of this type of supplier in the sustainable performance of the supply chain in which they are operating. Logistics operations need to be implemented based on sustainability criteria, as their share in energy consumption and, consequently, greenhouse gas emissions, is quite high. The US Energy Information Administration reported that freight modes roughly consume 39 percent of total world transportation energy consumption (IEO, 2016, p. 130). Furthermore, it has been reported that this amount will increase from 40 quadrillion British thermal units (Btu) in 2012 to approximately 60 quadrillion Btu in 2040. In the same vein, in terms of social aspects, the number of employees in this industry contributes, in most nations, three to five percent of the total workforce (Rashidi & Cullinane, 2019).

### *1.1.3. Evaluation approaches*

Similar to most other disciplines, the approaches applied for evaluating the performance of suppliers are categorised into qualitative and quantitative. In qualitative approaches, the performance of a set of potential suppliers is evaluated by experts in the field to select the best supplier. Similarly, in some cases, the buying company sends inspectors to the suppliers' production sites, and those inspectors evaluate the suppliers' performance based on their observations. In the quantitative approach, however, suppliers' performance based on each criterion – defined and determined by the purchasing department – is quantified and provided in numbers. Given the quantified data set gathered, the set of suppliers is ranked based on a mathematical and/or statistical model. The literature in the field of supplier selection is dominated by the quantitative approach, with quantitative models that are able to decrease subjectivity in the evaluation process.

On the other hand, the necessity to consider suppliers' environmental and social performance has made the traditional supplier evaluation process more complex, with multiple dimensions involved. In sustainable supplier selection, more data is needed; more constraints need to be considered, and sometimes conflicting objective functions need to be harmonised. Accordingly, researchers have developed diverse single and hybrid methods to establish an appropriate framework for evaluating the sustainability of a set of suppliers based on selected variables (Gimenez & Tachizawa, 2012; Genovese et al., 2013; Igarashi et al., 2013). There are a number of different methods, ranging from mathematical programming to multi-criteria analytical and artificial intelligence (AI) methods, and each has pros and cons. Some researchers have applied a single method to deal with sustainable supplier evaluation (e.g. Shaik & Abdul-Kader, 2011; Yakovleva et al., 2012; Wen et al., 2013), while other researchers have proposed hybrid methods,

combining two or more than two methods in a unified framework, to improve sustainability evaluation of suppliers (e.g. Bai et al., 2010; Kuo et al., 2010; Girubha et al., 2016).

### 1.2. Problem description

As discussed earlier, a company’s success is highly dependent on its suppliers’ performance. Any deviation from an appropriate decision imposes a high cost and risk to purchasing companies. For instance, decreasing market share is only one of the many negative consequences of this deviation. A kind of deviation from the appropriate supplier selection decision is any difference in the ranking of suppliers based on different methods/techniques/tools. These methods will be later discussed in theoretical views, Section 2. In other words, if each method provides a unique ranking of suppliers, then it implies that the performance of suppliers and, accordingly, their rankings, is partially influenced by the methods’ algorithm, and not by their real performance. More importantly, whether the critical decisions made based on the outcomes of these methods are changed needs to be investigated. This problem is further explained with an illustrative example and then discussed as to how these phenomena can affect strategic decisions in a supplier management system. Let us assume a buying company called XYZ asks its purchasing managers to rank five potential suppliers (A, B, C, D, and E) in the market for supplying a purchased item. The purchasing department chooses three evaluation methods to rank this set of suppliers. Figure 1 shows this illustrative example and outcomes of each model.

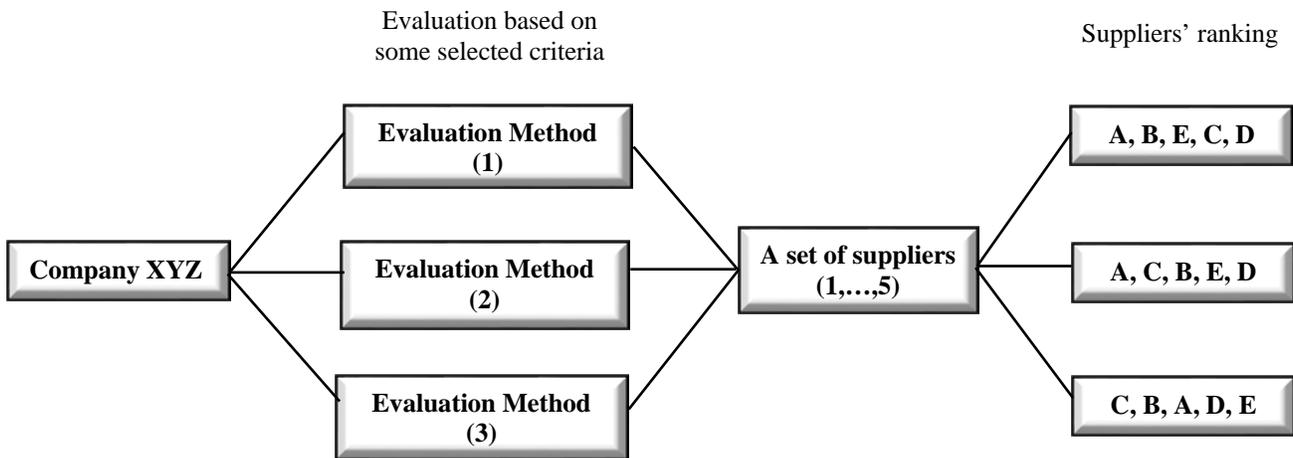


Figure 1. Supplier evaluation (an illustrative example)

#### 1.2.1. Sourcing problem

Based on the above example, company XYZ achieves three different types of supplier rankings by applying the three different evaluation methods. Next, the reliability of the above supplier rankings is reviewed based on diverse sourcing strategies. Generally, there are three types of sourcing strategies: single sourcing, dual sourcing, and multiple sourcing (Yu et al., 2009). Company XYZ faces some difficulty with single sourcing, since methods 1 and 2 determine

supplier A as the best, while supplier C has the best performance based on method (3). This problem becomes worse with dual sourcing since there is no consistency in the evaluation methods' outcomes for supplier rankings. Methods 1, 2, and 3 yield suppliers A and B, A and C, and C and B as the best suppliers, respectively. This problem remains for multiple sourcing, where the top three performers are different under each evaluation method. This example illustrates how rank reversal in evaluation methods can substantially affect a company's sourcing strategies.

### *1.2.2. Benchmarking problem*

In some cases, the evaluation methods are applied for supplier development purposes. The influential role of suppliers has obliged buying companies to establish long-term relationships with their suppliers. Therefore, evaluating and ranking suppliers for sourcing decisions is only one task of a purchasing department. For an efficient sourcing strategy, this task needs to be followed by monitoring and improving the performance of suppliers that are collaborating with the buying company. Benchmarking suppliers' performance helps to continuously monitoring their activities and establish a systematic supplier selection and improvement mechanism (Choy et al., 2002). In the benchmarking process, first, the most efficient suppliers are determined, and then suppliers with poor performance are prompted to imitate the best suppliers' performance suppliers in order to improve their efficiency (Forker & Mendez, 2001). Therefore, buying companies need to initially evaluate a set of suppliers, aiming to find the best possible alternatives for their sourcing strategy. Then, they need to periodically evaluate the efficiency of suppliers chosen to improve the performance of poor suppliers by implementing a suitable benchmarking policy.

Assume that company XYZ is implementing a program to benchmark the performance of poor suppliers against the best suppliers based on the ranking gained from the suppliers' evaluation methods depicted in Figure 1. Based on the benchmarking process, the best supplier first needs to be identified; that is supplier A based on methods 1 and 2, and supplier C based on method 3. This difference between determining the best supplier is problematic because company XZY cannot decide which supplier is the benchmark for poor performers. In the same vein, it is a challenge to identify suppliers with poor performance that need to imitate the best supplier's performance. More specifically, supplier C, with quite poor performance based on method 1, has reasonable performance in the outcomes of methods 2 and 3. Therefore, through this example, it can be seen that identifying the best and worst suppliers for implementing a benchmarking strategy may be different based on different evaluation methods.

### *1.3. The overall aim, sub-objectives, and research questions*

On the one hand, companies' decisions and policies are highly dependent on the outcome of the supplier evaluation process. Thus, selecting the most suitable technique among various methods is challenging, and any inconsistency between outcomes yielded from different methods is problematic for making purchasing decisions. Comparing the outcomes of various evaluation methods will inform decision makers about the challenges and difficulties they may face in applying different sourcing and benchmarking strategies. On the other hand, the importance of this

type of information has increased in recent years with the emergence of the sustainability concept and the role of suppliers in a company's sustainability performance. Two aspects of sustainable supplier selection constitute the core discussions of this research. First is the reliability and applicability of the results yielded by the widely applied methods in the literature for sustainable supplier selection. Second is the effect of the methods' outcomes on changing the purchasing, sourcing, and benchmarking strategies in a supply chain. According to Clark and Badiie (2010), the focus of any study is defined based on three principal elements: the content area, purpose/objective, and research questions. They created a hierarchy based on these three elements, so the content area, as the most general and broad element, encompasses the purpose/objective, which itself involves the research questions – the most specific and narrow element. In other words, the research questions fall into the objective of a study, and subsequently, the objective itself falls into the content area. The sustainable supplier selection introduced in Section 1.1 constitutes the content area of this research. Accordingly, the principal aim/objective of this dissertation is specifically:

***“To determine the extent to which supplier ranking is influenced by the method applied in sustainable supplier selection in general and particularly in the logistics industry, as well as how sourcing and benchmarking strategies can be modified based on the outcomes yielded by the evaluation methods.”***

This principal aim is, in fact, constituted with three sub-objectives. The ***first sub-objective*** seeks to investigate potential differences between the outcomes of the evaluation methods, if there are any. In other words, it highlights the extent to which the results of an evaluation method are reliable. The ***second sub-objective*** focuses on why sustainability in the logistics industry needs to be further investigated and in what way the previous research has studied the evaluation of sustainability of suppliers. The ***last sub-objective*** aims to investigate in what ways outcomes of the evaluation methods can change the decision made for sourcing and benchmarking. This sub-objective helps us further show the managerial implications of this research, or stated differently, the applicability of our findings in a real-world case situation.

Having the content area, main objective, and sub-objectives, the research questions can be defined as the last element in conducting this study. Research questions are more specific and help researchers achieve the main objective. Research questions can be stated in different forms, such as question, declarative, or hypothesis (Clark & Badiie, 2010). The main aim defined above, as the general framework for this research, is achieved by the following four specific research questions:

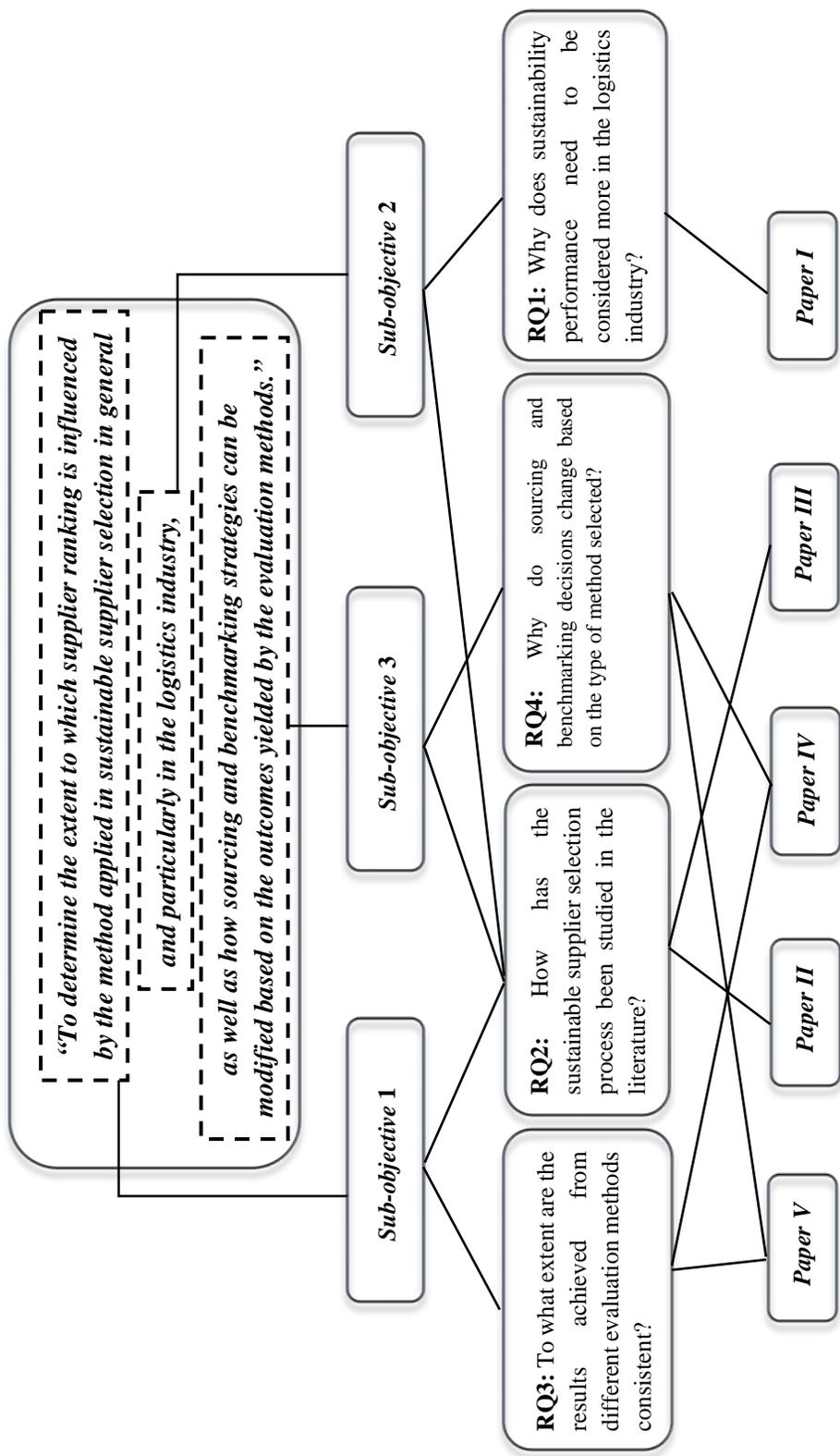
**RQ 1:** Why does sustainability performance need to be considered more in the logistics industry?

**RQ 2:** How has the sustainable supplier selection process been studied in the literature?

**RQ 3:** To what extent are the results achieved from different evaluation methods consistent?

**RQ 4:** Why do sourcing and benchmarking decisions change based on the type of method selected?

As a starting point for this study, **RQ1** explores the essential need for sustainability improvements in the logistics industry. In doing so, an exploration takes place to determine whether a competent logistics industry in terms of infrastructures and technology is also capable in terms of sustainability. Furthermore, the importance of environmental and social aspects in the logistics industry is highlighted by answering this question. **RQ2** helps find how sustainable supplier selection has been studied in the literature. Investigating previous studies enables the researcher to achieve a holistic view of the topic. In addition, the researcher benefits from this research question by systematically reviewing previous attempts and extracting potential research gaps. In addition to potential research gaps for future research, the widely applied methods and criteria for evaluating the sustainability performance of suppliers are explored. Doing so helps uncover which industries have been neglected by researchers in the field, industries that need to be explored more. This research question also has a central role in forming the main aim of this study. By **RQ3**, the outcomes of three widely applied methods (later, a detailed explanation regarding which methods and why is provided) are compared to discover the potential differences between their outcomes based on a common data set, if there are any. In other words, the consistency between the outcomes of methods is compared to answer this research question. The last research question, **RQ4**, aids in exploring the influence of the methods' outcomes on two strategic decisions in a sustainable supplier management system: sourcing and benchmarking. To answer this research question, the outcomes yielded from the selected methods are simultaneously applied in a sourcing and benchmarking empirical application to determine the differences. The interrelations between the main aim, sub-objectives and research questions with the papers are schematically indicated in Figure 2. This figure shows which paper contributes to which research question, and which research question contributes to which sub-objective and finally to the main objective. It is significant to note that **RQ2** is the only research question that contributes to all sub-objectives because for answering this question, two literature review-based papers are written that discover the methods applied in the field, the research gap in the logistics industry related to sustainable supplier selection, and the gap in investigating the consequences of inconsistency between the methods' outcomes on critical decisions in a sustainable supplier management system.



**Figure 2.** Interrelations between the main objective, the sub-objectives, the research questions, and the papers

#### *1.4. Summary of papers*

In addition to the interrelations between the research questions with the sub-objectives, Figure 2 shows the interrelations between the papers and research questions. The first paper, *Paper I*, contributes to one research question: **RQ1**. *Paper I* focuses on the importance of sustainability performance in the logistics industry at a macro level. The findings of this paper help to better understand the environmental and social concerns in the logistics industry as well as the role of sustainable operational performance at the national level. *Paper I*, through **RQ1**, contributes to the second sub-objective and helps explain the crucial role of sustainability in the logistics industry that needs to be implemented at both the micro and macro levels.

As seen in Figure 2, *Paper II* and *Paper III* contribute to **RQ2**, which is about the evolution of sustainable supplier selection as a research field in the literature. These two papers are literature review-based research and map the previous research in the field of sustainable supplier selection. The analysis performed in these papers shows which methods and criteria are the most applied in the field. Likewise, previous contributions and novelties introduced by other researchers in the field are summarized, and potential future research gaps are identified. The findings of these two papers contribute to all the sub-objectives through answering **RQ2**. More specifically, the findings justify the novelty of this research in the field of sustainable supplier selection. The findings imply that there has been a lack of comparison of widely applied methods as well as a lack of research to investigate how critical decisions in supplier management systems can change based on the differences (inconsistency) in the outcomes of different evaluation methods. Furthermore, the findings of these two papers support the findings of *Paper I* and contribute to the second sub-objective by justifying the need for this research in the logistics industry.

The last two papers, *Paper IV* and *Paper V*, simultaneously contribute to **RQ3** and **RQ4**. Based on the analysis performed in literature-review papers (*Paper II* and *Paper III*), three widely applied evaluation methods, i.e. Data Envelopment Analysis (DEA), Technique for Order of Preferences by Similarity to the Idea Solution (TOPSIS), and Analytical Hierarchy Process (AHP), are selected for achieving the main aim of this dissertation. These methods will be explained in detail in Sections 2 and 3. While these methods are compared in pairs in *Paper IV* and *Paper V*, DEA and TOPSIS, and DEA and AHP are compared in *Paper IV* and *Paper V*, respectively. Note that as we used a common data set for both papers, it is possible to compare the outcomes of TOPSIS and AHP as well, and there is no need for a separate paper for this purpose. To answer **RQ4**, in each of these papers, one critical decision in a supplier management system is selected to investigate the challenges managers may face in making these decisions based on the outcomes of these evaluation methods. Sourcing is discussed in *Paper IV*, and benchmarking is discussed in *Paper V*.

#### 1.4.1. Paper I

##### *“Evaluating the sustainability of national logistics performance using Data Envelopment Analysis”*

The only available and reliable index to show the logistics capabilities of each country is the Logistics Performance Index (LPI) calculated by the World Bank<sup>2</sup>. Based on this index, each country’s logistics industry performance is measured based on six functional, infrastructural, and technological criteria. However, the index fails to consider sustainability competency in the evaluation of a country’s logistics industry due to the lack of sustainability criteria in its calculation procedure. In this paper, the sustainability of operational logistics performance of a sample of 22 OECD<sup>3</sup> countries has been evaluated based on four sustainability criteria, i.e. energy consumption, greenhouse gas emissions, amount of total inland freight, and the ratio of those employed in the logistics industry to the total workforce in each country. Using a DEA-based model, each country was assigned an efficiency score based on sustainable operational logistics performance (SOLP). SOLP is a new index proposed in this paper, and its results are compared with the index provided by the World Bank, i.e. the LPI. Given the efficiency scores, the countries in the sample are ranked, and then the rank correlation coefficient between the index proposed in this paper and the LPI is discussed. The best and worst countries based on SOLP are identified, and the necessity of sustainability in the logistics industry is further investigated.

#### 1.4.2. Paper II

##### *“Techniques Applied for the Selection of Sustainable Suppliers: A Systematic Review of the Literature”*

There are two literature review-based articles in this study. The first one takes a subjective approach to map the literature in the field of sustainable supplier selection/evaluation, while the second one quantitatively analyses the content of papers in the literature. In *Paper II*, the key phrase “sustainable supplier selection” was searched in Google Scholar, Scopus, and Web of Science to extract the related published documents in the field. Through this initial search, a total of 847 documents in the field were extracted from the three search engines. After eliminating duplicates, 708 papers remained for consideration in the filtering process. A structured filtering process was applied to choose the most relevant papers and eliminate irrelevant papers for further analysis. Reading the abstracts, which was the first phase of the filtering process, reduced the number of papers from 708 to 127 papers. The full content of these 127 papers was reviewed, and only 37 final papers were chosen for further analysis. The publications’ evolution in the field of sustainable supplier selection, the criteria and methods applied, the number of suppliers under evaluation and applied variables in each paper, and the industries investigated are reviewed in this paper.

---

<sup>2</sup> <https://www.worldbank.org>

<sup>3</sup> Organization for Economic Co-operation and Development.

#### 1.4.3. Paper III

##### *“Applying the Triple Bottom Line in Supplier Selection: A Meta-Review of the State of the Art”*

Unlike *Paper II*, this paper applies a quantitative analysis to investigate the content of published papers in the field of sustainable supplier selection/evaluation. Using a bibliometric tool called Bibexcel and network software called Gephi, the related papers in the field are mapped and analysed. The other main difference between *Paper II* and *Paper III* is the domain of published documents covered. This literature review is a meta-review, with initial data of 15,393 documents searched in Scopus and Web of Science, based on 336 combinations of 21 keywords categorised into three groups. After eliminating duplicates, 4,882 documents remained for consideration in the filtering process. Reviewing the title and keywords, which was the first phase of the filtering process, reduced the number of documents from 4,882 to 1,328. The second phase, reading the abstracts, reduced the sample of related documents for further analysis to 746. As the last step of the filtering process, the full content of these 746 papers was reviewed to establish the final sample of papers for further analysis. Papers that failed to properly apply or address the triple-bottom-line perspective in supplier evaluation problems were excluded, and only 66 papers remained for the final analysis. In addition to the analysis in *Paper II*, several co-occurrence analyses, including co-word, co-methods and co-applications, were applied to extract key issues discussed in the field of sustainable supplier selection. The interrelations between applied methods in the literature for sustainable supplier selection are reviewed. The main trends, main contributing authors, and main contributing journals to the field are recognized. Based on a co-word analysis, the main research streams and topics addressed in the field are determined. Furthermore, a co-author analysis helps categorise the contributing authors and their respective applied methods. Co-citation analysis is conducted to map the main research clusters in the field. Finally, some potential future research directions are provided for further research in the field of sustainable supplier selection.

#### 1.4.4. Paper IV

##### *“A comparison of fuzzy DEA and fuzzy TOPSIS in sustainable supplier selection: Implications for sourcing strategy”*

In this paper, a comparative analysis is implemented between TOPSIS and DEA methods in a fuzzy environment. Given the fuzzy logic, the subjectivity of respondents' opinions is decreased. The data for a set of logistics service providers were gathered through a questionnaire survey. Based on the literature review-based articles, six criteria (cost, quality, energy consumption, environmental management system, labour health and work safety, and social responsibility) as the most-applied criteria are selected for evaluating sustainability performance of a set of logistics service providers. The data gathered are transformed into fuzzy data, and then the logistics service providers are ranked using both fuzzy TOPSIS and fuzzy DEA. Given the rankings of suppliers yielded from both fuzzy DEA and fuzzy AHP, a rank correlation coefficient is calculated to check differences, if there are any, between these two methods. The sourcing decision in procurement is

discussed based on the results yielded. Finally, these two methods are compared based on time and calculation complexities.

#### *1.4.5. Paper V*

##### *“AHP versus DEA for the gradual improvement of unsustainable suppliers: A comparative analysis”*

In this paper, DEA is compared with AHP for evaluating the sustainability performance of the same set of logistics service providers based on the same data set in *Paper IV*. Unlike the previous paper, this paper discusses the effect of each method's outcomes on benchmarking strategy instead of the sourcing decision. Based on the rankings yielded from both methods, the suppliers are classified into different clusters using the self-organizing map method in the R environment – a programming language for data analysis. The self-organizing map method exploits the similarities between the data set and the efficiency scores of suppliers achieved from DEA and AHP. Three different scenarios were selected for clustering the suppliers, including categorising suppliers into four, six, and eight clusters to investigate the potential differences between the two applied methods. Finally, the differences between the benchmarking frameworks proposed by each method, i.e. AHP and DEA, are discussed.

#### *1.5. Delimitations*

Small- and medium-size logistics service providers are excluded from this research. This does not mean that they have no impact on the logistics industry, or that the sustainability evaluation of these suppliers is unimportant. However, it was assumed that the respondents involved in the questionnaire survey are more acquainted with the well-known and big logistics service providers in the industry. A higher number of criteria involved in the evaluation process can increase the validity of the findings. The more criteria involved, the better the reliability of the results. However, it is not an easy task to select only a few criteria, as there is a wide range of diverse criteria applied in the literature. To keep the length of the questionnaire at a reasonable level, six criteria were chosen for this study. ‘Delivery’, ‘financial capability’, and ‘flexibility’ were also among the widely applied criteria for evaluating the economic performance of suppliers, but they were excluded in this study to keep the length of questionnaires short and increase the chance of higher response rate. In a similar vein, ‘greenhouse gas emission’ and ‘reuse/recycling’ were excluded for evaluating the environmental performance of the suppliers under assessment.

In addition to the three methods studied in this research, Analytical Network Process (ANP) was recognized as one of the most-applied methods in the field. However, ANP is the extended form of AHP, considering the interdependencies between criteria, while AHP does not consider these interdependencies. The algorithms of both methods are similar, and the only difference is that ANP considers more factors (interdependencies) in its calculation process. Therefore, it is not surprising to see some differences between the outcomes of these two methods, as it is not possible to feed both methods with a unique and common data set. Therefore, the ANP method was

excluded, as it needed different types of data compared to the other three methods applied in this study.

### ***1.6. Disposition of the dissertation***

The rest of this dissertation is unfolded as follows. Section 2 provides some related theoretical views, including supply chain management, logistics industry, procurement, supplier management, sustainability approach, and applied evaluation approaches for supplier selection. The methodology is described in Section 3. In this section, the research method, data for each paper, and measurement tools applied in this dissertation are explained in more detail. Section 4 provides the main findings of each paper. Finally, Section 5 concludes this research by answering the research questions and suggesting future research avenues.

## **2. Theoretical views**

---

*This section provides a review of some theoretical views related to the core topic of this dissertation. First, a holistic discussion of supply chain management and the logistics industry is provided, then procurement in the supply chain, sustainability in procurement and supplier management, and finally, evaluation methods for performance assessment are discussed. The theoretical views section help readers understand fundamental assumptions and perspectives in the research area under study.*

### **2.1. Supply chain management**

Supply chain management has different meanings, ranging from the operation of the flow of materials and products to a management philosophy or process. Mentzer et al. (2001) discussed how there is ambiguity in defining this concept in spite of its popularity. It is defined as a set of companies – such as raw materials producers, assemblers, transporters, distributors, and retailers – that convey materials forward (La Londe & Masters, 1994). In another definition, Christopher (2016) considered supply chain management as a whole, so the relationship between upstream and downstream companies in the chain is managed to deliver products and services to the final customer. Based on key activities of a supply chain proposed by Mentzer et al. (2001), throughout a supply chain, the behaviour of some companies is integrated, and accordingly, information, risks and rewards are shared within and among companies involved. The companies cooperate to address customers' needs based on common goals and objectives. This cooperation often creates a long-term strategic relationship between the parties involved. Production, procurement, marketing and sales, logistics, finance, and R&D are some of the key operations within a supply chain management system (Lambert & Cooper, 2000). According to Hugos (2018), each company in a supply chain, in addition to making decisions in cooperation with other companies in the chain, needs to make decisions individually in five areas, i.e. production, inventory, location, transportation, and information. Comparing the slow-moving mass market of the industrial era with current fast-moving markets, modern companies prefer to focus on their core activities and outsource non-core activities. Traditionally, companies preferred to keep more activities in-house.

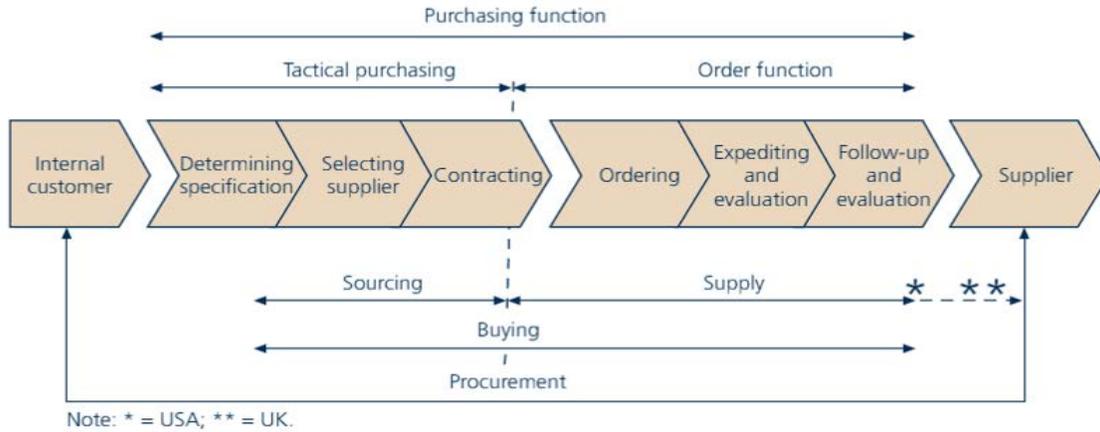
### **2.2. Logistics industry**

As mentioned earlier, one of the key drivers in each supply chain is a logistics operation. In business, through logistics management, forward and reverse flow and storage of goods, services, and information are implemented and controlled efficiently and effectively between the point of origin and the point of consumption (Council of Supply Chain Management Professionals, 2013). The logistics operation combines diverse activities such as information flow, packing, inventory, materials handling, warehousing, transportation, production, and even security. The logistics operation in each supply chain can be a combination of all above activities, or just some of them, depending on the structure of each supply chain. According to Investopedia, logistics management is defined as the process of acquiring essential resources for the production process, storing

productions, and finally transporting these productions to final customers. Logistics terms were initially used in the military and referred to the process of obtaining, storing, and moving equipment and supplies. However, the terms are now widely used in business. Due to the emergence of environmental issues and the importance of cost-saving in competitive markets, logistics management was re-defined to cover the reverse flow of goods as well for different purposes such as re-use, refurbishment, and re-manufacturing. (Ozsen et al., 2008; Rogers & Tibben-Lembke, 2001). In fact, the logistics operation starts from the point of origin to the point of consumption, and then it continues from the point of consumption to the point of origin for the purpose of waste management. According to Christopher (2016), a single plan for the flow of products and information within a business can be formed through the logistics operation. In addition, planning and coordinating all essential activities for achieving the desired level of delivered products/services and quality at the lowest possible cost constitutes the core of logistics management. Christopher (2016) defined logistics as the link between the supply base and the marketplace.

### **2.3. Procurement**

Procurement is a principal function within any organisation or supply chain, and supplier evaluation plays a pivotal role within this function (see Fig. 3). Procurement is defined as “*the management of the company’s external resources in such a way that the supply of all goods, services, capabilities and knowledge which are necessary for running, maintaining and managing the company’s primary and support activities is secured at the most favorable conditions covering the materials, information and money flows up to the end point of consumption*” (Van Weele, 2018, p. 7). The procurement function is considered a support activity in a supply chain. Support activities are those that support the completion of primary activities (Porter, 2008). Through direct purchasing, to buy raw materials for production operations or different types of logistics activities, the procurement function supports the primary activities of a business. However, also through indirect purchasing, to supply products and services for other support activities, procurement supports other support activities (Van Weele, 2018). The impact of procurement on both the financial and operational performance of a firm has been demonstrated (Cousins et al., 2006). Procurement is also recognized as one of the competitive advantages a firm can gain in today’s global business environment (Porter, 2008). Effective procurement could be regarded as a value-creating strategy, which is not an easy task to execute and not easily implemented by competitors in the market.



**Figure 3.** Purchasing process (Van Weele, 2018, p. 8)

### 2.3.1. Procuring logistics services

Nowadays, logistics operations constitute a remarkable share of services purchased in many manufacturing companies. Outsourcing logistics operations enables companies to introduce their innovative products and services quickly to the target market (Lai, 2004). Likewise, buying companies can achieve a wide range of benefits by purchasing their needed logistics operations from a third-party provider. These benefits include, but are not limited to, risk-sharing, access to world-class technology and innovations, reducing required capital investment, increasing adapting capability to new competing environments, reducing operational costs as well as challenges related to the operation, which enables managers to access resources that are not available in their company (Quinn & Hilmer, 1994; Persson & Virum, 2001). Buying companies select logistics service providers, as vital suppliers in the supply chain, to perform either all or part of their logistics operation. For instance, DHL, the world’s largest logistics specialist, helps its customers solve practical problems as well as reduce the cost of freight (DHL Global, 2019). It cooperates with companies operating in different industries such as automotive, consumer goods, food/groceries, retail, and healthcare. Some of DHL’s big customers are Ford Motor, Exxon Mobil, Hershey, Walmart, Xerox, Goodyear Tire & Rubber, Procter & Gamble, and Toys “R” US (*Logistics Quarterly Magazine*, 2019). Thus, it can be seen that not only small- and medium-size enterprises but also big manufacturers prefer to cooperate with third-party logistics service providers to manage their logistics operations. This implies that in today’s business environment, a large number of manufacturers need to evaluate the performance of available suppliers in the logistics market to choose the best for their pre-defined purposes. To do so, the buying company needs to establish an effective supplier management system in its procurement function.

### 2.3.2. *Supplier management*

Supplier management is considered a vital task in procurement management. Strategic supplier management is an investment in a transaction with an external party that results in benefits without increasing the actual cost of ownership (Carr & Pearson, 1999). Generally, supplier management encompasses three main processes, namely, supplier evaluation, supplier selection (Braglia & Petroni, 2000) and supplier development (Schiele, 2007). The first two processes, in fact, are decision phases in which the suppliers' performances are assessed, and then, the results yielded lead to the final choice. In fact, evaluation and selection are intertwined processes and can be considered the starting point of a strategic relationship established between buyers and suppliers.

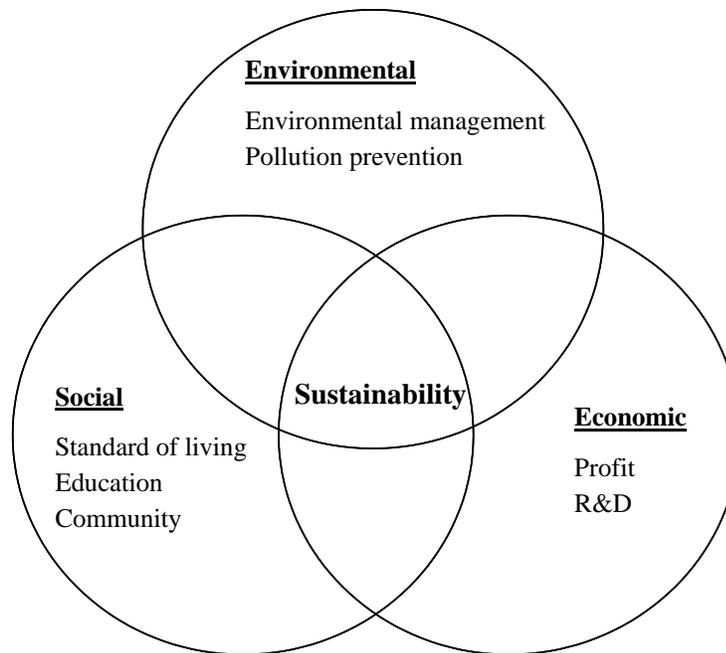
Supplier evaluation and selection commits the required resources to run the main operational processes while concurrently affecting product quality, production planning and control, cash flow requirements, cost reduction, and inventory management (Narasimhan, 1983; Ghodsypour & O'Brien, 1998). In addition, the process has an undeniable impact on the integration of the supply chain management system in the competitiveness of the whole supply chain and, more importantly, any organisation's overall performance (Chan et al., 2008; Chang et al., 2011; Junior et al., 2014). On the other hand, a failure in supplier evaluation undermines product quality, increases the total cost, leads to an excess or shortfall in inventory, decreases responsiveness, and brings many other disruptions into the supply chain. Supplier selection is considered selection of the most appropriate supplier(s) that is able to deliver the required amount of purchased goods/services within the pre-defined standards of quality, at the most favourable price, and with an acceptable level of delivery reliability. Through this process, potential suppliers are investigated, evaluated, and finally, one or several of them – depending on the buying company's sourcing strategy– are selected to secure the supply of goods/services requested by internal customers (Chopra & Meindl, 2015). The supplier selection process starts by defining some pre-qualifications based on purchase-order specifications, ranging from quality specifications to logistics and maintenance requirements. Then, a pool of potential suppliers – the bidder's long list – that may be able to deliver the goods/services is initiated. The buying company then gathers invaluable information about these nominated suppliers. In doing so, a request for information will be distributed to each supplier, and the buying company will implement a data-gathering process. Given the requested information and data, the buying company implements the evaluation and ranking process (Van Weele, 2018; Chopra & Meindl, 2015).

Using supplier development programs, the company can improve suppliers' level of performance. Supplier development practices can result in cost reduction, quality improvement, decreased lead times, delivery enhancement, and, finally, increased productivity levels (Modi & Mabert, 2007). Therefore, the purchasing company evaluates a set of suppliers and then selects one or several, depending on the sourcing strategy. In this way, the buying company creates a network of suppliers to continuously monitor, evaluate, and improve by implementing various supplier development programs.

#### ***2.4. Sustainability approach***

During the last three decades, the perspective on supplier evaluation has shifted from a traditional point of view to a contemporary approach. Traditionally, cost factors were predominantly applied to supplier selection, as the only criterion in the suppliers' evaluation process was the lower the cost, the better the rank of suppliers (Ho et al., 2010). As a consequence of the industrial era, purchasing managers have discovered that this type of one-dimensional evaluation perspective cannot guarantee the success of the supplier selection process and have started to look at the supplier selection process from a single-factor to a multi-criteria problem. Quality, delivery, manufacturing capability, service management, technology, etc. have been added to the supplier evaluation and selection process (Ho et al., 2010).

The need to integrate environmental management systems in supply chains has emerged among scholars and practitioners (Srivastava, 2007). External pressures have obliged companies to consider environmental and social issues in their supply chains (Wolf, 2014). These external pressures frequently originate with other stakeholders in society, such as diverse national and international organisations, media, and customers. However, external pressures were not the only reasons for considering sustainability in supply chains. In addition to contributing to an organisation's reputation, sustainability initiatives enable companies to reduce cost and risk, as well as measure value-creating activities in a value chain (Lubin & Esty, 2010; Wolf, 2014). Accordingly, environmental criteria have increasingly gained attention in the supplier evaluation process but remain less important than economic attributes. Companies have started to identify and manage different types of waste and pollution in their production process. After environmental considerations, social criteria have been brought into the main discussion with companies by consumers, shareholders, governmental and non-governmental organisations (Fig. 4). In fact, the social aspect of sustainability investigates a company's impact on the society in which it operates (Jung, 2017). As a consequence, green operations have been changed to sustainable ones that simultaneously consider the economic, environmental, and social aspects of supply chains (Seuring & Müller, 2008); the supplier evaluation process has not been excluded from these substantial changes. Figure 4 schematically shows where the sustainability concept is located and in which areas there is social and environmental sustainability. This figure also supports the findings of Seuring and Müller (2008) that the triple bottom lines need to be considered simultaneously for achieving sustainability.



**Figure 4.** Issues within the three pillars of sustainability (Rodriguez et al., 2002, p. 8)

#### *2.4.1. Sustainability in the logistics industry*

In a recent work, Evangelista et al. (2018) investigated the environmental sustainability of third-party logistics service providers based on a literature review from 2000 to 2016. They claimed that there are several reasons the environmental concerns of the logistics service sector have become more severe. First and foremost, the transport and logistics sector is one of the principal contributors to greenhouse gas emissions. Because there is tremendous energy consumption in this sector, the high amount of emissions is a big threat to global warming (IEO, 2016). Second, the number of goods transported has increased in recent years and will grow in future years (ITF, 2016). This is possibly rooted in the consumerism that surrounds our daily lives. According to an analysis performed by Evangelista et al. (2018), the number of articles published in the field of environmental sustainability of logistics service providers has increased from 2013. They linked this increasing trend in publications to the second phase of commitment of the Kyoto Protocol held in Japan, in which countries committed to reducing greenhouse gas emissions. Evangelista et al. (2018) also categorised the drivers for considering green indicators by logistics service providers based on content analysis. The content analysis showed five main groups of drivers, including economic-financial, governmental, market-related, organisational, and technological. In another type of categorisation, Colicchia et al. (2013) classified some principal environmental sustainability initiatives into two categorisations, namely intra-organisational and inter-organisational practices. The first categorisation includes distribution strategies and transport execution, warehousing and green building, reverse logistics, packaging management, and internal management, while the latter contains collaboration with customers and other types of external collaborations. The latter group of initiatives is considered a joint collaboration that logistics

service providers establish within different parties in a supply chain. According to Colicchia et al. (2013), the internal management initiative encompasses managers and operative staff. In fact, increasing managerial skills for considering environmental consequences and setting appropriate environmental targets as well as prioritising sustainability practices within a company are all internal management initiatives.

In addition to environmental sustainability, the social aspect needs to be considered in the supplier evaluation process; this issue has been widely neglected by both academicians and practitioners (Hutchins & Sutherland, 2008). The social aspect of suppliers' performance can be evaluated considering social criteria, e.g. social commitment, child labour, social management, occupational health and safety, working time, wages, training programs for employees, diversity and discrimination, employment relationship, stakeholder involvement, and corporate citizenship (Zimmer et al., 2016). These are only a few of the criteria to consider when evaluating the social aspect of suppliers' sustainability performance. Ghadimi et al. (2016) provided another categorisation of social criteria: local communities influence. The criteria in this categorisation represent social influences extended by a supplier to the local communities around its business. For instance, this includes to what extent suppliers contribute to housing, education, health, security, economic welfare, social cohesion, grants and donations, and supporting community projects. Today, a great interest in work safety and labour health exists in most countries, which obliges organisations to consider social sustainability in their operations, and logistics service providers are not excluded. According to Jung (2017), three main classifications of social criteria are applied in the literature for considering the social sustainability of logistics service providers. These classifications are society-, employee-, and management policy-related criteria. Society-related criteria are those that directly contribute to the well-being of societies, for instance, investigating the extent to which logistics service providers consider philanthropy or investment in a local community. Philanthropy is defined as "*a discretionary responsibility of a firm that involves choosing how it will voluntarily allocate resources to charitable or social service activities in order to reach marketing and other business related objectives for which there are no clear social expectations as to how the firm should perform*" (Ricks Jr, 2005, pp. 122). Employee-related criteria include average employee length of stay, minority hires, and average salary. Finally, organisational training processes, appraisal systems and development, human rights, and occupational health and safety are considered management policy-related social criteria.

In a comprehensive literature review, Selviaridis and Spring (2007) listed economic criteria, applied in the field, for choosing a logistics service provider; these include reliability, service quality, cost, financial stability, flexibility, and responsiveness to requests. These are general criteria used in the literature, while based on customers' needs, other criteria are added to the evaluation and selection process in some cases, and some criteria depend on the buying company's policy. For instance, some companies ask for references from those on the potential list of suppliers. References are companies the supplier has had previous collaboration with. There is no need to explain some of the criteria above, as it is clear which aspect of suppliers is evaluated by embedding

them into the evaluation process, e.g. service quality or financial stability. However, clarification about some of these criteria would be beneficial. Cost refers to the price of logistics services the supplier delivers to the buying company or the total cost of logistics outsourcing (Jharkharia & Shankar, 2007). The reliability of a logistics service provider refers to the reliability of different measures. It sometimes refers to the reliability of delivery and sometimes to the ability to perform the service agreed upon in the contract (Thai, 2013; So et al., 2006). Flexibility is another principal criterion buying companies use to evaluate the ability of a set of suppliers. Bowersox (1989) defined flexibility as a high level of responsiveness to customers' changing needs as well as responding quickly and efficiently to non-routine requests from customers. Responsiveness to request refers to suppliers' responsiveness after they receive a customer's request for quotation (Sivapornpunlerd & Setamanit, 2014).

### ***2.5. Modelling approaches for supplier evaluation***

As explained above, suppliers need to be evaluated and ranked based on some selected criteria using an appropriate approach. There are diverse decision-making approaches for ranking suppliers based on different preferences. Although absolute categorisation is not possible for these approaches, some researchers have classified these modelling approaches into various categorisations.

As one of the earliest attempts to review supplier selection approaches, Weber et al. (1991) identified and categorised supplier selection modelling approaches into three classifications: i) linear weighting models, ii) mathematical programming models, and iii) statistical/probabilistic approaches. In linear weighting models, the model assigns a weight to each criterion that is normally determined subjectively and aggregates the supplier's performance on each criterion based on the assigned weights. Finally, the suppliers are ranked based on the aggregated scores. In mathematical programming models, the supplier selection problem is formulated as a mathematical objective function that can be either maximised or minimised depending on managerial preferences. In addition, some constraints are embedded in the model to consider real-world limitations and increase the reliability of the results. The statistical/probabilistic approach considers the stochastic uncertainty that decision makers face in supplier selection; for instance, uncertainty in internal demand or order lead time may fluctuate or vary from one period to another.

De Boer et al. (2001) extended the above classification by Weber et al. (1991) by embedding the total cost of ownership and artificial intelligence-based approaches into the above categorisation. Total cost of ownership models deal with all quantifiable and related costs the company needs to pay for purchasing an item. In other words, the cost of goods/services is not limited to their final price; it also includes other types of cost, such as maintenance cost and spare parts cost, that are considered in the supplier evaluation. The second group of models is artificial intelligence-based models, which enable decision makers to extract invaluable information for a new supplier selection problem based on computer-aided systems. In these methods, the computer-aided systems are trained and developed by purchasing experts as well as by using historical data

from previous purchasing instances. This trained system can provide invaluable information to non-experts dealing with similar purchasing situations.

Chai et al. (2013) provided a new group of supplier evaluation techniques, known as a multi-criteria decision-making approach and stated that these techniques can be considered an independent group of models in addition to mathematical programming and artificial intelligence (AI) techniques. A multi-criteria decision-making approach provides a ranking for a set of alternatives based on several criteria chosen by experts. Although these models apply mathematical calculations in their algorithm, they cannot be considered mathematical programming because they do not constitute an objective function for the supplier selection problem.

As seen, the literature on methods for supplier evaluation is dominated by quantitative approaches. Although the share of qualitative approaches in the field is quite small, the contribution of these methods to the field needs to be asserted. These approaches have brought decision makers both advantages and disadvantages. The subjectivity of these methods has been found to be their main disadvantage, while their ability to consider unquantifiable criteria is considered one of their advantages.

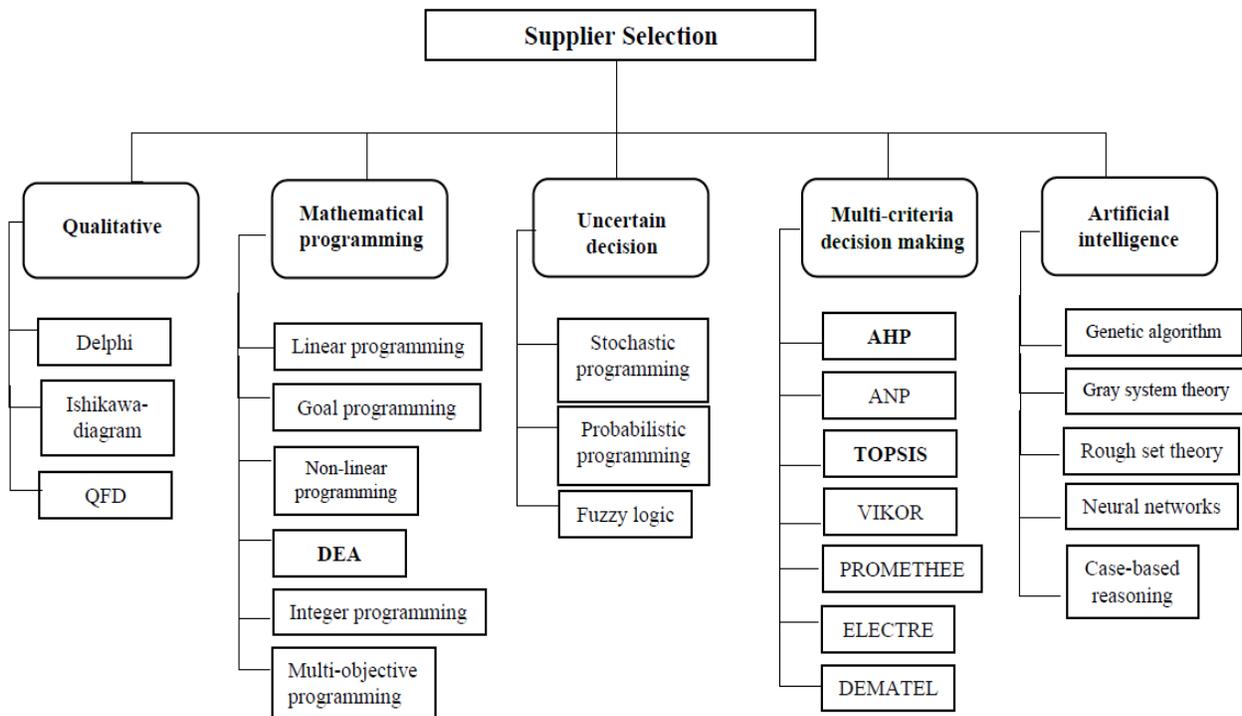
#### *2.5.1. Performance evaluation methods*

Sub-section 2.5 has highlighted that supplier selection approaches can be categorised into seven categorisations: 1) linear weighting, 2) mathematical programming, 3) statistical/probabilistic, 4) total cost of ownership, 5) artificial intelligence-based, 6) multi-criteria decision making, and 7) qualitative approaches. In this sub-section, the models under each of the above approaches will be discussed. Note that only the most-applied methods in the literature are provided in Figure 5. There is no common perspective among scholars for classifying supplier selection models tools into different approach categorisations (Chen, 2011; Bruno et al., 2012; Kannan et al., 2013; Brandenburg et al., 2014). Zimmer et al. (2016) proposed a diagram combining several classifications in the literature that is partially changed in this research (see Fig. 5). Although their diagram is open to dispute, it provides a general framework of most of the methods widely applied in the literature. For instance, they categorised Data Envelopment Analysis (DEA) as an analytical mathematical model, while some scholars consider it a mathematical programming technique (e.g. Chai et al., 2013). To avoid such ambiguities, Zimmer et al.'s (2016) diagram has been partially adapted to provide a more comprehensive framework for supplier evaluation methods. In the framework proposed for this research, DEA is classified under the mathematical programming categorisation, as it is a nonlinear-programming-based method. Second, the number of categorisations decreased from seven in Zimmer et al. (2016) to five herein. Linear weighting models are classified under the multi-criteria decision-making categorisation. This is also the case for the total cost of ownership models, with the difference being that it is difficult to embed these models in only one categorisation. Different models from different categorisations can calculate the total cost of ownership and subsequently rank some suppliers in a set under evaluation. Lastly, the statistical/probabilistic approach in Zimmer et al. (2016) is replaced by an uncertain decision

approach which, to the best of the author's knowledge, is a better label for this group of methods, as they all consider uncertainty in the decisional environment and are not necessarily a statistical or probabilistic method. The more significant methods in each approach category are depicted in Figure 5. The methods in bold format are the methods applied in this research, which will be explained in detail in the 'methodology' section.

#### *2.5.1.1. Qualitative methods*

As mentioned before, the supplier selection field is dominated by quantitative methods. However, it is important to note that the qualitative approach still needs to be considered one of the main approaches for supplier selection. The importance of this approach is increased when it can be seen that a significant number of scholars apply a qualitative approach to prioritise the most suitable criteria for their evaluation processes (Lee et al., 2009; Kuo et al., 2010). The Delphi method enables the analyst to collect a group of experts' opinions by utilising a series of questionnaires. In this method, feedback and information from the first questionnaire is sent to the respondents, and the process continues until the analyst gains some form of consensus in the respondents' answers (Dalkey & Helmer, 1963). The Ishikawa diagram (also known as a fishbone diagram) is another qualitative method applied for supplier selection (Ishikawa and Ishikawa, 1982). This diagram shows the main causes of a specific issue, for example, a defect in the production process or quality control. The major causes are depicted as bones in this diagram, and the sub-causes of each bone are also depicted. For instance, the Ishikawa diagram has been employed for evaluating suppliers' environmental performance (Enarsson, 1998). The last qualitative method in Figure 5 is quality function deployment (QFD), a method for considering customer preferences in the product development process (Cristiano et al., 2000). In this method, a matrix is defined which connects customers' requirements to the product engineering characteristics. Using this method, a quality matrix for purchased items is defined to satisfy the buying company's requirements. Then, the suppliers' assessment criteria are determined based on this matrix (Bevilacqua et al., 2006).



**Figure 5.** Supplier selection approaches and methods (Partially adapted from Zimmer et al., 2016, p. 1418)

### 2.5.1.2. Mathematical programming methods

Geoffrion (1976) stated that the main purpose of mathematical programming in strategic applications is to enable managers to embed insights into system behaviour. An optimal solution is achieved through a mathematical programming problem with given input data. A mathematical programming model generates a solution space called the feasible area, which is constructed when all constraints are satisfied (Zimmermann, 1983). Mathematical programming models can generally be categorised into two classifications: linear and nonlinear programming. Regardless of the linearity or nonlinearity of models, the mathematical programs can be formulated with different preferences. Goal programming was generated in the 1950s, when managers sought to satisfy several goals simultaneously. Thus, a set of goals based on selected attributes is defined, and the deviation variables are embedded in the model to evaluate alternative performances (Romero, 2014). Instead of several goals, two or more objective functions are formulated for a mathematical programming model. In a single-objective problem, an optimal solution in the feasible space which has the best value is determined by the model. However, in a multi-objective problem, one solution that optimizes one objective cannot guarantee optimisation in another objective. In fact, a trade-off needs to be found between several objectives (Weber & Ellram, 1993). Irrespective of the type of mathematical programming model, some or all variables can have integer values. This idea has constructed another widely applied trend in mathematical programming approaches called integer programming (Miller et al., 1960).

### *2.5.1.3. Uncertain decision methods*

In some real-world cases, the supplier selection decision needs to be made under uncertainty. To tackle this problem, a probability distribution is estimated to consider some unknown parameters. Stochastic programming is applied to estimate this probability distribution through modelling uncertainty optimisation (Chai et al., 2013). Stochastic or chance constraints programming can be categorised under mathematical programming, but here it is preferred to be in the uncertain situation approach. This method deals with a situation where the coefficients in formulation fluctuate rather than being constant (Kataoka, 1963). In some cases, the uncertainty is rooted in the subjectivity of expert opinions in the evaluation process. This can be handled by another method. This type of uncertainty has been extensively managed by the fuzzy logic theory proposed by Zadeh (1965), which is undoubtedly the most-used approach in the literature to tackle the problem of uncertainty in the supplier selection process (Govindan et al., 2015; Ghadimi et al., 2016).

### *2.5.1.4. Multi-criteria decision-making methods*

Multi-attribute utility theory was recognized by Holt (1998) as one of the popular approaches in the field of supplier selection. It is an extension of multi-attribute analysis, employing the concept of ‘utility’, which is the level of desirability or satisfaction. Some studies, such as Ho et al. (2010), state that all approaches described in Figure 5 are considered multi-criteria decision-making (MCDM). Of course, all of the above models can deal with multiple criteria/attributes. However, it is preferable to differentiate this approach from others in this study. This differentiation is in line with a literature review-based article published by Chai et al. (2013), where they also differentiate MCDM methods from other applied approaches in supplier selection.

Furthermore, a comprehensive categorisation of MCDM methods based on their inherent algorithms has been proposed by Chai et al. (2013). Analytical hierarchy process (AHP) and its extension analytical network process (ANP), proposed by Saaty (1987, 1996), are categorised as multi-attribute utility methods. The utility value generates a preference degree for each alternative/supplier that is the basis for ranking. The second categorisation is called outranking methods, including Elimination and Choice Expressing Reality (ELECTRE) proposed by Roy (1968) and Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE) proposed by Brans and Vincke (1985). Technique for Order Performance by Similarity to Ideal Solution (TOPSIS) proposed by Zelany (1974) and later developed by Hwang and Yoon (1981), and Multi-criteria Optimization and Compromise Solution (VIKOR), proposed by Duckstein and Opricovic (1980), are categorised into the compromise methods classification. These methods evaluate suppliers’ performance based on a closeness function to the ideal solution. The Decision-Making Trial and Evaluation Laboratory (DEMATEL) is another widely applied method that cannot fall into any of the above categorisations. It considers the influential interrelations among evaluation attributes.

#### 2.5.1.5. Artificial intelligence methods

The last category of approaches in the field of supplier selection is artificial intelligence (AI). This approach comes from the computer science domain and helps purchasing managers improve their supplier selection process. In one of the most recent studies on the implications of AI methods, Kaplan and Haenlein (2019, p. 15) defined AI as “*a system’s ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation*”. Based on this approach, machines can interpret and analyse data like humans to solve a problem or give invaluable input data to a decision-support system. AI methods are classified into three categorisations: i) analytical AI; ii) human-inspired AI; and iii) humanised AI. The first category applies cognitive intelligence, in which experience is used to support future decisions. In addition to cognitive intelligence, the second category also considers emotional intelligence, such as anger, surprise, etc. The third and most complex category brings social intelligence into the system in addition to cognitive and emotional intelligence. These methods are not available yet and will have a type of self-awareness while interacting with others (Kaplan & Haenlein, 2019). They also claim one common feature exists in AI methods, and that is learning from the past. Managers deal with three main situations when applying AI methods. In the first, both inputs and outputs are labelled, and the method only maps the given inputs to the given outputs. In the second, only the inputs are labelled. In this situation, the underlying structure from the data needs to be mapped by the method. In the third situation, an output variable that needs to be maximised is received by the system. Then, a series of decisions is made to influence the given output.



### 3. Methodology

---

*In this section, the research design, types of data applied, performance measurement tools for evaluating suppliers' performance, and finally, the validity and reliability of the results are discussed. Through the outline of the research design, the research approach implemented in this study is explained. Moreover, the main methods compared in this study, i.e. DEA, AHP, and TOPSIS, are explained more in detail. Finally, the research design is schematically depicted.*

#### **3.1. Research design**

Research design enables the researcher to achieve valid conclusions through a systematic research process (Brewer & Crano, 2000). Mitchell and Jolley (2012) stated that differences between science, pseudoscience, non-science, and nonsense can be distinguished through understanding research design. They claim that understanding research design is crucial for those who want to be able to read and evaluate research, be better thinkers, be scientifically literate, increase their marketability, and, most importantly, do their own research.

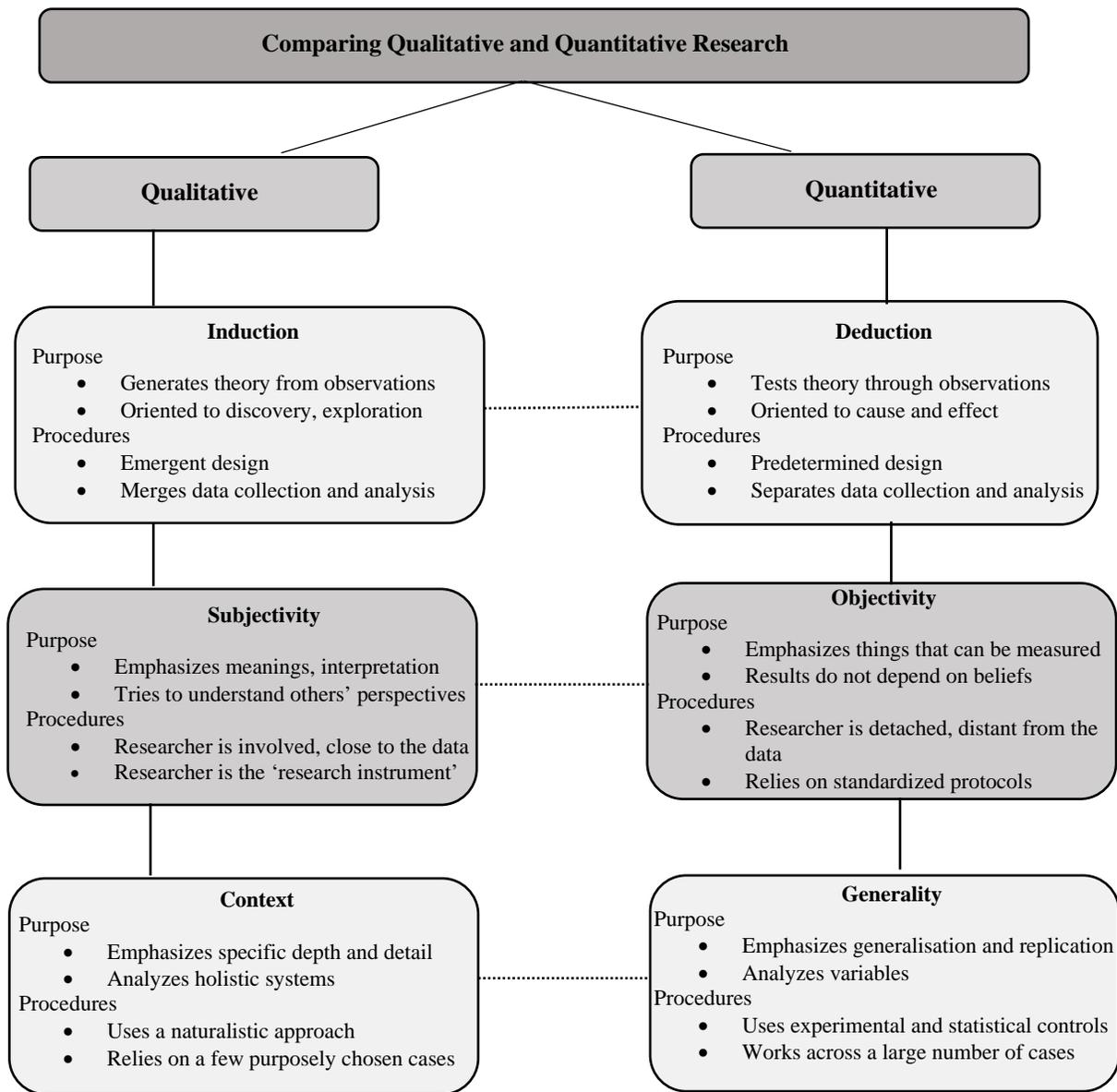
Having an appropriate research design, the analyst is able to differentiate original and second-hand scientific sources and analyse the conclusions independently. Apart from this, however, by accessing first-hand and reliable information, the analyst is able to evaluate the information achieved. Today, we are at risk of gaining misinformation that is widely available on the Internet, newspapers and television. Competency in evaluating research gives us the ability to take full advantage of reliable information. Moreover, a suitable research design enables the analyst to be a better thinker. It helps the researcher simultaneously develop critical thinking, problem-solving and decision-making skills. Becoming scientifically literate is considered another outcome of the research design, shedding light on the fact that it is crucial to understand how science works. All academics need to interpret scientific research if they want to make informed decisions. Exclusively relying on experts' opinions may lead to misinformation because these experts may belong to specific groups or organisations. Interestingly, when a researcher is able to understand the research design process, she/he makes herself/himself employable. Job-relevant information cannot be considered an inimitable ability because it can be acquired in different and easy ways such as the Internet or other open sources. Further, a competency in evaluating and differentiating good and bad research is not easily achieved, at least in a short period. Last but not least, another capability that can be attained through suitable research design is one's ability to conduct independent research. To achieve the above-mentioned abilities, a researcher needs to suitably construct and understand research design (Mitchell & Jolley, 2012).

### *3.1.1. Research method*

In explaining the research method, it is beneficial first to distinguish qualitative and quantitative methods. A differentiation between words and numbers is not a sensible logic to classify diverse methods into qualitative or quantitative categorisations. More specifically, the type of data cannot be the separational boundary between qualitative and quantitative methods because numerical data can be expressed as words, and vice versa. Morgan (2013) proposed a framework that more broadly differentiates qualitative from quantitative research (see Fig. 6).

As seen in Figure 6, Morgan (2013) compares qualitative and quantitative methods based on several distinctive key features. He claimed that ‘deduction’, ‘objectivity’, and ‘generality’ are the principal features for quantitative methods, while qualitative methods are ‘inductive’, ‘subjective’, and ‘contextual’. The differentiation between these two methods, based on Figure 6, is explained as follows.

Interrelations between theory and observations determine a research method’s inductivity or deductivity. In the inductive approach, theory is generated from observations. However, the opposite is true in the deductive approach, where the research begins with theories and hypotheses to be tested by observations (Dubois & Gadde, 2002; De Vaus & de Vaus, 2001). Moving from observations to theory is considered exploratory research, while the reverse movement links causes to effects. The procedures are also different in deductive and inductive research. In the inductive approach, data collection and analyses are combined with an ‘emergent’ design (Haverty, 2002), which means the research has evolved and shifted as it moves forward. In deductive research, however, the researcher is dealing with a predetermined design, and any deviation in the design is rare except for severe problems.



**Figure 6.** Comparing Qualitative and Quantitative Research (Morgan, 2013, p. 48)

Subjectivity in qualitative research applies to the interpretive researcher's role (Dixon-Woods et al., 2004). However, a quantitative method is objective because the researcher attempts to minimise his or her impact on the research (Yilmaz, 2013). In terms of procedures, in subjective methods, the researcher is close to the data and is a research instrument. The researcher is close to the data because he/she normally gathers the data. In addition, the researcher is an instrument because the researcher's beliefs and perspectives highly affect not only the research process but also the conclusions. On the other hand, a standardized protocol exists in objective research, which means if other researchers use the same process as you, then they achieve the same results. It is important to note that in the objective method, the researcher is distant from the data.

Based on the third distinctive feature in Figure 6, in qualitative research, a specific situation or set of people is monitored and analysed in depth; a holistic approach is required to consider as many variables as possible. However, in a quantitative approach, the emphasis is on generalisation and replication, as well as the analysis of variables. As analysis in the qualitative approach increases in detail and depth, a set of a few cases is chosen to achieve this purpose, while the quantitative approach highlights the importance of controlling ‘extraneous’ factors so the findings can be extended to other sets of people or settings. The above features highlight the contextual and generality features of qualitative and quantitative approaches.

To answer the research questions presented in the introduction, the performance of a set of suppliers needs to be evaluated. The evaluation outcomes then enable the analyst to rank the suppliers under evaluation. This implies that the performance measurement approach is the core component for answering the research questions. According to Neely et al. (1995, pp. 80), performance measurement is “*the process of quantifying the efficiency and effectiveness of action*”. Given this definition, through implementing a supplier evaluation process, suppliers’ performance is quantified based on a set of attributes/criteria/variables. One may discuss that this can also be implemented qualitatively, which would mean a set of suppliers is ranked based on subjective evaluation received from experts in the field. However, a big problem exists in qualitatively evaluating suppliers’ performance. Of course, experts can rank a set of suppliers; however, the differences between suppliers cannot be determined exactly and are subjective. The distance between the supplier in the third rank is difficult to distinguish from the supplier in the second rank, and the distance between the supplier in the second rank is difficult to distinguish from the first-ranked supplier. In today’s global business environment, with its extremely competitive market, minor differences matter and distinguish the performance of suppliers. The high discriminatory power in the essence of a quantitative approach enables analysts to more clearly observe the differences between suppliers and accordingly interpret such differences easier. Second, when the number of criteria or suppliers increases, it is difficult – if not impossible – to qualitatively evaluate the performance of a set of suppliers. While a few suppliers can be ranked qualitatively, to rank a larger set of suppliers (in this study, 19 logistics service providers), a quantitative method with high discriminatory power is needed.

To achieve reasonable answers for the research questions, the evaluation performance methods discussed in the literature need to be tested through empirical observation. In fact, an analytical approach is needed to investigate the extent to which the performance measurement approach helps buying companies efficiently evaluate suppliers. Indeed, the performance evaluation approach is not created based on the observations in this study; rather, an empirical application needs to be investigated to test the evaluation measurement approach. Furthermore, to gain reasonable answers for the questions, the interrelations between the algorithms of different methods and the rankings achieved for suppliers need to be addressed. Evaluating the performance of suppliers requires a separate data collection and analysis due to the number of criteria and suppliers, as discussed above. For the purpose of this research, the data had to be gathered and later entered into the models for

achieving the outcomes. This process cannot be performed simultaneously, as simultaneous evaluation requires an online evaluation system that is constantly fed with data. Thus, the quantitative approach was selected to gather and analyse the data separately. The quantitative method was also more helpful for the purpose of this research in terms of relying on standardized protocols and generality. It was necessary to use a protocol that is replicable and widely accepted by other scholars in the field. The quantitative approach was selected to evaluate the performance of logistics service providers, with some widely used methods that can be considered a standardized protocol and help other researchers replicate the results we achieved. Last but not least, it was necessary to evaluate a set of cases (suppliers), not just one or two. Ranking a few suppliers only cannot help answer the research questions because the ranking analysis was of interest.

### 3.2. Data

The vital role of data and the data-gathering process in any research is undeniable. Gathering the most appropriate data leads the researcher to applicable findings. To conduct this research, the required data were gathered in three different periods with three different approaches. For the purpose of *Paper I*, the data are mainly gathered from two well-known and international organisations, namely the OECD and the World Bank. In addition, some data were extracted from annual reports, titled “Energy Statistics for OECD Countries” provided by the International Energy Agency (IEA). These organisations provide diverse indicators of economic, environmental and social aspects for a wide range of countries. They gather information from member and non-member countries and publish annual reports in different disciplines. Accessibility to a wide range of data in different disciplines, in addition to well-established structures for measuring diverse indicators, increases the reliability of available data from these two databases. In the same vein, these datasets are continuously and dynamically updated, which increases the reliability of data. The organisations and reports provide different types of data, including time-series data, microdata, and geospatial data. There are several advantages to gathering the data from sources like the OECD and World Bank. First, the data are more reliable and valid, as they have been gathered by experts in international organisations. Second, the panel data for each criterion is available, enabling the researcher to compare performance in several consecutive periods. Third, in the event of missing data, the researcher can estimate a value for the missing data based on data available from the previous or next year. Figure 7 schematically describes the data sources for *Paper I*.

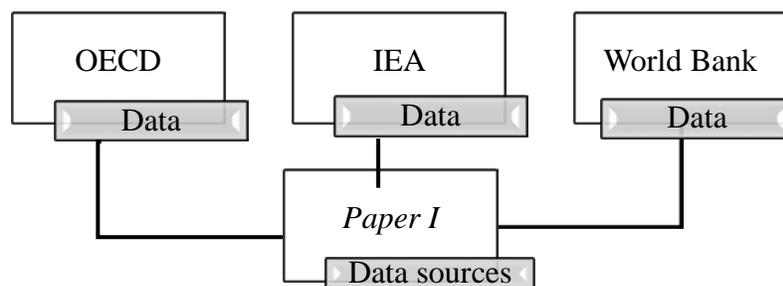
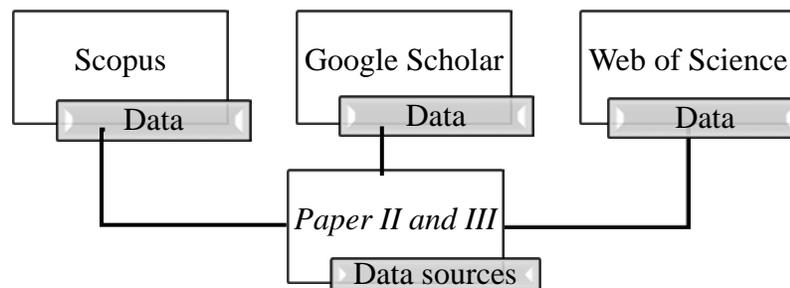


Figure 7. Data sources for *Paper I*

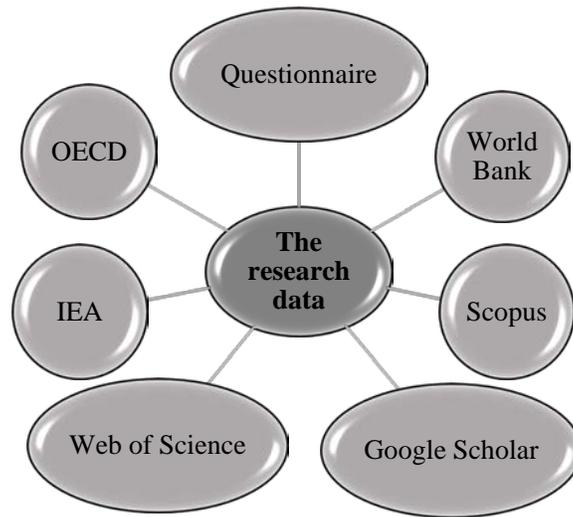
Owing to the different purposes of *Papers II and III* compared to *Paper I*, the type of data gathered is also different (see Fig. 8). For these two papers (literature review-based articles), the data were gathered from reliable search engines that play a vital role in today's research quality. For mapping the literature in the field, the data were gathered from Scopus, Web of Science, and Google Scholar. Of course, there are other search engines that can be used, but the three that were chosen have more unique features than the others. For example, the Google Scholar search engine encompasses a more extensive domain of scientific papers than any other search engine, while Scopus and Web of Science enable researchers to create a search profile in different formats with invaluable data for further bibliometric analysis.



**Figure 8.** Data sources for *Paper II* and *Paper III*

For *Papers IV* and *V*, the data were gathered using a survey questionnaire. Through a survey, a group of people is asked to answer questions in a specific field, based on their knowledge and opinions. For decades, the survey has been extensively used by researchers as a methodological approach for gathering data, and it is still widely applied by large organisations for data-gathering processes (Church & Waclawski, 2017). Like any other survey method, a questionnaire has advantages and disadvantages. The first advantage is its relatively low cost compared to other methods. In addition to a cost-saving advantage, a questionnaire enables researchers to gather a large amount of data in a short period (Lajunen & Summala, 2003). The researcher is able to simultaneously send the questionnaire to all respondents, which is impossible through a telephone or interview survey for a researcher. Time saving is another unique feature for this type of survey. For instance, in some cases a researcher needs less than one month from an initial mail-out to the final follow-up that needs to be sent to the sample of respondents. Furthermore, the respondents cannot be influenced by the researcher during the data-gathering process, unlike other methods such as interview, observation, etc. A questionnaire survey method also gives respondents enough time to answer questions, while in interviews, respondents do not have as much time to answer questions. Fife-Schaw (1995) stated that versatility can be considered another unique feature of the questionnaire, as it can be adapted to many diverse functions and disciplines. He also claimed that the data gathered through a questionnaire are of enough good quality for testing a hypothesis or advising on managerial implications. On the other hand, a questionnaire has disadvantages that a researcher needs to know about. When the respondent has any questions about the questionnaire, the researcher cannot help him/her, and the researcher cannot be sure that the questionnaire was

completed by the person who was supposed to do it (Marshall, 2005). In this research, the questionnaire method was selected because a large sample of respondents (practitioners active in procuring logistics services) was targeted for gathering data. It also enabled the author to gather the required data in a short period (almost one month). All data sources used in this research are depicted in Figure 9.



**Figure 9.** All data sources

### ***3.3. Measurement methods***

To achieve the main objective of this research, three main methods have been applied in the papers, including DEA, AHP, and TOPSIS (these methods are widely applied in sustainable supplier selection, based on the findings of *Paper II* and *Paper III*). They are briefly explained in this sub-section.

#### ***3.3.1. TOPSIS***

TOPSIS is categorised as one of the MCDM methods for choosing the best alternative among a set of alternatives. TOPSIS determines the best alternative based on its distance from the ideal point. The shorter the distance, the better the alternative (Zelany, 1974). Hwang and Yoon (1981) further investigated this distance from the ideal solution; they claimed the best alternative is determined based on the shortest distance from the positive solution and the farthest distance from the negative solution. A combination of both distances is considered to select the best alternative. Given this feature, the best and the worst alternatives are recognized simultaneously. There are six main steps in the TOPSIS procedure, according to Opricovic and Tzeng (2004). First, the normalised decision matrix needs to be calculated based on the data gathered. Second, the weights of criteria are multiplied by the normalised decision matrix for calculating the weighted normalised decision matrix. Third, based on the type of criteria – benefit or cost criteria – positive and negative

ideal solutions are determined. Fourth, using the n-dimensional Euclidean distance, the separation measure is calculated that represents each alternative's distance from the ideal solution. Fifth, given the separation measure, the relative closeness from the ideal solution is calculated for each alternative. In the last step, the alternatives are ranked based on their relative closeness as calculated in step 5.

### 3.3.2. *AHP*

AHP is another MCDM method widely applied in management science. This method was initially proposed by Saaty (1987) as a general theory of measurement. AHP enables analysts to deal with both tangible (objective) and intangible (subjective) ideas where both physical and psychological events are the core interest. Saaty described AHP as modelling the problem into a network structure where pairwise comparisons are being made to create the relations within this network. A priority score is calculated for each alternative based on the comparison matrix, and then all alternatives are ranked based on the scores calculated (Saaty, 2008). AHP simultaneously considers both decision makers' intuitive judgement and consistency of the comparison matrix provided by the decision makers (Saaty, 1987). Given the decision makers' intuitive judgement, AHP embeds the decision makers' behaviour – which evolves around their knowledge and experience – in its algorithm (Al-Harbi, 2001). The consistency ratio can be considered one of the admirable strengths of AHP, in which the consistency of the decision makers' judgement matrixes is calculated to check whether the comparisons between different alternatives follow a rational framework. This ratio is set to be equal or less than 10 percent. If the restriction for the consistency ratio is not satisfied, then the analyst needs to ask the decision makers to provide another new comparison judgement matrix or improve the inconsistency of the previous matrixes using different methods provided in the literature (e.g. Ergu et al. 2011).

### 3.3.3. *DEA*

Farrell (1957) initiated the beginnings of DEA, introducing non-parametric methods for removing ties in parametric techniques. Charnes et al. (1978) were inspired by Farrell's (1957) findings and proposed DEA method. In DEA, the fundamental functional relationship between inputs and outputs is not determined before the calculation (Seiford and Thrall, 1990). DEA is a method for evaluating the relative efficiency of decision-making units based on a mathematical modelling approach. Decision-making units can be companies, organisations, hospitals, universities, airports, that perform the same function under similar conditions to produce multiple outputs by consuming multiple inputs. Another strong point of DEA is that it needs no subjective weights from decision makers because the most appropriate weights are allocated to each input and output by the DEA model itself. Production economics helps managers with choices among several alternative production processes or technologies, resource allocation, outputs or enterprise selection. In fact, production economics enables managers to know how much and what to produce and what combination should be between inputs (resources) and outputs (final product). This can be generalised to any level, such as firm level, industry level, or even society level (Beattie et al.,

2009). In production economics, index number methods evaluate the productivity of units when it is assumed that all units are fully efficient. When this assumption is relaxed, those methods are practical which estimate frontier functions and evaluate efficiencies of units relative to this estimated efficiency frontier (Coelli et al., 2005). DEA is one of the methods that calculates units' efficiency, which refers to a comparison between observed and optimal values of input(s) and output(s) (Lovell, 1993). Many methods have been applied for frontier estimation; DEA is one that has been widely applied for the supplier selection problem (Coelli et al., 2005).

Figure 10 shows a holistic view of the research design, including the main objective, research method, research data, and measurement tools applied.

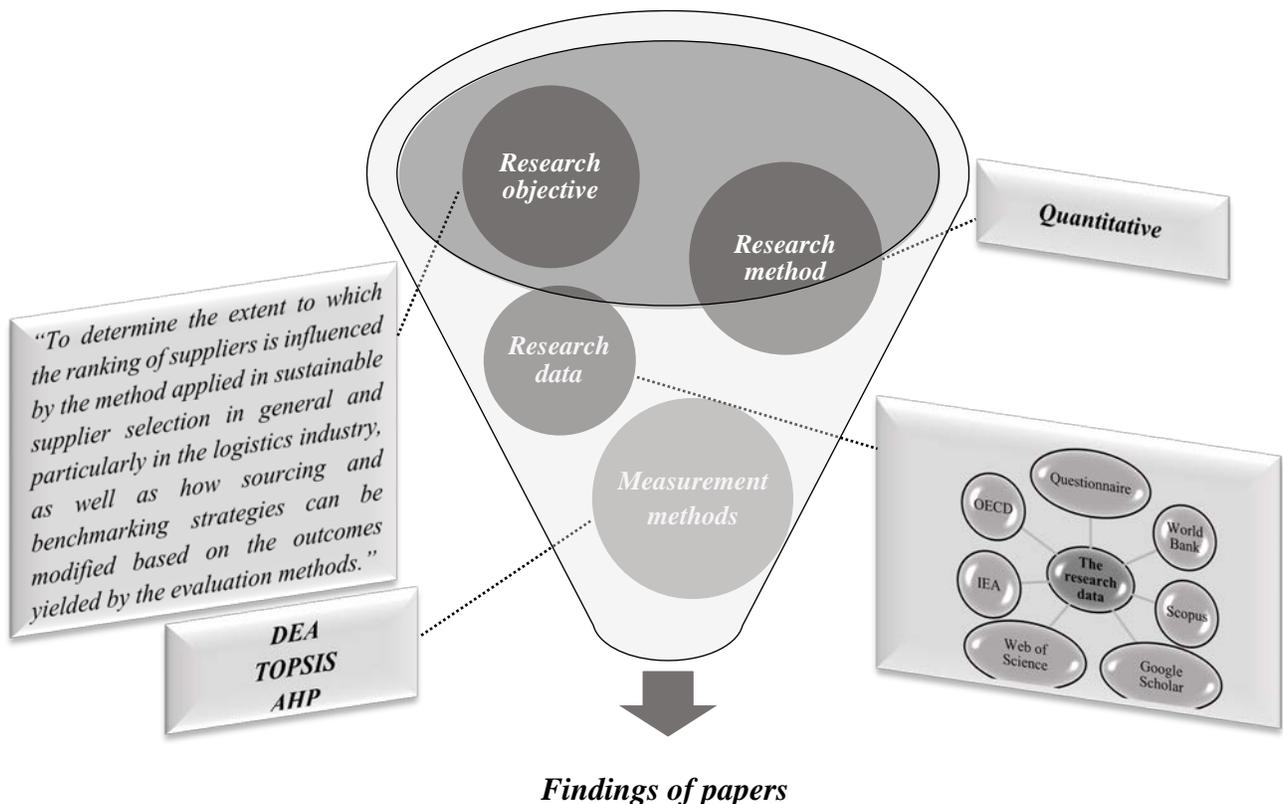


Figure 10. Conceptual framework of research design

### 3.4. Validity and reliability

Validity and reliability are inseparable parts of research. Each research needs to be valid, which means the extent to which it measures what it was intended to investigate (Chapelle, 1999). Reliability is different from validity and refers to the replicability of research (Daly et al., 2000); it reviews the assumption of whether the same results are achieved under the same condition. According to Varkevisser et al. (1991), there is no relationship between validity and reliability in research. Research can be valid but not reliable, and vice versa. The best scenario, of course, can occur when research is both reliable and valid. These two crucial concepts are reviewed for the

present research in this sub-section. According to Onwuegbuzie (2000), many researchers avoid discussing validity and reliability in their research, because they assume it would not be beneficial to discuss the weaknesses of their research. However, it is important to note that every research is at risk of being influenced by threats of internal and external validity and/or reliability. More specifically, some threats cannot be eliminated by the researcher. Thus, the threats to internal and external validity need to be discussed for three reasons: 1) when the reader is aware of different sources of invalidity, he/she can appropriately place the research findings in the proper context; 2) identifying these threats gives future direction to other researchers; and 3) when discussion of these threats becomes common among researchers, then it becomes possible to extract the most prevalent threats by conducting an analysis.

#### 3.4.1. Validity

According to Fitzner (2007) and Heale and Twycross (2015), validity needs to be discussed to check the internal and external validity of a research through *face validity*, *content validity*, *construct validity*, and *criterion validity*.

Face validity answers this question: Is the tool/instrument/measurement a good measure or not? The questionnaire used in this research was tested before sending it to the respondents. Experts in the field were asked to check whether the questionnaire contained any confusing phrasing or concepts. The structure of the questionnaire was checked with the researcher's main supervisor and accordingly revised in several consecutive steps. The questionnaire's structure was also compared with some similarly applied questionnaires in other research in the field. Content validity investigates the question: To what extent do the dimensions (variables and questions) in the measurement (questionnaire) appropriately measure the construct (sustainability of suppliers)? The variables used in the questionnaire were selected based on the comprehensive content analysis performed in *Paper II* and *Paper III*. Also, the selected variables were the most-used variables in other articles published in the field of sustainable supplier selection. Construct validity explores the extent to which a measurement (questionnaire) measures the intended construct (sustainability of suppliers). To check this type of validity, both *homogeneity* and *convergence* need to be investigated. In terms of homogeneity, when a measurement measures only one construct, it means the research has homogeneity evidence. This is the case in this study's questionnaire, as it only measures the sustainability performance of logistics service providers. In terms of convergence, research contains such evidence when the applied measurement outcomes are identical to other measurements. As discussed earlier, only one measurement (i.e. questionnaire) was used to gather the data, so in this study it is not possible to compare the outcomes of the applied measurement with another one. The criterion validity investigates the correlation calculated between two measurements, and this type of validity cannot be checked either herein, as only one measurement (i.e. questionnaire) was used to gather the data from the respondents.

### 3.4.2. *Reliability*

The consistency of the results achieved from research over time is considered reliability (Joppe, 2000). In another definition, Kirk et al. (1986) defined reliability as the reproducibility of results yielded by a measurement, regardless of the place and time it is carried out. Thus, it can be concluded that reliability evaluates the research to answer the question regarding the extent to which the research findings are independent of accidental circumstances. The reliability of this study needs to be discussed from two perspectives. First, based on the comparisons performed between the three selected methods (i.e. DEA, TOPSIS and AHP), this study's findings can be generalised to other methods discussed in the literature. This means that if other researchers use the same data set gathered in this research with other methods, they may observe some differences between the rankings of suppliers under evaluation. Because, as will be discussed later in the findings and conclusion section, only a perfect correlation value of 1 – which is rare – between the rankings of suppliers produces the same outcome for decision-making process. Any difference, even a minor one, can change the decisions made based on the outcomes of evaluation methods. Second, in terms of the reproducibility of the results yielded from the questionnaire, the results of this research are subject to some changes. If other researchers send the same questionnaire to the same participants or other recipients with similar expertise and experience, the ranking of suppliers can differ from the rankings yielded in this research due to the subjectiveness of experts' evaluation. However, this cannot decrease the reliability of this study because the results from this research are considered a counterexample for a situation in which researchers find similar rankings for a set of suppliers based on different evaluation methods.



## 4. Findings

---

*This section provides the findings of the papers. Explaining the findings enables a researcher to prepare the reader for the next section, which is the conclusion wherein the research questions are answered. It is significant to note that the findings of each paper are classified into two categorisations. One of those findings directly related to the main objective is discussed in the introduction. The others are those that are only related to the specific topic in each paper and may be unrelated if they are explained here. So, only findings that contribute to the main objective of this research are discussed, and not those that only contribute to each specific paper.*

### 4.1. Paper I

*“Evaluating the sustainability of national logistics performance using Data Envelopment Analysis”*

The results of this paper imply that functional, infrastructural, and technological competencies in the logistics industry cannot guarantee the sustainability of operational logistics performance in that industry. Correspondingly, the opposite is not true. In other words, a country can be among the top performers based on functional, infrastructural, and technological performance but ranked as a poor performer in terms of operational logistics sustainability performance. This negative correlation implies that an index representing the performance of a country’s logistics industry needs to consider some sustainability criteria in its calculation process rather than only technological, infrastructural, and functional competencies. This is in line with Vesilind et al. (2010), who stated that technology can help managers in sustainable development, but it is not the only requirement for being sustainable. A higher level of technology does not necessarily result in higher performance in sustainability. The technological and sustainability competencies cannot be replaced by each other; rather, they play complementary roles. It means all types of competencies need to be combined in a unified framework to measure the performance level of the logistics industry.

Given an improvement matrix consisting of all functional, infrastructural, and technological competencies, as well as all sustainable operational logistics competencies, diverse potential policies may be applied to the logistics industry for different purposes at the country level. The proposed policies encompass promoting economic development through greater international trade, rechannelling finance from the passenger to the freight sector, restructuring energy use templates to manage greenhouse gas emissions, or initiating recruitment programs to increase the employment rate at the national level. Regardless of the type of improvement policy each country selects, it is crucial to find a balance between logistics performance competencies and sustainability performance improvements.

#### **4.2. Paper II and Paper III**

*“Techniques Applied for the Selection of Sustainable Suppliers: A Systematic Review of the Literature”*

*“Applying the Triple Bottom Line in Supplier Selection: A Meta-Review of the State of the Art”*

The findings of these two literature-review articles demonstrate a significant increase in the number of published papers in the field of sustainable supplier selection from 2009 to 2018. This increasing trend supports the fact that supplier sustainability has become one of the main concerns for researchers in the field. From around four to five papers each year, since 2015, the number of published papers has increased to at least ten papers per annum.

The implemented analysis of case studies revealed two noteworthy findings. First, the countries of origin for applied case studies shows that researchers in both developed and developing countries aim to shed light on different aspects of the sustainable supplier selection process. It is delightful that developing countries, similar to developed ones, at least in academia, are trying to raise awareness of sustainable procurement among manufacturing companies. Such an increasing awareness of sustainable supplier selection and hence, sustainable procurement at the company or micro level, will result in a better sustainability performance at an aggregated or macro level. Second, some industries, such as automotive and electronics, dominate the other industries investigated in the field, including printing, oil, resin, packing, textile, and food industries. On the other hand, the analysis demonstrates that some industries, such as the logistics industry, have been widely neglected, despite the need for further improvements in occupational health and safety as well as decreasing CO<sub>2</sub> emissions in that industry.

Furthermore, the applied methods analysis shows that DEA, TOPSIS, and AHP are among the most widely applied methods in the literature for dealing with the sustainable supplier selection problem. These methods have been combined together or with other methods for proposing a new framework for handling the sustainable supplier selection process more efficiently. However, the applied methods analysis shows there is a lack of studies to compare the results yielded from applying diverse methods – based on a common data set – to investigate potential differences or similarities between these methods.

Moreover, the criteria analysis demonstrates that from the total number of criteria applied (1,219) in the final 66 analysed papers, roughly 24 percent are social criteria, while this value is 40 percent and 36 percent for economic and environmental criteria, respectively. In a similar case, in most papers, researchers consider more economic than environmental criteria and more environmental than social criteria in their evaluation processes. This difference between the number of criteria of triple-bottom-line in sustainability for evaluating suppliers' sustainability performance is questionable. Such a difference can cause bias in the evaluation process. For example, when the analyst selects more economic criteria, it means the analyst assigns more

importance to the economic criteria before even starting the data-gathering process. In such a situation, suppliers with better performance in economic criteria and inferior performance in social criteria are prioritised over suppliers with better performance in social criteria and inferior performance in economic criteria. The same scenario exists between environmental and social criteria. To tackle this problem, it is highly recommended to consider an equal number of criteria for each economic, environmental, and social aspect.

#### **4.3. Paper IV**

*“A comparison of fuzzy DEA and fuzzy TOPSIS in sustainable supplier selection:  
Implications for sourcing strategy”*

This paper has implemented a comparative analysis between TOPSIS and DEA in a fuzzy environment. The rankings achieved for suppliers (logistics service providers under evaluation) based on DEA and TOPSIS techniques reveals the presence of displacement in rankings. Based on the suppliers' ranking yielded by each model, Spearman's and Kendall's Tau rank correlation coefficients were calculated, and are 0.679 and 0.520, respectively, which are significant at the 0.01 level. This implies there is a positive rank correlation based on the outcomes of these two models, but it is not a strong correlation. Only two logistics service providers out of 19 gained the same ranking based on both applied techniques, i.e. roughly 0.11 percent. Despite the results of the rank correlation coefficients indicating a positive monotonic correlation between suppliers' ranking, some problems are due to the suppliers' ranking displacement.

One of these problems is the sourcing problem described earlier (Sub-section 1.2.1). The outcomes yielded imply that making a final decision for a sourcing strategy based on the outcomes of these two techniques may be problematic. Although in the empirical case in the paper, both techniques recognized one common supplier as the best one, the high percentage of suppliers' ranking displacement raises question. The presence of the rank-reversal phenomenon in the outcomes of these methods strengthens the claim that, in another application, the best supplier will be different and not the same based on each method. When this is the case, a buying firm will fail to implement a single sourcing strategy exclusively based on the outcomes of these techniques. In the case of multiple sourcing, the results show even more severe problems. The second-ranked supplier is different based on each method, and this continues for the third and lower rankings. In fact, there is a general inconsistency in the suppliers' rankings, which makes it impossible to select the top suppliers for multiple sourcing. This inconsistency implies that suppliers' rankings are greatly influenced by method selection and not their competencies and performance. Although a final decision for supplier selection exclusively based on the outcomes of the methods may be impossible, by applying DEA and TOPSIS, the outcomes of these techniques can support the final decision. In the case of single sourcing, the number of potential suppliers can be reduced to two (each model ranks one supplier as the best), which helps decision makers focus only on these two suppliers for further analysis. This policy is also practical for multiple sourcing, in which a long list of potential suppliers is replaced with a short list of suppliers based on the suppliers' ranking

each model achieves. Then these suppliers are further analysed for the final decision. Basing the final decision on the techniques' outcomes is problematic, but the outcomes can be used as a source of information to support the final decision.

Both methods were also analysed based on sensitivity to the number of suppliers under evaluation. Over time, the buying company may terminate its relationship with a supplier and/or establish a strategic relationship with a new supplier. Whatever the case, the methods must generate a consistent ranking of suppliers when the input and output values are kept unchanged. The consistency of the methods' outcomes was analysed based on two scenarios in which one supplier is excluded from, or included in, the evaluation process. The results reveal that fuzzy DEA is sensitive to the changes in the number of suppliers under evaluation, while fuzzy TOPSIS produces the same supplier rankings in both scenarios.

#### **4.4. Paper V**

*“AHP versus DEA for the gradual improvement of unsustainable suppliers:  
A comparative analysis”*

Based on the suppliers' rankings achieved from AHP and DEA methods, the Spearman's and Kendall's Tau rank correlation coefficient values are 0.961 and 0.871, both at the significance level 0.01, respectively. Despite this high value for the rank correlation coefficient between these two methods, the findings demonstrate that implementing a benchmarking strategy based on the outcomes of these methods may be challenging in some cases. Clustering the suppliers based on the similarities between their inputs, outputs and efficiency scores demonstrates that all suppliers are categorised in the same clusters based on both AHP and DEA scores. However, it is not possible to choose only one supplier as the worst performer for improvement programs based on the rankings yielded. AHP and DEA determine two different suppliers as the worst performers. Furthermore, the impossibility of determining only one best supplier in some clusters is the other side of the same coin. This difference in benchmarking strategy becomes more tangible when the suppliers are categorised in more clusters. Categorising suppliers in four clusters exhibits less inconsistency between the benchmarking strategies provided by each method, than does categorising suppliers in six or eight clusters. Each method provides a unique path of gradual improvement for the supplier under the improvement process, which makes decision-making more difficult in supplier management. Generally, AHP provides a benchmarking strategy with more improvement steps and less sudden large changes in supplier performance compared to DEA. Each type of benchmarking strategy for gradual improvement can be practical for different types of suppliers in supply chains. DEA offers a better improvement path for suppliers that are able access different types of resources required for implementing a faster improvement process, while AHP provides a gradual improvement process in more steps for suppliers that have limited access to the required resources and need more time for improving their performance.

## 5. Conclusion

---

*This section concludes this study. In doing so, first, the research questions are answered based on the findings of the papers. Second, research directions that can contribute to the findings of this research in the future are provided.*

### **5.1. Research question 1**

RQ1: Why does sustainability performance need to be considered more in the logistics industry?

What happens at the macro level (i.e. country level) is a consequence of the interrelationships/interactions between the main key players at the micro level in each industry. The level and type of interactions between the key players in the logistics industry can be mapped as macro level logistics performance. In fact, the outcome of a macro-level analysis in the logistics industry is not something other than the performance of key companies such as forwards, transporters, distributors, shippers, and other types of logistics service providers. The evaluation of the sustainability of operational performance in the logistics industry at a macro level revealed an essential need to further study this field of research. The findings imply that creating a network of technological achievements in combination with a high level of investments in logistical facilities does not necessarily bring sustainability to the logistics industry. Regardless of the rank of each country, whether top or poor performer, sustainability performance needs to be integrated – as a determinant measurement – in the evaluation of countries' logistics competencies. Although the poor performers need to act at once to improve their sustainability performance, it does not mean there is no need for top performers to improve their performance as well. Being that sustainability efficiency is an unbounded approach, and there is no best perfect performance of it, the top performers also need to continuously enhance their capabilities in terms of sustainability.

One possible way to improve a nation's sustainability in its logistics industry is to improve the performance of core players in that industry. One of the main groups of players in each logistics industry is the group of logistics service providers. These players' sustainability performance needs to be studied more by scholars in general, as this continuous monitoring can improve their performance at the micro level and thus contribute to national logistics performance at the macro level. A large number of manufacturers, distributors, and retailers work with one or more than one logistics service provider in their supply chain operations. Conducting this research with an empirical application for measuring the sustainability of a set of logistics service providers has yielded more tangible findings for buying companies wanting to apply the methods developed by academics for handling an appropriate sustainable supplier selection process.

## **5.2. Research question 2**

RQ 2: How has the sustainable supplier selection process been studied in the literature?

The findings illustrate that the number of papers published in this field has increased since 2015, a trend that will probably continue in future years due to the high importance of sustainability in both local and global supply chains. More specifically, due to the importance of sustainability performance for companies, academics accordingly are studying sustainable supplier selection instead of the traditional supplier selection process. In other words, academics are trying to show the applicability of their research for real-world situations in the context of sustainable supplier selection. Therefore, supplier selection in its traditional form is expected to be completely replaced by sustainable supplier selection in the near future. It does not make sense to only consider economic performance in choosing suppliers when the importance of considering sustainability aspects is being widely discussed.

The two literature-review papers in this study demonstrate that, in most cases, researchers have proposed a new standalone method for measuring suppliers' sustainability and provided a hybrid method that combines two or more than two methods in a unified framework. This implies that most of the contributions revolve around methods development rather than practical managerial implications for real-world cases. In a similar vein, it can be seen that most research has focused on different types of data gathering, involving different types of criteria in the evaluation process, considering uncertain evaluation environments, and sometimes just simple application of previously discussed methods proposed in a new context that can be any industry. For instance, there is a need to determine why, in spite of numerous investigations implemented in academia, many companies still do not use the methods proposed in the literature to select a supplier for buying goods or services or developing new frameworks by combining two or more methods. It would be beneficial to investigate the extent to which these methods can help managers in real-world situations and the extent to which these methods effectively feed the decision-making support system with reliable information and trustworthy ranking of suppliers.

## **5.3. Research question 3**

RQ 3: To what extent are the results achieved from different evaluation methods consistent?

The comparisons between the three methods selected for this study (i.e. DEA, AHP, and TOPSIS) reveal some differences between the outcomes yielded based on each method that can affect the reliability and applicability of these outcomes in a real-world situation. Note that although two methods were compared in each paper, the outcomes achieved in *Paper IV* and *Paper V* can also be compared, as the same data set was applied in both papers. Comparisons between the results show there is no perfect correlation between the rankings of suppliers achieved by each method. The findings of *Paper IV* demonstrate the rank reversal between the outcomes of TOPSIS and DEA methods. In the same vein, the findings of *Paper V* confirm the existence of rank-reversal phenomena between the outcomes of AHP and DEA methods. Likewise, comparing the rankings

of suppliers in both papers – i.e. *Paper IV* and *Paper V* – the rank displacement can be seen between the outcomes of TOPSIS and AHP methods. This contradicts the research implemented by Junior et al. (2014), which found a perfect rank correlation coefficient of five suppliers evaluated by fuzzy AHP and fuzzy TOPSIS. However, their findings can be questioned in two ways. First, they only compared AHP and TOPSIS under a fuzzy environment and not under any other type of data. Second, due to the small number of suppliers under evaluation, their findings are questionable. The probability of differences in supplier rankings can be increased by enlarging the number of suppliers under evaluation. In addition to the differences between the methods in the ranking of suppliers, a comparison between the rankings of suppliers based on the DEA technique in *Paper IV* and *Paper V* proves that there is not even a full correlation in supplier rankings using one method, but in a different calculation environment. The ranking of suppliers calculated by a fuzzy DEA model in *Paper IV* differs from the ranking of suppliers calculated by a non-fuzzy DEA model in *Paper V*. This means that not only the evaluation method, but also the calculation environment, whether there is crisp, interval, or fuzzy data, can influence the supplier rankings. This means the suppliers' ranking is sensitive to both the methods selected and the type of data used in the calculation process.

It has been discussed above that supplier rankings are not exclusively an outcome of their performance based on some selected criteria but rather partially influenced by the evaluation methods' algorithm. Despite the different outcomes between some of the widely applied methods in the literature, TOPSIS, AHP, and DEA are still the practical and useful for decision-support systems in supplier management than a subjective and/or unstructured evaluation process. However, it is problematic and challenging when managers or practitioners rely on only one method for evaluating suppliers' sustainability. Due to the differences discovered above between the outcomes of the methods, it is impossible to apply one method and select the best or worst supplier(s) based on the outcomes of the method selected. Of course, there is a positive rank correlation coefficient between the selected methods, but there is a general inconsistency in the supplier rankings. TOPSIS, DEA, and AHP are useful and practical methods, at least based on the literature, for evaluating suppliers' sustainability and providing support information to decision makers. However, decision makers need to further analyse suppliers' performance based on the primary outcomes achieved by applying these method, which can reduce the number of potential suppliers. For instance, a pool of the top ten suppliers based on the outcomes of each method can be selected for further analysis. Then, the buying company can evaluate the suppliers' performance by requesting more information or inspecting the suppliers' facilities.

#### **5.4. Research question 4**

RQ 4: Why do sourcing and benchmarking decisions change based on the type of method selected?

Categorising suppliers in different groups, sourcing decisions, creating or terminating strategic relationships based on the performance of suppliers, and benchmarking decisions are only a few strategic tasks in supplier management practice. Based on the findings, these strategic decisions – at least the sourcing and benchmarking decisions discussed in this study – are directly influenced and changed based on the method selected and the type of data applied in the sustainable supplier evaluation process. It has been discussed that only in case of positive and perfect rank correlation coefficient between the outcomes of two methods (i.e. value of 1) does the decision made by decision makers remain unchanged. Otherwise, with a rank correlation coefficient value less than 1, the sourcing and benchmarking decisions based on the suppliers' rankings can be modified. Thus, it is concluded that despite a positive correlation coefficient between the outcomes of the methods, reaching a unique decision would be difficult – if not impossible. However, it is crucial for researchers and especially managers in industries to be aware of the fact that the key strategies in procurement and supplier management (e.g. sourcing and benchmarking strategies) are influenced by the outcomes of each method and not only by the performance of suppliers. It means that making decisions on the final outsourcing and benchmarking policies based only on the outcomes of these methods can be considered an undesirable decision. Although the methods' outcomes provide different sourcing and benchmarking policies, these outcomes can help managers during the initial evaluation of suppliers.

Through an empirical application, the case study of logistics service providers proposed in *Paper IV* and *Paper V*, the difficulties surrounding sourcing and benchmarking strategies were discussed based on evaluating the sustainability performance of a set of logistics service providers. As many companies outsource their logistics services to one or several external parties known as logistics service providers, they need to handle the selection process carefully. Buying companies can use the methods discussed in this research to quantitatively evaluate their suppliers' sustainability competencies, considering the challenges discussed. Enhancing the effectiveness of an appropriate supplier management system in procuring logistics services enables buying companies to systematically evaluate the sustainability of their suppliers and continuously monitor and improve their performance. Although the outcomes of the methods are not fully correlated, applying different methods enables managers to create a network of logistics service providers that performs better than their competitors based on the outcomes of all methods. Given this network, the best supplier can be selected based on sustainability performance, but neither based exclusively on the outcomes of a single method, nor on a subjective framework.

### 5.5. Future research directions

More research at the macro level in the logistics industry considering sustainability aspects is beneficial for discovering countries' sustainability performance. Evaluation of sustainable operational logistics performance of countries in the first paper can be considered a starting point for considering sustainability performance in countries' logistics competency, calculated every other year by World Bank Group. To do so, a meta-dynamic analysis considering the data from several consecutive periods is beneficial to investigate the extent to which the result yielded in our paper is reliable and valid. A comparison between different periods reveals invaluable information for both policymakers and key players in the industry. In a similar vein, it is advantageous to embed more criteria in the evaluation process to help increase the discriminatory power of the method selected. Then, it is easier to discover the differences between the countries under evaluation. Furthermore, the claim that technological and infrastructural competencies in the logistics industry cannot guarantee the sustainability performance in that respective industry can be investigated to determine if it is true.

The analysis conducted in the two literature-review papers in this research has revealed some potential future research directions in the field of sustainable supplier selection that are briefly summarized here. First, the criteria analysis shows that the number of criteria considered in contemporary supplier evaluation (sustainable supplier selection) has slightly increased compared to traditional supplier evaluation. Traditionally, decision makers considered very few attributes for evaluating the performance of a set of suppliers under evaluation. The criteria analysis indicates that, in most cases, the number of criteria applied is less than ten. However, due to the computerisation of evaluation processes and access to a wide range of diverse software applications, it is easy to consider more criteria for measuring suppliers' competencies. Considering only a few sets of variables makes the evaluation process biased, in which the only suppliers selected are those that have acceptable performance based on these criteria, while they could be recognized as average performers if more criteria are considered in the evaluation process. Second, the comparison of other methods discussed by researchers in the literature – listed in Section 2 – can also be useful for academics and practitioners. More comparison between methods enables researchers to find out whether rank-reversal phenomena exist between all methods or the outcomes of some methods are perfectly correlated. Furthermore, some methods, such as stochastic analysis frontier (SFA), which has been used in many industries, can be applied more in the field and compared with other widely used methods for sustainable supplier selection. Third, in addition to comparison between the outcomes of methods, the sensitivity to the number of suppliers under evaluation analysis conducted in *Paper IV* needs to be implemented for other methods to investigate which method other than DEA is sensitive to the number of suppliers under evaluation. This type of analysis can reveal invaluable information for decision makers, as they will know to what extent the ranking of suppliers is changed or influenced when a supplier is added to, or excluded from, the evaluation process. Fourth, a misalignment is seen between the triple bottom lines of sustainability based on the criteria analysis performed in *Paper III*. Although the co-citation

analysis confirms that economic, environmental, and social aspects have been appropriately covered in the field, a misalignment between these three aspects still exists. This analysis shows that economic criteria are considered more in the evaluation process compared with environmental criteria; however, this difference is negligible. Oppositely, the difference between economic and social aspects is more tangible, where the analysis shows that considering economic criteria is far more than social variables in evaluating suppliers' sustainability performance. One of the main reasons behind this, in the author's opinion, is the existence of a well-defined justification for considering environmental variables based on their undeniable effects on economic efficiency, particularly cost reduction and quality enhancement. Thus, it is advantageous to conduct more research and investigate how companies can benefit from social sustainability performance based on increasing market share, cost reduction, or quality improvements.

Sourcing and benchmarking challenges in supplier management systems were discussed in *Paper IV* and *Paper V*. Investigating other key decisions that can be influenced by different outcomes from different methods is helpful for real-world situations. For instance, it is beneficial to discover, when a company wants to terminate its collaboration with a supplier, how the rank reversal present in the outcomes of different methods can change this strategic decision. Such strategic decisions can flourish or ruin any business, so decision makers need to be extremely careful with the information they gather from different analyses. Last but not least, the findings of this research, like any other research, need to be supported by the findings of other research in the field. To do so, it is crucial to conduct similar research with different data sets, variables, and a greater number of suppliers. Then, the results of this other research can be compared with the findings of the present research to find out the extent to which the results and conclusions discussed herein are viable and practical for both practitioners and academics. The more research, the more the reliability of the results yielded.

## 6. References

---

- Al-Harbi, K. M. A. S. (2001). Application of the AHP in project management. *International journal of project management*, 19(1), 19-27.
- Bai, C., Sarkis, J., & Wei, X. (2010). Addressing key sustainable supply chain management issues using rough set methodology. *Management Research Review*, 33(12), 1113-1127.
- Beattie, B. R., Taylor, C. R., & Watts, M. J. (2009). *The economics of production*, Second Ed. Krieger Publishing Company.
- Bevilacqua, M., Ciarapica, F. E., & Giacchetta, G. (2006). A fuzzy-QFD approach to supplier selection. *Journal of Purchasing and Supply Management*, 12(1), 14-27.
- Bhutta, M. K. S. (2003). Supplier selection problem: methodology literature review. *Journal of International Information Management*, 12(2), 5.
- Boström, M., Jönsson, A. M., Lockie, S., Mol, A. P., & Oosterveer, P. (2015). Sustainable and responsible supply chain governance: challenges and opportunities. *Journal of Cleaner Production*, 107, 1-7.
- Bowersox, D. J. (1989). *Leading edge logistics competitive positioning for the 1990s: comprehensive research on logistics organization strategy and behavior in North America*. Council of Logistics Management.
- Braglia, M., & Petroni, A. (2000). A quality assurance-oriented methodology for handling trade-offs in supplier selection. *International Journal of Physical Distribution & Logistics Management*, 30(2), 96-112.
- Brandenburg, M., Govindan, K., Sarkis, J., & Seuring, S. (2014). Quantitative Models for Sustainable Supply Chain Management: Developments and Directions. *European Journal of Operational Research*, 233(2), 299–312.
- Brans, J. P., & Vincke, P. (1985). Note—A Preference Ranking Organisation Method: (The PROMETHEE Method for Multiple Criteria Decision-Making). *Management science*, 31(6), 647-656.
- Brewer, M. B., & Crano, W. D. (2000). Research design and issues of validity. *Handbook of research methods in social and personality psychology*, 3-16.
- Bruno, G., Esposito, E., Genovese, A., & Passaro, R. (2012). AHP-based Approaches for Supplier Evaluation: Problems and Perspectives. *Journal of Purchasing and Supply Management*, 18(3), 159–172.
- Capon, N., Farley, J. U., & Hoenig, S. (1990). Determinants of financial performance: a meta-analysis. *Management science*, 36(10), 1143-1159.

- Carr, A. S., & Pearson, J. N. (1999). Strategically managed buyer–supplier relationships and performance outcomes. *Journal of operations management*, 17(5), 497-519.
- Chai, J., Liu, J. N., & Ngai, E. W. (2013). Application of decision-making techniques in supplier selection: A systematic review of literature. *Expert Systems with Applications*, 40(10), 3872-3885.
- Chan, F. T., Kumar, N., Tiwari, M. K., Lau, H. C., & Choy, K. L. (2008). Global supplier selection: a fuzzy-AHP approach. *International Journal of production research*, 46(14), 3825-3857.
- Chang, B., Chang, C. W., & Wu, C. H. (2011). Fuzzy DEMATEL method for developing supplier selection criteria. *Expert systems with Applications*, 38(3), 1850-1858.
- Chapelle, C. A. (1999). Validity in language assessment. *Annual review of applied linguistics*, 19, 254-272.
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, 2(6), 429-444.
- Chen, Y-J. (2011). Structured Methodology for Supplier Selection and Evaluation in a Supply Chain. *Information Sciences*, 181(9), 1651–1670.
- Chkanikova, O., & Mont, O. (2015). Corporate supply chain responsibility: drivers and barriers for sustainable food retailing. *Corporate Social Responsibility and Environmental Management*, 22(2), 65-82.
- Chopra, S., & Meindl, P. (2015). *Supply chain management: strategy, planning, and operation*, 6<sup>th</sup> ed. Harlow: Pearson Education Limited.
- Choy, K. L., Lee, W. B., & Lo, V. (2002). An intelligent supplier management tool for benchmarking suppliers in outsource manufacturing. *Expert Systems with applications*, 22(3), 213-224.
- Christopher, M. (2016). *Logistics & supply chain management*. Pearson UK.
- Church, A. H., & Waclawski, J. (2017). *Designing and using organizational surveys*. Routledge.
- Clark, V. L. P., & Badiie, M. (2010). Research questions in mixed methods research. *SAGE Handbook of mixed methods in social & behavioral research*, 275-304.
- Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis*. Springer Science & Business Media.
- Colicchia, C., Marchet, G., Melacini, M., & Perotti, S. (2013). Building environmental sustainability: empirical evidence from Logistics Service Providers. *Journal of Cleaner Production*, 59, 197-209.
- Council of Logistics Management (1999), 1999 Membership Roster, Oak Brook, IL: Council of Logistics Management, p. 506.
- Council of Supply Chain Management Professionals. (2013). *Supply chain management terms and glossary*.

- Cousins, P. D., Lawson, B., & Squire, B. (2006). An empirical taxonomy of purchasing functions. *International Journal of Operations & Production Management*, 26(7), 775-794.
- Cristiano, J. J., Liker, J. K., & White III, C. C. (2000). Customer-driven product development through quality function deployment in the US and Japan. *Journal of Product Innovation Management: an international publication of the product development & management association*, 17(4), 286-308.
- Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management science*, 9(3), 458-467.
- Daly, L. E., Bourke, G. J., & McGilvray, J. (2000). Interpretation and uses of medical statistics (p. 568). Oxford: Blackwell Science.
- Day, J. (2001). Nike: 'no guarantee on child labour'. The Guardian. Accessed (28/05/2019).
- De Boer, L., Labro, E., & Morlacchi, P. (2001). A review of methods supporting supplier selection. *European journal of purchasing & supply management*, 7(2), 75-89.
- De Vaus, D. A., & de Vaus, D. (2001). Research design in social research. Sage.
- DHL Global, Logistics, Supply chain solutions, <https://www.mydhl.dhl.com/en.html>. Accessed (21/09/2019).
- Dixon-Woods, M., Shaw, R. L., Agarwal, S., & Smith, J. A. (2004). The problem of appraising qualitative research. *BMJ Quality & Safety*, 13(3), 223-225.
- Dubois, A., & Gadde, L. E. (2002). Systematic combining: an abductive approach to case research. *Journal of business research*, 55(7), 553-560.
- Duckstein, L., & Opricovic, S. (1980). Multiobjective optimization in river basin development. *Water Resources Research*, 16(1), 14-20.
- Enarsson, L. (1998). Evaluation of suppliers: how to consider the environment. *International Journal of Physical Distribution & Logistics Management*, 28(1), 5-17.
- Ergu, D., Kou, G., Peng, Y., & Shi, Y. (2011). A simple method to improve the consistency ratio of the pair-wise comparison matrix in ANP. *European Journal of Operational Research*, 213(1), 246-259.
- Evangelista, P., Santoro, L., & Thomas, A. (2018). Environmental sustainability in third-party logistics service providers: A systematic literature review from 2000–2016. *Sustainability*, 10(5), 1627.
- Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A (General)*, 120(3), 253-281.

- Faruk, A. C., Lamming, R. C., Cousins, P. D., & Bowen, F. E. (2001). Analyzing, mapping, and managing environmental impacts along supply chains. *Journal of Industrial Ecology*, 5(2), 13-36.
- Fife-Schaw, C. (1995). Questionnaire design. *Research methods in psychology*, 174-193.
- Fitzner, K. (2007). Reliability and validity a quick review. *The Diabetes Educator*, 33(5), 775-780.
- Forker, B. L., & Mendez, D. (2001). An analytical method for benchmarking best peer suppliers. *International Journal of Operations & Production Management*, 21(1/2), 195-209.
- Genovese, A., Lenny Koh, S. C., Bruno, G., & Esposito, E. (2013). Greener supplier selection: state of the art and some empirical evidence. *International Journal of Production Research*, 51, 2868-2886.
- Geoffrion, A. M. (1976). The purpose of mathematical programming is insight, not numbers. *Interfaces*, 7(1), 81-92.
- Ghadimi, P., Azadnia, A. H., Heavey, C., Dolgui, A., & Can, B. (2016). A review on the buyer-supplier dyad relationships in sustainable procurement context: past, present and future. *International Journal of Production Research*, 54(5), 1443-1462.
- Ghodsypour, S. H., & O'Brien, C. (1998). A decision support system for supplier selection using an integrated analytic hierarchy process and linear programming. *International journal of production economics*, 56, 199-212.
- Gimenez, C. & Tachizawa, E. M. (2012). Extending sustainability to suppliers: a systematic literature review. *Supply Chain Management: An International Journal*, 17, 531-543.
- Girubha, J., Vinodh, S. & KEK, V. (2016). Application of Interpretative Structural Modelling integrated Multi Criteria Decision Making methods for sustainable supplier selection. *Journal of Modelling in Management*, 11(2), 358-388.
- Govindan, K., Rajendran, S., Sarkis, J., & Murugesan, P. (2015). Multi criteria decision making approaches for green supplier evaluation and selection: a literature review. *Journal of Cleaner Production*, 98, 66-83.
- Handfield, R. B., Cousins, P. D., Lawson, B., & Petersen, K. J. (2015). How can supply management really improve performance? A knowledge-based model of alignment capabilities. *Journal of Supply Chain Management*, 51(3), 3-17.
- Haverty, M. (2002). Information architecture without internal theory: An inductive design process. *Journal of the American society for information science and technology*, 53(10), 839-845.
- Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-based nursing*, 18(3), 66-67.

- Helin, S., & Babri, M. (2015). Travelling with a code of ethics: a contextual study of a Swedish MNC auditing a Chinese supplier. *Journal of Cleaner Production*, 107, 41-53.
- Ho, W., Xu, X., & Dey, P. K. (2010). Multi-criteria decision making approaches for supplier evaluation and selection: A literature review. *European Journal of operational research*, 202(1), 16-24.
- Hollos, D., Blome, C., & Foerstl, K. (2012). Does sustainable supplier co-operation affect performance? Examining implications for the triple bottom line. *International Journal of Production Research*, 50(11), 2968-2986.
- Holt, G. D. (1998). Which contractor selection methodology?. *International Journal of project management*, 16(3), 153-164.
- Huang, S. H., & Keskar, H. (2007). Comprehensive and configurable metrics for supplier selection. *International journal of production economics*, 105(2), 510-523.
- Hugos, M. H. (2018). Essentials of supply chain management. John Wiley & Sons.
- Hutchins, M. J., & Sutherland, J. W. (2008). An exploration of measures of social sustainability and their application to supply chain decisions. *Journal of cleaner production*, 16(15), 1688-1698.
- Hwang, C. L., & Yoon, K. (1981). Methods for multiple attribute decision making. In Multiple attribute decision making. Springer, Berlin, Heidelberg.
- Igarashi, M., de Boer, L., & Fet, A. M. (2013). What is required for greener supplier selection? A literature review and conceptual model development. *Journal of Purchasing and Supply Management*, 19(4), 247-263.
- International Energy Outlook (IEO), (2016). U.S. Energy Information Administration, Washington, DC 20585.
- International Transport Forum (ITF), (2016). The Carbon Footprint of Global Trade. Tackling Emissions from International Freight Transport, Paris, France, pp. 1–10.
- Ishikawa, K., & Ishikawa, K. (1982). Guide to quality control (Vol. 2). Tokyo: Asian Productivity Organization.
- Jahre, M., & Johan Hatteland, C. (2004). Packages and physical distribution: implications for integration and standardisation. *International Journal of Physical Distribution & Logistics Management*, 34(2), 123-139.
- Jharkharia, S., & Shankar, R. (2007). Selection of logistics service provider: An analytic network process (ANP) approach. *Omega*, 35(3), 274-289.
- Joppe, M. (2000). The Research Process. Retrieved February 25, 1998, from <http://www.ryerson.ca/~mjoppe/rp.htm>.

- Jung, H. (2017). Evaluation of third party logistics providers considering social sustainability. *Sustainability*, 9(5), 777.
- Junior, F. R. L., Osiro, L., & Carpinetti, L. C. R. (2014). A comparison between Fuzzy AHP and Fuzzy TOPSIS methods to supplier selection. *Applied Soft Computing*, 21, 194-209.
- Kannan, D., Khodaverdi, R., Olfat, L., Jafarian, A., & Diabat. A. (2013). Integrated Fuzzy Multi Criteria Decision Making Method and Multi-Objective Programming Approach for Supplier Selection and Order Allocation in a Green Supply Chain. *Journal of Cleaner Production*, 47, 355–367.
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who’s the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15-25.
- Kataoka, S. (1963). A stochastic programming model. *Econometrica*, 31(1/2), 181.
- Kirk, J., Miller, M. L., & Miller, M. L. (1986). Reliability and validity in qualitative research (Vol. 1). Sage.
- Klassen, R. D., & McLaughlin, C. P. (1996). The impact of environmental management on firm performance. *Management science*, 42(8), 1199-1214.
- Krause, D. R., Vachon, S., & Klassen, R. D. (2009). Special topic forum on sustainable supply chain management: introduction and reflections on the role of purchasing management. *Journal of Supply Chain Management*, 45(4), 18-25.
- Kuo, R. J., Wang, Y. C., & Tien, F. C. (2010). Integration of artificial neural network and MADA methods for green supplier selection. *Journal of Cleaner Production*, 18(12), 1161- 1170.
- La Londe, B. J., & Masters, J. M. (1994). Emerging logistics strategies: blueprints for the next century. *International journal of physical distribution & logistics management*, 24(7), 35-47.
- Lai, K. H. (2004). Service capability and performance of logistics service providers. *Transportation Research Part E: Logistics and Transportation Review*, 40(5), 385-399.
- Lajunen, T., & Summala, H. (2003). Can we trust self-reports of driving? Effects of impression management on driver behaviour questionnaire responses. *Transportation research part F: traffic psychology and behaviour*, 6(2), 97-107.
- Lambert, D. M., & Cooper, M. C. (2000). Issues in supply chain management. *Industrial marketing management*, 29(1), 65-83.
- Lee, A. H., Kang, H. Y., Hsu, C. F., & Hung, H. C. (2009). A green supplier selection model for high-tech industry. *Expert systems with applications*, 36(4), 7917-7927.
- Logistics Quarterly Magazine, <http://logisticsquarterly.com/issues/16-3/3pl2011/dhl.html>. Accessed (21/09/2019).

- Lovell, C. K. (1993). Production frontiers and productive efficiency. *The measurement of productive efficiency: Techniques and applications*, 3, 67.
- Lubin, D. A., & Esty, D. C. (2010). The sustainability imperative. *Harvard business review*, 88(5), 42-50.
- Mahler, D. (2007). The sustainable supply chain. *Supply Chain Management Review*, 11(8), 59-60.
- Markley, M. J., & Davis, L. (2007). Exploring future competitive advantage through sustainable supply chains. *International Journal of Physical Distribution & Logistics Management*, 37(9), 763-774.
- Marshall, G. (2005). The purpose, design and administration of a questionnaire for data collection. *Radiography*, 11(2), 131-136.
- Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., & Zacharia, Z. G. (2001). Defining supply chain management. *Journal of Business logistics*, 22(2), 1-25.
- Miller, C. E., Tucker, A. W., & Zemlin, R. A. (1960). Integer programming formulation of traveling salesman problems. *Journal of the ACM (JACM)*, 7(4), 326-329.
- Mitchell, M. L., & Jolley, J. M. (2012). *Research design explained*. Cengage Learning.
- Modi, S. B., & Mabert, V. A. (2007). Supplier development: Improving supplier performance through knowledge transfer. *Journal of operations management*, 25(1), 42-64.
- Morgan, D. L. (2013). *Integrating qualitative and quantitative methods: A pragmatic approach*. Sage publications.
- Narasimhan, R. (1983). An analytical approach to supplier selection. *Journal of supply chain management*, 19(4), 27-32.
- Neely, A., Gregory, M., & Platts, K. (1995). Performance measurement system design: a literature review and research agenda. *International journal of operations & production management*, 15(4), 80-116.
- Onwuegbuzie, A. J. (2000). Expanding the Framework of Internal and External Validity in Quantitative Research.
- Oprićović, S., & Tzeng, G. H. (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. *European journal of operational research*, 156(2), 445-455.
- Ozsen, L., Coullard, C. R., & Daskin, M. S. (2008). Capacitated warehouse location model with risk pooling. *Naval Research Logistics (NRL)*, 55(4), 295-312.

- Pant, D. R. (2005). A place brand strategy for the Republic of Armenia: 'Quality of context' and 'sustainability' as competitive advantage. *Place branding*, 1(3), 273-282.
- Persson, G., & Virum, H. (2001). Growth strategies for logistics service providers: a case study. *The International Journal of Logistics Management*, 12(1), 53-64.
- Porter, M. E. (2008). *Competitive advantage: Creating and sustaining superior performance*. Simon and Schuster.
- Quinn, J. B., & Hilmer, F. G. (1994). Strategic outsourcing. *MIT Sloan Management Review*, 35(4), 43.
- Rashidi, K., & Cullinane, K.P.B (2019). Evaluating the sustainability of national logistics performance using Data Envelopment Analysis. *Transport Policy*, 74, 35-46.
- Ricks Jr, J. M. (2005). An assessment of strategic corporate philanthropy on perceptions of brand equity variables. *Journal of Consumer Marketing*, 22(3), 121-134.
- Rodriguez, S. I., Roman, M. S., Sturhahn, S. C., & Terry, E. H. (2002). Sustainability assessment and reporting for the University of Michigan's Ann Arbor Campus. Center for Sustainable Systems, Report No. CSS02-04. University of Michigan, Ann Arbor, Michigan.
- Rogers, D. S., & Tibben-Lembke, R. (2001). An examination of reverse logistics practices. *Journal of business logistics*, 22(2), 129-148.
- Romero, C. (2014). *Handbook of critical issues in goal programming*. Elsevier.
- Roorda, M. J., Cavalcante, R., McCabe, S., & Kwan, H. (2010). A conceptual framework for agent-based modelling of logistics services. *Transportation Research Part E: Logistics and Transportation Review*, 46(1), 18-31.
- Roy, B. (1968). Classement et choix en présence de points de vue multiples. *Revue française d'informatique et de recherche opérationnelle*, 2(8), 57-75.
- Saaty, R. W. (1987). The analytic hierarchy process—what it is and how it is used. *Mathematical modelling*, 9(3-5), 161-176.
- Saaty, T. L. (1996). *Decision making with dependence and feedback: The analytic network process*. RWS Publication.
- Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *International journal of services sciences*, 1(1), 83-98.
- Schiele, H. (2007). Supply-management maturity, cost savings and purchasing absorptive capacity: Testing the procurement-performance link. *Journal of purchasing and supply management*, 13(4), 274-293.
- Seiford, L. M., & Thrall, R. M. (1990). Recent developments in DEA: the mathematical programming approach to frontier analysis. *Journal of econometrics*, 46(1), 7-38.

- Selviaridis, K., & Spring, M. (2007). Third party logistics: a literature review and research agenda. *The international journal of logistics management*, 18(1), 125-150.
- Seuring, S., & Müller, M. (2008). From a literature review to a conceptual framework for sustainable supply chain management. *Journal of cleaner production*, 16(15), 1699-1710.
- Shaik, M., & Abdul-Kader, W. (2011). Green supplier selection generic framework: a multi-attribute utility theory approach. *International Journal of Sustainable Engineering*, 4(01), 37-56.
- Shane, P. B., & Spicer, B. H. (1983). Market response to environmental information produced outside the firm. *The accounting review*, 58(3), 521.
- Sivapornpunlerd, N., & Setamanit, S. O. (2014). Supplier performance evaluation: a case study of thai offshore oil & gas exploration and production company. *ASBBS Proceedings*, 21(1), 647.
- So, S. H., Kim, J., Cheong, K., & Cho, G. (2006). Evaluating the service quality of third-party logistics service providers using the analytic hierarchy process. *JISTEM-Journal of Information Systems and Technology Management*, 3(3), 261-270.
- Srivastava, S. K. (2007). Green supply-chain management: a state-of-the-art literature review. *International journal of management reviews*, 9(1), 53-80.
- Story, L. (2007). Lead Paint Prompts Mattel to Recall 967,000 Toys. *The New York Times*. Accessed (28/05/2019).
- Thai, V. V. (2013). Logistics service quality: conceptual model and empirical evidence. *International Journal of Logistics Research and Applications*, 16(2), 114-131.
- Van Hoek, R. I. (2002). Using information technology to leverage transport and logistics service operations in the supply chain: an empirical assessment of the interrelation between technology and operations management. *International Journal of Information Technology and Management*, 1(1), 115-130.
- Van Weele, A. J. (2018). *Purchasing and supply chain management: Analysis, strategy, planning and practice*, 7<sup>th</sup> ed. Cengage Learning EMEA.
- Varkevisser, C. M., Pathmanathan, I., & Brownlee, A. (1991). *Designing and conducting health systems research projects: part 1, proposal development and fieldwork*. IDRC, Ottawa, ON, CA.
- Vesilind, P. A., Morgan, S. M., & Heine, L. G. (2010). *Introduction to environmental engineering-SI version*. Cengage Learning.
- Weber, C. A., & Ellram, L. M. (1993). Supplier selection using multi-objective programming: a decision support system approach. *International Journal of Physical Distribution & Logistics Management*, 23(2), 3-14.

- Weber, C. A., Current, J. R., & Benton, W. C. (1991). Vendor selection criteria and methods. *European journal of operational research*, 50(1), 2-18.
- Wen, L., Xu, L., & Wang, R. (2013). Sustainable supplier evaluation based on intuitionistic fuzzy sets group decision methods. *Journal of Information and Computational Science*, 10(10), 3209-3220.
- Wolf, J. (2014). The relationship between sustainable supply chain management, stakeholder pressure and corporate sustainability performance. *Journal of business ethics*, 119(3), 317-328.
- Yakovleva, N., Sarkis, J., & Sloan, T. (2012). Sustainable benchmarking of supply chains: the case of the food industry. *International Journal of Production Research*, 50(5), 1297-1317.
- Yilmaz, K. (2013). Comparison of quantitative and qualitative research traditions: Epistemological, theoretical, and methodological differences. *European Journal of Education*, 48(2), 311-325.
- Yu, H., Zeng, A. Z., & Zhao, L. (2009). Single or dual sourcing: decision-making in the presence of supply chain disruption risks. *Omega*, 37(4), 788-800.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3), 338-353.
- Zelany, M. (1974). A concept of compromise solutions and the method of the displaced ideal. *Computers & Operations Research*, 1(3-4), 479-496.
- Zimmer, K., Fröhling, M., & Schultmann, F. (2016). Sustainable supplier management—a review of models supporting sustainable supplier selection, monitoring and development. *International Journal of Production Research*, 54(5), 1412-1442.
- Zimmermann, H. J. (1983). Fuzzy mathematical programming. *Computers & operations research*, 10(4), 291-298.