

Set up a supply chain observatory through the comparison of multi-criteria parsimonious methods

Supply chain
observatory

Maria Angela Butturi

*Department of Science and Method for Engineering (DISMI),
Università Degli Studi di Modena e Reggio Emilia, Reggio Emilia, Italy, and*

Francesco Lolli and Rita Gamberini

*Department of Science and Method for Engineering (DISMI),
Università Degli Studi di Modena e Reggio Emilia, Reggio Emilia, Italy and
EN&TECH, Università Degli Studi di Modena e Reggio Emilia, Reggio Emilia, Italy*

Received 10 February 2023

Revised 22 November 2023

17 January 2024

Accepted 2 February 2024

Abstract

Purpose – This study presents the development of a supply chain (SC) observatory, which is a benchmarking solution to support companies within the same industry in understanding their positioning in terms of SC performance.

Design/methodology/approach – A case study is used to demonstrate the set-up of the observatory. Twelve experts on automatic equipment for the wrapping and packaging industry were asked to select a set of performance criteria taken from the literature and evaluate their importance for the chosen industry using multi-criteria decision-making (MCDM) techniques. To handle the high number of criteria without requiring a high amount of time-consuming effort from decision-makers (DMs), five subjective, parsimonious methods for criteria weighting are applied and compared.

Findings – A benchmarking methodology is presented and discussed, aimed at DMs in the considered industry. Ten companies were ranked with regard to SC performance. The ranking solution of the companies was on average robust since the general structure of the ranking was very similar for all five weighting methodologies, though simplified-analytic hierarchy process (AHP) was the method with the greatest ability to discriminate between the criteria of importance and was considered faster to carry out and more quickly understood by the decision-makers.

Originality/value – Developing an SC observatory usually requires managing a large number of alternatives and criteria. The developed methodology uses parsimonious weighting methods, providing DMs with an easy-to-use and time-saving tool. A future research step will be to complete the methodology by defining the minimum variation required for one or more criteria to reach a specific position in the ranking through the implementation of a post-fact analysis.

Keywords Benchmarking, Decision support systems, Competitive strategy

Paper type Research paper

1. Introduction

In today's competitive environment, characterized by high innovation, the competitive advantages can be constantly eroded. This means that companies should understand their

© Maria Angela Butturi, Francesco Lolli and Rita Gamberini. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>

The authors wish to acknowledge Makeitalia srl (<https://www.makeitalia.com>) for supporting this study and allowing its experts to participate in the survey.



reference market and industry performance standards as much as possible, in order to be able to evolve in line with it, without having to reinvent solutions that have already been proved to provide positive results (Dobrzykowski *et al.*, 2012). Companies that exploit benchmarking can improve their performance by setting realistic goals, continuously improving internal processes, enhancing their process thinking and innovation diffusion and making financial savings. It involves learning from competitors' successful experiences through methodically measuring, assessing, comparing and applying assimilated knowledge to improve performances (Krishnamoorthy and D'Lima, 2014). In the globalized market, companies' boundaries are unclear due to outsourcing and information technology (IT) (Balfaqih *et al.*, 2016). Thus it is commonly accepted that competition is no longer about competing with individual organizations, but rather with the supply chains (SCs) (Trkman *et al.*, 2010). This study therefore focuses on the performance of the SC.

There is a vast literature on SC performance and benchmarking (Hoek, 1998; Wong and Wong, 2008; Kailash *et al.*, 2017), highlighting different focuses and different metrics. According to (Maestrini *et al.*, 2017), the scientific literature often deals with partial and incomplete aspects of the overall performance of the SC. Most studies focus only on a few key aspects and rarely analyze the performance of the SC as a whole (Balfaqih *et al.*, 2016), and often it is not considered that different strategies and set of metrics are needed for different sectors (Jagan Mohan Reddy *et al.*, 2019). Despite the large number of models and frameworks developed for SC performance analysis and monitoring, they still face limitations in terms of practical applicability due to the competitive environment, the emergence of innovative technologies like digitalization and new sustainability-related requirements (Oubrahim *et al.*, 2022). In addition, supply chain management (SCM) functions can differ between developing and developed countries (Ramos *et al.*, 2022), especially regarding sustainability issues (Ali *et al.*, 2021; Ali *et al.*, 2023a, b).

Two fundamental aspects thus emerge from the academic discussion on the SC performance metrics (Bigliardi and Bottani, 2014; Hallikas *et al.*, 2021; Jha *et al.*, 2022). The first is that since the SC is a complex system, involving many actors with different strategic objectives, the goal of perfectly representing it through a defined series of parameters and metrics is very difficult. The second point is that the SC performance depends on the reference context (business sector, geographic positioning, available resources, business strategy, study objective, etc.). Moreover, the need for DMs to provide responses rapidly, taking into account the knowledge of the experts, is also highlighted (Oubrahim and Sefiani, 2022). In this context, the use of multi-criteria decision-making (MCDM) techniques allows for transforming the expertise of one or more DMs into a standardized and shared model, thus saving time and resources for selecting and ranking the alternatives available (Velasquez and Hester, 2013; Khan *et al.*, 2019).

One of the key points of multi-criteria models is the set of weights to assign to the criteria, as these weights represent their significance quantitatively and consequently influence the results of the analysis (Vinogradova *et al.*, 2018). The weights of the criteria can be elicited either directly or indirectly. Indirect elicitation methods make use of ranked (sorted) examples from which the criteria weights are obtained indirectly, i.e. they are implicitly given by the DM when assessing the ranking (class) of the examples taken as references. For indirect elicitation methods, the reader can refer to (Lolli *et al.*, 2019, 2022). In direct elicitation methods, the criteria weights are assessed directly by the DMs without using any pre-ranked (pre-sorted) example. As we aim to consider directly the knowledge of the experts, we focus solely on direct elicitation methods that can be divided further into subjective and objective methods. When reliable weights cannot be obtained from experts (e.g. if the DMs do not have sufficient expertise) objective methods can be used, which apply mathematical models to the decision matrix (e.g. the entropy method (Deng *et al.*, 2000)). In subjective methods the weights are based on the judgment of experts in the field, and this is our case.

The complexity of the SC performance metrics demands handling a high number of criteria, so, among the subjective weighting methods, we consider the so-called parsimonious weighting methods requiring the least possible number of decisions from the DM, reducing the cognitive effort of the DMs' while maintaining accuracy in the calculation of the weight values.

To address the aforementioned issues, this study presents an SC observatory, which is an assessment tool aimed at supporting companies in understanding their positioning in terms of SC performance from a benchmarking perspective. The tool was developed by focusing on manufacturers of automatic equipment for wrapping and packaging, which is a flourishing industry in Italy.

The tackled research questions are:

RQ1. Are weighting parsimonious methods effective for handling a numerous set of SC performance criteria from a benchmarking perspective, providing the DMs with an accurate, easy-to-use and time-saving tool?

RQ2. How to establish an SC observatory for a specific industrial sector?

This paper contributes to field of the overall SC performance evaluation by adopting a benchmarking methodology using parsimonious criteria weighting methods. The comparison of five parsimonious criteria weighting methods is implemented in order to analyze their performance and also their simplicity of use, comprehensibility of the requests and perception of the DMs. As highlighted by the performed literature review, and at the best of our knowledge, this comparison has never been made neither parsimonious methods have been adopted in the field of SC performance despite the high number of involved criteria. The numerical application demonstrates the robustness of the rankings, which cannot be generalized as being case-sensitive, but allows results to be used for improvement actions starting from a solid basis. Hence, we provide a theoretical contribution in the field of SC performance, where parsimonious methods are effective to deal with a high number of criteria and show a good appreciation by DMs.

The remainder of this paper is organized as follows. Firstly, a set of criteria for the evaluation of the SC performance was created, selecting from the literature the most important criteria pertaining to the automatic equipment for wrapping and packaging industry. Secondly, a ranking of companies was created using the technique for order preference by similarity to ideal solution (TOPSIS) MCDM method, after weighting the criteria according to an expert evaluation. Five different parsimonious subjective methodologies for determining the weights of the criteria are compared and their performances are discussed.

2. Literature review

The following literature review aims at showing how the researchers tried to address the evaluation of the SC performance through the application of MCDM methods and the limitations of previous studies. As the supply chain operation reference (SCOR) and the balanced scorecard (BSC) models, two of the most used SC evaluation models (Jagan Mohan Reddy *et al.*, 2019), showed some limitations since they do not include all SC functions and are scarcely flexible (Jha *et al.*, 2022), we limited our review to the studied using MCDM methods to include all the relevant performance indicators in a specific sector. Moreover, we investigated the academic studies comparing the subjective criteria weighting methods and specifically the parsimonious methods.

2.1 Multi-criteria methods to evaluate the supply chain performance

A complex hybrid exploratory three-phased MCDM model based on the "Decision making trial and evaluation laboratory" (DEMATEL) is used in Chand *et al.* (2020) to evaluate the SC

performance in the mining industry. DEMATEL is mainly used for analyzing cause and effect relationships among components of a system (Si *et al.*, 2018). It establishes the level of interdependence between the chosen criteria. The criteria are grouped into seven macro-categories: sustainability, order planning, collaboration capacity, operational performance, delivery performance, customer service level, costs and financial ratios. An integrated approach between the analytic hierarchy process (AHP) (Saaty, 1977) and the BSC is used by Varma *et al.* (2008) to assess the SC performance of the oil industry. Sufiyan *et al.* (2019) evaluate the performance of the food sector SC by integrating fuzzy DEMATEL and DEMATEL based-analytic network process (DANP) methods. They used 18 criteria, grouped into six categories: agility, sustainability, quality, level customer service, collaboration skills and SC efficiency. In some cases, the analysis of the performance focuses only on particular aspects of the SC, such as in the study by Arshinder *et al.* (2007), where collaboration and coordination skills are focused on using a fuzzy AHP method.

Most studies that address the SC performance are focused on sustainability (Ali *et al.*, 2020; Qorri *et al.*, 2022; Ali *et al.*, 2023a, b; Erol *et al.*, 2011) investigate the SC sustainability from environmental, social and economic perspectives, using a total of 37 criteria, which are weighted through the fuzzy entropy method, while the performance calculation is obtained through the fuzzy multi-attribute utility theory (FMAUT). The criteria considering the whole life cycle of the product are analyzed by Sarkis (2003) using the analytic network process (ANP) method. The sustainability of the service SC is analyzed using the Viekriterijumsko KOMPromisno Rangiranje (VIKOR) and ELimination Et Choix Traduisant la REalité (ELECTRE) models by Chithambaranathan *et al.* (2015) and Anas *et al.* (2018) use criteria defined by SCOR to evaluate the SC in the hospital environment, through a fuzzy AHP and best-worst method (BWM) for determining the weights of the criteria. Menon and Ravi (2022) use the AHP-TOPSIS methods to select sustainable suppliers in an electronics SC; the combined methods are used to evaluate the quantitative and qualitative data and manage the involved uncertainty. The social dimension of sustainability in warehousing operations is investigated by Ali and Kaur (2021) using the BWM with a hybrid approach; the study aims at understanding the effectiveness of corporate social responsibility in implementation of the 9 identified social sustainability practices.

The prioritization of 36 SC performance improvement indicators in the plastic manufacturing industry is implemented by Govindan *et al.* (2017) by using the fuzzy AHP method to improve the handling of human-based qualitative judgments; the sensitivity analysis examines the priority ranking of the indicators. A mixed SCOR-AHP approach is applied by Mañay *et al.* (2022) to study the performance of SCs at the floricultural sector.

The TOPSIS method is widely used in the SCM field (Behzadian *et al.*, 2012; Velasquez and Hester, 2013; Joshi *et al.*, 2011) use a Delphi-AHP-TOPSIS-based methodology to develop a benchmarking framework that evaluates the cold chain performance of a company, while (Kumar *et al.*, 2022) apply the hybrid AHP-fuzzy TOPSIS method to identify the critical performance criteria, responsible for the performance measurement of cold SC and suggest the best possible alternatives to improve it. The management performance of an electronic SC of an Indian automobile industry is evaluated through a hybrid approach using AHP-TOPSIS methods by Tyagi *et al.* (2014). A fuzzy TOPSIS method is used to evaluate the performance of the Indian petroleum SC in Kumar and Barua (2022).

2.2 Subjective criteria weighting methods comparison

There are several examples of comparisons between criteria weighting methods, which mainly consider subjective methods (Şahin, 2021). Although different aspects of the results are analyzed, the methods considered in the analysis tend to be the same. In fact, one of the main issues in comparing the methods for criteria weighting is that the “real” weight values

are lacking. Thus it is not possible to establish a standard for evaluating the validity of the weights assigned by each method (van Til *et al.*, 2014). It is therefore necessary to find other bases for the comparison of the methods.

Many of the studies use the most common criteria weighting methods, such as AHP and BWM, as terms of comparison to demonstrate the ability of other or newly developed methods to obtain the same results in a simpler way. van Til *et al.* (2014) compare the five-point scale, ranking, AHP and BWM assessment techniques from a total of 14 criteria and 60 DMs. The study shows that the choice of method in a group decision is not decisive, while it is for individual decisions. In the latter case pairwise comparisons are preferable, due to the greater ability in differentiating the criteria. Since some criteria impact more on the overall SC performance than others, the pairwise comparison of the AHP method, which ensures consistency among DMs when assigning the importance of a given factor over another, is preferred for the overall SC performance evaluation (Khan *et al.*, 2019). Eight subjective methods for determining subjective weights, including AHP and SMART (simple-multi-attribute utility theory), are compared by Németh *et al.* (2019) on the basis of the general complexity, the use of resources, software requirements and the possible distortions in the judgment due to the method setting. The results show that more complex methods make the judgment less distorted, but at the same time they require more resources. Thus the most appropriate method needs to be carefully selected on the basis of the reference context, considering the acceptable accuracy tolerance, the size of the group of DMs and the resources available.

AHP is used as a term of comparison also in a study conducted by Şahin (2021), where BMW and four objective methods are evaluated. The results of the study reveal similar weight values between the two subjective methods and similar weights between the four subjective methods. However, they differ when comparing the results obtained through subjective or objective methodologies, which is why the study suggests using a combined approach to obtain more reliable values, where possible.

Two other cases in which AHP is used as a term of comparison are the studies conducted by Riabacke *et al.* (2012) and by Pöyhönen and Hämäläinen (2001). In the first study different methods (including direct rating, point allocation, SWING, SMART) are evaluated according to three fundamental concepts: extraction (how the information is derived from the input provided by the DM), representation (the format in which the information derived from the DM is structured) and the interpretation (how the meaning is attributed to the information derived from the DM). The second study compares SMART, AHP, SWING, point allocation and TRADEOFF methods, assuming that these methods provide similar weights as they are based on the same theoretical foundations.

Some studies compare simple subjective methods, where pair comparisons are not required but, for example, a simple classification by importance or direct scoring is implemented (Bottomley and Doyle, 2001; Alfares and Duffuaa, 2008; Borchering *et al.*, 1991) compare the ratio method, SWING, TRADEOFF and pricing out method considering their internal consistency, the level of agreement between the weights obtained with the different methods and their external validity (the concordance between the obtained weights and judgments of the experienced managers). The simpler methods however need to be used carefully, as an inverse relationship is possible between the precision of the methods and their simplicity (Németh *et al.*, 2019).

Another area of interest concerns the subjective perception of the DM and the judgment biases found in the implementation of subjective methodologies. Weber and Borchering (1993) analyze SMART, SWING and TRADEOFF concluding that the weights could be influenced firstly by the choice of weighting method, the hierarchical structure of the problem, and the reference used. Thus, there is no way to determine the “true” weights of the criteria considered, since all weighting methods induce bias; a possible solution is therefore to rely on multiple evaluations. Other cognitive biases are anchoring bias, which is influenced

by the structure of the weighting method and the first judgment that is submitted to the DM (Buchanan and Corner, 1997; Rezaei, 2021), along with classification bias, loss aversion and status quo (Deniz, 2020).

3. Research methodology

A case study is used to demonstrate the set-up of an observatory to support benchmark analyses for companies belonging to the same industry. The observatory structure focuses on a set of quantitative criteria and is aimed at providing an exhaustive representation of companies' performance in the SC area.

The performance criteria of the SC were taken from the literature, and a group of experts selected the most important, each belonging to five macro-categories that the experts judged as the most significant for the chosen industry: purchasing, planning, internal logistics, transportation and quality. The experts (see Table 1) were asked to assign a weight to the selected criteria as well as to the macro-categories in order to obtain the criteria global weights. The TOPSIS method was then applied to calculate, for each company, an overall score combining the scores obtained from the set of criteria weighted by applying the aforementioned five subjective parsimonious methods for criteria weighting.

The research methodology is displayed in Figure 1.

3.1 Case study – the sample companies

The industry analyzed is the automatic equipment for wrapping and packaging industry, a key industrial sector in Italy and a leading sector worldwide, made up of 633 companies in Italy, 81% of which are small production units. The companies are mainly located in the north of Italy and particularly in the Emilia-Romagna region, which is also known as “packaging valley” as it makes half of the almost eight-billion-euro national turnover for the sector. Most of the production of wrapping and packaging machines is absorbed by the Food and Beverage customer sector (58.2%), followed by Pharmaceutical and Biomedical (17.4%), Cosmetic and Personal Care (4.5%) and Chemicals and Home (3.5%) [1].

The sample of selected companies includes 10 companies belonging to the chosen sector with a turnover of over 50 million euros and based in the Emilia-Romagna region (IT). The turnover distribution is shown in Figure 2.

	Percentage (%)
Male	58.3
Female	41.7
<i>Age range</i>	
25–30	41.7
31–35	50
36–40	8.3
<i>Years of experience</i>	
4–7	66.7
8–11	33.3
<i>Role</i>	
Project Manager	58.3
Customer Manager	41.7

Table 1. Age, years of experience and role of the experts participating in the study

Source(s): Table created by authors

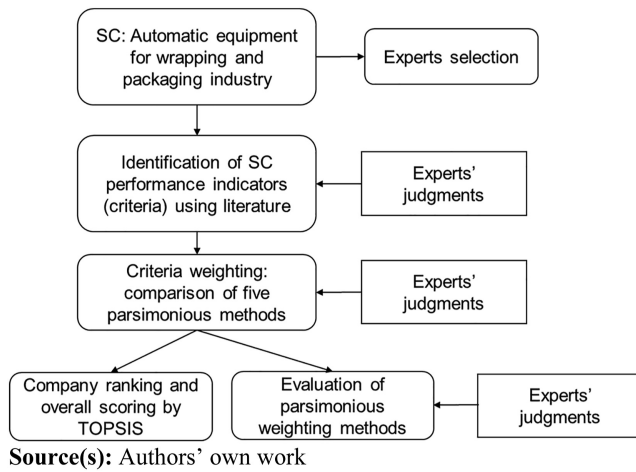
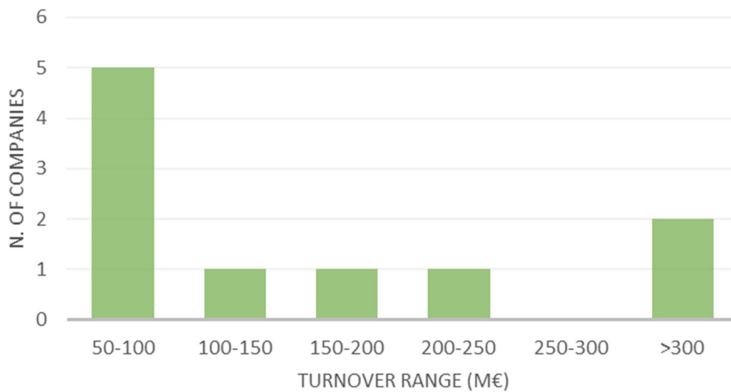


Figure 1.
Research methodology
used in this study



Source(s): Figure created by authors

Figure 2.
Turnover distribution
of the sample
companies

3.2 The decision-makers

Twelve professionals with significant experience (see Table 1) in the reference industry were asked to select the relevant criteria and evaluate them. To guarantee the quality of the expert's judgment, they were chosen among managers operating in SC design, development and management, ensuring a deep knowledge of the competitive environment and of the complexity of the SC performance evaluation.

After expressing their criteria weights, the experts gave a qualitative evaluation of the five parsimonious weighting methods applied.

3.3 Criteria selection guiding principles

The performance criteria of the SC were collected through a literature search carried out on the Scopus, WoS and Google Scholar databases. The literature survey was conducted searching for such keywords as "supply chain performance indicators", "supply chain performance criteria", "supply chain performance parameters", "supply chain performance metrics".

Based on the experts' knowledge of the SCM of the reference industry, the following hypotheses drove the selection of the criteria:

- (1) All the metrics had to refer to a one-year time period;
- (2) Each metric involving inbound materials only considered the goods and raw materials that directly contribute to obtain the finished product;
- (3) When referring to a warehouse, this always meant the raw materials warehouse, since, according to experts, the automatic machines industry generally adopts a make-to-order (MTO) or assemble-to-order (ATO) production. This is because often this type of company has a catalog of standard machines, to which specifications and adaptations can be added upon the customer's request. For this reason, warehouses dedicated to finished products were not considered relevant in our study;
- (4) The "Production" business function was not included among the considered macro-categories since it was considered too specific with respect to each company. The experts suggested that in this sector some companies produce each component of the machine, while others only deal with the assembly, others produce only some components, etc., Thus this business area is not significant for the purposes of this study. Furthermore, other metrics included, such as punctuality in the delivery of the finished product to the end customer (proposed in the "Planning" macro-category) indirectly require the correct management of production;
- (5) No rejected products were foreseen (rather, the hours for reworking a product not qualitatively acceptable can be considered);
- (6) No sent back products by the end customers were foreseen (the number of complaints and the costs incurred to carry out maintenance or repairs were considered at the end customer).

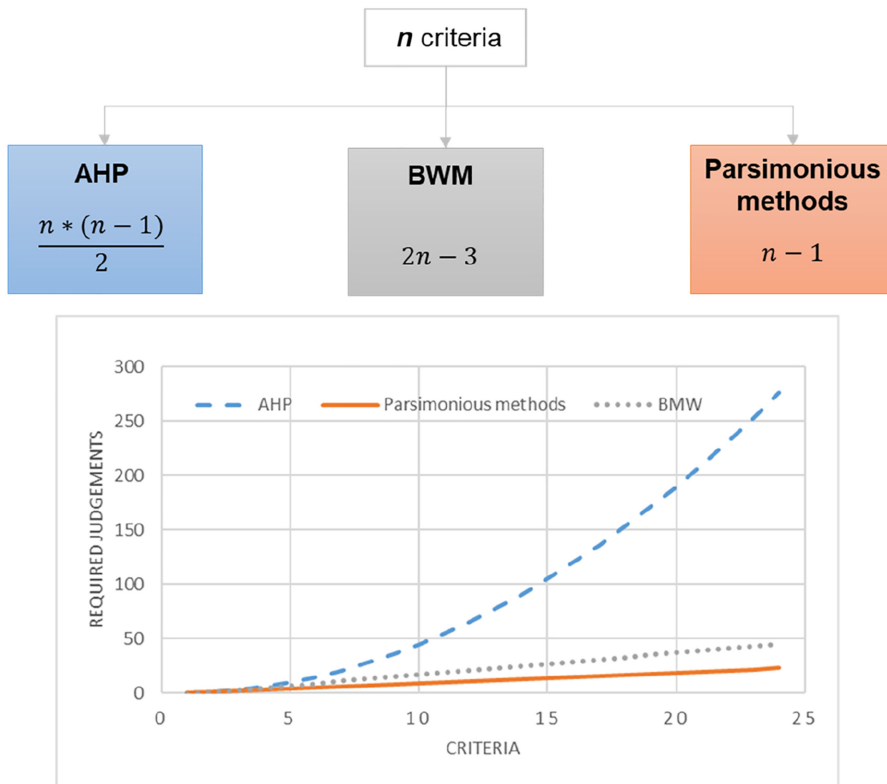
3.4 Criteria weighting – five parsimonious methods

Many methods have been developed for determining criteria weights based on expert judgments, with AHP and BWM (Best Worst Method) being the most common. AHP is one of the best known and most common methods (Ishizaka and Labib, 2011). It is based on pairwise comparisons made by expert DMs to create a priority scale. Although this is an easy and widely used method, the number of necessary pairwise comparisons increases rapidly with the increase in the number of criteria. It has also been shown that it is in fact almost impossible to perform consistent pairwise comparisons if more than nine criteria are present (Pamučar *et al.*, 2018). The validation of the results is then based on the degree of consistency of the matrix, which must not exceed 0.10.

Given n criteria, the BMW method (Rezaei, 2015) is aimed at improving some of the AHP flaws, reducing the number of pairwise comparisons from $n(n - 1)$ to $2n - 3$ and the probability of matrix inconsistency. It is based on the concept of identifying the best and the worst criterion and proceeds by comparing these with the other criteria; the values of the weights are then obtained through an optimization model.

Figure 3 compares the number of judgments required by the DMs in AHP, BWM and parsimonious weighting methods as a function of the number of criteria considered. Parsimonious methods significantly reduce the judgments required as the number of criteria increases over 10.

The experts were asked to evaluate the previously selected criteria using five parsimonious subjective methods. The five methods are (for details, see the papers of the cited authors):



Source(s): Figure created by authors

Figure 3. Number of judgments required as a function of the number of criteria to be weighted in AHP, BWM and parsimonious methods

- (1) Simplified-AHP: this method, proposed by [Benitez et al. \(2019\)](#), does not modify the AHP algorithm (developed by [Saaty \(1977\)](#)), but reduces the number of pairwise comparisons by selecting just a sample of n pairwise comparisons which while providing balanced and unbiased (incomplete) information, still produces consistent and robust decisions.
- (2) AHP-express ([Leal, 2020](#)), is another simplified method for the application of the AHP. Assuming that DMs make consistent evaluations, the number of pairwise comparisons is reduced to $n-1$ and the comparisons are carried out between a chosen criterion (normally the most significant one) and all the other criteria. The assumption of consistency in the evaluation is validated since the inconsistency occurs mainly in the comparisons between alternatives that are less significant. Thus taking the criterion deemed most important as the first term of comparison, the DM will pay more attention to the assignment of the score.
- (3) Full consistency method (FUCOM), developed by [Pamučar et al. \(2018\)](#). This is based on the comparison of $n-1$ pairs of criteria, ensuring maximum consistency thanks to the validation of the model by determining the deviation from full consistency. By respecting the transitivity conditions defined by specific constraints, it also eliminates the problems related to comparisons of redundant pairs.

- (4) Non-decreasing series at criteria significance levels (NDSL) developed by [Žižović et al. \(2020\)](#). This method, as with the previously described methods, requires n-1 pairwise comparisons and is based on the identification of the best criterion. It classifies the criteria into decreasing series from the most to the least important in order to group the criteria into significance levels, based on the judgment expressed by the DMs. In addition, unlike standard AHP, it obtains consistent results even with a number of criteria greater than nine. Preferences are not defined on the 1/9–9 scale, which limits the expression of preferences.
- (5) Level-based weight assessment (LBWA) developed by [Žižović and Pamučar \(2019\)](#). This method is based on the classification of criteria in levels of importance, requiring a reduced number of comparisons in pairs of criteria (n-1). Another advantage is that the algorithm does not become more complex as the number of criteria increases, which is why it can be used in contexts with a high number of criteria. The existence of the elasticity coefficient also enables further corrections of the weight values according to the preferences of the DMs.

The set of selected criteria was broken down into six sections: five sections consisted of the aforementioned macro-categories and the sixth section, SC, concerned the global weighting of the macro-categories. The DMs were asked to attribute weights to one or more sections of the set, according to their area of expertise in SC management; the judgments of three experts were assessed for each macro-category.

Then, the weights $w_{i,j,z}$ resulting from the judgments of the DMs, were calculated as the average of the judgments of the three DMs for the selected method:

$$w_{i,j,z} = \frac{\sum_{k=1}^K w_{i,j,k,z}}{K} \quad (1)$$

where $w_{i,j,k,z}$ represents the weight associated with the criterion j ($j = 1, \dots, J_j$) within the macro-category i ($i = 1, \dots, I$), defined by the DM k ($k = 1, \dots, K$) for the weighting method z ($z = 1, \dots, Z$).

The weights $W_{i,z}$ of the macro-categories are calculated as:

$$W_{i,z} = \frac{\sum_{k=1}^K W_{i,k,z}}{K} \quad (2)$$

where $W_{i,k,z}$ represents the weight associated with the macro-category i , defined by the DM k for the weighting method z .

The degree of discrepancy between the DMs D_i , within each macro-category i , was then measured by calculating the variance of the results according to [equation \(3\)](#):

$$D_i = \frac{\sum_{z=1}^Z \sum_{k=1}^K \sum_{j=1}^{J_i} \left(w_{i,j,k,z} - \bar{w}_{i,z} \right)^2}{Z \cdot K} \quad (3)$$

where $\bar{w}_{i,z} = \frac{\sum_{k=1}^K \sum_{j=1}^{J_i} w_{i,j,k,z}}{J_i \cdot K}$

3.5 Evaluation of parsimonious methods

The performances of the five parsimonious methods were evaluated both in terms of the perceptions of the DMs and by calculating the discriminating capacity of each method.

A questionnaire (available in [Appendix 1](#)) was given to the experts to assess their opinions regarding the time needed to respond, the comprehensibility of the method and their preferred method.

The discriminating capacity d_z of each method z was measured by calculating the average variation range for each proposed methodology, according to [equation \(4\)](#):

$$d_z = \frac{\sum_{K=1}^K \sum_{i=1}^l [\max(W_{i,j,k,z}) - \min(W_{i,j,k,z})]}{K.l} \quad (4)$$

where $\max(w_{i,j,z,k})$ and $\min(w_{i,j,z,k})$ are the maximum and minimum values respectively of the weights assigned by the DM k , for the macro-category i , according to the method z .

3.6 Company ranking and overall scoring by TOPSIS

TOPSIS was selected as the method for our case study due to its practicality and ease of use, speed of application and the standardization of the steps (the number of steps in fact does not change with any increase in criteria ([Velasquez and Hester, 2013](#))). It is based on the classification of alternatives based on the concept that the selected alternative must simultaneously have the shortest distance from the positive-ideal solution and the farthest distance from the negative-ideal solution. The positive-ideal solution maximizes the benefit criteria and minimizes the cost criteria, whereas the negative-ideal solution maximizes the cost criteria and minimizes the benefit criteria. This approach applies well to the benchmark concept that is the focus of this study. [Appendix 1](#) outlines steps in the TOPSIS implementation.

The global weights $w_{i,j,z}^*$ used as input for the TOPSIS methodology implementation are calculated as the product between the weight of the criterion $w_{i,j,z}$ ([eq. \(1\)](#)) and the weight of the macro-category $W_{i,z}$ ([eq. \(2\)](#)) to which it belongs:

$$w_{i,j,z}^* = w_{i,j,z} \times W_{i,z} \quad (5)$$

4. Results and discussion

4.1 Criteria set

According to the criteria selection guiding principles previously agreed by the experts, a total of 22 numerical KPIs were considered as relevant for the reference industry, categorized according to the five macro-categories, each one corresponding to a relevant company area: purchasing, planning, internal logistics, transportation and quality. The set of criteria chosen by the experts aims to be exhaustive, thus justifying the use of parsimonious weighting methods. The 22 criteria are presented in [Table 2](#).

The selected criteria can be both costs and benefits for the companies, as specified in column 4 of [Table 2](#).

4.2 Criteria weighting using the five parsimonious methods

[Figure 4](#) shows the comparison of the weights obtained after the weighting process using the five parsimonious criteria weighting methods. The graph includes the five macro-categories and the SC evaluation; the weights of the criteria belonging to the five macro-categories are calculated using [equation \(1\)](#), the weights of the macro-categories (SC) are calculated using [equation \(2\)](#). The calculated values and the weight assigned to the 22 criteria and to the macro-categories are available as supplementary material ([Appendix 3](#)).

The calculated weights can be used both to analyze the importance of the selected criteria and to compare the results obtained with the different parsimonious criteria weighting methods.

Macro-category	Criterion	Description	Type	References
C ₁ - Purchasing	C _{1,1} – Reduction of purchase costs compared to the previous year	Percentage reduction in the purchase cost of direct materials that the purchasing function has achieved compared to the previous year	Benefit	1, 2
	C _{1,2} – Reliability of the purchase budget	Percentage of reliability of the purchase budget (for direct materials), based on the expected budget and the actual purchase turnover	Benefit	3
C ₂ - Planning	C _{2,1} – Warehouse turnover index	Referred to the inbound warehouse	Benefit	1, 2, 4, 5, 6
	C _{2,2} – Inbound on time pieces (number)	The number of pieces (direct materials only) that arrived on time, as planned	Benefit	4, 6, 7
	C _{2,3} – Inbound on time pieces (value)	The value of pieces (direct materials only) that arrived on time, as planned	Benefit	4, 6, 7
	C _{2,4} – Outbound on time finished products (number)	The number of finished products that have been delivered on time to the end customer	Benefit	2, 4, 5, 6, 7
	C _{2,5} – Outbound on time finished products (value)	The value of finished products that have been delivered on time to the end customer	Benefit	2, 4, 5, 6, 7, 8
	C _{2,6} – Emergency purchases	The purchase value of direct materials purchased in emergency conditions (with a lead time shorter than the agreed one)	Cost	7
C ₃ – Internal logistics	C _{3,1} – Unitary cost of stocking the goods in the warehouse	The unitary cost of stocking direct materials in the warehouse dedicated to raw materials	Cost	4, 5, 6
	C _{3,2} – Inventory accuracy	The difference between the theoretical value of the warehouse inventory relating to raw materials (considering only direct materials) and the actual value	Benefit	5, 9
	C _{3,3} – Value of the inventory adjustments	The value of the direct materials adjusted during the inventories made on the raw material warehouse	Cost	5, 6, 8, 9, 10
	C _{3,4} – Warehouse saturation level	The average occupancy level of the raw material warehouse compared to its nominal capacity	Benefit	5
C ₄ – Transportation	C _{4,1} – Inbound transport costs	The costs incurred in inbound transport, compared to the purchase turnover of the goods for which the transport is actually paid	Cost	5, 10, 11
	C _{4,2} – Outbound transport costs	The costs incurred in outbound transport, compared to the purchase turnover of the goods for which the transport is actually paid	Cost	5, 10, 11
	C _{4,3} – Inbound transport disruptions	The percentage of inbound transports during which disruptions occur (delays, damage, logistical problems, etc.)	Cost	5, 6, 11, 12
	C _{4,4} – Outbound transport disruptions	The percentage of outbound transports during which disruptions occur (delays, damage, logistical problems, etc.)	Cost	5, 6, 11, 12

Table 2. Quantitative criteria for defining the performance of the supply chain

(continued)

Macro-category	Criterion	Description	Type	References
C ₅ – Quality	C _{5,1} – Cost of inbound non-conformities	The purchase value of direct non-compliant materials, compared to the total value of the purchase	Cost	6, 7
	C _{5,2} – Inbound non-compliant pieces	The number of pieces (direct materials only) found to be non-compliant, compared to the total number of pieces purchased	Cost	6
	C _{5,3} – Rework hours needed	The hours of reworking of a product required due to a detected non-compliance, compared to the standard hours of processing of that product	Cost	4
	C _{5,4} – Cost of required rework	The costs incurred for reworking of products due to a detected non-conformity	Cost	4, 6
	C _{5,5} – Customer complaints for non-compliance	The number of products for which a complaint has been made by a customer due to a detected non-compliance, compared to the total number of products sold	Cost	2, 4, 6, 8
	C _{5,6} – Cost of customer complaints	The costs incurred for maintenance or replacement of the product due to a non-conformity detected by the customer	Cost	2, 4, 6, 8

Note(s): References: (1) (Trkman *et al.*, 2010); (2) (Thakkar *et al.*, 2009); (3) (Vaidya and Hudnurkar, 2013); (4) (Bigliardi and Bottani, 2014); (5) (Anand and Grover, 2015); (6) (Vaidya and Hudnurkar, 2013); (7) (Gunasekaran *et al.*, 2004); (8) (Beamon, 1999); (9) (Fleisch and Tell Kamp, 2005); (10) (Varma *et al.*, 2008); (11) (Lai *et al.*, 2002); (12) (Wilson, 2007)

Source(s): Table created by authors

Table 2.

Regarding the macro-category Purchasing, the weights assigned to criteria C_{1,1} and C_{1,2} are very similar for all the considered five methods (only the weights identified using the LBWA methodology differ slightly from those obtained by the other methods: ~3% for criterion C_{1,1} and ~10% for criterion C_{1,2}). Therefore, the greater importance assigned by the experts to criterion C_{1,1} (the reduction of purchasing costs) over criterion C_{1,2} (the reliability of the purchase budget) is clear.

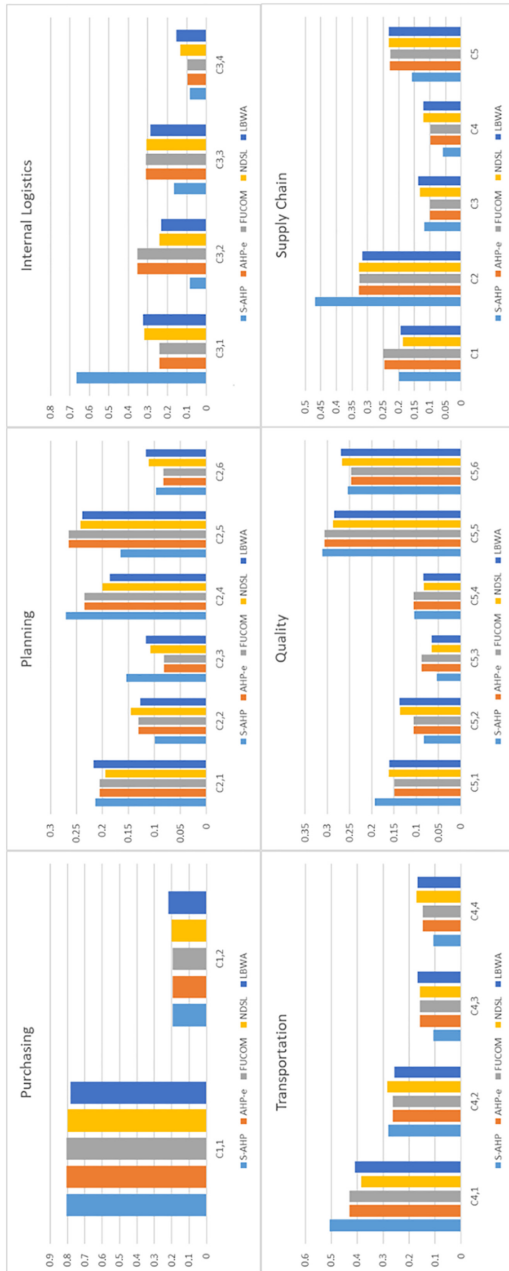
Considering Planning, a disagreement between simplified-AHP and the other methods can be seen for criteria C_{2,3}, C_{2,4} and C_{2,5}. Initially, the criteria C_{2,1}, C_{2,4} and C_{2,5} appear more significant than C_{2,2}, C_{2,3} and C_{2,6}.

The criteria for Internal Logistics show the greatest degree of disagreement among the DMs (the low reliability of the simplified-AHP data should be highlighted due to issues encountered by two experts in making their judgments with this methodology). The results show that criterion C_{3,4} can be considered as the least important.

On the other hand, Transportation showed the lowest degree of discrepancy between the DMs. The criteria are clearly distinguishable by their importance, with C_{4,1} criterion being the most significant followed by C_{4,2} and finally by C_{4,3} and C_{4,4}.

As regards Quality, identical weights for the AHP-express and FUCOM methodologies were obtained, as well as for the NDSL and LBWA methods. The greater importance given to criteria C_{5,5} and C_{5,6} is clearly highlighted, while the criteria C_{5,3} and C_{5,4} are less important.

Lastly, the analysis of the macro-categories (SC section) shows that there are very similar weights for all the methodologies, apart from category C₁ (Purchasing), where AHP-express and FUCOM show a slightly greater weight and category C₂ (Planning), for which the simplified-AHP methodology presents a much greater weight than the other methodologies.



Source(s): Figure created by authors

Figure 4. Comparison of the weights resulting from the application of the five parsimonious criteria weighting methods

Planning is given the most important, followed by Purchasing and Quality. The least importance was given to Internal Logistics and Transportation.

From the obtained results, it can be argued that methodologies with comparable structures yield comparable results, as observed by Pöyhönen and Hämäläinen (2001).

Supply chain
observatory

4.3 Discrepancy between decision-makers

The degree of discrepancy D_i for each macro-category i is calculated according to eq. (3) and presented in Table 3:

Purchasing shows the lowest degree of discordance, while the greatest discrepancy between the DMs was for Planning and Quality, which were the categories with the greatest number of criteria.

4.4 Evaluation of the five parsimonious subjective criteria weighting methods

4.4.1 Experts' perception. At the end of the criteria weighting procedure, each expert was asked to answer the evaluation questionnaire (Appendix 1) on the proposed methods in order to evaluate their perception.

The results are presented in Table 4.

Macro-category	Degree of discrepancy between decision-makers
Purchasing	0.19541
Planning	0.73300
Internal logistics	0.28118
Transportation	0.31761
Quality	0.67760
Supply Chain	0.47629

Source(s): Table created by authors

Table 3.
Degree of discrepancy
between decision-
makers for each macro-
category

<i>Which method was the quickest to apply?</i>	
s-AHP	41.67%
AHP-e/FUCOM	25.00%
NDSL	16.67%
LBWA	8.34%
<i>Which method was the quickest to understand?</i>	
s-AHP	33.33%
AHP-e/FUCOM	25.00%
NDSL	16.67%
LBWA	16.67%
<i>Which method enabled you to most accurately express your preference?</i>	
s-AHP	33.33%
AHP-e/FUCOM	41.67%
NDSL	0.00%
LBWA	25.00%
<i>Which method did you prefer overall?</i>	
s-AHP	41.67%
AHP-e/FUCOM	33.33%
NDSL	8.34%
LBWA	16.67%

Source(s): Table created by authors

Table 4.
Subjective preferences
of the experts
regarding the five
methods

The questionnaire results showed that simplified-AHP was the preferred method in terms of all the aspects considered, except for the perception of accurately selecting a preference, for which AHP-express and FUCOM methods were ranked highest.

The degree of appreciation of the DMs was not considered in previous studies (Luthra *et al.*, 2017; Mañay *et al.*, 2022), making our framework more attractive for business applications where DMs are subject to time constraints.

4.4.2 Discriminating capacity of each method. The discriminating capacity d_z of each method z was calculated according to eq. (4). The resulting values are shown in Table 5.

The results show that the method with the greatest discriminating capacity is the simplified-AHP. This may be due to the fact that the pairs of criteria to assign a judgment to are always different; on the other hand, in AHP-express or in FUCOM, the first term of comparison is always the most important criterion. This phenomenon can also be considered a bias since, taking the criterion deemed to be most important as the first term of comparison, the DM will pay more attention to the assignment of the score (Leal, 2020).

The AHP-express and FUCOM methods have almost identical values, since as seen previously, they often result in identical weights.

4.5 Company rankings and overall TOPSIS scoring

The data of the 10 sample companies were analyzed in order to assign each company with a score in relation to each of the selected criteria (Table 6).

4.5.1 TOPSIS implementation. Starting with the original score decision matrix (Table 6), the TOPSIS method was implemented (see Appendix 2).

The global weights to be used as input for the TOPSIS methodology were calculated according to equation (3) for each of the five weighting methods (Table 7).

The final ranking of the companies included in the study is shown in Table 8.

Table 8 compares the rankings obtained from the five subjective parsimonious weighting methods. As shown, identical results were obtained for the AHP-express and FUCOM methodologies. Almost the same resulting ranking was obtained for the other three methods (simplified-AHP, NDSL and LBWA), with a few inverted positions (between companies A1 and A9 in the first and second place) and between companies A3 and A6 in the 4th and 5th place for the simplified-AHP).

The consistency of the ranking results confirms the medium robustness of the solution (mainly due to the uncertainty between the first and second positions), which can be critical for creating a ranking of companies and providing a benchmark. The robustness of the solution as the weighting method varies determines the validation of the ranking procedure using the parsimonious weighting methods.

5. Conclusions

This study presents the development of a SC observatory. This is a practical tool aimed at supporting companies within the same industry in understanding their positioning in terms of SC performance from a benchmarking perspective.

Method	Discriminating capacity
s-AHP	0.34903
AHP-e	0.34761
FUCOM	0.34731
NDSL	0.31030
LBWA	0.30875

Table 5. Discriminating capacity calculated for the five methods

Source(s): Table created by authors

	C1.1	C1.2	C2.1	C2.2	C2.3	C2.4	C2.5	C2.6	C3.1	C3.2	C3.3	C3.4	C4.1	C4.2	C4.3	C4.4	C5.1	C5.2	C5.3	C5.4	C5.5	C5.6
A1	311319	77.5438	6.9922	81.0762	73.6006	86.3983	96.5519	12.9772	8.5882	85.0811	4.4987	68.7456	3.4685	0.9985	14.5337	26.787	7.2757	4.3572	1.9129	8.5195	2.7111	2.6441
A2	-3.3283	84.4855	7.1918	71.67	88.6399	72.3531	74.7679	26.7683	11.7402	78.3909	18.1321	56.945	3.5233	3.4717	22.3463	7.7464	28.1708	0.1952	4.5212	9.8916	8.7018	4.8474
A3	0.8786	89.5683	9.9687	88.1053	77.6556	82.7572	81.3198	2.464	15.6796	91.1456	17.2218	95.9745	3.0614	0.9787	15.6402	24.8897	21.7308	5.6795	0.4371	6.8181	3.7776	7.198
A4	-4.5326	75.6016	5.9174	81.8637	71.7201	94.1234	84.2061	19.3188	14.8486	86.8828	18.0712	54.0353	6.8548	8.9322	21.5427	14.5105	28.435	4.585	1.3211	1.358	3.0862	5.7205
A5	-4.4318	83.8628	7.5362	84.2101	81.2605	77.9129	82.6035	18.0521	9.3189	86.8584	19.8765	54.7897	8.744	6.5484	19.9797	11.565	6.2232	6.4808	3.3218	0.6877	3.3436	6.8204
A6	1.0813	80.4502	6.7504	95.9989	75.8913	93.523	94.7172	19.9653	9.7596	83.6834	2.7318	63.2998	8.9648	0.6638	28.9872	6.9902	14.0472	9.5145	8.2271	8.0948	8.6111	7.3837
A7	-0.5008	79.5622	8.3594	95.095	71.7413	99.0326	72.572	9.7622	10.1035	76.5123	19.0745	86.8534	3.3167	8.0916	9.618	15.9224	6.7003	6.9606	9.248	2.4318	3.5593	3.8945
A8	2.3183	93.1194	6.6198	86.3948	73.2028	84.0285	85.9097	2.1544	17.792	82.2927	4.5777	98.2994	9.7438	1.1549	0.7104	1.4234	27.2281	9.7455	2.9024	7.8979	2.8418	7.4707
A9	4.4001	89.8946	6.626	72.7733	94.1585	86.2387	74.2207	25.4595	19.4045	73.4841	1.8825	53.0782	4.396	9.2826	8.2505	25.4194	22.4037	8.9278	1.1498	7.4359	7.1551	2.2124
A10	-3.521	82.9369	11.6959	99.5867	79.5458	94.82	83.5959	2.6101	1.6044	78.4255	7.5579	64.9827	1.9536	6.2836	22.5095	6.0333	16.2009	9.5968	9.791	5.433	7.2023	5.6331

Source(s): Table created by authors

Table 6.
Scores associated with
each criterion for each
company (A1 to A10) in
the sample

	S-AHP	AHP-e	Global weights FUCOM	NDSL	LBWA
C1,1	0.159873	0.197906	0.201588	0.148851	0.151291
C1,2	0.038587	0.047767	0.048656	0.037213	0.042192
C2,1	0.100164	0.067224	0.066762	0.063493	0.068778
C2,2	0.046223	0.042826	0.042531	0.04744	0.040036
C2,3	0.072336	0.026735	0.026551	0.035297	0.036686
C2,4	0.126604	0.076929	0.0764	0.065279	0.058516
C2,5	0.077372	0.086798	0.086201	0.079023	0.075384
C2,6	0.045039	0.026975	0.026789	0.03619	0.036665
C3,1	0.078977	0.02436	0.024179	0.042398	0.044752
C3,2	0.009868	0.035422	0.035158	0.03179	0.032248
C3,3	0.019747	0.031068	0.030837	0.040664	0.039843
C3,4	0.009868	0.009763	0.009691	0.017853	0.021324
C4,1	0.029579	0.042104	0.041911	0.046936	0.04948
C4,2	0.016355	0.025806	0.025688	0.034698	0.031146
C4,3	0.006263	0.015556	0.015485	0.019364	0.020444
C4,4	0.006263	0.01455	0.014483	0.020892	0.02021
C5,1	0.030232	0.033909	0.033742	0.037623	0.03708
C5,2	0.01315	0.024249	0.024129	0.03166	0.031729
C5,3	0.008537	0.020103	0.020004	0.015291	0.014983
C5,4	0.016383	0.024058	0.02394	0.019463	0.019364
C5,5	0.048694	0.069787	0.069444	0.066592	0.065768
C5,6	0.039922	0.056081	0.055805	0.06201	0.062072

Table 7.
Global weights assigned to the criteria, calculated according to equation (3)

Source(s): Table created by authors

Ranking	Simplified-AHP	AHP-express	FUCOM	NDSL	LBWA
1	A1	A9	A9	A1	A1
2	A9	A1	A1	A9	A9
3	A8	A8	A8	A8	A8
4	A6	A3	A3	A3	A3
5	A3	A6	A6	A6	A6
6	A7	A7	A7	A7	A7
7	A10	A10	A10	A10	A10
8	A2	A2	A2	A2	A2
9	A5	A5	A5	A5	A5
10	A4	A4	A4	A4	A4

Table 8.
Company rankings according to the TOPSIS method

Source(s): Table created by authors

A case study is used to demonstrate the set-up of the observatory, involving twelve professionals with expertise in automatic equipment for the wrapping and packaging industry. These experts selected and evaluated a set of relevant criteria, identifying the company areas that could be improved using MCDM techniques. Ten companies were ranked with their respective positioning in terms of SC performance, comparing five parsimonious subjective weighting criteria methods selected to reduce the cognitive efforts of the DMs.

As previously mentioned, the ranking solution of the companies obtained was on average robust, since despite some changes in the order, the general structure of the ranking was very similar for all the weighting methodologies. It is possible to conclude that methodologies with a similar structure identify similar solutions. With regard to the subjective perceptions of the DMs, simplified-AHP was considered the best method in almost all aspects: faster to carry out and more quickly understood. The only aspect for which other methods were better is the

perceived accuracy in giving preferences; in this case, the AHP-express and FUCOM methods prevailed. Simplified AHP showed the greatest ability to discriminate between the criteria importance, although it sometimes provided weights that were visibly higher or lower than the average. This may be due to the fact that the pairs of criteria to assign a judgment to are always different; that is not the case with AHP-express or FUCOM, where the first term of comparison is always the most important criterion, leading the DM to pay more attention to the assignment of the score. These considerations should be taken into account when choosing a method.

The main theoretical implications of this study concern the robustness of the obtained results when applying different parsimonious weighting methods, which are strongly recommended in cases of a high number of criteria (Pamučar *et al.*, 2018). Each of the method can be used for improvement actions starting from a solid basis.

The obtained results have some practical and managerial implications, which help us respond to RQ2. With a few adjustments, such as shifting the weights and criteria, the presented general methodology can be applied to many industries and take into account evolving SC strategies and policies. So, it paves the way for models of resource rationalization from the perspective of continuous improvement, according to a competitive logic aimed at a whole sector.

To give an answer to RQ1, we observe that the use of parsimonious weighting methods has shown good accuracy in the calculation of the weight values and a good degree of appreciation by DMs, enhancing the appeal of our system for corporate applications where DMs must work under time restrictions.

However, the study has some limitations. The choice of the criteria does not include innovative SC aspects such as digitalization-related indicators or sustainability-related indicators; this highlights the need for a wider discussion with the experts to make the set of indicators as exhaustive and inclusive as possible (Mishra *et al.*, 2018). This improvement will positively impact not only the SC performance but also the whole society, improving sustainability at the global level.

As future research directions, the next step of our research foresees the application of the post factum analysis to the obtained ranking to achieve the minimum variation on one or more criteria to allow companies to reach a specific position. This can be very useful in the context of competitive benchmarking: once a company has learned its position in the ranking, it is essential to understand which business areas and which indicators to invest in, in order to improve its positioning compared to competing companies.

We also plan to define a set of qualitative criteria in order to use a combined qualitative-quantitative framework in the evaluation process.

Note

1. Source: UCIMA (Unione Costruttori Italiani Macchine Automatiche per il Confezionamento e l'Imballaggio), <https://www.ucima.it/servizio/analisi-centro-studi-mecs/notizia/1014/macchine-packaging-2021-da-record-per-il-settore/>

References

- Alfares, H.K. and Duffuaa, S.O. (2008), "Determining aggregate criteria weights from criteria rankings by a group of decision makers", *International Journal of Information Technology and Decision Making*, Vol. 7 No. 4, pp. 769-781, doi: [10.1142/S0219622008003174](https://doi.org/10.1142/S0219622008003174).
- Ali, S.S. and Kaur, R. (2021), "Effectiveness of corporate social responsibility (CSR) in implementation of social sustainability in warehousing of developing countries: a hybrid approach", *Journal of Cleaner Production*, Vol. 324, September, 129154, doi: [10.1016/j.jclepro.2021.129154](https://doi.org/10.1016/j.jclepro.2021.129154).
- Ali, S.S., Kaur, R., Ersöz, F., Altaf, B., Basu, A. and Weber, G.W. (2020), "Measuring carbon performance for sustainable green supply chain practices: a developing country scenario", *Central European Journal of Operations Research*, Vol. 28 No. 4, pp. 1389-1416, doi: [10.1007/s10100-020-00673-x](https://doi.org/10.1007/s10100-020-00673-x).

- Ali, S.S., Ersöz, F., Kaur, R., Altaf, B. and Weber, G.W. (2021), "A quantitative analysis of low carbon performance in industrial sectors of developing world", *Journal of Cleaner Production*, Vol. 284, 125268, doi: [10.1016/j.jclepro.2020.125268](https://doi.org/10.1016/j.jclepro.2020.125268).
- Ali, S.S., Kaur, R., Persis, D.J., Saha, R., Pattusamy, M. and Sreedharan, V.R. (2023a), "Developing a hybrid evaluation approach for the low carbon performance on sustainable manufacturing environment", *Annals of Operations Research*, Vol. 324 Nos 1-2, pp. 249-281, doi: [10.1007/s10479-020-03877-1](https://doi.org/10.1007/s10479-020-03877-1).
- Ali, S.S., Kaur, R. and Khan, S. (2023b), "Evaluating sustainability initiatives in warehouse for measuring sustainability performance: an emerging economy perspective", *Annals of Operations Research*, Vol. 324 Nos 1-2, pp. 461-500, doi: [10.1007/s10479-021-04454-w](https://doi.org/10.1007/s10479-021-04454-w).
- Anand, N. and Grover, N. (2015), "Measuring retail supply chain performance: theoretical model using key performance indicators (KPIs)", *Benchmarking*, Vol. 22 No. 1, pp. 135-166, doi: [10.1108/BIJ-05-2012-0034](https://doi.org/10.1108/BIJ-05-2012-0034).
- Anas, Z., Sefiani, N., Reklouli, K. and Azzouzi, H. (2018), "A SCOR based performance evaluation system of the hospital supply chain", *2018 International Colloquium on Logistics and Supply Chain Management (LOGISTIQUA)*, Tangier, Morocco, 2018, pp. 182-187, doi: [10.1109/LOGISTIQUA.2018.8428273](https://doi.org/10.1109/LOGISTIQUA.2018.8428273).
- Arshinder, Kanda, A. and Deshmukh, S.G. (2007), "Coordination in supply chains: an evaluation using fuzzy logic", *Production Planning and Control*, Vol. 18 No. 5, pp. 420-435, doi: [10.1080/09537280701430994](https://doi.org/10.1080/09537280701430994).
- Balfaqih, H., Nopiah, Z.M., Saibani, N. and Al-Nory, M.T. (2016), "Review of supply chain performance measurement systems: 1998-2015", *Computers in Industry*, Vol. 82, pp. 135-150, doi: [10.1016/j.compind.2016.07.002](https://doi.org/10.1016/j.compind.2016.07.002).
- Beamon, B.M. (1999), "Measuring supply chain performance", *International Journal of Operations and Production Management*, Vol. 19 No. 3, pp. 275-292, doi: [10.1108/01443579910249714](https://doi.org/10.1108/01443579910249714).
- Behzadian, M., Khanmohammadi Otagh Sara, S., Yazdani, M. and Ignatius, J. (2012), "A state-of-the-art survey of TOPSIS applications", *Expert Systems with Applications*, Vol. 39 No. 17, pp. 13051-13069, doi: [10.1016/j.eswa.2012.05.056](https://doi.org/10.1016/j.eswa.2012.05.056).
- Benitez, J., Carpitella, S., Certa, A. and Izquierdo, J. (2019), "Sampling of pairwise comparisons in decisions making", *Mathematical Modelling in Engineering and Human Behavior*, Instituto Universitario de Matemática Multidisciplinar, Valencia, available at: <https://riunet.upv.es/bitstream/handle/10251/180550/BenitezCarpitellaCerta%20-%20Sampling%20of%20pairwise%20comparisons%20in%20decision-making.pdf?sequence=1>
- Bigliardi, B. and Bottani, E. (2014), "Supply chain performance measurement: a literature review and pilot study among Italian manufacturing companies", *International Journal of Engineering, Science and Technology*, Vol. 6 No. 3, p. 1, doi: [10.4314/ijest.v6i3.1s](https://doi.org/10.4314/ijest.v6i3.1s).
- Borcherding, K., Eppel, T. and von Winterfeldt, D. (1991), "Comparison of weighting judgements in multiattribute utility measurement", *Management Science*, Vol. 37 No. 12, pp. 1603-1619, doi: [10.1287/mnsc.37.12.1603](https://doi.org/10.1287/mnsc.37.12.1603).
- Bottomley, P.A. and Doyle, J.R. (2001), "A comparison of three weight elicitation methods: good, better, and best", *Omega*, Vol. 29 No. 6, pp. 553-560, doi: [10.1016/S0305-0483\(01\)00044-5](https://doi.org/10.1016/S0305-0483(01)00044-5).
- Buchanan, J.T. and Corner, J.L. (1997), "The effects of anchoring in interactive MCDM solution methods", *Computers and Operations Research*, Vol. 24 No. 10, pp. 907-918, doi: [10.1016/S0305-0548\(97\)00014-2](https://doi.org/10.1016/S0305-0548(97)00014-2).
- Chand, P., Thakkar, J.J. and Ghosh, K.K. (2020), "Analysis of supply chain performance metrics for Indian mining & earthmoving equipment manufacturing companies using hybrid MCDM model", *Resources Policy*, Vol. 68, April, 101742, doi: [10.1016/j.resourpol.2020.101742](https://doi.org/10.1016/j.resourpol.2020.101742).
- Chithambaranathan, P., Subramanian, N., Gunasekaran, A. and Palaniappan, P. (2015), "Service supply chain environmental performance evaluation using grey based hybrid MCDM approach", *International Journal of Production Economics*, Vol. 166, pp. 163-176, doi: [10.1016/j.ijpe.2015.01.002](https://doi.org/10.1016/j.ijpe.2015.01.002).
- Deng, H., Yeh, C.H. and Willis, R.J. (2000), "Inter-company comparison using modified TOPSIS with objective weights", *Computers and Operations Research*, Vol. 27 No. 10, pp. 963-973, doi: [10.1016/S0305-0548\(99\)00069-6](https://doi.org/10.1016/S0305-0548(99)00069-6).

-
- Deniz, N. (2020), "Cognitive biases in MCDM methods: an embedded filter proposal through sustainable supplier selection problem", *Journal of Enterprise Information Management*, Vol. 33 No. 5, pp. 947-963, doi: [10.1108/JEIM-09-2019-0285](https://doi.org/10.1108/JEIM-09-2019-0285).
- Dobrzykowski, D., Hong, S.W., Jungbae Roh, J. and Park, K. (2012), "Evolving benchmarking practices: a review for research perspectives", *Benchmarking: An International Journal*, Vol. 19 Nos 4-5, pp. 444-462, doi: [10.1108/14635771211257945](https://doi.org/10.1108/14635771211257945).
- Erol, I., Sencer, S. and Sari, R. (2011), "A new fuzzy multi-criteria framework for measuring sustainability performance of a supply chain", *Ecological Economics*, Vol. 70 No. 6, pp. 1088-1100, doi: [10.1016/j.ecolecon.2011.01.001](https://doi.org/10.1016/j.ecolecon.2011.01.001).
- Fleisch, E. and Tellkamp, C. (2005), "Inventory inaccuracy and supply chain performance: a simulation study of a retail supply chain", *International Journal of Production Economics*, Vol. 95 No. 3, pp. 373-385, doi: [10.1016/j.ijpe.2004.02.003](https://doi.org/10.1016/j.ijpe.2004.02.003).
- Govindan, K., Mangla, S.K. and Luthra, S. (2017), "Prioritising indicators in improving supply chain performance using fuzzy AHP: insights from the case example of four Indian manufacturing companies", *Production Planning and Control*, Vol. 28 Nos 6-8, pp. 552-573, doi: [10.1080/09537287.2017.1309716](https://doi.org/10.1080/09537287.2017.1309716).
- Gunasekaran, A., Patel, C. and McGaughey, R.E. (2004), "A framework for supply chain performance measurement", *International Journal of Production Economics*, Vol. 87 No. 3, pp. 333-347, doi: [10.1016/j.ijpe.2003.08.003](https://doi.org/10.1016/j.ijpe.2003.08.003).
- Hallikas, J., Immonen, M. and Brax, S. (2021), "Digitalizing procurement: the impact of data analytics on supply chain performance", *Supply Chain Management*, Vol. 26 No. 5, pp. 629-646, doi: [10.1108/SCM-05-2020-0201](https://doi.org/10.1108/SCM-05-2020-0201).
- Hoek, R.I.V. (1998), "Measuring the unmeasurable"- measuring and improving performance in the supply chain", *Supply Chain Management*, Vol. 3 No. 4, pp. 187-192, doi: [10.1108/13598549810244232](https://doi.org/10.1108/13598549810244232).
- Ishizaka, A. and Labib, A. (2011), "Review of the main developments in the analytic hierarchy process", *Expert Systems with Applications*, Vol. 38 No. 11, pp. 14336-14345, doi: [10.1016/j.eswa.2011.04.143](https://doi.org/10.1016/j.eswa.2011.04.143).
- Jagan Mohan Reddy, K., Neelakanteswara Rao, A. and Krishnanand, L. (2019), "A review on supply chain performance measurement systems", *Procedia Manufacturing*, Vol. 30, pp. 40-47, doi: [10.1016/j.promfg.2019.02.007](https://doi.org/10.1016/j.promfg.2019.02.007).
- Jha, A., Sharma, R., Kumar, V. and Verma, P. (2022), "Designing supply chain performance system: a strategic study on Indian manufacturing sector", *Supply Chain Management*, Vol. 27 No. 1, pp. 66-88, doi: [10.1108/SCM-05-2020-0198](https://doi.org/10.1108/SCM-05-2020-0198).
- Joshi, R., Banwet, D.K. and Shankar, R. (2011), "A Delphi-AHP-TOPSIS based benchmarking framework for performance improvement of a cold chain", *Expert Systems with Applications*, Vol. 38 No. 8, pp. 10170-10182, doi: [10.1016/j.eswa.2011.02.072](https://doi.org/10.1016/j.eswa.2011.02.072).
- Kailash, Saha, R.K. and Goyal, S. (2017), "Systematic literature review of classification and categorisation of benchmarking in supply chain management", *International Journal of Process Management and Benchmarking*, Vol. 7 No. 2, pp. 183-205, doi: [10.1504/IJPMB.2017.083105](https://doi.org/10.1504/IJPMB.2017.083105).
- Khan, S.A., Chaabane, A. and Dweiri, F. (2019), "A knowledge-based system for overall supply chain performance evaluation: a multi-criteria decision making approach", *Supply Chain Management*, Vol. 24 No. 3, pp. 377-396, doi: [10.1108/SCM-06-2017-0197](https://doi.org/10.1108/SCM-06-2017-0197).
- Krishnamoorthy, B., D'Lima, C. and Lima, N. (2014), "Benchmarking as a measure of competitiveness", *International Journal of Process Management and Benchmarking*, Vol. 4 No. 3, pp. 342-359, doi: [10.1504/IJPMB.2014.063240](https://doi.org/10.1504/IJPMB.2014.063240).
- Kumar, S. and Barua, M.K. (2022), "Exploring and measure the performance of the Indian petroleum supply chain", *International Journal of Productivity and Performance Management*, Vol. 71 No. 6, pp. 2114-2139, doi: [10.1108/IJPPM-12-2020-0640](https://doi.org/10.1108/IJPPM-12-2020-0640).
- Kumar, N., Tyagi, M. and Sachdeva, A. (2022), "Depiction of possible solutions to improve the cold supply chain performance system", *Journal of Advances in Management Research*, Vol. 19 No. 1, pp. 106-138, doi: [10.1108/JAMR-10-2020-0285](https://doi.org/10.1108/JAMR-10-2020-0285).

-
- Lai, K., Ngai, E.W. and Cheng, T.C. (2002), "Measures for evaluating supply chain performance in transport logistics", *Transportation Research E: Logistics and Transportation Review*, Vol. 38 No. 6, pp. 439-456, doi: [10.1016/S1366-5545\(02\)00019-4](https://doi.org/10.1016/S1366-5545(02)00019-4).
- Leal, J.E. (2020), "AHP-express: a simplified version of the analytical hierarchy process method", *MethodsX*, Vol. 7, 100748, doi: [10.1016/j.mex.2019.11.021](https://doi.org/10.1016/j.mex.2019.11.021).
- Lolli, F., Balugani, E., Ishizaka, A., Gamberini, R., Butturi, M.A., Marinello, S. and Rimini, B. (2019), "On the elicitation of criteria weights in PROMETHEE-based ranking methods for a mobile application", *Expert Systems with Applications*, Vol. 120, pp. 217-227, doi: [10.1016/j.eswa.2018.11.030](https://doi.org/10.1016/j.eswa.2018.11.030).
- Lolli, F., Balugani, E., Butturi, M.A., Coruzzolo, A.M., Ishizaka, A., Marinelli, S. and Romano, V. (2022), "A decision support system for the selection of insulating material in energy retrofit of industrial buildings: a new robust ordinal regression approach", *IEEE Transactions on Engineering Management*, Vol. 71, pp. 1-12, doi: [10.1109/TEM.2022.3160055](https://doi.org/10.1109/TEM.2022.3160055).
- Luthra, S., Govindan, K., Kannan, D., Mangla, S.K. and Garg, C.P. (2017), "An integrated framework for sustainable supplier selection and evaluation in supply chains", *Journal of Cleaner Production*, Vol. 140, pp. 1686-1698, doi: [10.1016/j.jclepro.2016.09.078](https://doi.org/10.1016/j.jclepro.2016.09.078).
- Maestrini, V., Luzzini, D., Maccarrone, P. and Caniato, F. (2017), "Supply chain performance measurement systems: a systematic review and research agenda", *International Journal of Production Economics*, Vol. 183, May 2016, pp. 299-315, doi: [10.1016/j.ijpe.2016.11.005](https://doi.org/10.1016/j.ijpe.2016.11.005).
- Mañay, L.O.R., Guaita-Pradas, I. and Marques-Perez, I. (2022), "Measuring the supply chain performance of the floricultural sector using the SCOR model and a multicriteria decision-making method", *Horticulturae*, Vol. 8 No. 2, p. 168, doi: [10.3390/horticulturae8020168](https://doi.org/10.3390/horticulturae8020168).
- Menon, R.R. and Ravi, V. (2022), "Using AHP-TOPSIS methodologies in the selection of sustainable suppliers in an electronics supply chain", *Cleaner Materials*, The Author(s), 100130, doi: [10.1016/j.clema.2022.100130](https://doi.org/10.1016/j.clema.2022.100130).
- Mishra, D., Gunasekaran, A., Papadopoulos, T. and Dubey, R. (2018), "Supply chain performance measures and metrics: a bibliometric study", *Benchmarking*, Vol. 25 No. 3, pp. 932-967, doi: [10.1108/BIJ-08-2017-0224](https://doi.org/10.1108/BIJ-08-2017-0224).
- Németh, B., Molnár, A., Bozóki, S., Wijaya, K., Inotai, A., Campbell, J.D. and Kaló, Z. (2019), "Comparison of weighting methods used in multicriteria decision analysis frameworks in healthcare with focus on low-and middle-income countries", *Journal of Comparative Effectiveness Research*, Vol. 8 No. 4, pp. 195-204, doi: [10.2217/cer-2018-0102](https://doi.org/10.2217/cer-2018-0102).
- Oubrahim, I. and Sefiani, N. (2022), "Supply chain performance measurement systems: benefits and drawbacks", *International Journal of Latest Engineering and Management Research (IJLEMR)*, Vol. 7 No. 9, pp. 24-28, doi: [10.56581/ijlemr.7.9.24-28](https://doi.org/10.56581/ijlemr.7.9.24-28).
- Oubrahim, I., Sefiani, N. and Haponen, A. (2022), "Supply chain performance evaluation models: a literature review", *Acta Logistica*, Vol. 9 No. 2, pp. 207-221, doi: [10.22306/al.v9i2.298](https://doi.org/10.22306/al.v9i2.298).
- Pamučar, D., Stević, Ž. and Sremac, S. (2018), "A new model for determining weight coefficients of criteria in MCDM models: full Consistency Method (FUCOM)", *Symmetry*, Vol. 10 No. 9, pp. 1-22, doi: [10.3390/sym10090393](https://doi.org/10.3390/sym10090393).
- Pöyhönen, M. and Hämäläinen, R.P. (2001), "On the convergence of multiattribute weighting methods", *European Journal of Operational Research*, Vol. 129 No. 3, pp. 569-585, doi: [10.1016/S0377-2217\(99\)00467-1](https://doi.org/10.1016/S0377-2217(99)00467-1).
- Qorri, A., Gashi, S. and Kraslawski, A. (2022), "A practical method to measure sustainability performance of supply chains with incomplete information", *Journal of Cleaner Production*, Vol. 341, January, 130707, doi: [10.1016/j.jclepro.2022.130707](https://doi.org/10.1016/j.jclepro.2022.130707).
- Ramos, E., Coles, P.S., Chavez, M. and Hazen, B. (2022), "Measuring agri-food supply chain performance: insights from the Peruvian kiwicha industry", *Benchmarking*, Vol. 29 No. 5, pp. 1484-1512, doi: [10.1108/BIJ-10-2020-0544](https://doi.org/10.1108/BIJ-10-2020-0544).

-
- Rezaei, J. (2015), "Best-worst multi-criteria decision-making method", *Omega (United Kingdom)*, Vol. 53, pp. 49-57, doi: [10.1016/j.omega.2014.11.009](https://doi.org/10.1016/j.omega.2014.11.009).
- Rezaei, J. (2021), "Anchoring bias in eliciting attribute weights and values in multi-attribute decision-making", *Journal of Decision Systems*, Vol. 30 No. 1, pp. 72-96, doi: [10.1080/12460125.2020.1840705](https://doi.org/10.1080/12460125.2020.1840705).
- Riabacke, M., Danielson, M. and Ekenberg, L. (2012), "State-of-the-art prescriptive criteria weight elicitation", *Advances in Decision Sciences*, Vol. 2012, pp. 1-24, doi: [10.1155/2012/276584](https://doi.org/10.1155/2012/276584).
- Saaty, T.L. (1977), "A scaling method for priorities in hierarchical structures", *Journal of Mathematical Psychology*, Vol. 15 No. 3, pp. 234-281, doi: [10.1016/0022-2496\(77\)90033-5](https://doi.org/10.1016/0022-2496(77)90033-5).
- Şahin, M. (2021), "A comprehensive analysis of weighting and multicriteria methods in the context of sustainable energy", *International Journal of Environmental Science and Technology*, Vol. 18 No. 6, pp. 1591-1616, doi: [10.1007/s13762-020-02922-7](https://doi.org/10.1007/s13762-020-02922-7).
- Sarkis, J. (2003), "A strategic decision framework for green supply chain management", *Journal of Cleaner Production*, Vol. 11 No. 4, pp. 397-409, doi: [10.1016/S0959-6526\(02\)00062-8](https://doi.org/10.1016/S0959-6526(02)00062-8).
- Si, S.L., You, X.Y., Liu, H.C. and Zhang, P. (2018), "DEMATEL technique: a systematic review of the state-of-the-art literature on methodologies and applications", *Mathematical Problems in Engineering*, Vol. 2018 No. 1, pp. 1-33, doi: [10.1155/2018/3696457](https://doi.org/10.1155/2018/3696457).
- Sufiyan, M., Haleem, A., Khan, S. and Khan, M.I. (2019), "Evaluating food supply chain performance using hybrid fuzzy MCDM technique", *Sustainable Production and Consumption*, Vol. 20, pp. 40-57, doi: [10.1016/j.spc.2019.03.004](https://doi.org/10.1016/j.spc.2019.03.004).
- Thakkar, J., Kanda, A. and Deshmukh, S.G. (2009), "Supply chain performance measurement framework for small and medium scale enterprises", *Benchmarking*, Vol. 16 No. 5, pp. 702-723, doi: [10.1108/14635770910987878](https://doi.org/10.1108/14635770910987878).
- Trkman, P., McCormack, K., de Oliveira, M.P.V. and Ladeira, M.B. (2010), "The impact of business analytics on supply chain performance", *Decision Support Systems*, Vol. 49 No. 3, pp. 318-327, doi: [10.1016/j.dss.2010.03.007](https://doi.org/10.1016/j.dss.2010.03.007).
- Tyagi, M., Kumar, P. and Kumar, D. (2014), "A hybrid approach using AHP-TOPSIS for analyzing e-SCM performance", *Procedia Engineering*, Vol. 97, pp. 2195-2203, doi: [10.1016/j.proeng.2014.12.463](https://doi.org/10.1016/j.proeng.2014.12.463).
- Vaidya, O. and Hudnurkar, M. (2013), "Multi-criteria supply chain performance evaluation: an Indian chemical industry case study", *International Journal of Productivity and Performance Management*, Vol. 62 No. 3, pp. 293-316, doi: [10.1108/17410401311309195](https://doi.org/10.1108/17410401311309195).
- van Til, J., Groothuis-Oudshoorn, C., Lieferink, M., Dolan, J. and Goetghebeur, M. (2014), "Does technique matter; a pilot study exploring weighting techniques for a multi-criteria decision support framework", *Cost Effectiveness and Resource Allocation*, Vol. 12 No. 1, pp. 1-11, doi: [10.1186/1478-7547-12-22](https://doi.org/10.1186/1478-7547-12-22).
- Varma, S., Wadhwa, S. and Deshmukh, S.G. (2008), "Evaluating petroleum supply chain performance: application of analytical hierarchy process to balanced scorecard", *Asia Pacific Journal of Marketing and Logistics*, Vol. 20 No. 3, pp. 343-356, doi: [10.1108/13555850810890093](https://doi.org/10.1108/13555850810890093).
- Velasquez, M. and Hester, P. (2013), "An analysis of multi-criteria decision making methods", *International Journal of Operations Research*, Vol. 10 No. 2, pp. 56-66.
- Vinogradova, I., Podvezko, V. and Zavadskas, E.K. (2018), "The recalculation of the weights of criteria in MCDM methods using the Bayes approach", *Symmetry*, Vol. 10 No. 6, pp. 1-18, doi: [10.3390/sym10060205](https://doi.org/10.3390/sym10060205).
- Weber, M. and Borcherding, K. (1993), "Behavioral influences on weight judgments in multiattribute decision making", *European Journal of Operational Research*, Vol. 67 No. 1, pp. 1-12, doi: [10.1016/0377-2217\(93\)90318-H](https://doi.org/10.1016/0377-2217(93)90318-H).
- Wilson, M.C. (2007), "The impact of transportation disruptions on supply chain performance", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 43 No. 4, pp. 295-320, doi: [10.1016/j.tre.2005.09.008](https://doi.org/10.1016/j.tre.2005.09.008).
- Wong, W.P. and Wong, K.Y. (2008), "A review on benchmarking of supply chain performance measures", *Benchmarking*, Vol. 15 No. 1, pp. 25-51, doi: [10.1108/14635770810854335](https://doi.org/10.1108/14635770810854335).
-

Žižović, M. and Pamučar, D. (2019), "New model for determining criteria weights: level based weight assessment (lbwa) model", *Decision Making: Applications in Management and Engineering*, Vol. 2 No. 2, pp. 126-137, doi: [10.31181/dmame1902102z](https://doi.org/10.31181/dmame1902102z).

Žižović, M., Pamučar, D., Čirović, G., Žižović, M.M. and Miljković, B.D. (2020), "A model for determining weight coefficients by forming a non-decreasing series at criteria significance levels (NDSL)", *Mathematics*, Vol. 8 No. 5, p. 745, doi: [10.3390/math8050745](https://doi.org/10.3390/math8050745).

Appendix

Appendix 1

The questionnaire includes 4 closed-ended questions, given that the experts had to choose one of the 5 parsimonious methods used. The AHP-express and FUCOM methods were grouped into a single point, since the judgments and steps required by the DMs for the method to work are the same.

- (1) Which method was the quickest to apply?
- (2) Which method was the quickest to understand?
- (3) Which method enabled you to most accurately express your preference?
- (4) Which method did you prefer overall?

Appendix 2

TOPSIS implementation steps

This appendix presents a summary of the TOPSIS method (Behzadian *et al.*, 2012) for the ranking of a number of alternatives (the 10 companies in this study).

In the presented case study, 10 alternatives are considered (i.e. 10 companies), evaluated over 22 criteria.

Step 1: Normalization

$x_{i,j}$ being the score of the i -th alternative with respect to the j -th criterion, the normalized decision matrix is built according to the (1.1):

$$r_{ij} = \frac{x_{ij}}{\sqrt{\left(\sum x_{ij}^2\right)}} \text{ for } i = 1, \dots, m; j = 1, \dots, n \quad (1.1)$$

Step 2: Weighting

Starting from the assigned weights w_j to the j -th criterion, the weighted normalized decision matrix is calculated (1.2):

$$v_{ij} = w_j \cdot r_{ij} \quad (1.2)$$

Step 3: Identifying the positive-ideal A^* (1.3) and negative-ideal A' solutions (1.4)

$$A^* = \{v_1^*, \dots, v_n^*\}, \text{ where } v_j^* = \{\max(v_{ij}) \text{ if } j \in J; \min(v_{ij}) \text{ if } j \in J'\}, \quad (1.3)$$

$$A' = \{v_1', \dots, v_n'\}, \text{ where } v_j' = \{\min(v_{ij}) \text{ if } j \in J; \max(v_{ij}) \text{ if } j \in J'\}, \quad (1.4)$$

where J is the set of benefit criteria and J' the set of the cost criteria.

Step 4: Calculation of the separation of each alternative from positive-ideal and negative-ideal solutions

The separation from the positive-ideal solution is (1.5):

$$S_i^* = \sqrt{\sum (v_j^* - v_{ij})^2}, i = 1, \dots, m \tag{1.5}$$

The separation from the negative-ideal solution is (1.6):

$$S_i' = \sqrt{\sum (v_j' - v_{ij})^2}, i = 1, \dots, m \tag{1.6}$$

Step 5: Calculation of the similarity to the ideal solution C_i^* (1.7)

$$C_i^* = \frac{S_i'}{S_i^* + S_i'}, 0 < C_i^* < 1 \tag{1.7}$$

Step 6: Ranking preference order

The best alternative has C_i^* closest to 1.

Appendix 3

Criteria weighting using five parsimonious methods

Tables A1 to A5 of Appendix 3 present the numerical results of the criteria weighting process implemented by the experts according to the five methods. The simplified AHP method is referred to as “S-AHP”, and the AHP-express method as “AHP-e”.

The resulting weights calculated using equation (1) are summarized in Table A7, while the weights of the macro-categories, calculated with eq. (2), are shown in Table A8.

	E1					E2					E3				
	S-AHP	AHP-e	FUCOM	NDSL	LBWA	S-AHP	AHP-e	FUCOM	NDSL	LBWA	S-AHP	AHP-e	FUCOM	NDSL	LBWA
$C_{1,1}$	0.875	0.875	0.875	0.838	0.833	0.875	0.875	0.875	0.875	0.813	0.667	0.667	0.667	0.688	0.700
$C_{1,2}$	0.125	0.125	0.125	0.163	0.167	0.125	0.125	0.125	0.125	0.188	0.333	0.333	0.333	0.313	0.300

Table A1.
Weights of the criteria belonging to the macro-category purchase as assigned by the three experts

	E1					E2					E3				
	S-AHP	AHP-e	FUCOM	NDSL	LBWA	S-AHP	AHP-e	FUCOM	NDSL	LBWA	S-AHP	AHP-e	FUCOM	NDSL	LBWA
$C_{2,1}$	0.3553	0.25	0.25	0.24	0.2815	0.0417	0.1087	0.1087	0.1221	0.1265	0.2455	0.2571	0.2571	0.2209	0.2448
$C_{2,2}$	0.1184	0.25	0.25	0.24	0.1877	0.0417	0.0652	0.0652	0.1067	0.1107	0.1364	0.0771	0.0771	0.0889	0.0816
$C_{2,3}$	0.0132	0.0833	0.0833	0.08	0.0866	0.0417	0.0652	0.0652	0.0925	0.0984	0.4091	0.0964	0.0964	0.1516	0.1632
$C_{2,4}$	0.3553	0.25	0.25	0.24	0.2252	0.375	0.3261	0.3261	0.2625	0.2214	0.0818	0.1286	0.1286	0.0969	0.1088
$C_{2,5}$	0.0395	0.0833	0.0833	0.08	0.0938	0.375	0.3261	0.3261	0.2775	0.2953	0.0818	0.3857	0.3857	0.3681	0.3264
$C_{2,6}$	0.1184	0.0833	0.0833	0.12	0.1251	0.125	0.1087	0.1087	0.1387	0.1476	0.0455	0.0551	0.0551	0.0736	0.0753

Table A2.
Weights of the criteria belonging to the macro-category planning as assigned by the three experts

Source(s): Table created by authors

Table A3.

Weights of the criteria belonging to the macro-category internal logistics as assigned by the three experts

	E1					E2					E3				
	S-AHP	AHP-e	FUCOM	NDSL	LBWA	S-AHP	AHP-e	FUCOM	NDSL	LBWA	S-AHP	AHP-e	FUCOM	NDSL	LBWA
C3,1	-	0.2	0.2	0.1789	0.1977	0.6667	0.3947	0.3947	0.5101	0.4988	-	0.1316	0.1316	0.2695	0.2752
C3,2	-	0.2667	0.2667	0.3187	0.2373	0.0833	0.3947	0.3947	0.0945	0.0959	-	0.3947	0.3947	0.3055	0.367
C3,3	-	0.4	0.4	0.3611	0.3955	0.1667	0.1316	0.1316	0.2527	0.2494	-	0.3947	0.3947	0.3055	0.2202
C3,4	-	0.1333	0.1333	0.1413	0.1695	0.0833	0.0789	0.0789	0.1428	0.1559	-	0.0789	0.0789	0.1195	0.1376

Source(s): Table created by authors

Table A4.

Weights of the criteria belonging to the macro-category transportation as assigned by the three experts

	E1					E2					E3				
	S-AHP	AHP-e	FUCOM	NDSL	LBWA	S-AHP	AHP-e	FUCOM	NDSL	LBWA	S-AHP	AHP-e	FUCOM	NDSL	LBWA
C4,1	0.6429	0.5607	0.5607	0.3611	0.3916	0.5	0.375	0.375	0.4274	0.459	0.375	0.3529	0.3529	0.3667	0.3733
C4,2	0.2143	0.1869	0.1869	0.3187	0.2937	0.25	0.25	0.25	0.2117	0.1967	0.375	0.3529	0.3529	0.3236	0.28
C4,3	0.0714	0.1121	0.1121	0.1413	0.1469	0.125	0.1875	0.1875	0.1805	0.1721	0.125	0.1765	0.1765	0.1548	0.1867
C4,4	0.0714	0.1402	0.1402	0.1789	0.1678	0.125	0.1875	0.1875	0.1805	0.1721	0.125	0.1176	0.1176	0.1548	0.16

Source(s): Table created by authors

Table A5.

Weights of the criteria belonging to the macro-category quality as assigned by the three experts

	E1					E2					E3				
	S-AHP	AHP-e	FUCOM	NDSL	LBWA	S-AHP	AHP-e	FUCOM	NDSL	LBWA	S-AHP	AHP-e	FUCOM	NDSL	LBWA
C5,1	0.2353	0.225	0.225	0.2823	0.2742	-	0.0909	0.0909	0.1126	0.1194	0.15	0.1299	0.1299	0.0903	0.088
C5,2	0.1176	0.15	0.15	0.1867	0.1828	-	0.0909	0.0909	0.1126	0.1194	0.05	0.0779	0.0779	0.109	0.1099
C5,3	0.0588	0.05	0.05	0.0415	0.0422	-	0.1364	0.1364	0.0912	0.0896	0.05	0.0779	0.0779	0.0645	0.0628
C5,4	0.0588	0.05	0.05	0.0415	0.0439	-	0.1364	0.1364	0.1367	0.1343	0.15	0.1299	0.1299	0.0728	0.0733
C5,5	0.4706	0.45	0.45	0.3734	0.3656	-	0.2727	0.2727	0.2735	0.2687	0.15	0.1948	0.1948	0.2119	0.2199
C5,6	0.0588	0.075	0.075	0.0747	0.0914	-	0.2727	0.2727	0.2735	0.2687	0.45	0.3896	0.3896	0.4515	0.4461

Source(s): Table created by authors

Table A6.

Weights of the five macro-categories as assigned by the three experts

	E1					E2					E3				
	S-AHP	AHP-e	FUCOM	NDSL	LBWA	S-AHP	AHP-e	FUCOM	NDSL	LBWA	S-AHP	AHP-e	FUCOM	NDSL	LBWA
C1	0.32	0.47368	0.47368	0.29029	0.31088	0.07692	0.10949	0.1232	0.13148	0.13365	-	0.15385	0.15385	0.13642	0.13592
C2	0.32	0.23684	0.23684	0.26796	0.2487	0.61538	0.43796	0.43121	0.44018	0.42769	-	0.30769	0.30769	0.27203	0.27184
C3	0.16	0.07895	0.07895	0.14558	0.15544	0.07692	0.14599	0.14374	0.14673	0.14256	-	0.07692	0.07692	0.10579	0.1165
C4	0.04	0.05263	0.05263	0.06809	0.07772	0.07692	0.08759	0.08624	0.08384	0.08225	-	0.15385	0.15385	0.21374	0.20388
C5	0.16	0.15789	0.15789	0.22808	0.20725	0.15385	0.21898	0.21561	0.19776	0.21385	-	0.30769	0.30769	0.27203	0.27184

Source(s): Table created by authors

	S-AHP	AHP-e	FUCOM	NDSL	LBWA	Supply chain observatory
C1,1	0.805567	0.805567	0.805567	0.8	0.781933	
C1,2	0.194433	0.194433	0.194433	0.2	0.218067	
C2,1	0.214167	0.205267	0.205267	0.194333	0.2176	
C2,2	0.098833	0.130767	0.130767	0.1452	0.126667	
C2,3	0.154667	0.081633	0.081633	0.108033	0.116067	
C2,4	0.2707	0.2349	0.2349	0.1998	0.185133	
C2,5	0.165433	0.265033	0.265033	0.241867	0.2385	
C2,6	0.0963	0.082367	0.082367	0.110767	0.116	
C3,1	0.6667	0.2421	0.2421	0.3195	0.3239	
C3,2	0.0833	0.352033	0.352033	0.239567	0.2334	
C3,3	0.1667	0.308767	0.308767	0.306433	0.288367	
C3,4	0.0833	0.097033	0.097033	0.134533	0.154333	
C4,1	0.505967	0.429533	0.429533	0.385067	0.407967	
C4,2	0.279767	0.263267	0.263267	0.284667	0.2568	
C4,3	0.107133	0.1587	0.1587	0.158867	0.168567	
C4,4	0.107133	0.148433	0.148433	0.1714	0.166633	
C5,1	0.19265	0.1486	0.1486	0.161733	0.160533	
C5,2	0.0838	0.106267	0.106267	0.1361	0.137367	
C5,3	0.0544	0.0881	0.0881	0.065733	0.064867	
C5,4	0.1044	0.105433	0.105433	0.083667	0.083833	
C5,5	0.3103	0.305833	0.305833	0.286267	0.284733	
C5,6	0.2544	0.245767	0.245767	0.266567	0.268733	

Source(s): Table created by authors

Table A7.
Weights of the criteria calculated using equation (1)

	S-AHP	AHP-e	FUCOM	NDSL	LBWA	
C1	0.19846	0.245673	0.250243	0.186063	0.193483	
C2	0.46769	0.327497	0.325247	0.326723	0.316077	
C3	0.11846	0.10062	0.09987	0.1327	0.138167	
C4	0.05846	0.098023	0.097573	0.12189	0.121283	
C5	0.156925	0.228187	0.227063	0.232623	0.23098	

Source(s): Table created by authors

Table A8.
Weights of the macro-categories calculated using equation (2)

Corresponding author

Maria Angela Butturi can be contacted at: mariaangela.butturi@unimore.it

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com