MULTIPLE CRITERIA DECISION MAKING

ANNUAL

VOL. 17 (2022)

UNIVERSITY OF ECONOMICS IN KATOWICE 2022

Printed by

EXDRUK Spółka Cywilna Wojciech Żuchowski, Adam Filipiak
ul. Rysia 6, 87-800 Włocławek
e-mail: biuroexdruk@gmail.com, www.exdruk.com

ISSN 2084-1531

Edition: 30 copies

Original version of the MCDM is the paper version

Publishing House of the University of Economics in Katowice
ul. 1 Maja 50, 40-287 Katowice, tel. +48 32 25 77 633
www.wydawnictwo.ue.katowice.pl, e-mail: wydawnictwo@ue.katowice.pl
Facebook: @wydawnictwoueekatowice

Contents

Part I
Special issue

Multiple Criteria Decision Aid: Advances in Theory and Applications

Guest Editor: Alessio Ishizaka

FROM THE GUEST EDITOR

This special issue entitled Multiple Criteria Decision Aid: Advances in Theory and Applications offers a selection of papers presented and discussed at the 26th International MCDM Conference held in Portsmouth (United Kingdom) on 26th June-1st July 2022. It was also open to the MCDA community at large. I would like to thank the Editor-in-Chief, Professor Tadeusz Trzaskalik, University of Economics in Katowice, for his support. We also thank the authors for choosing this issue to submit their papers to, and the referees for their rigorous reviews and their comments which improved the quality of papers. This special issue presents theoretical research results and interesting applications reflecting the utility of using the multicriteria approaches. Four papers have been accepted to the special issue.

Being green is today a necessity to save our planet. Yen-Tsang Chen in his paper "REVISITING GREEN SUPPLIER SELECTION PUBLICATIONS FROM THE LAST DECADE (2010-2022): A STRUCTURED REVIEW AND BIBLIOMETRIC STUDY" performs an up-to-date literature review on green supplier selection. 462 papers has been reviewed and the main methods and theories have been highlighted.

Decisions are often taken as a snapshot in a static world. However, our world is constantly changing. Francisco Salas-Molina et al. in their paper "IMPLICATIONS OF PARAMETER SELECTION IN DYNAMIC MULTIOBJECTIVE MODELS IN ECONOMICS AND FINANCE", have investigated techniques for dynamic decisions and then applied it on the economic and finance sector.

An interesting application has been proposed by Pedro Nunes Lopes Neto et al. in the paper "RANKING OF LTE CELLS BASED ON KEY PERFORMANCE INDICATORS USING MCDM METHODS". As the data trafic is constantly growing, the maintenance of the telecommunication network is of high importance. In this paper, the authors propose a method based on multi-attribute theory and AHP to detect faulty cells in a network.

Giving weights to criteria is always difficult. The task is even more difficult if we do not have experts available. To avoid this problem, Chris Tofallis in his paper "OBJECTIVE WEIGHTS FOR SCORING: THE AUTOMATIC DEMOCRATIC METHOD" proposes an objective method based on DEA and a regression to obtain objective weights.

This special issue shows a strong relationship between theoretical and methodological developments in MCDA. It also shows the potential offered by MCDA to solve real-world case problems. Therefore, we recommend this issue to the MCDA community. We hope that the researchers will find this collection of papers useful from both methodological and application perspectives.

Alessio Ishizaka

Alessio Ishizaka is distinguished professor at Neoma Business School, France. He was the head of supply chain, information systems and decision making department from 2019 to 2022. He was Full Professor in Decision Analysis, research lead and Founding Deputy Director of the Centre of Operations Research and Logistics (CORL) at the Portsmouth Business School of the University of Portsmouth, UK. He received his PhD from the University of Basel (Switzerland). He worked successively for the University of Exeter (UK), University of York (UK), Audencia Grande Ecole de Management Nantes (France) and University of Portsmouth (UK). He has been visiting professor at the Università del Sannio, Politecnico di Torino, Università degli Studi di Trento, INSA Strasbourg, Université de Lorraine, Universität Mannheim, Università degli Studi di Modena e Reggio Emilia, Universität der Bundeswehr Hamburg, Université d'Aix-Marseille, Università degli Studi di Torino, Università degli Studi della Tuscia and Università degli Studi di Padova. His research is in the area of decision analysis, where he has published more than 130 papers. He is regularly involved in large European funded projects. He has been the chair, co-organiser and guest speaker of several conferences on this topic. Alongside his academic activities, he acts as a consultant for companies in helping them to take better decisions. He has written the key textbooks *Multicriteria Decision Analysis: Methods and Software* (2013) and *Multi-Criteria Decision-Making Sorting Methods* (2023).

MULTIPLE CRITERIA DECISION MAKING
Vol. 17 2022

Yen-Tsang Chen*

REVISITING GREEN SUPPLIER SELECTION PUBLICATIONS FROM THE LAST DECADE (2010-2022): A STRUCTURED REVIEW AND BIBLIOMETRIC STUDY

DOI: 10.22367/mcdm.2022.17.01

Received: 16.06.2022 | Revised: 22.11.2022 | Accepted: 1.03.2023.

Abstract

Almost ten years have passed since some seminal structured literature reviews about multi-criteria decision-making for green supplier selection were published. We aimed to investigate the evolution of intellectual structures in this field through a structured literature review and bibliometric analysis using publications between 2010 and 2022. We noted that mathematical and analytical approaches are still dominating, and the complexity of the methods has increased. Bibliometrically, their theoretical foundation and techniques are the same despite the change of leading papers over time. Our contribution consist in extending earlier studies and discussing the evolution of the field.

Keywords: green supplier selection, multi-criteria decision-making, structured review, bibliometric analysis.

1 Introduction

The importance of selecting a proper supplier has already been demonstrated. A good supplier could help organizations to achieve superior monetary performance, efficient strategy implementation, higher quality, or better reputation (Dobos and Vörösmarty, 2019; Ellram, 1990; Famiyeh and Kwarteng, 2018; Kannan and Tan, 2002; Kaufmann, Mesching and Reimann, 2014). In order to contribute to the efficiency of supplier selection, academics have extensively investigated this organizational task in various aspects, such as: alignment

* NEOMA Business School − Campus de Reims: Reims, e-mail: yen-tsang.chen@neoma-bs.fr, ORCID: 0000-0001-9057-737X.

of sourcing and business strategy (Chen, 2011), supplier selection criteria (Choi and Hartley, 1996; van der Rhee, Verma and Plaschka, 2009; Weber, Current and Benton, 1991), process and decision making (Kaufmann, Carter and Buhrmann, 2012; Riedl et al., 2013), optimization modeling (Ho, Xu and Dey, 2010; Xia and Wu, 2007) and sustainability in supplier selection (dos Santos, Godoy and Campos, 2019; Ehrgott et al., 2011; Kannan, 2018).

Until the end of the 1990s, the supplier selection process mainly employed conventional operational and strategical criteria such as quality, cost, delivery, and flexibility (Choi and Hartley, 1996; Ellram, 1990; Weber, Current and Benton, 1991). However, since the late 1990s, given the positive impact of sustainability on firm performance (Rao and Holt, 2005), sustainability concerns are getting more and more noticed in supply chain management and supplier selection.

Despite the importance of sustainability for the organization, relatively few papers studied green supplier selection until 2010. For instance, Igarashi, de Boer and Fet (2013) found only 60 papers focused on green supplier selection while reviewing the publications from 1991-2011; Genovese et al. (2013) collected 28 papers for their review of publications from 1997-2010; and Wetzstein et al. (2016) analyzed only 25 papers dealing with green supplier selection from 248 papers researching supplier selection.

More recently, Schramm, Cabral and Schramm (2020) analyzed 82 papers that investigated green supplier selection, published in the last three decades. They reported the multi-criteria decision-making (MCDM) methods mainly used to support green supplier selection. Despite their study's extensiveness, it was impossible to determine how the intellectual structure evolved from early research on green supplier selection to more recent studies. At the same time, the intellectual structure could be observed in earlier articles, such as Igarashi, de Boer and Fet (2013) and Genovese et al. (2013). Considering that no studies updated these earlier review papers with the last decade's data (2010-2022), our study investigates how the intellectual structure evolved from 2010 to 2022 in green supplier selection and how green supplier selection has developed in green supply chain management.

Methodologically, to answer our research question objectively, we adopted a Structured Literature Review (SLR) as suggested by Thomé, Scavarda and Scavarda (2016), which differs from the traditional literature review by providing a clear and well-defined process. As part of our SLR, we employed a bibliometric analysis using a sample of articles published from 2010 to 2022. From the bibliometric analysis, we could describe our sample articles quantitatively and obtain a citation network, bibliographic coupling, and a co-citation network, which could allow us to infer the intellectual structure of the field. To support our bibliometric analysis, we used VOSviewer (van Eck and Waltman, 2017; Yu et al., 2020) to analyze the sample of manuscripts built through *the ISI Web of Science* and *Scopus* indexer.

Through an updated database, our study extended the earlier structured literature reviews; we identified the principal authors in green supplier selection, the prominent publications, the proximity of the leading publications, and the evolution of the intellectual structure of the last decade.

To organize this document, the rest of our manuscript is structured into: (2) a Literature review, (3) a Methodology, (4) Results and discussions, and (5) Final considerations.

2 Literature review

2.1 Structured literature review publications on green supplier selection

Supplier selection is a subject that has been studied since the 1960s, the seminal work of Dickson (1966) proposed a list with 23 supplier selection criteria that companies commonly use. This list of 23 criteria was updated later considering factors such as operational strategy (Weber, Current and Benton, 1991), industry (Choi and Hartley, 1996), nature of the product to be purchased (van der Rhee, Verma and Plaschka, 2009) or purchasing process (Scott, Burke and Szmerekovsky, 2018). Apart from supplier selection criteria, according to Wetzstein et al. (2016), research in supplier selection could be classified into six significant streams, where green supplier selection is one of them.

Although green supplier selection is one of the mainstreams in supplier selection, studies about selecting suppliers incorporating environmental and social/ethical criteria and the related process are relatively recent. Noci (1997) observed that the necessity of organizations to improve their environmental performances led to the necessity of considering these factors in supply chain management, thus impacting supplier selection. Following Noci (1997), the process of selecting a supplier considering these environmental and social/ethical criteria is called green supplier selection. It is not so different from the traditional supplier selection apart from the inclusion of sustainable factors in the process (Govindan et al., 2015; Noci, 1997; Qin, Liu and Pedrycz, 2017).

As for the literature on green supplier selection, some structured literature reviews served as guidelines for the research on this topic. Igarashi, Boer and Fet (2013) mapped the literature on green supplier selection. It classified how the articles are distributed in research methodology, theories, stages of the purchasing process, and what environmental criteria the articles treated. According to their study, most of the studies in green supplier selection are concentrated on the criteria formulation and final decision process, since these are the significant points that differ from the traditional selection. They also proposed a conceptual model of green supplier selection based on supply context, process, tools, and strategy alignment. However, they did not aim to demonstrate the intellectual structure underneath the green supplier selection criteria, nor how this structure had evolved for the period they investigated (1991-2011).

Genovese et al. (2013) conducted another relevant structured literature review in this area. These authors analyzed 28 papers, and they noted that 24 used a mathematical approach to investigate this subject and suggested that the availability of a waste management system is the most frequently occurring environmental criterion, followed by green design capability and environmental staff training and involvement. In addition to primary green supplier selection criteria, Genovese et al. (2013) noticed that, over time, studies tend to move from theoretical framework approaches involving only green criteria (Noci, 1997) to synthetic models where green are combined with the traditional criteria (Lee et al., 2009).

More recently, Zimmer, Fröhling and Schultmann (2016) conducted a structured literature review using a sample of 143 papers published from 1997 to 2014. They also observed that most of the publications in this field focused on the study of evaluation and final selection of suppliers. In contrast to earlier studies, Zimmer, Fröhling and Schultmann (2016) focused only on publications that bring models supporting green supplier selection, and they found that 62.2% used a combined model, such as linear programming and AHP or AHP and VIKOR, to support green supplier selection. Zimmer and colleagues also proposed a detailed mapping of green criteria employed by their sample papers, and they classified the selection criteria into three main categories: economic, environmental, and social. However, social criteria are less employed than the first two categories.

Similarly to Zimmer, Fröhling and Schultmann (2016), Schramm, Cabral and Schramm (2020) mapped structurally 82 papers dealing with mathematical approaches to support green supplier selection published between 1990 and 2019. Like previous studies, Schramm, Cabral and Schramm (2020) also observed that most of their analyzed papers integrate more than two methods. According to Schramm, Cabral and Schramm (2020), combining more methods can bring more robust results. However, the methods employed should avoid the high cognitive demand of the decision-makers. Unlike previous studies, Schramm, Cabral and Schramm (2020) did not investigate the green criteria used in their papers.

3 Methodology

Structured literature review has been widely employed to understand and organize the publications in a research field. For instance, Üsdisken and Pasadeos (1995) employed this approach to investigate differences in organization studies between US and European researchers; Burgess, Singh and Koroglu (2006) mapped the supply chain management publications between 1985 and 2003; Pilkington and Meredith (2009) studied the evolution of intellectual structure in operations management from 1980 to 2006. Song et al. (2019) investigated how classroom communication research evolved in the education field between 1999 and 2018.

There are two approaches to conducting a structured literature review. The first one is a more qualitative approach, such as those employed by Burgess, Singh and Koroglu (2006), Zimmer, Fröhling and Schultmann (2016), and Schramm, Cabral and Schramm (2020). In this approach, the researchers analyze a sample of papers, classify them according to several criteria and infer the theoretical paradigms existing in the sample, the dominant research methodology and approaches, the main research streams, definitions of research terminologies, and possible research gaps. Another approach is based on bibliometric investigation, such as those applied by Üsdisken and Pasadeos (1995), Pilkington and Meredith (2009), and Song et al. (2019).

In contrast to the qualitative approach, using bibliometric analysis, it is possible to analyze a larger sample. It uses the bibliographic data of a sample of publications to build the intellectual structure of the field (Zupic and Čater, 2015). Among the five significant metrics in the bibliometric analysis: keywords, citation, co-citation, bibliographic, and coauthor analyses, our study will employ the first four metrics to analyze our sample of publications. From these four bibliographic metrics, we could identify: a) the main topics treated by the sample articles (keyword analysis), b) the relatedness of the sample articles (bibliographic coupling and citation analysis), c) the relatedness of the references of the sample papers (co-citation analysis).

Keyword analysis counts the number of times that each keyword supplied by authors appears in the sample article and the number of times they appear together. From the frequency of the keywords co-occurrence, it is possible to identify the main topics treated by the sample papers.

Citation analysis counts the number of times each sample publication was cited and the number of times a sample paper cited other papers from the sample. This analysis assumes that the higher the citation of a paper in the sample, the more influential it is in the field. As opposed to citation analysis, the co-citation analysis deals with the bibliographic references of our sample paper. This analysis is defined by the frequency at which two bibliographic references of the sample papers are cited together. This analysis assumes that papers are cited together when they have: a) similar theoretical foundation regardless of their positioning and/or b) complementary ideas. From the co-citation analysis, it could be possible to identify theoretical streams, concepts, models, or research methodologies (Pilkington and Meredith, 2009; Small, 1973).

The fourth metric that we adopted in our study is bibliographic coupling. This indicator counts the common references shared by two papers of the analyzed sample. From this indicator, it is expected that the more common references two publications share, the more similar they are (Zupic and Čater, 2015).

However, bibliometrics analysis is based on formal communication among scientific productions; therefore, the proximity of the publications and authors does not consider informal communications such as technical reports, exchanges among authors in conferences, events, or personal aspects.

3.1 Sampling and data treatment

For two reasons, we used the indexers ISI Web of Science (WoS) and Scopus to create the sample of articles for our study. First, it is one of the most reliable scientific publication databases (Yang et al., 2013), and second, this platform could provide the information to elaborate the bibliometric study we needed.

Our sample is limited to English articles published from January 2010 to June 2022. The keywords that we used to search the publications were: “supplier selection” + “green”; “supplier selection” + “sustainability”. Since these two indexers bring articles not only from social studies, we limited our search to the areas related to business management and sustainability, such as *environmental science*, *green sustainable science technology*, *operations research management science*, *environmental engineering*, *industrial engineering*, *manufacturing engineering*, *environmental studies*, *management*, *electrical*, *electronic engineering*, *multidisciplinary engineering*, *multidisciplinary science*, *business, chemical engineering*, *civil engineering*, *public environmental*, *occupational health*, *transportation*, *mechanical engineering*, *material science textile*, *regional urban planning*, *ethics*, and *public administration*. From our search, we first screened the abstracts of all articles and removed all those unrelated to our subject. Then we removed overlapping papers, obtaining a sample of 942 articles.

Before data treatment, we used the OpenRefine application to standardize the keywords supplied by the authors, for instance, “analytic hierarchy process” to “AHP”. However, we did not reinterpret the keywords; for instance, if the article used “sustainable supplier selection”, we kept it as it was; even another article used “green supplier selection”. We also standardized and corrected the bibliographic references of our sample articles since there were discrepancies across the references of the sample articles when referring to papers or books, for instance, different descriptions of the same author (Barney, B. Jay or Barney, J) or different editions of the same book or incorrect year of a cited reference. To run the bibliometric analysis, we employed the VOSviewer application (van Eck and Waltman, 2017; Yu et al., 2020).

4 Results and discussion

4.1 Description of the sample

From our sample of 942 articles, we noted that the top 10 journals published about green supplier selection are responsible for more than 30% of total production from 2010 to 2022 (Table 1), which means that the articles on this subject are widely spread in a significant number of journals (318 journals) and not restricted to those dedicated to sustainability. Among the top 10 journals, the *Journal of Cleaner Production* has the largest number of publications on this topic, which is somehow expected. One interesting observation is the fifth place of the *International Journal of Production Economics*, the top publisher among multidisciplinary journals. Our finding is in line with the bibliometric study of Fahimnia, Sarkis and Davarzani (2015), and the *Journal of Cleaner Production* remained the leading source of articles related to the green supply chain subjects.

Table 1: Journals that publish the largest number of papers on green supplier selection

Rank	Journal	# publications	Cumulated %
1	Journal of Cleaner Production	85	9.0%
2	Sustainability	66	16.0%
3	Computers & Industrial Engineering	25	18.7%
4	Journal of Intelligent & Fuzzy Systems	24	21.2%
5	International Journal of Production Economics	23	23.7%
6	International Journal of Production Research	19	25.7%
7	Mathematical Problems in Engineering	19	27.7%
8	Mathematics	18	29.6%
9	Symmetry	18	31.5%
10	Applied Soft Computing	16	33.2%

Regarding the number of publications per year, from Figure 1 it can be seen that the number of publications dealing with green supplier selection has increased consistently since 2010, indicating that this area has still many research opportunities, either as regards the methodology or the supplier selection process.

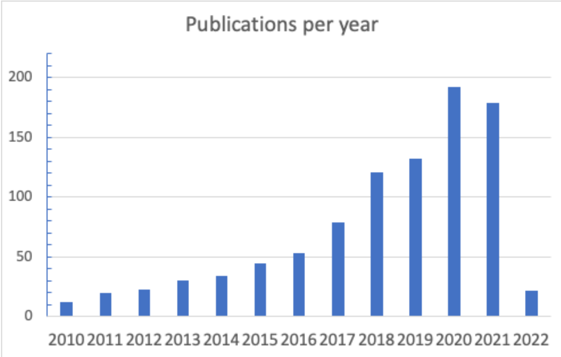

Figure 1: Number of publications per year

Regarding authorship, when considering only the first author, from Table 2 we can see that the ten most publishing authors are responsible for approximately 7% of total publications; therefore, we infer that contributions to the field are distributed, with a significant number of researchers (737 authors for 942 articles). We would like to remind that, for the counting of authorship, we considered only the first authors; while numerous prominent authors appear in several articles as second or third ones, such as Sarkis, Joseph (Dou and Sarkis, 2010), Kannan, Devika (Awasthi, Govindan and Gold, 2018) or Wei, Guiwu (Tang, Wei and Gao, 2019).

Table 2: Number of citations per author

Rank	Authors	# publications
1	Krishankumar, Raghunathan	8
2	Wei, Guiwu	8
3	Fallahpour, Alireza	7
4	Govindan, Kannan	7
5	Tavana, Madjid	7
6	Yazdani, Morteza	7
7	Kannan, Devika	6
8	Wang, Jie	6
9	Amindoust, Atefeh	5
10	Ghadimi, Pezhman	5

Regarding the principal authors whom Fahimnia, Sarkis and Davarzani (2015) noted in their green supply chain management study between 1996 and 2013, we noted that none of those top 10 authors appeared in our top 10 list. However, the top 10 authors identified by Fahimnia, Sarkis and Davarzani (2015) frequently appeared in the cited references and as second or third authors of our sample articles. We inferred that the different set of leading authors we obtained is due to the difference in period and the central theme of our sample articles as opposed to theirs. While Fahimnia, Sarkis and Davarzani (2015) analyzed the articles centered on green supply chain management between 1992 and 2012, we focused on green supplier selection between 2010 and 2022.

To analyze our sample qualitatively, we selected the 100 most cited papers. From these articles, in terms of research methodology, we noted that only a few used empirical methods (survey or case study), while 87% used analytical methods (one or more methods combined). This result reflects the finding of previous studies (Genovese et al., 2013; Igarashi, de Boer and Fet, 2013).

Table 3: Publications and research methods

# publications	Methods
3	Case study
6	Survey (OLS or Structural equation modeling or Factor analysis)
4	Literature review
87	Analytical methods (AHP, ANP, TOPIS, DEA, DEMATEL, VIKOR, Fuzzy AHP, Fuzzy ANP, etc.)

The number of studies based on analytical methods is not surprising since most papers are related to engineering. In these papers, green supplier selection focused on the operational approach; for instance, green supplier selection and order allocation problem (Hamdan and Cheaitou, 2017), green supplier selection using objective operational factors such as quality rejection, cost, late delivery, and greenhouse gas emission (Shaw et al., 2012) or green supply chain management practices through sustainable supplier selection (Kannan, de Sousa Jabbour and Jabbour, 2014).

In addition, given the wide range of analytical methods and the possibility of their combinations, it is possible to explore the green supplier selection process with a multitude of approaches, for instance, a single approach such as the application of ANP for offshoring strategy based on green supplier selection (Dou and Sarkis, 2010); multiple objective mixed-integer linear programming for green supply chain management using operational and strategical factors (Mota et al., 2018) or a combination of multiple approaches, such as AHP and fuzzy linear multi-objective linear programming (Shaw et al., 2012); fuzzy DEMATEL, fuzzy ANP and fuzzy TOPSIS (Büyüközkan and Çifçi, 2012); DEA, ANP and artificial neural network (Kuo, Wang and Tien, 2010); ANP and AHP (Sarkis, Meade and Presley, 2012) or ANP and QFD for green supplier selection (Tavana, Yazdani and Di Caprio, 2017). In addition, we also observed that researchers combined multiple methodologies to overcome the limitations of specific methods and find consistent results (Schramm, Cabral and Schramm, 2020).

Regarding the green supplier selection criteria, our analyzed sample papers suggested that in the green supplier selection process, the traditional selection factors, such as cost, quality, delivery, and flexibility should still be included (Arabsheybani, Paydar and Safaei, 2018; Dou and Sarkis, 2010; Hamdan and Cheaitou, 2017; Shaw et al., 2012; Tang and Wei, 2018a; Trapp and Sarkis, 2016; Wang, Wei and Wei, 2018; Wang et al., 2019). The adoption of traditional selection criteria could be explained by the transaction cost economy, where the company's primary aim is profit maximization (Hashemi, Karimi and Tavana, 2015). In addition to cost, quality and delivery, other traditional criteria can be considered, such as technical capability, manufacturing capability, financial status (Kuo and Lin, 2012), supplier reputation/ geographic location (Memari et al., 2019), shareholder, public and customer orientations (Reuter, Goebel and Foerstl, 2012). According to prior studies (Noci, 1997; Stević et al., 2020), green supplier selection, from the process perspective, can be seen as supplier selection that formally introduces the sustainability factors into the selection process.

According to our sample of papers, there are mainly two significant sustainability groups of factors: environmental and social/ethical criteria, which are similar to those observed by Zimmer, Fröhling and Schultmann (2016). The environmental criteria form an extensive list that involve factors such as greenhouse gas emission/ CO2 emission/ Carbon footprint (Govindan and Sivakumar, 2016; Huang et al., 2016; Kumar, Jain and Kumar, 2014; Shaw et al., 2012); energy usage/ resource consumptions/ waste minimization/ waste disposal (Agrawal, Singh and Murtaza, 2016; Kumar, Rahman and Chan, 2017; Shaw et al., 2012); environmental risk (Song, Ming and Liu, 2017); eco-design/ green image/ green principle/ green product/ green innovation (Che, 2010; Hashemi, Karimi and Tavana, 2015; Shen et al., 2013; Song, Ming and Liu, 2017; Tavana, Yazdani and Di Caprio, 2017; Zhang and Xu, 2015); green practices/ green certification/ ISO 14001/ EMAS (Fallahpour et al., 2017; Freeman and Chen, 2015; Hatami-Marbini et al., 2017; Kannan, 2018; Kannan, de Sousa Jabbour and Jabbour, 2014; Tseng and Chiu, 2013); reverse logistics/ reduce/ recycling/ reuse (Senthil, Srirangacharyulu and Ramesh, 2014; Tavana, Yazdani and Di Caprio, 2017; Yazdani et al., 2017); and environmental management system (Arabsheybani, Paydar and Safaei, 2018; Luthra et al., 2017; Senthil, Srirangacharyulu and Ramesh, 2014; Su et al., 2016; Tavana, Yazdani and Di Caprio, 2017; Yazdani et al., 2017). These factors are not necessarily used together but will depend on the organization's strategies and objectives (Demirtas and Üstün, 2008; Kumar, Rahman and Chan, 2017; Shaw et al., 2012).

The social/ethical criteria in green supplier selection are not as extensive as the environmental ones. Therefore they are less frequently used than conventional and environmental criteria (Stević et al., 2020; Zimmer, Fröhling and

Schultmann, 2016). For instance, Amindoust et al. (2012) used the rights of employees, rights of stakeholders, work safety, and labor health as environmental criteria to select green suppliers; Goren (2018) included occupational health and safety among environmental criteria; Bai et al. (2019) used those employed by Amindoust et al. (2012) plus community influence, contractual stakeholder's influence, occupational health education, training, and safety management system. As opposed to the previous authors, Hatami-Marbini et al. (2017) employed social criteria factors such as discrimination exposure risks, child labor practice risks, and corruption exposure. Besides those social criteria, some authors used ethical factors such as ethical behavior of suppliers' top management, incentives, implementation of a code of conduct, and obedience to authority (Goebel et al., 2012); formalization of ethical culture (Reuter, Goebel and Foerstl, 2012) or respect of human rights, underage labor, long working hours, feminist labor issue and organizational, legal responsibilities (Kumar et al., 2014). Remembering that these social or ethical factors could be used solely as the driver of green selection supplier criteria (Goebel et al., 2012; Reuter, Goebel and Foerstl, 2012) or combined them with other environmental criteria (Amindoust et al., 2012).

Concerning the evolution of research on green supplier selection, we did not observe the pattern suggested by Genovese et al. (2013), where publications that focus on the theoretical framework are narrower in their scope of supplier selection criteria, which means that they focus on environmental ones, while synthetic models combine traditional and environmental criteria. For instance, dos Santos, Godoy and Campos (2019) and Zhang and Xu (2015) used only green criteria in their modeling to evaluate green supplier performance. We observed that combining green supplier selection criteria with conventional ones depends on the researcher's approach: narrower vs wider and strategical vs operational.

4.2 Bibliometric analysis

For the bibliometric analysis, we separated our sample into two periods to evaluate possible changes in intellectual structure (2010-2015; 2016-2022). We started by analyzing the keywords, citation network, bibliographic coupling, and co-citation networks. It is worthwhile to remember that the citation network analyzes how influential each article of our sample papers is and how these influential articles are related. The bibliographic coupling analyzes the relatedness of the sample articles based on how many references they share. The co-citation network analyzes the relatedness of the cited references based on how often they are cited together.

4.2.1 Keyword analysis

From the keyword analysis, we could observe a proliferation of keywords from the first period to the second. This phenomenon is expected since publications increase consistently from 2010 to 2022. In both periods, the keyword is consistent with our search. We noted that the central theme is supplier selection. Still, in the first period, it is related to green supply chain management, while in the second, the term green supplier selection (GSS) appeared significantly related to supplier selection. In addition, we noticed that the supplier selection process is treated as a multicriteria decision-making problem in both periods, but with many more alternative methods associated.

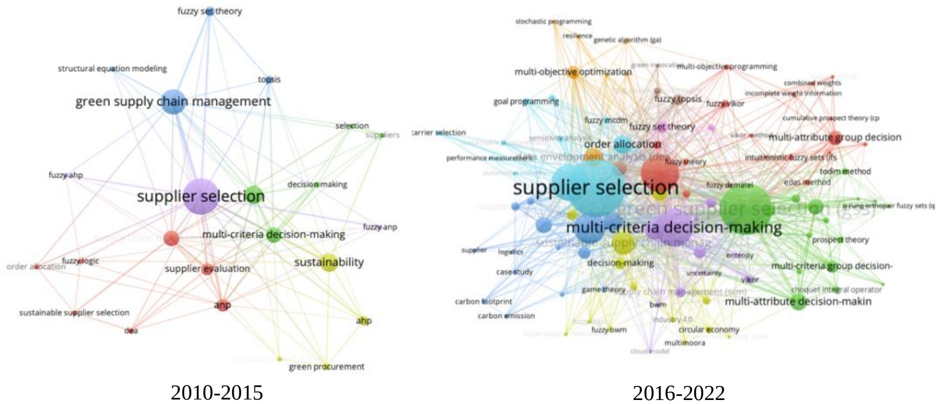

Figure 2: Keyword analysis

4.2.2 Citation analysis

The five most cited papers during the period 2010-2015 are: Govindan, Jafarian and Nourbakhsh (2015), Govindan, Khodaverdi and Jafarian (2013), Buyukozkan and Çifçi (2012), Bai and Sarkis (2010), and Kannan et al. (2013), each one with more than 500 citations (see appendix). However, from Figure 3, we can see that among these top-cited articles, Kuo, Wang and Tien (2010) played a central role since our top-cited papers cited it. Our top-cited papers cited Kuo, Wang and Tien (2010), because this article justifies the importance of sustainable supplier selection in green supplier management. This paper is also essential due to the integration of several methods for supplier selection (Artificial Neural Network, DEA, MADA, and ANP). In addition, Kuo, Wang and Tien (2010) bring an extensive list of supplier selection criteria that include traditional, environmental, and, especially, social ones, as social criteria are not considered often in the studies at that moment (Amindoust et al., 2012).

In the same fashion as Fahimnia, Sarkis and Davarzani (2015), who compiled an extensive literature review about green supply chain management through a bibliometric study and pointed out a possible green supply chain management typology, research methodologies, and critical research areas, Govindan et al. (2015) is also highly influential in this area of research, because it is an extensive literature review paper. However, Govindan, Jafarian and Nourbakhsh (2015) discussed the Multi-criteria decision-making in green supplier selection based on the methodology (individual vs. integrated), and they mapped the selection criteria.

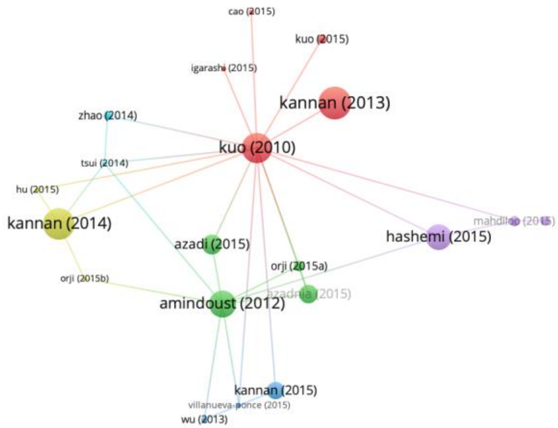

Figure 3: Citation network (2010-2015)

The top five cited papers in the period 2016-2021 are Luthra et al. (2017), Qin, Liu and Pedrycz (2017), Awasthi, Govindan and Gold (2018), Stević et al. (2020) and Banaeian et al. (2018). Since the second period is recent, these papers have at least 230 citations. From Figure 4, we observe that Luthra et al. (2017), Qin, Liu and Pedrycz (2017), and Stević et al. (2020) assumed the central roles in our citation network. Our sample often cites Luthra et al. (2017), because they applied − in a very instructive way − the integration of two commonly used methods in multi-criteria decision-making, AHP, and VIKOR, to the green supplier selection in the Indian automobile industry; in addition, they included social criteria in their supplier selection, which until then occurred very rarely. To eliminate the limitations of TOPSIS, Qin, Liu and Pedrycz (2017) extended the TODIM (Interactive and Multicriteria Decision-Making) into the fuzzy environment. Similarly, Stević et al. (2020) discussed how each of the previous papers contributed to the field by extending the commonly used methods; they also proposed a new method called Measurement of Alternatives and Ranking according to COmpromise Solution (MARCOS).

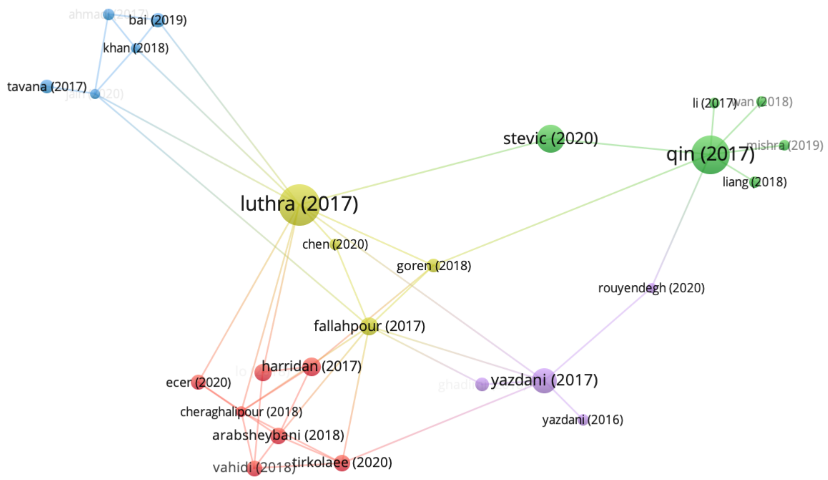

Figure 4: Citation network (2016-2021)

4.2.3 Bibliographic coupling

To complement the citation analysis, we also conducted a bibliographic coupling analysis, which assessed the relatedness of the papers based on the number of shared references. The rationale behind this analysis is that the more references two publications share, the more similar they are.

From the bibliographic coupling analysis, we can note that in our sample of papers from 2010-2015 (Figure 5), there are four clusters of papers. The first small group is composed of empirically-oriented papers, related mainly to sustainable supply chain and supply chain management (Harms, Hansen and Schaltegger, 2013; Paulraj, 2011). The second group focused on commonly used methods in MCDM (AHP, ANP, DEMATEL, etc.) and their extensions applied to the sustainable supplier selection (Bai and Sarkis, 2010; Dai and Blackhurst, 2012; Govindan et al., 2015; Govindan, Jafarian and Nourbakhsh, 2015; Hsu et al., 2013). The third group is related mainly to applying fuzzy concepts and their integration/extension to those commonly used MCDM methods (Büyüközkan, 2012; Büyüközkan and Çifçi, 2011; Govindan, Khodaverdi and Jafarian, 2013; Tseng and Chiu, 2013). The fourth group is related to the sustainable supplier selection (Ehrgott et al., 2011; Goebel et al., 2012) by applying diverse MCDM methods and their extensions, such as ANP, fuzzy AHP, or integration of artificial neural networks to the MADA (Freeman and Chen, 2015; Hashemi, Karimi and Tavana, 2015; Kuo, Wang and Tien, 2010; Wu, Hsieh and Chang, 2013).

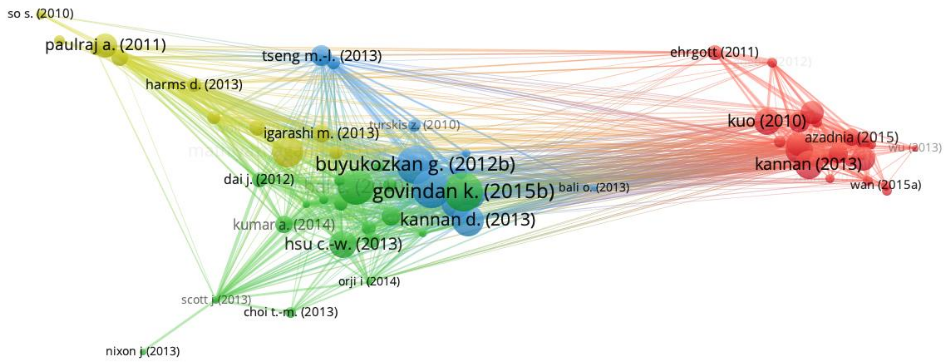

Figure 5: Analysis of bibliographic coupling (2010-2015)

Concerning the bibliographic coupling of the second period, Figure 6 also demonstrated that publications of 2016-2021 could be grouped into four major groups. The first group focuses mainly on the development of integration of common MCDM methods such as AHP and VIKOR, the extension of TODIM, or the proposition of MARCOS (Luthra et al., 2017; Qin, Liu and Pedrycz, 2017; Stević et al., 2020). The second cluster focuses on the Pythagorean fuzzy set and its extensions (Tang and Wei, 2018b; Wan, Jin and Dong, 2018; Wei et al., 2018). Similarly to the first period, there is a group of papers focusing on applying fuzzy concepts to MCDM methods for the green supplier selection (Awasthi and Kannan, 2016; Guo et al., 2017; Memari et al., 2019; Wu et al., 2019) and another group that employs diverse methods (analytical and empirical) to investigate green supplier selection (Huang et al., 2016; Jabbarzadeh, Fahimnia and Sabouhi, 2018; Kumar, 2019; Su et al., 2016).

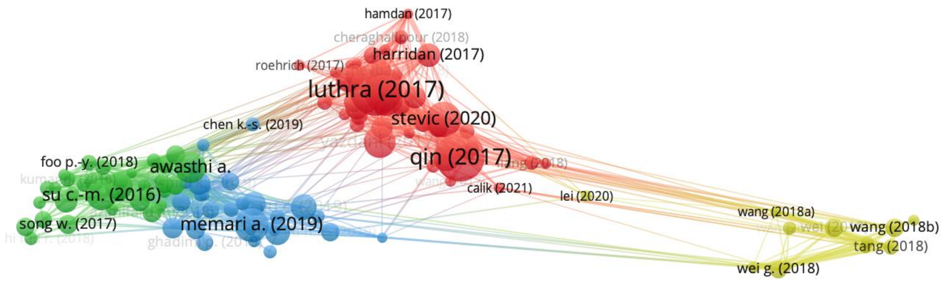

Figure 6: Analysis of bibliographic coupling (2016-2021)

4.2.4 Co-citation analysis

From the co-citation analysis, we could observe that for the period 2010-2015 (Figure 6), the intellectual structure that supported our sample papers was composed mainly of three groups of references. The first group is related to the concept of sustainable supply chain management, its definition, and how it relates to the organizational performance (Rao, 2002; Rao and Holt, 2005; Sarkis, 2003; Srivastava, 2007). The second theme of the intellectual structure is related to supplier selection, and includes: selection criteria (Dickson, 1966; Weber, Current and Benton, 1991), integration of environmental criteria in the supplier selection (Humphreys, Wong and Chan, 2003), and definition of green supplier selection (Noci, 1997). The third theme of the intellectual structure is associated with instruction-oriented references that apply commonly used MCDM methods in green supplier selection, such as the employment of ANP (Hsu and Hu, 2009), AHP (Handfield et al., 2002), concepts of fuzzy set theory (Zadeh, 1965; Zimmermann, 2011) and its application in green supplier selection, including fuzzy AHP (Lee et al., 2009), Fuzzy TOPSIS (Govindan, Khodaverdi and Jafarian, 2013), or integration of several fuzzy methods (Büyüközkan and Çifçi, 2012).

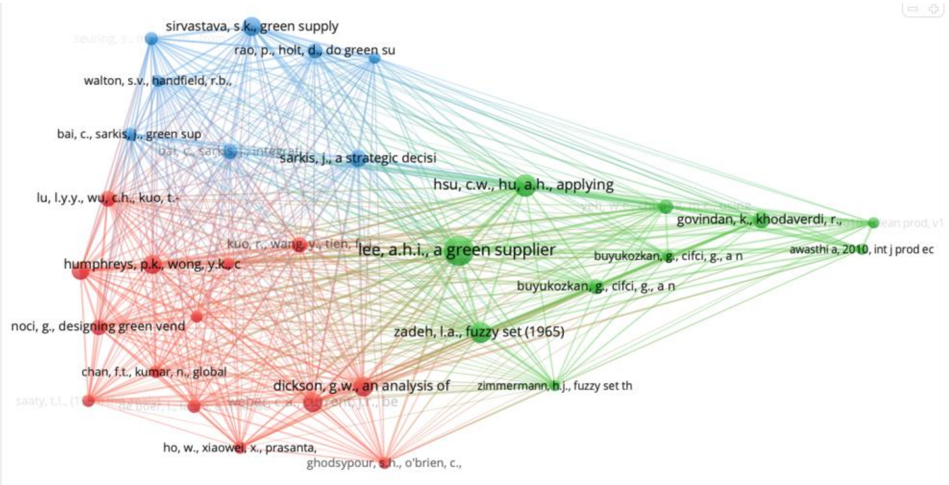

Figure 7: Co-citation network of the bibliographic references of publications in 2010-2015

From the analysis of the co-citation network of the sample papers from 2016-2022 (Figure 7), our first observation is an increase in the number of nodes in this network, which suggests an increase in the number of references co-cited. This augmentation is expected, since the number of publications in 2016-2022 increased. By comparing the intellectual structure of the co-citation of both periods, we could note that the themes found in the first period occur also in the second one. In green supplier selection studies it is essential to define its relationship to green supply chain management; therefore, we noticed a cluster of papers dedicated to supporting this topic (Carter and Rogers, 2008; Sarkis, 2003; Seuring and Müller, 2008; Srivastava, 2007). Another repeated cluster is the one that explained the methodology fundamentals, such as AHP, ANP, and Fuzzy set theory (Atanassov, 1994; Saaty, 1980; Yager, 2013; Zadeh, 1965). The papers in this cluster are highly co-cited with references that employed MCDM methods to study green supplier selection.

By contrast to the previous period, this time the cluster about supplier selection and the one about the application of MCDM methods in supplier selection are the same. This tendency suggests that MCDM and supplier selection are highly related, and it became the central reference for studies in this field. (Büyüközkan and Çifçi, 2012; Dickson, 1966; Govindan, Khodaverdi and Jafarian, 2013; Govindan, Jafarian and Nourbakhsh, 2015; Kuo, Wang and Tien, 2010; Lee et al., 2009; Weber, Current and Benton, 1991). In addition to this central cluster, the last cluster of intellectual structure that we identified is related mainly to the integration of multiple MCDM methods in the green supplier selection (Amindoust et al., 2012; Kuo, Wang and Tien, 2010; Luthra et al., 2017; Memari et al., 2019; Qin, Liu and Pedrycz, 2017).

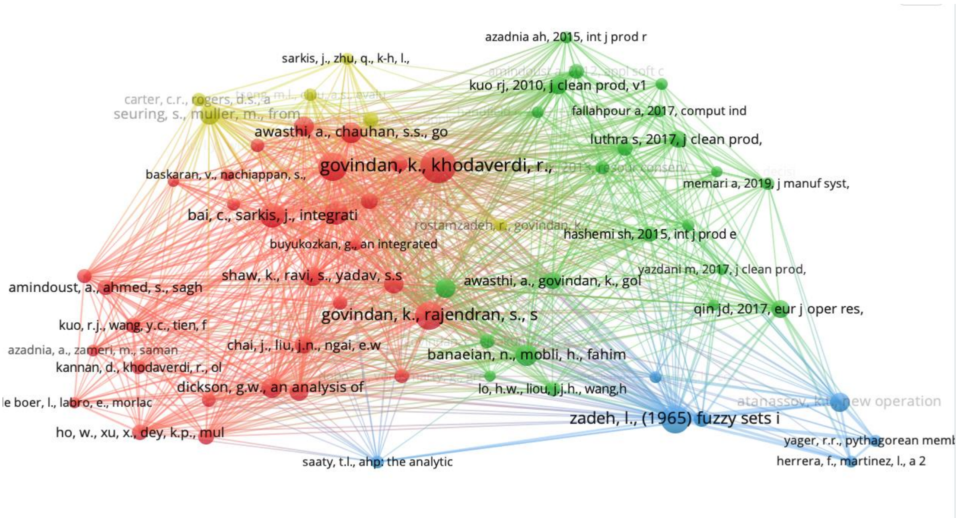

Figure 8: Co-citation network of the bibliographic references of publications in 2016-2021

5 Final considerations

Using a sample of 942 publications, our study analyzed structurally and bibliometrically the intellectual patterns of publications in green supplier selection. We compared our results to the existing structured literature review papers (Fahimnia, Sarkis and Davarzani, 2015; Genovese et al., 2013; Igarashi, de Boer and Fet, 2013; Schramm, Cabral and Schramm, 2020; Zimmer, Fröhling and Schultmann, 2016) and updated the results with papers published between 2010 and 2022. Practically, our paper serves as a picture of the current state of the field and can serve as a map for other researchers to start their investigations in green supplier selection.

Our results suggest that research in green supplier selection maintained the same pattern over the last decade (Genovese et al., 2013; Igarashi, de Boer and Fet, 2013), when the majority of papers used mathematical and analytical models, such as AHP, ANP, DEA, TOPIS, VIKOR, Linear programming, Fuzzy theory, Grey system theory, etc. (Zimmer, Fröhling and Schultmann, 2016). Our results also agreed with earlier studies, suggesting that combining those methods would increase the models' robustness and consistency or their application in the fuzzy environment (Qin, Liu and Pedrycz, 2017). Regarding the area and journal of publications, we found no discrepancies with the early studies, where the *Journal of Cleaner Production* is still the leading publisher in this field.

From the bibliographic coupling, we observed that our sample, in both periods, can be grouped into four main streams. Likewise, our co-citation analysis suggests that the intellectual structures in both periods demonstrated similar patterns. Both periods have a group of references that serve as methodological foundations (Saaty, 1980; Zadeh, 1965; Zimmermann, 2011), a group for green supply chain management concepts (Carter and Rogers, 2008; Sarkis, 2003; Seuring and Müller, 2008; Srivastava, 2007), a group for supplier selection criteria and integration of environmental factors in supplier selection (Dickson, 1966; Humphreys, Wong and Chan, 2003; Noci, 1997; Weber, Current and Benton, 1991) and a group of application of MCDM methods and its developments (Amindoust et al., 2012; Govindan, Khodaverdi and Jafarian, 2013; Hsu and Hu, 2009; Humphreys, Wong and Chan, 2003; Kuo, Wang and Tien, 2010; Lee et al., 2009; Luthra et al., 2017; Memari et al., 2019; Qin, Liu and Pedrycz, 2017; Stević et al., 2020).

5.1 Limitations and future studies

As with any structured literature review and bibliometric analysis, the first limitation is related to the choice of the sample of publications. We worked with ISI Web of Science and SCOPUS to make our study more comprehensive. The advantage of bibliometric analysis is its ability to analyze a significant number of papers. However, negative citations, where the citing article criticizes the cited publication, as well as some harmful citation practices, such as self-citation and self-team citation, can eventually alter the results of the metrics or the qualitative interpretation of the results. Nevertheless, bibliometric analysis is still a very reliable and objective method for analyzing the literature (Lim et al., 2009; Okubo, 1997; Zupic and Čater, 2015).

Our sample did not cover publications before 2010, hence we are not sure what the influence of those papers was in our bibliometric analysis. Therefore, we suggest that future studies create a sample of papers from 1999-2020, similarly to Schramm, Cabral and Schramm (2020), and investigate it through bibliometric metrics.

References

Agrawal S., Singh R.K., Murtaza Q. (2016), *Outsourcing Decisions in Reverse Logistics: Sustainable Balanced Scorecard and Graph Theoretic Approach*, Resources, Conservation and Recycling, 108, 41-53.

Amindoust A., Ahmed S., Saghafinia A., Bahreininejad A. (2012), *Sustainable Supplier Selection: A Ranking Model Based on Fuzzy Inference System*, Applied Soft Computing, 12(6), 1668-1677.

Arabsheybani A., Paydar M.M., Safaei A.S. (2018), *An Integrated Fuzzy MOORA Method and FMEA Technique for Sustainable Supplier Selection Considering Quantity Discounts and Supplier's Risk*, Journal of Cleaner Production, 190, 577-591.

Atanassov K.T. (1994), *New Operations Defined over the Intuitionistic Fuzzy Sets*, Fuzzy Sets and Systems, 61(2), 137-142.

Awasthi A., Govindan K., Gold S. (2018), *Multi-tier Sustainable Global Supplier Selection Using a Fuzzy AHP-VIKOR Based Approach*, International Journal of Production Economics, 195, 106-117.

Awasthi A., Kannan G. (2016), *Green Supplier Development Program Selection Using NGT and VIKOR under Fuzzy Environment*, Computers & Industrial Engineering, 91, 100-108.

Bai C., Kusi-Sarpong S., Badri Ahmadi H., Sarkis J. (2019), *Social Sustainable Supplier Evaluation and Selection: A Group Decision-support Approach*, International Journal of Production Research, 57(22), 7046-7067.

Bai C., Sarkis J. (2010), *Integrating Sustainability into Supplier Selection with Grey System and Rough Set Methodologies*, International Journal of Production Economics, 124(1), 252-264.

Banaeian N., Mobli H., Fahimnia B., Nielsen I.E., Omid M. (2018), *Green Supplier Selection Using Fuzzy Group Decision Making Methods: A Case Study from the Agri-food Industry*, Computers & Operations Research, 89, 337-347.

Burgess K., Singh P.J., Koroglu R. (2006), *Supply Chain Management: A Structured Literature Review and Implications for Future Research*, International Journal of Operations & Production Management, 26(7), 27.

Büyüközkan G. (2012), *An Integrated Fuzzy Multi-criteria Group Decision-making Approach for Green Supplier Evaluation*, International Journal of Production Research, 50(11), 2892-2909.

Büyüközkan G., Çifçi G. (2011), *A Novel Fuzzy Multi-criteria Decision Framework for Sustainable Supplier Selection with Incomplete Information*, Computers in Industry, 62(2), 164-174.

Büyüközkan G., Çifçi G. (2012), *A Novel Hybrid MCDM Approach Based on Fuzzy DEMATEL, Fuzzy ANP and Fuzzy TOPSIS to Evaluate Green Suppliers*, Expert Systems with Applications, 39(3), 3000-3011.

Carter C.R., Rogers D.S. (2008), *A Framework of Sustainable Supply Chain Management: Moving toward New Theory*, International Journal of Physical Distribution & Logistics Management, 38(5), 360-387.

Che Z.-H. (2010), *Using Fuzzy Analytic Hierarchy Process and Particle Swarm Optimisation for Balanced and Defective Supply Chain Problems Considering WEEE/RoHS Directives*, International Journal of Production Research, 48(11), 3355-3381.

Chen Y.-J. (2011), *Structured Methodology for Supplier Selection and Evaluation in a Supply Chain*, Information Sciences, 181(9), 1651-1670.

Choi T.Y., Hartley J.L. (1996), *An Exploration of Supplier Selection Practices across the Supply Chain*, Journal of Operations Management, 14(4), 333-343.

Dai J., Blackhurst J. (2012), *A Four-phase AHP–QFD Approach for Supplier Assessment: A Sustainability Perspective*, International Journal of Production Research, 50(19), 5474-5490.

Demirtas E.A., Üstün Ö. (2008), *An Integrated Multiobjective Decision Making Process for Supplier Selection and Order Allocation*, Omega, 36(1), 76-90.

Dickson G.W. (1966), *An Analysis of Vendor Selection Systems and Decisions*, Journal of Purchasing, 2(1), 5-17.

Dobos I., Vörösmarty G. (2019), *Inventory-related Costs in Green Supplier Selection Problems with Data Envelopment Analysis (DEA)*, International Journal of Production Economics, 209, 374-380.

dos Santos B.M., Godoy L.P., Campos L.M.S. (2019), *Performance Evaluation of Green Suppliers Using Entropy-TOPSIS-F*, Journal of Cleaner Production, 207, 498-509.

Dou Y., Sarkis J. (2010), *A Joint Location and Outsourcing Sustainability Analysis for a Strategic Offshoring Decision*, International Journal of Production Research, 48(2), 567-592.

Ehrgott M., Reimann F., Kaufmann L., Carter C.R. (2011), *Social Sustainability in Selecting Emerging Economy Suppliers*, Journal of Business Ethics, 98(1), 99-119.

Ellram L.M. (1990), *The Supplier Selection Decision in Strategic Partnerships*, Journal of Purchasing and Materials Management, 26(4), 8-14.

Fahimnia B., Sarkis J., Davarzani H. (2015), *Green Supply Chain Management: A Review and Bibliometric Analysis*, International Journal of Production Economics, 162, 101-114.

Fallahpour A., Olugu E.U., Musa S.N., Wong K.Y., Noori S. (2017), *A Decision Support Model for Sustainable Supplier Selection in Sustainable Supply Chain Management*, Computers & Industrial Engineering, 105, 391-410.

Famiyeh S., Kwarteng A. (2018), *Supplier Selection and Firm Performance: Empirical Evidence from a Developing Country's Environment*, International Journal of Quality & Reliability Management, 35, 690-710.

Freeman J., Chen T. (2015), *Green Supplier Selection Using an AHP-Entropy-TOPSIS Framework*, Supply Chain Management: An International Journal, 20(3), 327-340.

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(10), 2868-2886.

Goebel P., Reuter C., Pibernik R., Sichtmann C. (2012), *The Influence of Ethical Culture on Supplier Selection in the Context of Sustainable Sourcing*, International Journal of Production Economics, 140(1), 7-17.

Gonzalez P., Sarkis J., Adenso-Diaz B. (2008), *Environmental Management System Certification and Its Influence on Corporate Practices Evidence from the Automotive Industry*, International Journal of Operations & Production Management, 28(11-12), 1021-1041. https://doi.org/Doi 10.1108/01443570810910179

Gören H.G. (2018), *A Decision Framework for Sustainable Supplier Selection and Order Allocation with Lost Sales*, Journal of Cleaner Production, 183, 1156-1169.

Govindan K., Jafarian A., Nourbakhsh V. (2015), *Bi-objective Integrating Sustainable Order Allocation and Sustainable Supply Chain Network Strategic Design with Stochastic Demand Using a Novel Robust Hybrid Multi-objective Metaheuristic*, Computers & Operations Research, 62, 112-130.

Govindan K., Khodaverdi R., Jafarian A. (2013), *A Fuzzy Multi Criteria Approach for Measuring Sustainability Performance of a Supplier Based on Triple Bottom Line Approach*, Journal of Cleaner Production, 47, 345-354.

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.

Govindan K., Sivakumar R. (2016), *Green Supplier Selection and Order Allocation in a Low-carbon Paper Industry: Integrated Multi-criteria Heterogeneous Decision-making and Multi-objective Linear Programming Approaches*, Annals of Operations Research, 238(1-2), 243-276.

Guo Z., Liu H., Zhang D., Yang J. (2017), *Green Supplier Evaluation and Selection in Apparel Manufacturing Using a Fuzzy Multi-criteria Decision-making Approach*, Sustainability, 9(4), 650.

Hamdan S., Cheaitou A. (2017), *Supplier Selection and Order Allocation with Green Criteria: An MCDM and Multi-objective Optimization Approach*, Computers & Operations Research, 81, 282-304.

Handfield R., Walton S.V., Sroufe R., Melnyk S.A. (2002), *Applying Environmental Criteria to Supplier Assessment: A Study in the Application of the Analytical Hierarchy Process*, European Journal of Operational Research, 141(1), 70-87.

Harms D., Hansen E.G., Schaltegger S. (2013), *Strategies in Sustainable Supply Chain Management: An Empirical Investigation of Large German Companies*, Corporate Social Responsibility and Environmental Management, 20(4), 205-218.

Hashemi S.H., Karimi A., Tavana M. (2015), *An Integrated Green Supplier Selection Approach with Analytic Network Process and Improved Grey Relational Analysis*, International Journal of Production Economics, 159, 178-191.

Hatami-Marbini A., Agrell P.J., Tavana M., Khoshnevis P. (2017), *A Flexible Cross-efficiency Fuzzy Data Envelopment Analysis Model for Sustainable Sourcing*, Journal of Cleaner Production, 142, 2761-2779.

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.

Hsu C.-W., Hu A.H. (2009), *Applying Hazardous Substance Management to Supplier Selection Using Analytic Network Process*, Journal of Cleaner Production, 17(2), 255-264.

Hsu C.-W., Kuo T.-C., Chen S.-H., Hu A.H. (2013), *Using DEMATEL to Develop a Carbon Management Model of Supplier Selection in Green Supply Chain Management*, Journal of Cleaner Production, 56, 164-172.

Huang Y., Wang K., Zhang T., Pang C. (2016), *Green Supply Chain Coordination with Greenhouse Gases Emissions Management: A Game-theoretic Approach*, Journal of Cleaner Production, 112, 2004-2014.

Humphreys P.K., Wong Y.K., Chan F.T.S. (2003), *Integrating Environmental Criteria into the Supplier Selection Process*, Journal of Materials Processing Technology, 138(1-3), 349-356.

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.

Jabbarzadeh A., Fahimnia B., Sabouhi F. (2018), *Resilient and Sustainable Supply Chain Design: Sustainability Analysis under Disruption Risks*, International Journal of Production Research, 56(17), 5945-5968.

Kannan D. (2018), *Role of Multiple Stakeholders and the Critical Success Factor Theory for the Sustainable Supplier Selection Process*, International Journal of Production Economics, 195, 391-418.

Kannan D., de Sousa Jabbour A.B.L., Jabbour C.J.C. (2014), *Selecting Green Suppliers Based on GSCM Practices: Using Fuzzy TOPSIS Applied to a Brazilian Electronics Company*, European Journal of Operational Research, 233(2), 432-447.

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.

Kannan V.R., Tan K.C. (2002), *Supplier Selection and Assessment: Their Impact on Business Performance*, Journal of Supply Chain Management, 38(3), 11-21.

Kaufmann L., Carter C.R., Buhrmann C. (2012), *The Impact of Individual Debiasing Efforts on Financial Decision Effectiveness in the Supplier Selection Process*, International Journal of Physical Distribution & Logistics Management, 42(5), 411-433.

Kaufmann L., Meschnig G., Reimann F. (2014), *Rational and Intuitive Decision-making in Sourcing Teams: Effects on Decision Outcomes*, Journal of Purchasing and Supply Management, 20(2), 104-112.

Kumar A., Jain V., Kumar S. (2014), *A Comprehensive Environment Friendly Approach for Supplier Selection*, Omega, 42(1), 109-123.

Kumar D. (2019), *Buyer-supplier Relationship Selection for a Sustainable Supply Chain: A Case of the Indian Automobile Industry*, International Journal of the Analytic Hierarchy Process, 11(2), 215-227.

Kumar D., Rahman Z., Chan F.T.S. (2017), *A Fuzzy AHP and Fuzzy Multi-objective Linear Programming Model for Order Allocation in a Sustainable Supply Chain: A Case Study*, International Journal of Computer Integrated Manufacturing, 30(6), 535-551.

Kumar D.T., Palaniappan M., Kannan D., Shankar K.M. (2014), *Analyzing the CSR Issues Behind the Supplier Selection Process Using ISM Approach*, Resources, Conservation and Recycling, 92, 268-278.

Kuo R.J., Lin Y.J. (2012), *Supplier Selection Using Analytic Network Process and Data Envelopment Analysis*, International Journal of Production Research, 50(11), 2852-2863.

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.

Lee A.H.I., 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.

Lim A., Ma H., Wen Q., Xu Z., Cheang B. (2009), *Distinguishing Citation Quality for Journal Impact Assessment*, Communications of the ACM, 52(8), 111-116.

Luthra S., Govindan K., Kannan D., Mangla S.K., Garg C.P. (2017), *An Integrated Framework for Sustainable Supplier Selection and Evaluation in Supply Chains*, Journal of Cleaner Production, 140, 1686-1698.

Memari A., Dargi A., Jokar M.R.A., Ahmad R., Rahim A.R.A. (2019), *Sustainable Supplier Selection: A Multi-criteria Intuitionistic Fuzzy TOPSIS Method*, Journal of Manufacturing Systems, 50, 9-24.

Mota B., Gomes M.I., Carvalho A., Barbosa-Povoa A.P. (2018), *Sustainable Supply Chains: An Integrated Modeling Approach under Uncertainty*, Omega, 77, 32-57.

Noci G. (1997), *Designing 'Green' Vendor Rating Systems for the Assessment of a Supplier's Environmental Performance*, European Journal of Purchasing & Supply Management, 3(2), 103-114.

Okubo Y. (1997), *Bibliometric Indicators and Analysis of Research Systems: Methods and Examples*, OECD Publications.

Paulraj A. (2011), *Understanding the Relationships between Internal Resources and Capabilities, Sustainable Supply Management and Organizational Sustainability*, Journal of Supply Chain Management, 47(1), 19-37.

Pilkington A., Meredith J. (2009), *The Evolution of the Intellectual Structure of Operations Management − 1980-2006: A Citation/Co-citation Analysis*, Journal of Operations Management, 27(3), 185-202.

Qin J., Liu X., Pedrycz W. (2017), *An Extended TODIM Multi-criteria Group Decision Making Method for Green Supplier Selection in Interval Type-2 Fuzzy Environment*, European Journal of Operational Research, 258(2), 626-638.

Rao P. (2002), *Greening the Supply Chain: A New Initiative in South East Asia*, International Journal of Operations & Production Management, 22(6), 632-655.

Rao P., Holt D. (2005), *Do Green Supply Chains Lead to Competitiveness and Economic Performance?* International Journal of Operations & Production Management, 25, 898-916.

Reuter C., Goebel P., Foerstl K. (2012), *The Impact of Stakeholder Orientation on Sustainability and Cost Prevalence in Supplier Selection Decisions*, Journal of Purchasing and Supply Management, 18(4), 270-281.

Riedl D.F., Kaufmann L., Zimmermann C., Perols J.L. (2013), *Reducing Uncertainty in Supplier Selection Decisions: Antecedents and Outcomes of Procedural Rationality*, Journal of Operations Management, 31(1-2), 24-36.

Saaty T.L. (1980), *AHP: The Analytic Hierarchy Process* [in:] *New York* (Vol. 324), McGraw-Hill.

Sarkis J. (2003), *A Strategic Decision Framework for Green Supply Chain Management*, Journal of Cleaner Production, 11(4), 397-409.

Sarkis J., Meade L.M., Presley A.R. (2012), *Incorporating Sustainability into Contractor Evaluation and Team Formation in the Built Environment*, Journal of Cleaner Production, 31, 40-53.

Sarkis J., Talluri S. (2002), *A Model for Strategic Supplier Selection*, Journal of Supply Chain Management, 38(4), 18-28.

Schramm V.B., Cabral L.P.B., Schramm F. (2020), *Approaches for Supporting Sustainable Supplier Selection − A Literature Review*, Journal of Cleaner Production, 273, 123089.

Scott M.A., Burke G., Szmerekovsky J. (2018), *"Do as I Do and Not as I Say": Exploring Price-Oriented Maverick Buying during Supplier Selection*, Decision Science, 49(1), 25-64.

Senthil S., Srirangacharyulu B., Ramesh A. (2014), *A Robust Hybrid Multi-criteria Decision Making Methodology for Contractor Evaluation and Selection in Third-party Reverse Logistics*, Expert Systems with Applications, 41(1), 50-58.

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.

Shaw K., Shankar R., Yadav S.S., Thakur L.S. (2012), *Supplier Selection Using Fuzzy AHP and Fuzzy Multi-objective Linear Programming for Developing Low Carbon Supply Chain*, Expert Systems with Applications, 39(9), 8182-8192.

Shen L., Olfat L., Govindan K., Khodaverdi R., Diabat A. (2013), *A Fuzzy Multi Criteria Approach for Evaluating Green Supplier's Performance in Green Supply Chain with Linguistic Preferences*, Resources, Conservation and Recycling, 74, 170-179.

Small H. (1973), *Co-citation in the Scientific Literature: A New Measure of the Relationship between Two Documents*, Journal of the American Society for Information Science, 24(4), 265-269.

Song W., Ming X., Liu H.-C. (2017), *Identifying Critical Risk Factors of Sustainable Supply Chain Management: A Rough Strength-relation Analysis Method*, Journal of Cleaner Production, 143, 100-115.

Song Y., Chen X., Hao T., Liu Z., Lan Z. (2019), *Exploring Two Decades of Research on Classroom Dialogue by Using Bibliometric Analysis*, Computers and Education, 137(April), 12-31. https://doi.org/10.1016/j.compedu.2019.04.002

Srivastava S.K. (2007), *Green Supply-chain Management: A State-of the-art Literature Review*, International Journal of Management Reviews, 9(1), 53-80.

Stević Ž., Pamučar D., Puška A., Chatterjee P. (2020), *Sustainable Supplier Selection in Healthcare Industries Using a New MCDM Method: Measurement of Alternatives and Ranking According to COmpromise Solution (MARCOS)*, Computers & Industrial Engineering, 140, 106231. https://doi.org/https://doi.org/10.1016/j.cie.2019.106231

Su C.-M., Horng D.-J., Tseng M.-L., Chiu A.S.F., Wu K.-J., Chen H.-P. (2016), *Improving Sustainable Supply Chain Management Using a Novel Hierarchical grey-DEMATEL Approach*, Journal of Cleaner Production, 134, 469-481.

Tang X., Wei G. (2018a), *Models for Green Supplier Selection in Green Supply Chain Management with Pythagorean 2-tuple Linguistic Information*, Ieee Access, 6, 18042-18060.

Tang X., Wei G. (2018b), *Some Generalized Pythagorean 2-tuple Linguistic Bonferroni Mean Operators in Multiple Attribute Decision Making*, Journal of Algorithms & Computational Technology, 12(4), 387-398.

Tang X., Wei G., Gao H. (2019), *Models for Multiple Attribute Decision Making with Interval-valued Pythagorean Fuzzy Muirhead Mean Operators and Their Application to Green Suppliers Selection*, Informatica, 30(1), 153-186.

Tavana M., Yazdani M., Di Caprio D. (2017), *An Application of an Integrated ANP–QFD Framework for Sustainable Supplier Selection*, International Journal of Logistics Research and Applications, 20(3), 254-275.

Thomé A.M., Scavarda L., Scavarda A. (2016), *Conducting Systematic Literature Review in Operations Management*, Production Planning & Control, 27, 1-13. https://doi.org/10.1080/09537287.2015.1129464

Trapp A.C., Sarkis J. (2016), *Identifying Robust Portfolios of Suppliers: A Sustainability Selection and Development Perspective*, Journal of Cleaner Production, 112, 2088-2100.

Tseng M.-L., Chiu A.S.F. (2013), *Evaluating Firm's Green Supply Chain Management in Linguistic Preferences*, Journal of Cleaner Production, 40, 22-31.

Usdiken B., Pasadeos Y. (1995), *Organizational Analysis in North-America and Europe − A Comparison of Cocitation Networks*, Organization Studies, 16(3), 503-526.

van der Rhee B., Verma R., Plaschka G. (2009), *Understanding Trade-offs in the Supplier Selection Process: The Role of Flexibility, Delivery, and Value-added Services/Support*, International Journal of Production Economics, 120(1), 30-41.

Van Eck N.J., Waltman L. (2017), *Citation-based Clustering of Publications Using CitNetExplorer and VOSviewer*, Scientometrics, 111(2), 1053-1070.

Wan S.-P., Jin Z., Dong J.-Y. (2018), *Pythagorean Fuzzy Mathematical Programming Method for Multi-attribute Group Decision Making with Pythagorean Fuzzy Truth Degrees*, Knowledge and Information Systems, 55(2), 437-466.

Wang J., Gao H., Wei G., Wei Y. (2019), *Methods for Multiple-attribute Group Decision Making with q-rung Interval-valued Orthopair Fuzzy Information and Their Applications to the Selection of Green Suppliers*, Symmetry, 11(1), 56.

Wang J., Wei G., Wei Y. (2018), *Models for Green Supplier Selection with Some 2-tuple Linguistic Neutrosophic Number Bonferroni Mean Operators*, Symmetry, 10(5), 131.

Weber C.A., Current J.R., Benton W.C.C. (1991), *Vendor Selection Criteria and Methods*, European Journal of Operational Research, 50(1), 2-18.

Wei G., Lu M., Tang X., Wei Y. (2018), *Pythagorean Hesitant Fuzzy Hamacher Aggregation Operators and Their Application to Multiple Attribute Decision Making*, International Journal of Intelligent Systems, 33(6), 1197-1233.

Wetzstein A., Hartmann E., Benton jr. W.C., Hohenstein N.-O. (2016), *A Systematic Assessment of Supplier Selection Literature – State-of-the-art and Future Scope*, International Journal of Production Economics, 182, 304-323.

Wu C.-M., Hsieh C.-L., Chang K.-L. (2013), *A Hybrid Multiple Criteria Decision Making Model for Supplier Selection*, Mathematical Problems in Engineering, 2013(4).

Wu Q., Zhou L., Chen Y., Chen H. (2019), *An Integrated Approach to Green Supplier Selection Based on the Interval Type-2 Fuzzy Best-worst and Extended VIKOR Methods*, Information Sciences, 502, 394-417.

Xia W., Wu Z. (2007), *Supplier Selection with Multiple Criteria in Volume Discount Environments*, Omega, 35(5), 494-504.

Yager R.R. (2013), *Pythagorean Membership Grades in Multicriteria Decision Making*, IEEE Transactions on Fuzzy Systems, 22(4), 958-965.

Yang L., Chen Z., Liu T., Gong Z., Yu Y., Wang J. (2013), *Global Trends of Solid Waste Research from 1997 to 2011 by Using Bibliometric Analysis*, Scientometrics, 96(1), 133-146.

Yazdani M., Chatterjee P., Zavadskas E.K., Zolfani S.H. (2017), *Integrated QFD-MCDM Framework for Green Supplier Selection*, Journal of Cleaner Production, 142, 3728-3740.

Yu Y., Li Y., Zhang Z., Gu Z., Zhong H., Zha Q., Yang L., Zhu C., Chen E. (2020), *A Bibliometric Analysis Using VOSviewer of Publications on COVID-19*, Annals of Translational Medicine, 8(13), 816.

Zadeh L.A. (1965), *Fuzzy Sets*, Information and Control, 8(3), 338-353.

Zhang X., Xu Z. (2015), *Hesitant Fuzzy QUALIFLEX Approach with a Signed Distance-based Comparison Method for Multiple Criteria Decision Analysis*, Expert Systems with Applications, 42(2), 873-884.

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. (2011), *Fuzzy Set Theory − and Its Applications*, Springer Science & Business Media.

Zupic I., Čater T. (2015), *Bibliometric Methods in Management and Organization*, Organizational Research Methods, 18(3), 429-472.

MULTIPLE CRITERIA DECISION MAKING
Vol. 17 2022

Francisco Salas-Molina[*]
David Pla-Santamaria[**]
Ana Garcia-Bernabeu[***]
Javier Reig-Mullor[****]

IMPLICATIONS OF PARAMETER SELECTION IN DYNAMIC MULTIOBJECTIVE MODELS IN ECONOMICS AND FINANCE

DOI: 10.22367/mcdm.2022.17.02

Received: 18.07.2022 | Revised: 12.12.2022 | Accepted: 1.03.2023.

Abstract

Time is a key variable in the field of economics and finance. However, most of the classic approaches to economic problems are static. In this paper, we first review the existing literature on the use of multiobjective techniques to control dynamic systems within the area of economics and finance. We also tackle the question of which measure should we use to evaluate alternative solutions. To this end, we elaborate on the meaning added by the selection of a parameter in a family of distance functions used to evaluate alternative solutions.

Keywords: time, dynamic systems, economics, distance function, review.

1 Introduction

Multiple criteria decision making (MCDM) problems are characterized by the presence of several conflicting objectives that are considered simultaneously. We

[*] Universitat Politècnica de València, Ferràndiz y Carbonell, s/n, 03801, Alcoy, Spain, e-mail: frasamo@upv.es, ORCID: 0000-0002-1168-7931.
[**] Universitat Politècnica de València, Ferràndiz y Carbonell, s/n, 03801, Alcoy, Spain, e-mail: dplasan@upv.es, ORCID: 0000-0003-1563-3997.
[***] Universitat Politècnica de València, Ferràndiz y Carbonell, s/n, 03801, Alcoy, Spain, e-mail: angarber@upv.es, ORCID: 0000-0003-3181-7745.
[****] Universidad Miguel Hernández de Elche, Av. de la Universidad, s/n, 03202, Elche, Spain, e-mail: javier.reig@umh.es, ORCID: 0000-0002-1595-0581.

formulate each relevant aspect as an objective function and we use multiobjective optimization to find the best solutions. MCDM dates back to the works by Pareto at the end of the 19th century, but the field has grown very fast during the last decades. Some general works on MCDM include, but are not limited to, Yu (1985), Steuer (1986), Romero (1991), Ballestero and Romero (1998), Erhgott (2005), Jones and Tamiz (2010), and Zopounidis and Pardalos (eds., 2010).

Within the context of MCDM, there is a group of problems in which time is a key variable in addition to the criteria under consideration. Time is particularly relevant in the field of economics and finance. However, most of the classic approaches to economic problems are static as in Ballestero and Romero (1998). In this paper, we focus on multiobjective techniques used to control dynamic systems. By dynamic multiobjective problems we mean multiperiod problems in which we want to optimize a set of objective functions over time. This definition includes problems in which we want to optimize the final or cumulative state of criteria, as in Caballero et al. (1998), but also the deviation of the trajectory of these criteria over time with respect to a given reference as described in Wierzbicki (1988). As a result, the first goal of this paper is to review relevant papers in multiobjective control within the fields of economics and finance. We restrict ourselves to economic and financial models because time series and multistage problems are ubiquitous in these areas.

An important research question arises when dealing with multiobjective control problems: which measure should we use to evaluate alternative solutions? If we adhere to dynamic goal programming, we should minimize the sum of deviations for each time step. However, other alternatives suggest the use of maximum absolute deviations or percentage deviations. Most of the alternatives are indeed special cases of the Minkowski distance function when a parameter is set to a particular integer value. As a second goal of the paper, we discuss the implications of selecting this parameter, extending the work by Gonzalez-Pachon and Romero (2016) in terms of the meaning added to the process. More precisely, we argue that there is an implicit selection of a decision-making principle when this parameter is set to a value. Finally, we reformulate multiobjective control problem as a constrained norm approximation problem. This reformulation presents the advantage of being convex and then solvable for any order of the norm (value of the parameter) using state-of-the-art convex optimization algorithms.

Summarizing, the contributions of this paper are twofold:

1. A review of relevant papers in multiobjective dynamic models in economics and finance.
2. An analysis of the decision-making principles that underlie the selection of a parameter in the Minkowski distance function.

In addition to this introduction, the structure of this paper is as follows. Section 2 reviews a set of the most relevant papers in multiobjective dynamic control. Section 3 discusses the implications of selecting a parameter in multi-objective dynamic control problems. Section 4 provides concluding remarks.

2 A review of multiobjective dynamic models in economics and finance

We can set the origins of a formal treatment of dynamic systems in the book by Bellman (1957). In this book, Bellman provided an introduction to the mathematical theory of multistage decision processes and introduced the notion of dynamic programming (DP) to describe the subject matter. Other basic concepts, such as the Bellman equation to derive optimal policies in a recursive manner, paved the way to recent advances in economics. For instance, Sargent and Ljungqvist (2000) used several recursive methods to study macroeconomics while Hansen and Sargent (2013) studied recursive models of dynamic linear economies. Briefly, recursive models break a multistage problem into small pieces by forming a sequence of time-dependent problems.

Following the definition by Kall and Wallace (1994), dynamic problems are characterized by stages or time steps indexed by t, the state x_t at time t, the decision taken u_t at time t, the transformation of the system from the current state and the decision taken to the next state, the return r_t (x_t, u_t) obtained at time t, the set X of feasible decisions, and the overall objective function F which depends on the returns r_t for the whole planning horizon T. The consideration of time as a key variable in a decision-making problem adds a new level of complexity to the problem. Indeed, Bellman (1957) refers to dynamic problems as multidimensional maximization problems. In this paper, we argue that a natural way to deal with multidimensional problems is to use multiple criteria decision making (MCDM).

MCDM covers a wide range of techniques as described in Yu (1985), Steuer (1986), Romero (1991), Ballestero and Romero (1998), Erhgott (2005), Jones and Tamiz (2010), and Zopounidis and Pardalos (eds., 2010). Among them, Goal Programming (GP) initially proposed by Charnes and Cooper (1957) is one of the most widely used techniques. The introduction of dynamic features into the problem led to the development of Dynamic Goal Programming (DGP). Yu and Leitmann (1974) considered a dynamic multiobjective decision problem in which the concepts of non-dominated solutions were extended to a dynamic context. The use of trajectories over the planning horizon that play the role of a reference signal for optimization purposes was proposed by Wierzbicki (1980; 1988).

Daellenbach and de Kluyver (1980) introduced a multiobjective dynamic programming (MODP) technique as an extension of dynamic programming concepts. Levary (1984) proposed a scalarization approach by means of GP. Later on, Li and Haimes (1989) highlighted the development of the research area and reviewed both the concepts and the works in relation to theory and practice of MODP. On the other hand, Opricovic (1993) developed a compromise programming method (Zeleny, 1973) by minimizing the distance to the ideal solution within a dynamic context and with application to water reservoir management.

In the proceedings of two multiple criteria decision-making conferences, Trzaskalik (1997a; 1997b) discussed several aspects, such as monotonicity and separability in a multiple criteria context. Caballero et al. (1998) described an approach with dynamic target values to control not only the final values of the objective functions, but also their evolution along the planning horizon. Discrete dynamic programming with partially ordered criteria set was also considered by Trzaskalik and Sitarz (2002; 2007). More recently, Trzaskalik (2022) provided novel theoretical results on the possibility of finding the best multistage policies using Bellman's optimality principle and the multicriteria bipolar method in which two sets of references points are determined.

Zopounidis and Doumpos (2013) analyzed the importance of multicriteria decision systems for financial problems and reviewed the most relevant papers in two main areas of financial decision support, namely, portfolio selection and corporate performance evaluation. In what follows, we adopt a similar approach, but focusing only on the dynamic aspects of multicriteria decision-making models in finance.

In what follows, we pay special attention to the dynamic portfolio selection problem. Probably the most studied problem in multicriteria financial decision-making is the portfolio selection problem, due to the pioneering work by Markowitz (1952). However, the classical mean-variance model by Markowitz (1952) is a one-period model. This fact is critical because investors are usually concerned with cumulative results over a period of time and optimal decisions for a single period may be suboptimal in a multiperiod framework according to Estrada (2010). To overcome this limitation, Kelly (1956) in the context of gambling and Latane (1959) in the context of investing proposed a multiperiod framework with cumulative results which is equivalent to the maximization of the geometric mean of returns.

Mossin (1968) proposed an extension of the one-period model to a multiperiod framework following a dynamic programming approach and acknowledging first that Tobin (1965) appeared to be one of the first authors to make an attempt in this direction. What is most interesting in Mossin (1968) is the definition of the single-period problem and the multiperiod problem that we reproduce here for clarity:

> "By a single-period model is meant a theory of the following structure: The investor makes his portfolio decision at the beginning of a period and then waits until the end of the period when the rate of return on his portfolio materializes. He cannot make any intermediate changes in the composition of his portfolio. The investor makes his decision with the objective of maximizing expected utility of wealth at the end of the period (final wealth)" (Mossin, 1968, p. 216).

> "By a multiperiod model is meant a theory of the following structure: The investor has determined a certain future point in time (his horizon) at which he plans to consume whatever wealth he has then available. He will still make his investment decisions with the objective of maximizing expected utility of wealth at that time. However, it is now assumed that the time between the present and his horizon can be subdivided into *n* periods (not necessarily of the same length), at the end of each of which return on the portfolio held during the period materializes and he can make a new decision on the composition of the portfolio to be held during the next period" (Mossin, 1968, p. 220).

Instead of maximizing expected utility functions of the terminal wealth and/or multiperiod consumption, Li and Ng (2000) proposed an analytical method for the mean-variance formulation to find the multiperiod optimal portfolio policy. Zhou and Li (2000) also used the mean-variance formulation to select portfolios in a continuous framework using a stochastic linear-quadratic model. This line of work was later extended by Basak and Chabakauri (2010), Wang and Zhou (2020), Dai et al. (2021) and many others.

More recently, Ben Abdelazziz et al. (2020) also proposed a stochastic dynamic multiobjective model for sustainable decision-making with applications in sustainable portfolio management with two stocks and two criteria (return and sustainability), and also in a workforce allocation problem in an economy with two sectors.

A novel line of research has recently arised from the application of the multiobjective dynamic techniques derived from the portfolio selection problem to the cash management problem by Salas-Molina, Pla-Santamaria and Rodriguez-Aguilar (2018a), Salas-Molina, Pla-Santamaria and Rodríguez-Aguilar (2018b), Salas-Molina, Rodríguez-Aguilar and Pla-Santamaria (2018) and Salas-Molina (2019). In this area of research, Sethi and Thomson (1970; 2000) proposed an optimal control theory approach to the cash management problem that has been recently extended by Bhaya and Kaszkurewicz (2022) in a single-objective context.

3 Implications of parameter selection in multiobjective dynamic control models

The main goal of this section is to provide a way to add meaning to multiobjective dynamic models by selecting a particular form of the objective function used for optimization purposes. To this end, we first formulate a general dynamic goal program in which a parametric distance function is used to find the best solutions. Second, we analyze the implications of selecting a key parameter in this distance function in terms of the implicit decision-making principle derived from this choice. We illustrate the implications by means of the analysis of the most important cases.

Let us start with the classical GP formulation by Charnes and Cooper (1957):

$$min \sum_{i=1}^{q}(w_i^{+}\delta_i^{+} + w_i^{-}\delta_i^{-}) \tag{1}$$

subject to:

$$g_i(\boldsymbol{u},\boldsymbol{x}) + \delta_i^{-} - \delta_i^{+} = b_i \tag{2}$$

$$\delta_i^{-}, \delta_i^{+} \geq 0 \tag{3}$$

$$u \in S \tag{4}$$

where we consider the positive $\boldsymbol{\delta_i^{+}}$ and negative deviations $\boldsymbol{\delta_i^{-}}$ of q different goals achievements measured by $\boldsymbol{g_i}(\boldsymbol{u},\boldsymbol{x})$ from targets $\boldsymbol{b_i}$. Goal achievements depend on control actions in vector $\boldsymbol{u}$ subject to some feasibility set S and states in vector $\boldsymbol{x}$.

By including time as a key variable in the previous GP formulation, we are dealing with a multiobjective control problem described as the minimization of deviations with respect to some dynamic targets or trajectories as proposed, for instance, by Wierzbicki (1988) and Caballero et al. (1998). We are dealing with a dynamic goal program (DGP):

$$\min \sum_{i=1}^{q}\sum_{t=1}^{n}(w_i^{+}\delta_{it}^{+} + w_i^{-}\delta_{it}^{-}) \tag{5}$$

subject to:

$$g_{it}(\boldsymbol{u},\boldsymbol{x}) + \delta_{it}^{-} - \delta_{it}^{+} = b_{it} \tag{6}$$

$$\delta_{it}^{-}, \delta_{it}^{+} \geq 0 \tag{7}$$

$$\boldsymbol{u} \in S \tag{8}$$

We can now move one step further by considering parameter p in the DGP formulation:

$$\min \left[\sum_{\mathrm{i}=1}^{\mathrm{q}} \sum_{t=1}^{n} (w_i^{+} \delta_{it}^{+})^p + (w_i^{-} \delta_{it}^{-})^p\right]^{1/\mathrm{p}} \tag{9}$$

subject again to equations (6), (7) and (8). The use of this parameter allows us to increase the degree of generality and, at the same time, to add meaning to the optimization process. We increase the degree of generality because we are able to consider not only linear deviations but also quadratic or maximum deviations as in the case of the Chebyshev variant of the classical linear GP formulation. Furtheremore, we are adding meaning to the optimization process, because by setting p, we are implicitly selecting a decision-making principle as we elaborate it next.

3.1 Case p = 1, linear control and the principle of maximum efficiency

For simplicity of notation, let us assume that all goal functions $\boldsymbol{g_i}(\boldsymbol{u}, \boldsymbol{x})$ are equally-weighted normalised non-negative linear functions of states in vector $\boldsymbol{x}$ and controls in vector $\boldsymbol{u}$ that are subject to a given set of constraints. As a result, when we set $\boldsymbol{p} = \mathbf{1}$ in equation (9), we are indeed minimizing the sum of absolute deviations. And this minimization can be viewed as the application of the principle of maximum efficiency (Gonzalez-Pachon and Romero, 2016), because we focus on the sum of achievements disregarding particular deviations in favour of the sum (majority) of deviations.

$$\min \sum_{i=1}^{q} \sum_{t=1}^{n} |g_{it}(\boldsymbol{u}, \boldsymbol{x}) - b_{it}| \tag{10}$$

In this case, we apply a multiobjective linear control of a set of q goals determined by a set of dynamic targets (or trajectories) over a planning horizon of n time steps.

3.2 Case p = ∞, minimax control and the principle of maximum fairness

Now consider the case when we set $\boldsymbol{p} = \infty$ in equation (9). In this case, we minimize the maximum absolute deviations and this minimization can be viewed as the application of the principle of maximum fairness, because we focus on the worst observation as suggested by Gonzalez-Pachon and Romero (2016).

$$\min \left[\sum_{i=1}^{q} \sum_{t=1}^{n} |g_{it}(\boldsymbol{u}, \boldsymbol{x}) - b_{it}|^{\infty}\right]^{1/\infty} \rightarrow \min \max(|g_{it}(\boldsymbol{u}, \boldsymbol{x}) - b_{it}|) \tag{11}$$

As a result, we apply a minimax control of a set of q goals determined by a set of dynamic targets (or trajectories) over a planning horizon of n time steps. The main implication here is that we are making decisions based on a single observation and this could be a problem for long time horizons.

3.3 Case p = 2, quadratic control and the principle of balance

When $\boldsymbol{p} = \mathbf{2}$, we minimize the Euclidean distance between a reference signal (dynamic targets) and goal achievement. This minimization can be viewed as the application of the principle of balance because we are somewhere in between the cases $\boldsymbol{p} = \mathbf{1}$ and $\boldsymbol{p} = \infty$.

$$\min \left[\sum_{i=1}^{q} \sum_{t=1}^{n} (g_{it}(\boldsymbol{u}, \boldsymbol{x}) - b_{it})^2\right]^{1/2} \tag{12}$$

In this case, we apply a quadratic control of a set of q goals determined by a set of dynamic targets (or trajectories) over a planning horizon of n time steps. This approach represents a compromise between the principle of maximum efficiency when $\boldsymbol{p} = \mathbf{1}$ (the rule of the majority) and the principle of maximum fairness when $\boldsymbol{p} = \infty$ (the rule of the most disadvantaged).

3.4 Case p = 0, geometric control and the principle of limited compensability

There is another case which is not so common in the literature but which leads to another important decision-making principle, namely, the principle of limited compensability. It can be shown that when $\boldsymbol{p} = \mathbf{0}$, equation (9) is equivalent to considering the product of deviations. This approach implies the principle of limited compensability because we limit the offset between bad performance in one deviation with superior performance in other deviations:

$$\left[\sum_{i=1}^{q} \sum_{t=1}^{n} |g_{it}(\boldsymbol{u}, \boldsymbol{x}) - b_{it}|^{p}\right]^{1/\mathrm{p}} \rightarrow \prod_{i=1}^{q} \prod_{t=1}^{n} |g_{it}(\boldsymbol{u}, \boldsymbol{x}) - b_{it}| \tag{13}$$

As a result, we apply a geometric control of a set of q goals determined by a set of dynamic targets (or trajectories) over a planning horizon of n time steps. In this case, it is convenient to set the targets to the anti-ideal values and maximize the product of deviations with respect to these anti-ideal values, because otherwise, a single null deviation would lead to a minimum of the functional.

3.5 Combination of decision-making principles

As suggested by Gonzalez-Pachon and Romero (2016), we can use parameter $\boldsymbol{\lambda}$ to produce a combination of decision-making principles. For instance, we can consider a weighted combination of the principle of maximum efficiency and the principle of maximum fairness by minimizing the following functional:

$$\lambda\mathcal{L}_1 + (1-\lambda)\mathcal{L}_\infty \tag{14}$$

where $\boldsymbol{\mathcal{L}_1}$ and $\boldsymbol{\mathcal{L}_\infty}$ are, respectively, the parametric distance function in equation (9) when $\boldsymbol{p = 1}$ and $\boldsymbol{p = \infty}$. We can extend this approach by considering any value of *p* in the range between zero and infinity as a potential representation of an additional decision-making principle. For instance, $\boldsymbol{p = 3}$ can be viewed, at least in theory, as a new principle that lies between the principle of maximum efficiency and the principle of maximum fairness and one step further of the principle of balance represented by $\boldsymbol{p = 2}$.

3.6 Solving the problem

It is obvious that when $p > 1$, we are dealing with a non-linear problem that may result in difficulties to find the optimal policies. However, the minimization of a sum of deviations of q goals over planning horizon *n* raised to *p* and the whole sum raised to *1/p* is equivalent to the minimization of the *p*-norm of a vector of dimension $n \cdot q$:

$$\min \left\| [\delta_{11}, \delta_{12}, \dots, \delta_{nq}] \right\|_p \tag{15}$$

If the minimization of the *p*-norm of a vector of deviations with respect to dynamic targets is subject to a set of linear constraints, we are dealing with a constrained norm approximation problem. Fortunately, this problem is convex and can be solved for any value of p using state-of-the-art convex optimization algorithms such as CVXPY within CPLEX or Gurobi (Boyd and Vandenberghe, 2004). When we consider non-linear goals or constraints, we need to apply some heuristics to solve the problem.

4 Concluding remarks

The main research question addressed in this paper is whether we can add meaning to the optimization process in multiobjective control. To this end, we consider dynamic goal programming that usually deals with linear-quadratic control problems as a starting point to propose a more general approach. This approach is based on the selection of a parameter of the Minkowski distance function. Extending previous works on the subject, we show that the selection of this parameter implies the use of multiple decision-making principles that may help practitioners to motivate the use of objective functions to derive control policies.

We also highlight the point that any value of p can be interpreted as representative of any decision-making principle. In order to deal with non-linearity of some of the decision-making principles, we suggest the use of constrained norm approximation methods to solve a general multiobjective control problem.

References

Ballestero E., Romero C. (1998), *Multiple Criteria Decision Making and Its Applications to Economic Problems*, Kluwer Academic Publishers, Dordrecht.

Basak S., Chabakauri G. (2010), *Dynamic Mean-variance Asset Allocation*, The Review of Financial Studies, 23(8), 2970-3016.

Bellman R. (1957), *Dynamic Programming*, Princeton University Press, Princeton.

Ben Abdelaziz F., Colapinto C., La Torre D., Liuzzi D. (2020), *A Stochastic Dynamic Multiobjective Model for Sustainable Decision Making*, Annals of Operations Research, 293(2), 539-556.

Bhaya A., Kaszkurewicz E. (2022), *The Generalized Cash Balance Problem: Optimization-based One Step Ahead Optimal Control*, International Transactions in Operational Research (in press).

Boyd S., Vandenberghe L. (2004), *Convex Optimization*, Cambridge University Press, Cambridge.

Caballero R., Gómez T., González M., Rey L., Ruiz F. (1998), *Goal Programming with Dynamic Goals*, Journal of Multi-Criteria Decision Analysis, 7(4), 217-229.

Charnes A., Cooper W.W. (1957), *Management Models and Industrial Applications of Linear Programming*, Management Science, 4(1), 38-91.

Daellenbach H.G., De Kluyver C.A. (1980), *Note on Multiple Objective Dynamic Programming*, Journal of the Operational Research Society, 591-594.

Dai M., Jin H., Kou S., Xu Y. (2021), *A Dynamic Mean-Variance Analysis for Log Returns*, Management Science, 67(2), 1093-1108.

Ehrgott M. (2005), *Multicriteria Optimization* (Vol. 491), Springer Science & Business Media, Berlin.

Estrada J. (2010), *Geometric Mean Maximization: An Overlooked Portfolio Approach?* The Journal of Investing, 19(4), 134-147.

González-Pachón J., Romero C. (2016), *Bentham, Marx and Rawls Ethical Principles: In Search for a Compromise*, Omega, 62, 47-51.

Hansen L.P., Sargent T.J. (2013), *Recursive Models of Dynamic Linear Economies* [in:] *Recursive Models of Dynamic Linear Economies*, Princeton University Press, Princeton.

Jones D., Tamiz M. (2010), *Practical Goal Programming* (Vol. 141), Springer, Berlin.

Kall P., Wallace S.W. (1994), *Stochastic Programming* (Vol. 6), Wiley, Chichester.

Kelly J. (1956), *A New Interpretation of Information Rate*, Bell System Technical Journal, 35, 917-926.

Latane H. (1959), *Criteria for Choice among Risky Ventures*, Journal of Political Economy, 67, 144-155.

Levary R.R. (1984), *Dynamic Programming Models with Goal Objectives*, International Journal of Systems Science, 15(3), 309-314.

Li D., Haimes Y.Y. (1989), *Multiobjective Dynamic Programming: The State of the Art*, Control, Theory and Advanced Technology, 5(4), 471-483.

Li D., Ng W.L. (2000), *Optimal Dynamic Portfolio Selection: Multiperiod Mean-Variance Formulation*, Mathematical Finance, 10(3), 387-406.

Markowitz H. (1952), *Portfolio Selection*, Journal of Finance, 7, 77-91.

Mossin J. (1968), *Optimal Multiperiod Portfolio Policies*, The Journal of Business, 41(2), 215-229.

Opricović S. (1993), *Dynamic Compromise Programming with Application to Water Reservoir Management*, Agricultural Systems, 41(3), 335-347.

Romero C. (1991), *Handbook of Critical Issues in Goal Programming*, Pergamon Press, Oxford.

Salas-Molina F., Pla-Santamaria D., Rodriguez-Aguilar J.A. (2018a), *A Multi-objective Approach to the Cash Management Problem*, Annals of Operations Research, 267(1-2), 515-529.

Salas-Molina F., Pla-Santamaria D., Rodríguez-Aguilar J.A. (2018b), *Empowering Cash Managers through Compromise Programming* [in:] *Financial Decision Aid Using Multiple Criteria* (149-173), Springer, Cham.

Salas-Molina F., Rodríguez-Aguilar J.A., Pla-Santamaria D. (2018), *Boundless Multiobjective Models for Cash Management*, The Engineering Economist, 63(4), 363-381.

Salas-Molina F. (2019), *Selecting the Best Risk Measure in Multiobjective Cash Management*, International Transactions in Operational Research, 26(3), 929-945.

Sargent T.J., Ljungqvist L. (2000), *Recursive Macroeconomic Theory*, Massachusetss Institute of Technology, Massachusetss.

Sethi S.P., Thompson G.L. (1970), *Applications of Mathematical Control Theory to Finance: Modeling Simple Dynamic Cash Balance Problems*, Journal of Financial and Quantitative Analysis, 5(4-5), 381-394.

Sethi S.P., Thompson G.L. (2000), *Optimal Control Theory: Applications to Management Science and Economics* (2nd Edition), Springer, Berlin.

Steuer R.E. (1986), *Multiple Criteria Optimization. Theory, Computation and Applications*, John Wiley & Sons, New York.

Tobin J. (1965), *The Theory of Portfolio Selection* [in:] *The Theory of Interest Rate* (3-51), Macmillan, London.

Trzaskalik T. (1997a), *Dynamic Goal Programming Models* [in:] *Advances in Multiple Objective and Goal Programming* (111-119), Proceedings of the 2nd International Conference on Multi-Objective Programming and Goal Programming (Torremolinos, Spain), Springer, Berlin.

Trzaskalik T. (1997b), *Multiple Criteria Discrete Dynamic Programming* [in:] *Multiple Criteria Decision Making* (202-211), Proceedings of the 12th International Conference (Hagen, Germany), Springer, Berlin.

Trzaskalik T. (2022), *Multiobjective Dynamic Programming in Bipolar Multistage Method*, Annals of Operations Research, 311(2), 1259-1279.

Trzaskalik T., Sitarz S. (2002), *Dynamic Discrete Programming with Partially Ordered Criteria Set* [in:] *Multiple Objective and Goal Programming* (186-195), Physica, Heidelberg.

Trzaskalik T., Sitarz S. (2007), *Discrete Dynamic Programming with Outcomes in Random Variable Structures*, European Journal of Operational Research, 177(3), 1535-1548.

Wang H., Zhou X.Y. (2020), *Continuous-time Mean-variance Portfolio Selection: A Reinforcement Learning Framework*, Mathematical Finance, 30(4), 1273-1308.

Wierzbicki A.P. (1980), *A Methodological Guide to Multiobjective Optimization* [in:] *Optimization Techniques* (99-123), Springer, Berlin, Heidelberg.

Wierzbicki A.P. (1988), *Dynamic Aspects of Multi-objective Optimization* [in:] A. Lewandowski, V. Volkovich (eds.), *Multiobjective Problems of Mathematical Programming* (154-174), Lecture Notes in Econ. and Math. Syst. No. 351, Springer, Berlin.

Yu P.L. (1985), *Multiple-criteria Decision Making: Concepts, Techniques, and Extensions*, Plenum Press, New York.

Yu P.L., Leitmann G. (1976), *Nondominated Decisions and Cone Convexity in Dynamic Multicriteria Decision Problems* [in:] *Multicriteria Decision Making and Differential Games* (61-72), Springer, Boston.

Zeleny M. (1973), *Compromise Programming* [in:] *Multiple Criteria Decision Making* (262-301), University of South Carolina Press.

Zhou X.Y., Li D. (2000), *Continuous-time Mean-variance Portfolio Selection: A Stochastic LQ Framework*, Applied Mathematics and Optimization, 42(1), 19-33.

Zopounidis C., Doumpos M. (2013), *Multicriteria Decision Systems for Financial Problems*, Top, 21(2), 241-261.

Zopounidis C., Pardalos P.M., eds. (2010), *Handbook of Multicriteria Analysis*, Springer Science & Business Media, Berlin.

MULTIPLE CRITERIA DECISION MAKING
Vol. 17 2022

Pedro Nunes Lopes Neto*
José Celso Freire Junior**
Celso Eduardo Tuna***

RANKING OF LTE CELLS BASED ON KEY PERFORMANCE INDICATORS USING MCDM METHODS

DOI: 10.22367/mcdm.2022.17.03

Received: 1.08.2022 | Revised: 8.01.2023 | Accepted: 1.03.2023.

Abstract

The growth in worldwide data traffic and user subscriptions in mobile telecommunication networks makes it increasingly difficult to manage network performance in an environment already containing multiple radio access technologies. Despite the rise of 5G, LTE remains the dominant technology, and new cells are installed daily to support traffic growth and new services such as voice over LTE. Detecting faulty cells in the network is one of the main concerns of operators. Self-organizing networks have been introduced to deal with this problem, and their self-healing functionality has improved cell fault management. Nonetheless, faulty cell detection remains challenging, and most of the tasks involved are still done manually. This paper introduces a new method of faulty cell detection in an LTE radio access network, applying multiple criteria methods to this problem. The cells are ranked based on selected key performance indicators, using the multi-attribute utility theory to construct a utility function. The analytic hierarchy process is used to define weights for the criteria.

Keywords: long-term evolution, multiple criteria methods, radio access network performance management, self-organizing networks.

* São Paulo State University, Campus of Guaratinguetá, SP, 12516-410, Brazil, e-mail: nunes.lopes@unesp.br.

** São Paulo State University, Electrical Engineering Department, Campus of Guaratinguetá, SP, 12516-410, Brazil, e-mail: jose-celso.freire@unesp.br, ORCID: 0000-0002-2519-9808.

*** São Paulo State University, Chemistry and Energy Department, Campus of Guaratinguetá, SP, 12516-410, Brazil, e-mail: celso.tuna@unesp.br, ORCID: 0000-0002-2020-7063.

1 Introduction

Despite the uncertainty caused by the COVID-19 pandemic, mobile subscriptions continue to grow globally, bolstered by the new 5G NR (5th Generation New Radio) radio access technology (RAT). Nonetheless, 4G LTE (4th Generation Long Term Evolution) remains the dominant RAT by subscription, and voice over LTE (VoLTE) service enables interoperable voice and communication services on 4G and 5G devices. VoLTE adoption also accelerates the decommissioning of 2G and 3G networks, freeing frequencies for use by new LTE bands. Meanwhile, the year-on-year mobile network data traffic growth remains at around 50%, driven by the rising number of smartphone subscriptions and an increasing average data volume per subscription (Jejdling, 2020).

While the data traffic continues to grow and VoLTE service continues to expand, new technologies, such as dynamic spectrum sharing (DSS), allow LTE and 5G to share the same carrier (Nory, 2019). Carrier aggregation between the two RATs is driving operators to expand their LTE access network capacity even more, adding more layers and, in the end, more cells to the existing network. The complexity of managing and operating such networks is forcing operators to auto-mate many operational processes to remain competitive. Self-organizing networks (SONs) were introduced to reduce the operating expenditures associated with managing the increased number of cells by reducing the need for manual network planning, configuration, and optimization (3GPP TS 32.500, 2020). SON functionalities (described in the next section) can be classified as self-configuration, self-optimization, or self-healing.

Barco, Lazaro, and Munoz (2012) point out that there are few studies on self-healing (Sallent et al., 2011; Hu et al., 2010), and emphasize the complexity of cell fault detection problems. They are usually revealed not through a highly anomalous value of one key performance indicator (KPI) but through slightly abnormal values of several KPIs. Szilagyi and Novaczki (2012) further point out that self-healing studies focus mainly on simple use cases, such as detecting complete cell outages. This paper proposes a new decision process to contribute to the literature on the self-healing of networks.

This process weighs multiple indicators and ranks the cells by their performance, filtering the most degrading ones to facilitate network operation and management. It is based on the Multiple Criteria Decision Methods (MCDM), defined by Obayiuwana and Falowo (2017) as an advanced technique of optimization research for resolving decision problems with multiple criteria using a more robust, explicit, rational, and efficient decision-making process.

According to Obayiuwana and Falowo (2017), MCDM methods have been used primarily for network selection decisions in situations where different RATs coexist. MCDM methods have rarely been applied to other RAT-related problems. Moreover, MCDM methods have been used in cases with a limited number of alternatives, such as RATs or small groups of cells. They have not been used for detecting fault cells while considering the complete network as an alternative.

This paper proposes a new application of MCDM methods in an LTE radio access network to detect and rank faulty cells based on key performance indicators (KPIs), considering all the cells of a given network.

The paper is organized as follows. Section 2 introduces the concepts and theories used, outlining the current decision methods used in radio access networks. Section 3 presents the proposed approach, highlighting its novelty and contributions to multiple criteria decision problems. Section 4 presents some results from applying the proposed method in an actual LTE network. Finally, Section 5 summarizes the contributions of this paper.

2 Theoretical and conceptual background

This section provides the background for the concepts and theories used in the paper. The idea of self-organizing networks is described, with particular attention to self-healing. Cell fault detection and key performance indicators are also discussed.

2.1 Self-organizing network concepts

In 2008, the Next Generation Mobile Networks (NGMN) Alliance, an open forum founded by major mobile network operators, defined the requirements and recommendations for implementing self-organizing networks (Next generation, 2008). This allowed the automation of some network planning, configuration, and optimization processes through SON functionalities (3GPP TS 32.500, 2020). The functionalities indicated by NGMN were self-configuration, self--optimization, fault management, and fault correction (subsequently renamed self-healing). Later, the 3rd Generation Partnership Project (3GPP), which provides reports and specifications for cellular telecommunications technologies, introduced SON in its standards as a fundamental element for LTE deployment (Barco, Lazaro, and Munoz, 2012) and defined the main SON functionalities based on the NGMN requirements (3GPP TS 32.500, 2020). The main SON functionalities are summarized by Barco, Lazaro, and Munoz (2012):

1. Self-configuration: includes functions for network deployment and configuration of its parameters. Thanks to autoconfiguration, network elements can start autonomously, run setup routines, and configure initial parameters.

2. Self-optimization: responsible for auto-tuning parameters, which should be dynamically recalculated when traffic and network conditions change. Self-optimization includes tuning parameters related to the list of neighboring cells, traffic balance, handover, and coverage.
3. Self-healing: includes functions to cope with service degradations or outages, including fault detection and diagnosis and mechanisms for outage compensation.

The first two functionalities are well-documented, and some functions, such as automatic neighbor relation (ANR) and node auto-connectivity, were even used in the first LTE deployments. Self-configuration reduces costs and accelerates cell deployment in the network, while self-optimization provides operational cost savings through energy saving or load-balancing optimization. On the other hand, studies on self-healing are scarce, as it is the most complex of the three domains due to the variety of vendors, software versions, and hardware types coexisting in a single network (Szilagyi and Novaczki, 2012). The existing studies on self-healing are incomplete, dealing only with certain straightforward self-healing aspects in specific scenarios, such as detecting complete cell outages (Barco, Lazaro, and Munoz, 2012). However, new studies on automatic fault detection and diagnosis can also reduce the cost of managing networks.

2.2 Self-healing concepts

As described by 3GPP, self-healing aims to solve or mitigate the faults that can be translated automatically by triggering appropriate recovery actions. The self-healing function consists of two parts: the monitoring part and the healing process part (3GPP TS 32.541, 2020). In the first part (shown in Figure 1), the trigger condition of self-healing (TCoSH), which could be either an alarm or a key performance indicator, is monitored; when the TCoSH is reached, a particular action is triggered to prevent or mitigate the specific fault. This article focuses on detecting a cell fault during the monitoring phase of the self-healing process.

Detecting and solving cell faults is one of the main concerns for network operators and vendors. Self-healing is required when a cell degrades, impacting the rest of the network. This kind of cell is called a problematic cell, and each operator and vendor uses a different indicator to identify the cell fault symptoms. A symptom is a measurement whose observed value helps identify a fault. Symptoms include key performance indicators, alarms, online measurements, and drive tests (Barco, Lazaro, and Munoz, 2012).

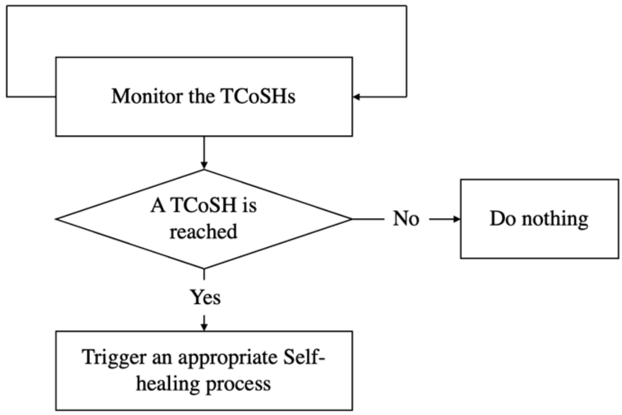

Figure 1: The monitoring part of the self-healing function
Source: Adapted from: 3GPP TS 32.541 (2020).

2.3 Cell fault detection

In general, alarms can identify a cell fault only in critical cases, i.e., software and hardware failures, transmission problems, or downtimes. A single fault may generate multiple alarms, and several different faults can trigger a single alarm. Furthermore, alarm messages cannot always be transmitted when a cell loses connectivity or stops sending information. Also, on many occasions, a cell fault does not generate any alarm. This can be caused by poor radio conditions, i.e., inadequate coverage, shadowing, or external interference (Barco, Lazaro, and Munoz, 2012), or else by incorrect configuration. Therefore, key performance indicators are the main inputs used by RAT experts to detect a cell fault and are used as criteria for the decision process proposed in this paper.

Barco, Lazaro, and Munoz (2012) propose a self-healing reference model, which is the basis for this paper's cell fault detection process. According to them, fault detection is responsible for identifying the problematic cells to be healed, including cells with service outage (cell outage detection) and cells with service degradation (cell degradation detection). A possible simple method of detecting a cell fault consists of setting thresholds for some KPIs. However, gradual degradations cannot be detected simply by a threshold, especially if proactive rather than reactive detection is carried out. Therefore, the authors state that an algorithm should be developed that considers all relevant KPIs and uses appropriate decision logic to determine whether an outage or degradation has occurred.

2.4 Key performance indicators

For managing purposes, to monitor the network's overall performance and compare the performance in different areas or periods, an operator needs to measure the statistical network performance periodically. Statistical data sampling can be performed regularly, i.e., daily, weekly, or monthly (3GPP TS 32.421, 2015). Performance data are collected and recorded by network elements, following a schedule established by the network element manager. These data are used to evaluate performance in five areas: network traffic, configuration, resource access, resource availability, and quality of service (QoS) (3GPP TS 32.401, 2018). QoS indicators measure the network performance the end user is expected to experience. These are the measurements considered in this paper.

Data performance is measured through specific parameters, or performance indicators (PIs), defined by each equipment vendor and used to monitor the current network status and performance. This enables prompt action to control, when necessary, the performance and resources of the network and services. A radio access network can have hundreds of PIs. Often measurements are taken simply because they are available, not because they are meaningful. It should be noted that the complete range of network status information and PIs is not necessary to manage the network. One of the challenges of managing networks is understanding which data are critical for supporting specific objectives (ITU-T Recommendation E.419, 2006). PIs representing the essential network performance measurements are called key performance indicators (KPIs).

For an LTE RAT, 3GPP defines six categories of KPIs (3GPP TS 32.450, 2019). All except for the last one can be used to measure the QoS. The categories are the following:

1. Accessibility KPIs: used to measure the availability of service within specified tolerances and other given conditions when requested by the user (ITU-T Recommendation E.800, 2008).
2. Retainability KPIs: used to measure the abnormal interruptions of service (ITU-T Recommendation E.800, 2008).
3. Mobility KPIs: used to measure how LTE mobility functionality is working.
4. Integrity KPIs: used to measure the data integrity, ensuring that data have not been altered in an unauthorized manner (ITU-T Recommendation E.800, 2008).
5. Energy efficiency KPIs: used to measure data energy efficiency in LTE network elements.
6. Availability KPIs: used to measure the percentage of times when the cell is considered available.

2.5 Multiple Criteria Decision Methods

Multiple Criteria Decision Methods have a relatively short history as a discipline. Their foundations were laid between 1950 and 1960, and they became the dominant paradigm in decision analysis and decision support in the presence of multiple evaluation dimensions (Zavadskas, Turskis and Kildienė, 2014). MCDM has been one of the fastest-growing problem areas in many disciplines, where a set of alternatives needs to be evaluated in terms of several criteria (Triantaphyllou, 2010). Nonetheless, there is no single well-defined methodology that one could follow from the beginning to the end of a decision-making process. When dealing with objects that can only be described and compared using multiple characteristics, aggregating them is a significant problem. The aggregation aims to synthesize the (usually contradictory) features of the objects to achieve a goal, such as choosing among the objects, rank-ordering them, sorting them into categories, and so on (Bouyssou et al., 2006).

MCDM methods use a wide range of approaches to solving the problems mentioned above. They can be broadly classified into two categories (cf. Figure 2): discrete MCDM or discrete MADM (multi-attribute decision-making) and continuous MODM (multi-objective decision-making) methods. MODM methods are associated with problems where alternatives are not predetermined. The goal is to design the best/optimal choice considering a set of well-defined design constraints and a set of quantifiable objectives. Thus, MODM methods deal with the design process, and the number of alternatives is infinite (continuous). On the other hand, discrete MCDM/MADM methods deal with discrete and predetermined options described by discrete determined criteria sets (Zavadskas, Turskis and Kildienė, 2014).

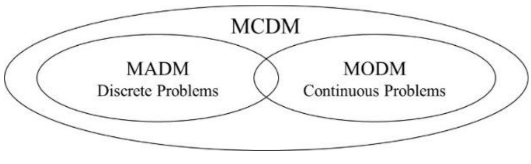

Figure 2: Broad classification of MCDM methods

Source: Zavadskas, Turskis and Kildienė (2014).

MCDM methods have been primarily used in radio access networks to address discrete problems of network selection in heterogeneous wireless networks (HWN). Decision-making problems have become more complex since the advent of the third-generation (3G) radio access technology WCDMA (wideband code division multiple access), specified in Release 99 by 3GPP in 1999, which allowed higher data rates and facilitated the significant increase in the number of mobile devices. Furthermore, mobile devices with advanced capabilities saw a massive proliferation with the evolution to 4G and the LTE RAT, specified in Release 8 by 3GPP in 2008. The imminent deployment of 5G will introduce yet another new RAT, which will coexist with the current RATs. Hence, the selection of the best network becomes essentially an MCDM process (Paul and Falowo, 2017). Obayiuwana and Falowo (2017) review and classify the most significant MCDM algorithms used to solve the network decision-making problems for HWNs.

On the other hand, Yeryomin and Seitz (2016) evaluated different algorithms used in the multiple criteria network selection problem, including simple additive weighting (SAW), weighted product model (WPM), elimination and choice expressing reality (ELECTRE), the technique for order of preference by similarity to ideal solution (TOPSIS), grey relational analysis (GRA), optimization and compromise solution (VIKOR), and the analytic hierarchy process (AHP). Some other critical applications of MCDM methods in those problems can be found in Pervaiz (2010), Sasirekha and Ilanzkumaran (2013), and Nguyen-Vuong et al. (2013). At the same time, Alhabo and Zhang (2018) predict that the introduction of 5G and the increasing demand for mobile data will lead to a selection scheme that considers different users with different priorities and preferences.

LTE evolution and increased data traffic have also introduced new decision-making problems that MCDM methods could address. While network selection can be considered a vertical handover, LTE is designed to support user mobility, even at high speed, moving from one cell to another during active service sessions (Nathaniel et al., 2014), which is considered a horizontal handover. Horizontal handover can also contribute to an effective load balancing for the optimum use of network resources. Recognizing this, Nathaniel et al. (2014) used a new MCDM approach to create a framework with a decision algorithm to solve the load-balancing problem in LTE. Furthermore, Dudnikova et al. (2015) introduced another innovative approach for MCDM while considering the problem generated by densely deployed heterogeneous networks with significant network energy consumption increments. To deal with this situation, they proposed using grey relational analysis and the analytic hierarchy process (MCDM tools) to find the number of base stations to switch off to maximize energy savings.

Therefore, different MCDM methods have been long applied in cellular networks and RATs, primarily for the network selection problem. However, the increasing complexity generated by new technologies, the colossal data traffic, the growing number of mobile subscriptions, and the cumulative number of cells installed in the network over the years make this scenario a fertile ground for new applications of MCDM methods to solve new decision problems.

The present study proposes a new application of the MCDM methods MAUT and AHP to solve the decision problem where faulty LTE cells need to be detected and ranked in a self-healing system in a network with non-predetermined alternatives and a vast number of options.

2.6 The Analytic Hierarchy Process

The AHP, first introduced by Saaty (2013; 1990), is a decision-making process based on the innate human ability to use information and experience to estimate relative magnitudes through paired comparisons. These comparisons are used to construct ratio scales of various dimensions, arranged in a hierarchic structure that allows for a systematic procedure to organize basic reasoning and intuition by breaking a problem down into smaller constituent parts. Thus, the AHP leads from superficial pairwise comparison judgments to the priorities in the hierarchy (Saaty, 2013).

2.7 The utility function

The utility function is a way of measuring the desirability of preferring different objects called alternatives. The utility score is the degree of well-being each of those alternatives provides to the decision-maker. The utility function comprises various criteria that assess an alternative's global utility. For each criterion, the decision-maker assigns a marginal utility score. One advantage to defining utility functions is that the options of the decision problem receive a global score. The marginal utility scores of the criteria are aggregated to yield the global utility score. This score makes it possible to compare all options and rank them from best to worst, with equal rankings permitted. A bad score on one criterion can be compensated by a good score on another (Ishizaka and Nemery, 2013). This approach is called the whole aggregation approach.

Ishizaka and Nemery state that if the utility function for each criterion (a representation of the perceived utility given the performance of the option on a specific criterion) is known, then the multi-attribute utility theory (MAUT) is recommended (Ishizaka and Nemery, 2013).

2.8 The Multi-Attribute Utility Theory

One of the most readily understandable approaches to decision analysis is multi-attribute utility analysis (MAUT) by Keeney and Raiffa (Keeney and Raiffa, 1993; Rupprecht et al., 2017). MAUT is based on the hypothesis that every decision-maker tries to optimize, consciously or implicitly, a function that aggregates all their points of view. It means that the decision maker's preferences can be represented by a function called the utility function *U*. Each alternative of set *A* is evaluated based on function *U* and receives a utility score *U*(*a*) (an example is shown in Figure 3). This utility score allows all alternatives to be ranked from best to worst (Ishizaka and Nemery, 2013). The preference and indifference relations among the other options of *A* are thus defined as follows:

$$\forall\, a, b\ \in A\colon a\ \mathrm{P}\ b\ \Leftrightarrow U(a) > U(b)\colon a\ is\ preferred\ to\ b \tag{1}$$

$$\forall\, a, b\ \in A\colon a\ \mathrm{P}\ b\ \Leftrightarrow U(a) = U(b)\colon a\ and\ b\ are\ indifferent \tag{2}$$

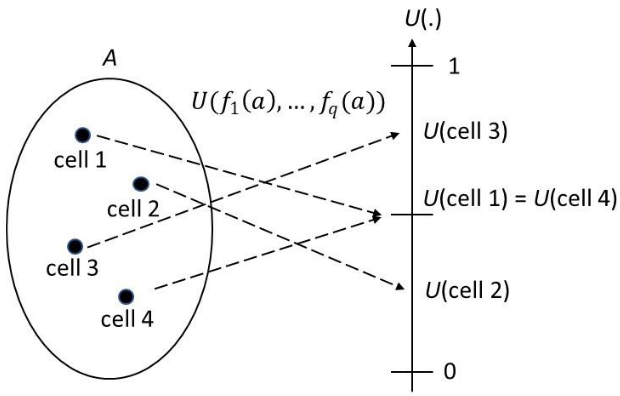

Figure 3: Representation of the set *A* ranking of the MAUT model

Source: Adapted from: Ishizaka and Nemery (2013, p. 82).

The utility function is defined using the additive model, the most popular and widely used model. In this model, the simple weighted sum is a particular case where U_j are all linear functions (Ishizaka and Nemery, 2013). The utility score corresponds to the following:

$$\forall\, a_i \in A\colon U(a_i) = U(f_1(a_i), \ldots, f_q(a_i)) = \sum\nolimits_{j=1}^{q} U_j(f_j(a_i)) \cdot \omega_j \tag{3}$$

where q is the number of criteria, ω_j is the weight of criterion f_j, and $U_j(f_j) \geq 0$. In general, they satisfy the normalization constraint (Ishizaka and Nemery, 2013):

$$\sum\nolimits_{j=1}^{q} \omega_j = 1 \tag{4}$$

The marginal utility function has the property that the best alternative on a specific criterion has a marginal utility score of 1, and the worst option has a score of 0 on the same criterion.

3 The proposed cell ranking method

We propose an algorithm to rank the LTE cells of a given RTA on the basis of their general performance, measured through the most relevant KPIs. The cells are ranked from the lowest-performing ones to a predefined threshold to facilitate cell fault management and self-healing processes by reducing the number of cells to be managed and healed. The main difficulty with ranking all the cells in a network is the vast number of alternatives (thousands of LTE cells). Thus, we propose using discrete MCDM methods to solve a problem where the other options are numerous and non-predetermined (faulty cells) with a set of quantifiable objectives (selected KPIs) that could be classified as a continuous MODM problem. Therefore, based on the broad classification of MCDM methods presented by Zavadskas (Zavadskas, Turskis and Kildienė, 2014), the present problem could be associated with a new category of problems, which Zavadskas had not considered, at the intersection of discrete and continuous problems (see Figure 4).

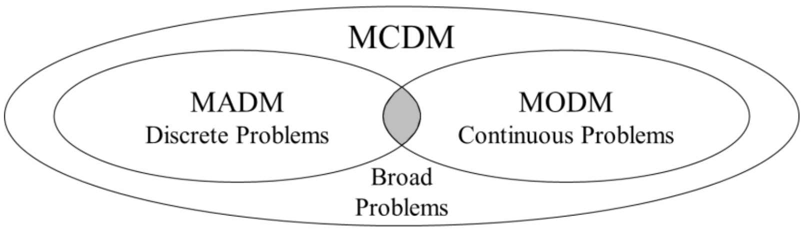

Figure 4: New proposed classification of MCDM methods

Source: Adapted from: Zavadskas, Turskis and Kildienė (2014).

The proposed method involves three steps, which are described below.

3.1 KPI selection

This first step is to select the RAT KPIs with the most significant impact on the end-user experience, considering both data and voice indicators. Selection is based on two factors: the motivation to improve the user experience (3GPP TS 28.404, 2020) and the difficulty of expressing it objectively and mathematically. To make this selection, it is essential to establish a relationship between user expectations and the QoS KPIs (Vaser and Forconi, 2015). Therefore, from the five categories of KPIs defined in (3GPP TS 32.450, 2019) and classified as QoS KPIs in this article, eight individual KPIs were selected from the three categories with the most significant impact on the end user: accessibility, retainability, and mobility. The Telemanagement Forum (TMF) Wireless Services Measurement Handbook GB923 (The Open Group, 2004) states that voice and data networks have been provisioned separately, and KPIs have been considered independently for each service. All the KPIs included in the proposed method are calculated as the ratio of two or more performance counters, so they are all normalized, ranging from 0 to 1. The KPIs selected and their formulas are described below.

Accessibility KPIs

E-RAB accessibility is a measurement showing the probability that an end user would be provided with an E-RAB (evolved UTRAN radio access bearer) on request (ITU-T Recommendation E.419, 2006). This type of KPI is perceived by the end user in data service as a connection delay and has a high impact on voice service, as it is perceived as service unavailability. Therefore, two KPIs of this type were selected:

1. Data E-RAB accessibility: a KPI that shows the probability success rate for E-RAB establishment:

$$ACC\_ERAB\_DATA = \frac{Number\ of\ successful\ data\ ERAB\ establishments}{Number\ of\ received\ data\ ERAB\ establishments\ attempts} \quad (5)$$

2. Data VoLTE E-RAB accessibility: a KPI that shows the probability success rate for VoLTE E-RAB establishment:

$$ACC\_ERAB\_VoLTE = \frac{Number\ of\ successful\ VoLTE\ ERAB\ establishments}{Number\ of\ received\ VoLTE\ ERAB\ establishments\ attempts} \quad (6)$$

Retainability KPIs

E-RAB retainability is a measurement that shows how often an end user loses an E-RAB in an abnormal way when the E-RAB is used (ITU-T Recommendation E.419, 2006). The end user in data service also perceives this type of KPI as a connection delay since the service needs to be re-established. It also seriously affects voice service, as it interrupts the voice call. 3GPP defines retainability as abnormal E-RAB releases per session time in seconds. However, this paper measures the ratio of normal E-RAB releases to the total number of E-RAB releases to be consistent with the other indicators, ranging from 0 (no success) to 1 (100% success). Therefore, two KPIs of this type are selected:

1. Data E-RAB retainability: a KPI that shows the rate of the number of normally released E-RABs with data in a buffer:

$$RET\_ERAB\_DATA = \frac{Number\ of\ normally\ released\ data\ ERABs}{Number\ of\ total\ released\ data\ ERABs} \quad (7)$$

2. Data VoLTE E-RAB retainability: a KPI that shows the rate of the number of normally released VoLTE E-RABs with data in a buffer:

$$RET\_ERAB\_VoLTE = \frac{Number\ of\ normally\ released\ VoLTE\ ERAB}{Number\ of\ total\ released\ VoLTE\ ERAB} \quad (8)$$

As for defining an abnormal E-RAB release with end-user impact, a release of the E-RAB is only considered abnormal if the eNodeB assumes that data is waiting for transfer in any of the buffers (ITU-T Recommendation E.419, 2006).

Mobility KPIs

LTE mobility is a measurement showing how LTE mobility functionality works (ITU-T Recommendation E.419, 2006). 3GPP includes handovers with both intra- and inter-LTE frequencies in the same KPI. However, in this paper, they are considered separate KPIs since they affect data and VoLTE services differently. Handover failures can cause delays in data transfers and call degradation, affecting the end user's perception. Four KPIs of this type are selected:

1. Intra-frequency handover: a KPI showing how E-UTRAN mobility functionality works within the same LTE frequency:

$$INTRA\_HO\_DATA = \frac{Number\ of\ successful\ intra\ frequency\ HO}{Number\ of\ intra\ frequency\ HO\ attempts} \quad (9)$$

2. Inter-frequency handover: a KPI that shows how E-UTRAN mobility functionality is working between different LTE frequencies:

$$INTER\_HO\_DATA = \frac{Number\ of\ successful\ inter\ frequency\ HO}{Number\ of\ inter\ frequency\ HO\ attempts} \quad (10)$$

Another process considered when evaluating problems related to mobility is the single radio voice call continuity (SRVCC), which is the continuity between voice calls in VoLTE and circuit-switched access (WCDMA or GSM RATs) (3GPP TS 23.216, 2020). The SRVCC procedure can be considered a particular case of handover, starting when the coverage or quality of the VoLTE call is poor. The session is transferred to a different RAT to keep the call active.

The SRVCC procedure consists of two steps: SRVCC preparation and SRVCC execution. SRVCC preparation does not directly impact the end user but can indicate a fault scenario. SRVCC execution strongly affects the end user, generating voice call interruptions. Therefore, both KPIs related to SRVCC are selected:

3. SRVCC preparation: a KPI that shows the success rate of the first step of SRVCC, preparing the SRVCC handover, starting when the user device receives the handover command (Qualcomm Technologies Inc., 2012):

$$SRVCC\_PREP = \frac{Number\ of\ successful\ SRVCC\ Preparation}{Number\ of\ SRVCC\ Preparation\ attempts} \tag{11}$$

4. SRVCC execution: a KPI that shows the success rate of the second step, which happens when the user device executes the handover after success in the previous step (Qualcomm Technologies Inc., 2012):

$$SRVCC\_EXE = \frac{Number\ of\ successful\ SRVCC\ Execution}{Number\ of\ SRVCC\ Execution\ attempts} \tag{12}$$

3.2 Weight definition

Weights must be defined for each of the selected KPIs to construct the utility function. AHP was chosen as it relies on simple hierarchic structures to represent decision problems (Saaty, 2013). The weights are found by calculating scores (or priorities, as they are called in AHP) based on the pairwise comparisons provided by the user (Ishizaka and Nemery, 2013).

To define the weights for the selected KPIs, AHP is implemented in three steps, following the procedure described by Dudnikova et al. (2015).

1. The problem is decomposed into its constituent parts, or criteria, which are the KPIs described in the previous subsection (summarized in Table 1).

Table 1: Selected KPIs

Category	Service	KPI
Accessibility	Data	ACC_ERAB_DATA
	Voice	ACC_ERAB_VoLTE
Retainability	Data	RET_ERAB_DATA
	Voice	RET_ERAB_VoLTE
Mobility	Data	INTRA_HO_DATA
	Data	INTER_HO_DATA
	Voice	SRVCC_PREP
	Voice	SRVCC_EXE

2. A relative importance value is assigned to each criterion by pairwise comparison. The fundamental scale, or the Saaty scale, defined in Saaty (2013), is used to rank the judgments introduced in Table 2. In LTE networks, voice is a data service (Voice over LTE − VoLTE). It follows the strictest quality criteria, as voice is susceptible to delay, jitter, and loss (The Open Group, 2004). Hence, the method proposed in this paper assigns higher importance to voice KPIs than to the other KPIs.

Table 2: Fundamental or Saaty Scale

Importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
3	Moderate importance	Experience and judgment slightly favor one activity over another
5	Strong importance	Experience and judgment strongly favor one activity over another
7	Very strong importance	An activity is favored very strongly over another; its dominance is demonstrated in practice
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation
2, 4, 6, 8	Intermediate values between adjacent judgments	

Source: Saaty (2013).

The quantified judgments about pairs of criteria are represented by the following $j \times j$ matrix A:

$$A = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1j} \\ 1/a_{12} & 1 & \cdots & a_{2j} \\ \vdots & \vdots & \vdots & \vdots \\ 1/a_{j1} & 1/a_{2j} & \cdots & 1 \end{bmatrix} \tag{13}$$

3. The eigenvector w of matrix A is calculated using the geometric mean method (Pervaiz and Bigham, 2009), and the relative weights of the factors (ω_j) are derived from the components of the normalized eigenvector (Dudnikova et al., 2015):

$$w_j = (\Pi a_{jj})^{1/q} \tag{14}$$

$$\omega_j = \frac{w_j}{\sum_{j=1}^{q} w_j} \tag{15}$$

The matrix A, the eigenvector, and the relative weights calculated from the presented formulas are shown in Table 3.

Table 3: Pairwise KPI Comparison, calculated eigenvectors and relative weights for each KPI

	RET_ERAB _VoLTE	ACC_ERAB _VoLTE	RET_ERAB _DATA	ACC_ERAB _DATA	INTER_HO _DATA	INTRA_HO _DATA	SRVCC _PREP	SRVCC _EXE	*w*	ω
RET_ERAB_ VoLTE	1	3	6	5	9	9	9	9	5.2650	0.4276
ACC_ERAB_ VoLTE	1/3	1	3	7	5	4	6	5	2.8373	0.2304
RET_ERAB_ DATA	1/6	1/5	1	3	6	5	2	2	1.3643	0.1108
ACC_ERAB_ DATA	1/5	1/7	1/3	1	3	3	2	2	0.8748	0.0710
INTER_HO_ DATA	1/9	1/5	1/6	1/3	1	5	2	2	0.6296	0.0511
INTRA_HO_ DATA	1/9	1/4	1/5	1/3	1/5	1	2	2	0.4429	0.0360
SRVCC_ PREP	1/9	1/6	½	1/2	1/2	1/2	1	9	0.5652	0.0459
SRVCC_ EXE	1/9	1/5	½	½	½	1/2	1/9	1	0.3339	0.0271

As expected, voice KPIs have the highest relative weights, and the retainability of VoLTE calls (RET_ERAB_VoLTE) is the most significant weight.

Since comparisons performed in AHP are subjective, judgment errors are inevitable and must be detected by verifying the consistency rate (*CR*) of *A* before selecting the weight values. The *CR* is calculated as follows:

$$CR = \frac{CI}{RI};\ CI = \frac{\lambda_{max} - q}{q - 1} \tag{16}$$

where *CI* is a consistency index, representing the deviation of the maximum eigenvalue of matrix A (λ_{max}) from the number of criteria used in the comparison process (q). *RI* is a random index, the average *CI* of a randomly generated reciprocal matrix. All *RI* values for different matrix dimensions are provided in Saaty (2013). If *CR* = 0, the matrix is perfectly consistent. If *CR* ≤ 0.1, the evaluated weight values are acceptable (Dudnikova et al., 2015). The λ_{max} is calculated as follows:

$$\lambda_{max} = \left[\sum_{j=1}^{q} a_{j1} \quad \cdots \quad \sum_{j=1}^{q} a_{jq}\right] \cdot \begin{bmatrix} \omega_1 \\ \cdots \\ \omega_j \end{bmatrix} \tag{17}$$

In the present problem, λ_{max}= 8.9276, *CI* = 0.1325, *RI* = 1.41, and *CR* = = 0.0940. Therefore, *CR* ≤ 0.1, and the obtained relative weights are consistent.

3.3 Utility function construction

As explained in Section 2, when the utility function for each criterion is known, the multi-attribute utility theory (MAUT) is recommended (Ishizaka and Nemery, 2013). It is the case for the present problem, where the criteria are the selected KPIs, each having a defined function. As all selected KPIs are ratios from the interval [0,1]. Hence, the MCDM method MAUT can construct the utility function.

The proposed method uses the simple weighted sum to construct the utility function for each LTE cell, considering the eight KPIs selected as criteria a_i, the relative weights obtained using the AHP method, and the number of fails for each KPI. The number of fails is necessary to avoid assigning a high score to a cell with degraded KPIs, but a low number of fails due to low traffic.

The function is then normalized by dividing the weighted sum by the sum of weights multiplied by the fails for each KPI, as shown below:

$$\forall a_i \in A: U(a_i) = \frac{\sum_{j=1}^{q} U_j(f_j(a_i)) \cdot \omega_j \cdot Fails_j}{\sum_{j=1}^{q} \omega_j \cdot Fails_j} \tag{18}$$

The number of fails of each KPI is calculated as a difference between the number of attempts and the number of successes for each indicator:

$$Fails_j = Number\ of\ attempts_j - Number\ of\ successes_j \tag{19}$$

4 Results

The proposed method has been applied in a real LTE network from a Brazilian telecommunications operator. 4925 cells from different equipment vendors covering a region of west-central Brazil were selected to verify the results. First, statistical data sampling of the eight selected KPIs was performed, aggregating the counters in a 24-hour base, and the utility function U (Formula 18) was applied to the cells. Then, the cells' utility functions were ranked in reverse order, from worst to best, to quickly identify the most degraded cells. The results are presented in a dashboard, with the cells labeled, starting with Cell 0 to Cell 4925.

The KPIs values in the dashboard are classified into three ranges for easier visual monitoring:

1) critical: from 0 to 0.50, indicating the most critical values;
2) alarming: from 0.50 to 0.99, indicating intermediate values;
3) OK: from 0.99 to 1, indicating the highest values.

The utility values follow the same classification, ranging from 0 to 1, with critical values between 0 and 0.50.

As an example of the results obtained, Table 5 reproduces the dashboard for a specific day, showing the first 12 results. The utility function allows the eight KPIs of the cells to be aggregated into a single indicator, facilitating ranking of the cells. Furthermore, only 0.20% of the cells have values below 0.50, significantly reducing the number of critical cells from the selected universe that need to be managed and healed, and highlighting the worst cells in terms of QoS, which are the main objectives of the proposed model.

On the analysis day, Cell 3083 was ranked as the worst cell, as SRVCC preparation performed very poorly, followed by SRVCC execution and handover KPIs. However, the cell had no active alarms or other operational problems and was not identified by traditional fault management. Crucially, although Cell 3083 was not the one with the most fails, its impact on the network was huge, as the fails were concentrated in VoLTE mobility, which the end user would have perceived as voice quality degradation. Table 4 shows all Cell 3083 measurements used to calculate its utility function, as detailed in Formula (20):

$$U(3083) = \frac{\sum_{j=1}^{q} U_j(f_j(3083)) \cdot \omega_j \cdot Fails_j}{\sum_{j=1}^{q} \omega_j \cdot Fails_j} \tag{20}$$

where:

$\sum_{j=1}^{q} U_j(f_j(3083)) \cdot \omega_j \cdot Fails_j = 0.9923 \cdot 0.4276 \cdot 1 + 1 \cdot 0.2304 \cdot 0 + 0.9996 \cdot$
$\cdot\, 0.1108 \cdot 10 + 0.9998 \cdot 0.0710 \cdot 3 + 0.9870 \cdot 0.0511 \cdot 76 + 0.9852 \cdot 0.0360 \cdot 119 +$
$+\, 0.0026 \cdot 0.0459 \cdot 2267 + 0.8333 \cdot 0.0271 \cdot 1$

and:

$\sum_{j=1}^{q} \omega_j \cdot Fails_j = 0.4276 \cdot 1 + 0.2304 \cdot 0 + 0.1108 \cdot 10 + 0.0710 \cdot 3 + 0.0511 \cdot$
$\cdot\, 76 + 0.0360 \cdot 119 + 0.0459 \cdot 2267 + 0.0271 \cdot 1$

Table 4: Measurements from Cell 3083 used to calculate its utility function

KPI	Success	Attempts	Fails	Value	Weights
VoLTE_RET	129	130	1	0.9923	0.4276
VoLTE_ACC	130	130	0	1.0000	0.2304
RET_ERAB	25776	25786	10	0.9996	0.1108
ACC_ERAB	19505	19508	3	0.9998	0.0710
HO_INTRA	7912	8031	76	0.9870	0.0511
HO_INTER	5762	5838	119	0.9852	0.0360
SRVCC_PREP	6	2273	2267	0.0026	0.0459
SRVCC_EXE	5	6	1	0.8333	0.0271

Another significant result from applying the proposed model is shown by the analysis of Cells 0574 and 0573. They are neighbor cells and present a value of zero in VoLTE KPIs. However, they did not have VoLTE fails, which indicates no traffic on that service, even if the KPIs impacting their utility function were from data handover. However, the lack of VoLTE traffic can also indicate a configuration failure and should be investigated by cell performance managers.

An overall analysis of the results for the worst cells of the network can give engineers valuable insight into its health, as it aggregates the most relevant LTE radio KPIs. For example, the KPIs values of the worst cells from Table 5 show that the network problems are concentrated in handover and SRVCC indicators, meaning that mobility is the key issue to address.

Table 5: Utility function results (U) and ranking applied in cells of an actual LTE network

Cell	U	Fails	RET_ERAB _VoLTE	ACC_ERAB _VoLTE	RET_ERAB _DATA	ACC_ERAB _DATA	INTER_HO _DATA	INTRA_HO _DATA	SRVCC _PREP	SRVCC _EXE
Cell 3083	0.0886	2477	0.9923	1.0000	0.9996	0.9998	0.9870	0.9852	0.0026	0.8333
Cell 2806	0.1190	3056	0.9967	0.9882	0.9995	0.9972	0.9876	0.9815	0.0029	1.0000
Cell 1411	0.2465	7761	0.9957	1.0000	0.9979	0.9963	0.1931	0.9863	1.0000	1.0000
Cell 1802	0.3054	14905	0.9883	0.9881	0.9991	0.9993	0.2823	0.9426	1.0000	1.0000
Cell 2217	0.3072	14822	0.9952	1.0000	0.9994	0.9993	0.9860	0.2809	0.0000	0.0000
Cell 1834	0.3161	1735	0.9985	0.9977	0.9996	0.9985	0.9793	0.9895	0.0061	1.0000
Cell 0574	39.7838	506	0.0000	0.0000	0.9993	0.9997	0.5068	0.3742	0.0000	0.0000
Cell 0862	40.3638	124	0.9969	0.9973	0.9986	0.9997	0.9971	0.9311	0.0829	0.9810
Cell 0573	41.8881	204	0.0000	0.0000	0.9987	0.9992	0.7213	0.3468	0.0000	0.0000
Cell 1417	42.4690	420	1.0000	0.9939	0.9994	0.9986	0.3808	0.9668	1.0000	1.0000
Cell 4501	44.2235	362	1.0000	0.9951	0.9994	0.9925	0.9916	0.3857	0.0000	0.0000
Cell 1349	45.9742	586	0.9931	0.9946	0.9992	0.9964	0.3743	0.9870	1.0000	1.0000

5 Conclusion

This paper proposes using an innovative application of MCDM methods for radio access network analysis and cell failure management by ranking the worst-performing LTE cells. This is the first attempt to use discrete MCDM methods to solve problems with quantifiable objectives where the alternatives are numerous and not predetermined.

As shown in the previous section, we could identify very poorly performing cells with no active operational alarms that could not be identified through traditional fault management. By ranking cells based on a utility function U that aggregates the main radio QoS KPIs, the proposed method automatically indicates the global worst cells to be repaired, improving network quality more efficiently. Furthermore, the utility function U can filter cells based on performance objectives to be analyzed and repaired by network engineers. This approach may reduce the time consumed in identifying faulty cells that affect the end-user performance, improving the perceived LTE network performance.

Repairing the most critical performance cells quickly and efficiently helps operators and cell optimization service providers with network performance management. It satisfies quality requirements set by the government or by other inspection agencies. The weights defined by the AHP method can also be adapted to the operators' needs − for example, by switching priorities from voice to data or mobility − making the method customizable.

The proposed method may also be used to rank a group of cells (e.g., clusters or cities), aggregating the selected KPIs and calculating the utility function for each defined group, helping to identify performance variations in that group. Furthermore, the method can be adapted to rank cells of other radio access technologies, such as 3G (WCDMA) and 5G NR (New Radio), selecting the most important KPIs for each technology and applying the weights and the utility function. Hence, the method described in this paper is a framework that can be adapted to different performance management systems.

The above advantages could be verified in a live LTE network. The time to detect a failing cell and the number of non-detected failing cells in the network were significantly reduced. Furthermore, the weights and KPIs prioritization can be changed according to the customer's priorities, being a flexible framework that fits network management. Some disadvantages of the method were also perceived during the tests, as sleeping cells, cells hanging resources, and low--traffic cells could not be well-detected. Cells off also become unreachable and undetected by the method, which doesn't replace traditional faulty cell detection systems.

The proposed approach can contribute significantly to cell performance management in radio access networks. We have presented a new method of KPI aggregation to rank the worst-performing LTE cells based on MCDM methods. This paper also contributes to the MCDM literature, introducing its methods to SON functionalities and applying them to a large set of non-predetermined options.

References

3GPP TS 23.216 (2020), *Single Radio Voice Call Continuity (SRVCC); Stage 2*, Release 16.

3GPP TS 28.404 (2020), *Quality of Experience (QoE) Measurement Collection; Concepts, Use Cases and Requirements*, Release 16.

3GPP TS 32.401 (2018), *Performance Management (PM); Concept and Requirements*, Release 15.

3GPP TS 32.421 (2015), *Subscriber and Equipment Trace; Trace Concepts and Requirements*, Release 11.

3GPP TS 32.450 (2019), *Key Performance Indicators (KPI) for Evolved Universal Terrestrial Radio Access Network (E-UTRAN): Definitions*, Release 15.

3GPP TS 32.500 (2020), *Self-organizing Networks (SON); Concepts and Requirements*, Release 16.

3GPP TS 32.541 (2020), *Self-organizing Networks (SON); Self-healing Concepts and Requirements*, Release 16.

Alhabo M., Zhang L. (2018), *Multi-criteria Handover Using Modified Weighted TOPSIS Methods for Heterogeneous Networks*, IEEE Access, 6, 40547-40558.

Barco R., Lazaro P., Munoz P. (2012), *A Unified Framework for Self-healing in Wireless Networks*, IEEE Commun. Mag., 50(12), 134-142 (December).

Bouyssou D., Marchant T., Pirlot M., Tsoukiàs A., Vincke P. (2006), *Evaluation and Decision Models with Multiple Criteria: Stepping Stones for the Analyst*, Springer, New York, NY, USA.

Dudnikova A., Dini P., Giupponi L., Panno D. (2015), *Multi-criteria Decision for Small Cell Switch off in Ultra-Dense LTE Networks*, 13th International Conference on Telecommunications (ConTEL), 1-8, Graz, Austria.

Hu H., Zhang J., Zheng X., Yang Y., Wu P. (2010), *Self-configuration and Self-optimization for LTE Networks*, IEEE Commun. Mag., 48(2), 94-100 (February).

Ishizaka A., Nemery P. (2013), *Multi-Criteria Decision Analysis: Methods and Software*, John Wiley & Sons Ltd., New York, NY, USA.

ITU-T Recommendation E.419 (2006), *Business Oriented Key Performance Indicators for Management of Networks and Services*, February.

ITU-T Recommendation E.800 (2008), *Definitions of Terms Related to Quality of Service*, September.

Jejdling F. (2020), *Ericsson Mobility Report*, www.ericsson.com/en/mobility-report, Telefonaktiebolaget LM Ericsson, Stockholm, Sweden (November).

Keeney R.L., Raiffa H. (1993), *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, Cambridge University Press, Cambridge, UK.

Nathaniel S., Ariffin S.H.S., Farzamnia A., Adegboyega A.J. (2014), *Multi-criteria Load Balancing Decision Algorithm for LTE Network*, 4th International Conference on Engineering Technology and Technopreneuship (ICE2T), 57-62, Kuala Lumpur, Malaysia.

Next generation (2008), *Mobile Networks Recommendation on SON and O&M Requirements*, Release (December 5).

Nguyen-Vuong Q., Agoulmine N., Cherkaoui E.H., Toni L. (2013), *Multicriteria Optimization of Access Selection to Improve the Quality of Experience in Heterogeneous Wireless Access Networks*, IEEE Transactions on Vehicular Technology, 62(4), 1785-1800 (May).

Nory R. (2019), *New WID on N.R. Dynamic Spectrum Sharing (DSS)*, 3GPP TSG RAN Meeting #86, Tdoc RP-193260, Ericsson (December).

Obayiuwana E., Falowo O.E. (2017), *Network Selection in Heterogeneous Wireless Networks Using Multicriteria Decision-making Algorithms: A Review*, Wireless Networks, 23(8), 2617-2649.

Paul U., Falowo O.E. (2017), *Efficient RAT-selection for Group Calls Using Intuitionistic Fuzzy TOPSIS in Heterogeneous Wireless Networks*, 2017 IEEE AFRICON, 365-370, Cape Town, South Africa.

Pervaiz H. (2010), *A Multi-criteria Decision Making (MCDM) Network Selection Model Providing Enhanced QoS Differentiation to Customers*, MCIT'2010: International Conference on Multimedia Computing and Information Technology, 5444854, 49-52.

Pervaiz H., Bigham J. (2009), *Game Theoretical Formulation of Network Selection in Competing Wireless Networks: An Analytic Hierarchy Process Model*, Third International Conference on Next Generation Mobile Applications, Services and Technologies, 292-297, Cardiff, UK.

Qualcomm Technologies Inc. (2012), *VoLTE with SRVCC: The Second Phase of Voice Evolution for Mobile LTE Devices*, White Paper, www.qualcomm.com/media/documents/files/srvcc-white-paper.pdf, San Diego, CA, USA (October).

Rupprecht F.A., Soni V., Schmidt C., Ravani B., Ebert A., van der Veer G. (2017), *An Approach for Evaluating Collaborative Software Environments Based on Integration of House of Quality with Multi-attribute Utility Theory*, 9th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 45-54, Munich, Germany.

Saaty T.L. (1990), *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*, 2nd ed., RWS Publications, Pittsburgh, PA, USA.

Saaty T.L. (2013), *Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process*, RWS Publications, Pittsburgh, PA, USA.

Sallent O., Pérez-Romero J., Sánchez-González J., Agustí R., Díaz-guerra M.A., Henche D., Paul D. (2011), *A Roadmap from UMTS Optimization to LTE Self-optimization*, IEEE Commun. Mag., 49(6), 172-182 (June).

Sasirekha V., Ilanzkumaran M. (2013), *Heterogeneous Wireless Network Selection Using FAHP Integrated with TOPSIS and VIKOR*, 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, 399-407, Salem, India.

Szilagyi P., Novaczki S. (2012), *An Automatic Detection and Diagnosis Framework for Mobile Communication Systems*, IEEE Transactions on Network and Service Management, 9(2), 184-197 (June).

The Open Group (2004), *Enterprise Perspective*, SLA Management Handbook, 4, Berkshire, UK.

Triantaphyllou E. (2010), *Multi-Criteria Decision Making Methods: A Comparative Study*, Springer, New York, NY, USA.

Vaser M., Forconi S. (2015), *QoS KPI and QoE KQI Relationship for LTE Video Streaming and VoLTE Services*, 9th International Conference on Next Generation Mobile Applications, Services and Technologies, 318-323, Cambridge, UK.

Yeryomin Y., Seitz J. (2016), *Enhanced Multi-criteria-based Path Selection Algorithm for Heterogeneous Networks*, Eighth International Conference on Ubiquitous and Future Networks (ICUFN), 804-809, Vienna, Austria.

Zavadskas E.K., Turskis Z., Kildienė S. (2014), *State of Art Surveys of Overviews on MCDM/MADM Methods*, Technological and Economic Development of Economy, 20(1), 165-179.

MULTIPLE CRITERIA DECISION MAKING
Vol. 17 2022

Chris Tofallis*

OBJECTIVE WEIGHTS FOR SCORING: THE AUTOMATIC DEMOCRATIC METHOD

DOI: 10.22367/mcdm.2022.17.04

Received: 27.07.2022 | Revised: 13.12.2022 | Accepted: 1.03.2023.

Abstract

When comparing performance (of products, services, entities, etc.), multiple attributes are involved. This paper deals with a way of weighting these attributes when one is seeking an overall score. It presents an objective approach to generating the weights in a scoring formula which avoids personal judgement. The first step is to find the maximum possible score for each assessed entity. These upper bound scores are found using Data Envelopment Analysis. In the second step the weights in the scoring formula are found by regressing the unique DEA scores on the attribute data. Reasons for using least squares and avoiding other distance measures are given. The method is tested on data where the true scores and weights are known. The method enables the construction of an objective scoring formula which has been generated from the data arising from all assessed entities and is, in that sense, democratic.

Keywords: multi-attribute decision making, weighting, ranking, performance measurement, composite indices, data envelopment analysis.

1 Introduction: Why have a formula for performance or efficiency?

We are often interested in comparing the performance or efficiency of entities, be they consumer products, services, people, organisations, etc. We naturally prefer results from techniques which are readily comprehensible because this provides greater confidence. Furthermore, this allows us to communicate the results more easily and to persuade others. Having a simple formula is therefore

* University of Hertfordshire: Hatfield, GB, e-mail: ormsresearch@gmail.com, ORCID: 0000-0001-6150-0218.

helpful. It also provides transparency: it clearly shows how the various factors are weighted and combined. The challenge that we shall investigate is deciding what these weights should be.

The factors or attributes associated with the entities are often combined to produce a multi-dimensional or composite index and the weights chosen for these criteria are clearly crucial to the results that follow. A startling illustration of this is given by Decanq and Lugo (2013, p. 16) describing the work of Becker et al. (1987):

> "The authors studied the quality of life in 329 metropolitan areas of the U.S. by ordering them according to standard variables such as quality of climate, health, security, and economical performance. The authors find that, depending on the weighting scheme chosen, there were 134 cities that could be ranked first, and 150 cities that could be ranked last. Moreover, there were 59 cities that could be rated either first or last, using the same data, but by selecting alternative weighting schemes".

Data envelopment analysis (DEA) provides a way of deriving *separate* attribute weights for each entity so as to maximize its score, and thereby overcomes the issue of subjectivity. An entity assessed with a DEA score would be pleased with its allocated weights since no other weights could provide a higher score. However, some would argue that this removes a degree of comparability because each entity has its own weights and so comparison is not on the same basis. Hence, there is a perceived need, at least in some quarters, for a common set of weights that can be applied to all the entities being compared (Liu and Peng, 2008; Ramon, Ruiz and Sirvent, 2012; Kritikos, 2017). Liu and Lu (2010, p. 453) give an example of such a situation: "DEA usually provides a group of performance leaders that can be used as benchmarks for those who are outside the leading group. The leaders are of equal significance under the original DEA methodology. Some applications, however, expect unambiguous, preferably ranked, performance leaders. For example, when applying DEA to compare the performance of R&D (research and development) organizations, one prefers a ranked list in order to correctly reward the R&D organizations and more importantly to allocate precious resources to organizations with better performance". Ruiz and Sirvent (2016, p. 8) point out that "The DMUs [decision making units] involved in production processes often experience similar circumstances, so benchmarking analyses in those situations should identify common referents and establish common best practices (...). In particular, this means that input and output weights should be common to all units in the evaluations, in contrast to DEA".

By having common weights one can produce rankings based on the score from the formula. The media are often keen to present results from rankings, and so this is one way of gaining public attention, the attention of policy makers, and decision makers. Rankings are sometimes based on aggregating expert opinions, but these may be heavily influenced by historical reputations which may be less relevant today. In this regard Rosenthal and Weiss (2017, p. 136) look at academic journals in the field of business and notes that “While we believe that subjective comparisons and rankings are very important (…) we also believe that subjective rankings represent opinions that are often slow to change”.

In magazines and newspapers, journalists will discuss shifting rank positions and comment on possible causes. Examples include the annual Human Development Index ranking of countries produced by the United Nations Development Programme, and the various national and international ‘league tables’ of universities which exist around the world. Despite being roundly criticised by academics, the latter attract a great deal of interest because of the power that they wield: potential students are influenced by them when selecting where to study, and academics seeking a position may also be guided by them. For their part, university executives scour these tables to see if they can trumpet recent successes when their position rises, or try to see which factors they need to improve if it falls. Another example is the Multidimensional Poverty Index, which is produced by the Oxford Poverty and Human Development Initiative (OPHI). They point out that a dashboard showing multiple measures does not reflect the multiple deprivations that some people face at the same time, whereas a single index can combine such information.

Other reasons for using a common set of weights relate to issues regarding the direct use of DEA. Perhaps the most serious issue from a managerial perspective is attaching zero (or epsilon) weights to criteria, thereby effectively hiding that factor from the assessment of that entity. Hence it is not unusual to have many DMUs with a score of 100%. It can be proved that if a unit has (uniquely) the largest value of a particular output then it is automatically efficient in some DEA models (e.g. the BCC or Additive DEA models) – and this is irrespective of how large (and possibly wasteful) its consumption of inputs! (Ali, 1993; Jahanshahloo et al., 2005). Similarly, if a unit has (uniquely) the smallest value of an input then it is automatically 100% efficient in some DEA models – irrespective of how low its output levels are! So, if there are m inputs and s outputs, up to $m + s$ units could be efficient in this way, which is an unfortunate artefact of such models.

Roll and Golany (1993, p. 99) point out that “it is usually deemed inappropriate to accord widely differing weights to the same factor, when assessing different DMUs”. Indeed, it is this weight flexibility that leads to poor discrimination between entities following a DEA analysis, with many achieving a 100% score. This is especially true when the number of DMUs being compared is small, and/or the number of factors (inputs, outputs, performance measures) is large.

A graphical representation may assist in understanding how DEA may accord widely differing weights. In Figure 1 all units are assumed to have the same input value; the four upper DMUs all produce more than 20 units of the output on the vertical axis, whereas the fifth DMU produces one unit only. Nevertheless, it is rated 100% under DEA due to the fact that it has the highest level of the output on the horizontal axis. Such units are sometimes referred to as mavericks. Moreover, any DMUs just behind this one will have scores close to 100%. The relative weights accorded to the two outputs will clearly differ greatly as indicated by the slope of the line-segment on the right compared to the slopes of the other sections of the frontier.

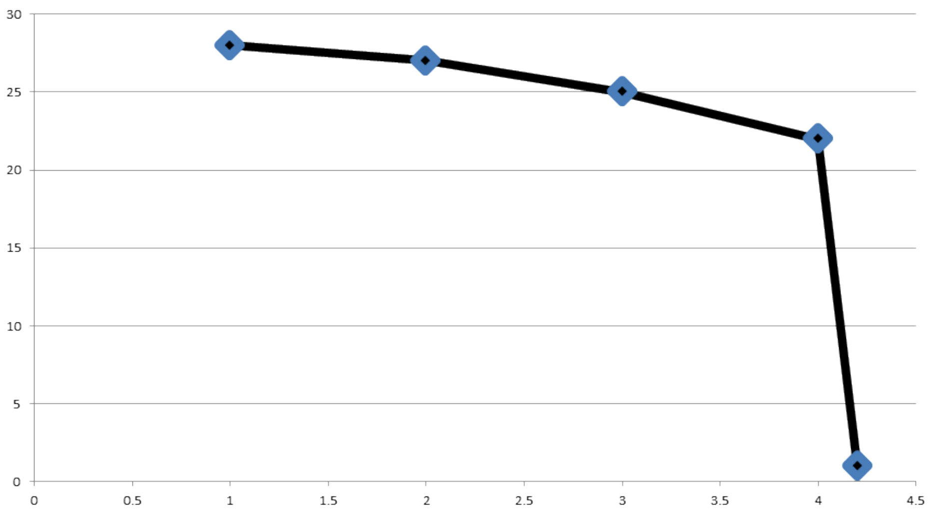

Figure 1: Frontier units in output space

2 Common weights and DEA

Perhaps the first paper to propose a common set of weights (CSW) in the DEA context is Cook, Roll and Kazakov (1990), which mentions it as a direction for further analysis. This was followed by Roll, Cook and Golany (1991, p. 6), who motivate their work by pointing out that using a single set of weights "(…) is the usual approach in all engineering, and most economic, efficiency analyses. In these cases it is assumed that all important factors affecting performance are included in the measurement system, and there is no need (nor wish) to allow for additional, individual, circumstances”. However, they point out that in reality some factors will not have been included in the analysis, and that there may also be differences in goals and missions. Roll, Cook and Golany (1991, p. 6) then propose an interesting interpretation:

> “A possible meaning of efficiencies computed with a CSW, in the context of DEA, is that such values represent the part of a DMU’s performance which can be explained when assuming uniformity of circumstances. The difference between the efficiency measured with an »individual« set of weights and that obtained with a CSW may indicate the effects of special circumstances under which a DMU operates”.

Among the approaches that they suggest is to run a conventional DEA and then take the average weight attached to each factor across all the DMUs, or to take some average of the highest and lowest weight. This is a move away from DEA as ‘self-appraisal’, and towards ‘peer appraisal’. This is the philosophy behind ‘cross-efficiency’; here the evaluation of an entity is some ‘average’ of the evaluations arising from applying the optimal weights of all the entities to the one being assessed. The problem with this is that DEA weights are not unique – one can obtain the same DEA score using different weights. To quote Thompson, Dharmapala and Thrall (1993, p. 383) “Zero values in optimal solutions for the primal or dual LP programs are indicators of degeneracy and multiple optimal solutions. Some analysts misleadingly use »the« instead of »a« or »an« for a DEA solution; and accordingly, they may fail to recognize the existence of multiple optimal solutions”. The non-uniqueness of DEA weights is a long-running issue with the cross-efficiency method (Wu, Sun and Liang, 2020), and has led to a plethora of sophisticated devices for dealing with them. Indeed, Table 1 in Contreras, Lozano and Hinojosa (2021) lists 44 cross-efficiency approaches. For the sake of simplicity we choose to avoid this route.

Roll and Golany (1993) presented further ideas for determining common weights. One suggestion was to choose the CSW which maximizes the average score of the DMUs. Another was to determine the CSW which maximized the number of efficient DMUs. We shall not attempt here to review the extensive literature on CSW, as this has been carried out by Aldamak and Zolfaghari (2017), which contains 113 references − there is also a review of ranking models in DEA by Lotfi et al. (2013) with 104 references, see also Adler, Friedman and Sinuany-Stern (2002), and Moghaddas and Vaez-Ghasemi, 2017. Nevertheless, we now mention those papers which bear some similarity to the work we shall present.

Kao and Hung (2005) treated the DEA scores as the ideal solution to be used in a compromise programming formulation to identify the common weights: this involves finding a solution which is as close as possible to the ideal point. The three most common distance measures were used: city-block (or Manhattan), Euclidean, and Chebyshev (L_1, L_2, and L_∞ metrics). They recommended investigating all three of these and then to make a subjective judgement. Surprisingly, they reported that replacing all the DEA scores by 100% led to the same CSW for the L_1 metric. We now explain why this happens: The objective in the L_1 metric is to find a common set of weights (u,v) so as to minimize the sum of absolute differences between the DEA scores (E_i*) and those from the CSW formula:

$$\text{Min } \Sigma\, |E_i^* - E_i\,(u,v)| \tag{1}$$

The DEA scores are optimal and therefore should not be exceeded by using any other set of weights, so $E_i^* \geq E_i\,(u,v)$, which implies that the objective simplifies to:

$$\text{Min } \Sigma[E_i^* - E_i\,(u,v)] \text{ i.e. Min [constant} - \Sigma\, E_i\,(u,v)] \tag{2}$$

Thus the actual DEA scores are removed from the problem and do not affect the resulting common set of weights. This is rather disheartening and leads us to reject this L_1 approach. (The objective function also simplifies to maximising the sum of scores.)

Zohrehbandian, Makui and Alinezhad (2010) adapted Kao and Hung's approach to avoid nonlinearities by restricting attention to the L_1 and L_∞ metrics, and by considering an additive DEA model. Pre-dating both of these papers is Despotis (2002), who combines the L_1 and L_∞ distance measures using a weighting parameter which is left to the user to set. This was later applied to the Human Development Index (Despotis, 2005).

Cook and Zhu (2007) consider the view that one should minimize the deviation for the unit whose final score is furthest from its original DEA score. This is an L_∞ or Chebyshev metric and so we have a minimax approach. This view could be supported on the grounds that this unit is the most disadvantaged. However, one needs to ask why there is a large deviation in score. It may be that this unit is a maverick as in Figure 1 and is using weights which differ considerably from the rest of the data set, or even had zero weights on a number of variables and so was able to hide the extent of its inefficiency; thus the more realistic (common) weights caused its score to decline the most. It therefore becomes highly questionable whether such a poor performing unit should be so influential in determining the weights for all other units. It also appears to be 'undemocratic' that a single unit should hold such power over the others. It is for these reasons that we do not adopt the L_∞ metric.

3 The automatic democratic method

In this paper we propose applying least squares regression to the DEA scores (E_i) as the dependent variable. For reasons stated above we do not use the L_1 and L_∞ fitting approaches. The underlying attribute data are used as the explanatory variables in a functional form that reflects the efficiency measure used in DEA when finding the optimal score for each assessed entity. We shall allow for both input (more is worse) and output (more is better) variables. One way of measuring performance efficiency is the ratio: sum of weighted outputs (y) to sum of weighted inputs (x), which is the original CCR model for DEA (Charnes, Cooper and Rhodes, 1978). We emphasise that any other functional form (additive, multiplicative etc.) can be used as the performance metric – our approach is not restricted in this way. For our ratio form the following least squares objective function would apply:

$$\text{Minimize } \sum_{i=1}^{n}\left[\frac{\sum_{r=1}^{s} u_r y_{ri}}{\sum_{j=1}^{m} v_j x_{ji}} - E_i\right]^2 \tag{3}$$

where u_r is the coefficient for output r and v_j is the coefficient for input j, these are the common weights to be determined by the regression, and the E_i are the numerical scores previously obtained from DEA for each unit i. The idea is to produce a handy compact formula for the scores. This can be described as an Automatic Democratic Approach since each entity is *initially* represented by its own optimal score, with individual weights emphasising its best aspects, with the *final* set of weights and scores being deduced objectively from these initial scores by regression.

Given that DEA scores are designed to show each entity in the best possible light, and are therefore optimistic, we feel that such scores ought to be treated as upper bounds, and not exceeded. We thus take measures to deal with this issue. Two ways will be considered:

1. Use constrained regression with one-sided residuals. The constraints which force all residuals to have the same sign guarantee that DEA scores (as the dependent variable) will not be exceeded.
2. Use ordinary regression followed by an adjustment which ensures the formula score does not exceed the DEA score. This can be done by subtracting a constant (equal to the largest over-estimate) from all scores; however, this runs the risk of making very low scores become negative. A better approach is to rescale i.e. divide the formula scores by a factor which ensures our requirement. The factor will be the largest instance of the ratio (DEA score)/(Formula score). This approach also has the benefit of retaining proportionality between the formula scores.

It should be noted that the regression we are discussing here differs from the commonly used second stage regressions where DEA scores are related to other variables (e.g. environmental, contextual, or otherwise explanatory) not used in estimating the DEA scores. Such models sometimes employ a censored normal Tobit specification to deal with the fact that there is a concentration of observations with a 100% score. Such points are an artefact of the DEA method of identifying a frontier: attaching a score of 100% to some observations should not make us think there is anything absolute or fixed about them, since DEA is always an indication of ‘relative’ performance.

We stress that we are using regression for descriptive purposes – to provide a formula which approximates a set of scores. There is no assumption regarding the underlying probability distribution of the scores, and no statistical inference is being made.

3.1 Illustrative application

It has been claimed that “Data Envelopment Analysis (DEA) is the most commonly used approach for evaluating healthcare efficiency” (Hollingsworth, 2008; Gajewski et al., 2009; Matawie and Assaf, 2010), and we shall use an example from this field to illustrate our method. The data (Cooper, Seiford and Tone, 2006, p. 169) cover 14 general hospitals and have two inputs: nurses and doctors. The two outputs are inpatient and outpatient numbers (see Table 1). Applying DEA using the CCR model we obtain the scores shown in Table 2. Also shown are those weights which are zero or extremely small. What is disturbing is that half of the hospitals have a zero weight on the number of outpatients! Also, four hospitals place zero weight on doctors. These variables are effectively being ignored in assessing those hospitals. Commenting on this, Cooper, Seiford and Tone (2006, p. 169) state: “This means that the inefficient hospitals are very likely to have a surplus in doctors and a shortage of outpatients”. There are even three hospitals which place zero weights on both doctors and outpatients, and so are being assessed purely on the ratio of inpatients per nurse, with the other variables ignored. This highlights how DEA, when used on its own, can be problematic when evaluating performance.

Table 1: Data on 14 general hospitals

Hospital	Doctors	Nurses	Outpatients	Inpatients
A	3008	20 980	97 775	101 225
B	3985	25 643	135 871	130 580
C	4324	26 978	133 655	168 473
D	3534	25 361	46 243	100 407
E	8836	40 796	176 661	215 616
F	5376	37 562	182 576	217 615
G	4982	33 088	98 880	167 278
H	4775	39 122	136 701	193 393
I	8046	42 958	225 138	256 575
J	8554	48 955	257 370	312 877
K	6147	45 514	165 274	227 099
L	8366	55 140	203 989	321 623
M	13 479	68 037	174 270	341 743
N	21 808	78 302	322 990	487 539

Table 2: DEA scores based on the CCR method applied to 14 hospitals

Hospital	DEA score	Doctor weight	Nurse weight	Outpatient weight	Inpatient weight
A	0.955		0		
B	1				
C	1				
D	0.702			0	
E	0.827	0		0	
F	1				
G	0.844			0	
H	1		4.6 E-08	1.1 E-08	
I	0.995	0			
J	1				
K	0.913		7.4 E-08	0	
L	0.969			0	
M	0.786	0		0	
N	0.974	0		0	

The table also shows those weights on inputs and outputs that turn out to be zero or extremely small.

We now apply the proposed automatic democratic method. We regress the DEA scores as the dependent variable on the underlying input and output data, as explanatory variables. The ordinary least squares (OLS) regression model is:

$$\text{DEA score} = [u_1 \text{ Outpatients} + u_2 \text{ Inpatients}]/[v_1 \text{ Doctors} + v_2 \text{ Nurses}] + \text{residual} \quad (4)$$

This is a nonlinear regression in the coefficients (or weights u, v), but is easily solved using standard statistical software such as SPSS.

The resulting formula is:

$$E = [\text{Outpatients} + 4.94 \text{ Inpatients}] / [52.18 \text{ Doctors} + 24.56 \text{ Nurses}] \quad (5)$$

where we have set the coefficient for Outpatients to be 1 for ease of interpretation; there is no loss of generality as the scores remain unchanged when all coefficients are divided by the same number.

First we are pleased to observe that there are no zero weights. Inspecting the output weights shows that an inpatient has a greater weight than an outpatient, which is as expected. Likewise, for the inputs, a doctor has a greater weight than a nurse; again this is what we would expect.

Table 3: Comparing ranks using DEA, OLS (Ordinary Least Squares), and a constrained regression formula

Hospital	A	***B***	***C***	D	E	***F***	G	***H***	I	***J***	K	L	M	N
OLS rank	10	6	2	14	12	3	11	7	4	*1*	8	5	13	9
DEA rank	9	*1*	*1*	14	12	*1*	11	*1*	6	*1*	10	8	13	7
Constrained regression	11	9	2	14	12	3	10	6	5	*1*	7	4	13	8

DEA-efficient hospitals are underlined.

In Table 3 we compare the ranks based on DEA with those arising from the formula. Most of the hospitals have the same or similar rank, which is reassuring. However, we need to look at those cases where there is a marked difference and see if this makes sense. The largest change occurs with hospital H which was 100% efficient under DEA but is now ranked 7^{th} under OLS. To explain this decline let us compare it with C which has rank 2. Note that the CCR model assumes constant returns to scale. H treats 15% more inpatients and just 2% more outpatients, but H is outperformed by C because H uses 45% more nurses to achieve this. Inspection of the weights in Table 2 shows that H has placed a negligible weight on nurses, thereby downplaying this high resource usage.

Next consider hospital B which was DEA efficient, but now has rank 6. Let us try and understand this by comparing it with C (ranked second). C has 5% more nurses and 8.5% more doctors than B, and so we would expect its processing of inpatients to be greater by this order, but it is actually processing 29% more inpatients than B. This disproportionate difference helps explain why C now outperforms B.

Next we turn to the (unconstrained) OLS scores provided by the formula. Given that it is an approximation based on regression, it is inevitable that some scores will be above or below their DEA scores, as can be seen in Table 4. As DEA scores are viewed as an optimistic upper bound, there is no requirement with formula scores which are below these. However, something needs to be done if the formula provides a score higher than the corresponding DEA value. This is fixed by a simple rescaling. For each hospital we calculate the ratio (DEA score)/(Formula score), then find the maximum ratio. We then divide all the formula scores by this maximum ratio. This ensures that DEA scores are never exceeded whilst also maintaining proportionality: all scores remain the same in relative terms, so that ratios of scores are unchanged.

Table 4: Comparison of DEA and formula scores

Hospital	A	B	C	D	E	F	G	H	I	J	K	L	M	N
DEA score	0.96	1.00	1.00	0.70	0.83	1.00	0.84	1.00	1.00	1.00	0.91	0.97	0.79	0.97
OLS Formula score	0.89	0.93	1.09	0.67	0.85	1.05	0.86	0.90	1.01	1.09	0.89	1.00	0.78	0.89
OLS Formula score (rescaled)	0.81	0.85	0.99	0.61	0.78	0.96	0.79	0.83	0.93	1.00	0.82	0.92	0.72	0.82
Constrained regression score	0.80	0.83	1.00	0.65	0.77	0.96	0.82	0.84	0.92	1.00	0.83	0.94	0.75	0.83

Our alternative approach for ensuring that formula scores do not exceed DEA scores is to use regression with constraints to ensure the residuals are one-sided, i.e. all have the same sign. The resulting formula for this is:

$$E = [\text{Outpatients} + 60\ \text{Inpatients}] / [628\ \text{Doctors} + 279\ \text{Nurses}] \qquad (6)$$

We note the relative weights for inpatients and outpatients are now 60 to 1, which is much higher than before. For doctors and nurses the relative weights are only slightly altered. From the bottom two rows of Table 3 we see that all but one of the ranks are the same or differ by one from the earlier formula; and the bottom two rows of Table 4 indicate similar scores using the two formulae. It is interesting to compare the constrained regression results with those from the earlier formula by looking at deviations from the DEA scores. Note that all the deviations have the same sign. The largest residual was 0.175 in both cases. More interesting was that the mean deviation using constrained regression was 0.0725, whereas it was higher, at 0.082, using the OLS formula with rescaling. To understand why this should be the case provides a useful insight which enables us to choose between these approaches. We expect the unconstrained OLS ‘predictions’ to be scattered above and below the DEA values in a roughly symmetric random manner. However, the effect of the subsequent adjustment by rescaling will extend the deviation associated with those points which were already under-predicted, i.e. they will be dragged down even further. This effect will, in general, cause constrained regression to provide a closer fit than OLS followed by rescaling because it takes into account the one-sided condition at the same time as the optimal fitting process.

Another way to improve the fit is to include a constant or intercept in the model formula. For the constrained regression this causes the mean deviation to fall to 0.039, which is quite a drop from 0.0725. The largest residual was 0.169. For OLS with rescaling the mean deviation was 0.066 and the largest deviation was 0.157.

We point out that whilst we have assumed a formula which is similar in form to that used to assess the units in DEA, this does not have to be the case. One could achieve an even closer fit using a more flexible form. A ratio of linear forms assumes constant rates of marginal substitution, and no curvature in the frontier. But a ratio of quadratics (or other nonlinear forms) would allow for non-constant substitution rates and curved frontiers.

3.2 Testing the method where the true weights and scores are known

In general, we would not know the 'true' weights or 'true' scores. However, one can set up a situation where there is an assumed pre-specified relationship and then generate data. An example of this is given by Bowlin et al. (1985), where they considered hospitals being assessed on three outputs: the number of regular patients (RP), number of severe patients (SP), and number of teaching units (TU; student nurses, interns, etc.), while the input was the total cost (in $'000), which was assumed to be related to the outputs according to:

$$\text{Efficient Cost} = 0.5TU + 0.13368\ RP + 0.17474\ SP \tag{7}$$

assuming the hospital is operating efficiently. They generated data for seven efficient hospitals and eight hospitals operating inefficiently (Table 5). The true score or efficiency can be found from the ratio Efficient Cost/Actual Cost, i.e.:

$$\text{True Score} = [0.5TU + 0.13368\ RP + 0.17474\ SP]/\ \text{Cost} \tag{8}$$

Table 5: Test data on 15 hospitals with true scores compared with estimated scores

HOSPITAL	Teaching units	Regular Patients	Severe Patients	Cost $'000	True Score	Automatic-Democratic Score
H1	50	3000	2000	775.5	1	0.995
H2	50	2000	3000	816.6	1	0.992
H3	100	2000	3000	841.6	1	0.992
H4	100	3000	2000	800.5	1	0.995
H5	50	3000	3000	950.3	1	0.994
H6	100	2000	5000	1191.05	1	0.990
H7	50	10000	2000	1711.3	1	1.000
H8	100	3000	2000	884.75	0.91	0.900
H9	50	2000	3000	841.6	0.97	0.963
H10	100	10000	2000	2036.3	0.85	0.852
H11	50	5000	3000	1362.6	0.89	0.890
H12	100	3000	3000	1070	0.91	0.905
H13	50	4000	5000	1491.1	0.96	0.955
H14	100	3000	3000	1070	0.91	0.905
H15	50	3000	2000	898.7	0.86	0.859

Source: Bowlin et al. (1985).

DEA was applied to the data in Table 5 to maximise the score (u_1TU + u_2RP + + u_3SP) / Cost, for each hospital. We then applied constrained least squares regression to obtain common weights. The resulting scoring formula was found to be:

$$\text{Automatic Democratic Score} = [0.488\ TU + 0.13420\ RP + 0.17243\ SP] / \text{Cost}. \quad (9)$$

We immediately notice that the weights are very close to those in the True Score formula above. The associated scores are shown in the last column of Table 5 and can be seen to be extremely close to the true values.

Thanassoulis (1993) applied regression directly to the above data using the cost as the dependent variable, and found:

$$\text{Cost} = 1.1054\ TU + 0.14811\ RP + 0.16198\ SP \quad (10)$$

These coefficients are further away from those in the underlying equation (7). This is to be expected because regression, when used alone, will model *average* behaviour, not relative performance. It is also worth noting that applying regression directly to the data is restricted to situations where there is a single dependent variable, and so cannot be applied when there are both multiple input and multiple output variables. By contrast, in our two-step approach we have used DEA to assess performance (a single variable score), and then used regression to model this performance (Tofallis, 2013).

4 Conclusion

Our aim has been to provide a straightforward method for producing a formula for performance when there are no agreed weights on the underlying criteria. Initially, scores are obtained using DEA (which, by itself, does not generate a scoring formula). These are scores that each entity would be pleased with, since DEA uses the weights which maximizes its score. Consequently, DEA scores are sometimes felt to be unrealistically high, so we view these as optimistic, and treat them as upper bounds when constructing the scoring formula. The DEA scores are regressed on the underlying attribute data in order to objectively generate a common set of weights (the regression coefficients). We thus have an 'automatic democratic' approach in that data from all entities being assessed play a part in obtaining the weights.

Whilst the DEA literature contains much work on obtaining common weights, the method presented here is simpler and more direct. For example, much research has been done using cross-efficiency (see Wu, Sun and Liang, 2020, for a review), which involves aggregating scores using weights from self and peer perspectives; however, the fact that such weights are not unique requires additional methodological complications, such as secondary goals, in order to deal with this issue. By focusing on DEA scores (which are unique) rather than the associated weights (which are not) we avoid complications arising in many other approaches. Furthermore, using least squares regression brings the benefits of a well-known and widely used technique to provide a way of generating a compact formula for performance. Use of a formula also helps overcome the lack of discrimination associated with DEA, where many units have unreasonably high or even 100% efficiency as a result of DEA's extreme weight flexibility.

Our view is that the DEA scores are optimistic upper bounds, and so we have ensured this condition is upheld. Two ways were investigated for achieving this. We compared results using constrained regression with those from OLS followed by rescaling, and found that the former provides a closer fit. An explanation was provided for this effect: Constrained regression, by definition, fits as close as possible while maintaining the required condition; whereas making an adjustment after OLS does not carry the same 'best fit' property because the optimisation and constraints are not dealt with simultaneously. This insight allows us to reject the rescaling approach. We also considered the Manhattan and Chebyshev distance metrics as alternatives to least squares for regression and identified reasons for not adopting them.

The method we have presented has the attractive property of allowing a ranking of entities without requiring the imposition of arbitrary choices or restrictions on weights. Some of the DEA ranking literature is focused on achieving what is called a 'full ranking', that is one in which there are no ties. But it has to be realised that, if a formula is to be used, there will always be the possibility of two units with different attribute levels achieving an equivalent overall score.

The functional form of the scoring formula is not restricted by our approach. Although we have used a ratio of weighted outputs to weighted inputs as our illustrative example, it is of course possible to apply regression to scores generated by other models. For example, the BCC (Banker, Charnes, Cooper) model has an additional parameter to model variable returns to scale, and so the associated regression equation would provide a closer fit to the DEA scores. There are also additive and multiplicative DEA models which can be used for generating additive (linear) and multiplicative scoring formulae.

Finally, a valuable benefit of the proposed approach is the transparency of using a simple, and objectively constructed formula, for performance evaluation.

References

Adler N., Friedman L., Sinuany-Stern Z. (2002), *Review of Ranking Methods in the Data Envelopment Analysis Context*, European Journal of Operational Research, 140(2), 249-265.

Aldamak A.M., Zolfaghari S. (2017), *Review of Efficiency Ranking Methods in Data Envelopment Analysis*, Measurement, 106, 161-172.

Ali A.I. (1993), *Streamlined Computation for Data Envelopment Analysis*, European Journal of Operational Research, 64(1), 61-67.

Becker R.A., Denby L., McGill R., Wilks A.R. (1987), *Analysis of Data from the Places Rated Almanac*, The American Statistician, 41(3), 169-186.

Bowlin W.F., Charnes A., Cooper W.W., Sherman H.D. (1985), *Data Envelopment Analysis and Regression Approaches to Efficiency Estimation and Evaluation*, Annals of Operations Research, 2, 113-138.

Charnes A., Cooper W.W., Rhodes E. (1978), *Measuring the Efficiency of Decision Making Units*, European Journal of Operational Research, 2(6), 429-444.

Contreras I., Lozano S., Hinojosa M.A. (2021), *A DEA Cross-efficiency Approach Based on Bargaining Theory*, Journal of the Operational Research Society, 72(5), 1156-1167.

Cook W.D., Roll Y., Kazakov A. (1990), *A DEA Model for Measuring the Relative Efficiency of Highway Maintenance Patrols*, INFOR: Information Systems and Operational Research, 28(2), 113-124.

Cook W.D., Zhu J. (2007), *Within-group Common Weights in DEA: An Analysis of Power Plant Efficiency*, European Journal of Operational Research, 178(1), 207-216.

Cooper W.W., Seiford L.M., Tone K. (2006), *Introduction to Data Envelopment Analysis and Its Uses: With DEA-solver Software and References*, Springer Science & Business Media.

Decancq K., Lugo M.A. (2013), *Weights in Multidimensional Indices of Wellbeing: An Overview*, Econometric Reviews, 32(1), 7-34.

Dehnokhalaji A., Hallaji B., Soltani N., Sadeghi J. (2017), *Convex Cone-based Ranking of Decision-making Units in DEA*, OR Spectrum, 39(3), 861-880.

Despotis D.K. (2002), *Improving the Discriminating Power of DEA: Focus on Globally Efficient Units*, Journal of the Operational Research Society, 53(3), 314-323.

Despotis D.K. (2005), *A Reassessment of the Human Development Index via Data Envelopment Analysis*, Journal of the Operational Research Society, 56(8), 969-980.

Gajewski B.J., Lee R., Bott M., Piamjariyakul U., Taunton R.L. (2009), *On Estimating the Distribution of Data Envelopment Analysis Efficiency Scores: An Application to Nursing Homes' Care Planning Process*, Journal of Applied Statistics, 36(9), 933-944.

Hollingsworth B. (2008), *The Measurement of Efficiency and Productivity of Health Care Delivery*, Health Economics, 17(10), 1107-1128.

Jahanshahloo G.R., Memariani A., Lotfi F.H., Rezai H.Z. (2005), *A Note on Some of DEA Models and Finding Efficiency and Complete Ranking Using Common Set of Weights*, Applied Mathematics and Computation, 166(2), 265-281.

Kao C., Hung H.T. (2005), *Data Envelopment Analysis with Common Weights: The Compromise Solution Approach*, J Operational Research Soc., 56(10), 1196-203.

Kritikos M.N. (2017), *A Full Ranking Methodology in Data Envelopment Analysis Based on a Set of Dummy Decision Making Units*, Expert Systems with Applications, 77, 211-225.

Liu F.H.F., Peng H.H. (2008), '*Ranking of Units on the DEA Frontier with Common Weights'*, Computers and Operations Research, 35(5), 1624-1637.

Liu J.S., Lu W.M. (2010), *DEA and Ranking with the Network-based Approach: A Case of R&D Performance*, Omega, 38(6), 453-464.

Lotfi F.H., Jahanshahloo G.R., Khodabakhshi M., Rostamy-Malkhlifeh M., Moghaddas Z., Vaez--Ghasemi M. (2013), *A Review of Ranking Models in Data Envelopment Analysis*, Journal of Applied Mathematics, 2013, 1-20.

Matawie K.M., Assaf A. (2010), *Bayesian and DEA Efficiency Modelling: An Application to Hospital Foodservice Operations*, Journal of Applied Statistics, 37(6), 945-953.

Moghaddas Z., Vaez-Ghasemi M. (2017), *Ranking Models in Data Envelopment Analysis Technique* [in:] F.H.Z. Lotfi, S.E. Najafi, H. Nozari (eds.), *Data Envelopment Analysis and Effective Performance Assessment* (265-311), IGI Global.

Ramón N., Ruiz J.L., Sirvent I. (2012), *Common Sets of Weights as Summaries of DEA Profiles of Weights: With an Application to the Ranking of Professional Tennis Players*, Expert Systems with Applications, 39(5), 4882-4889.

Roll Y., Cook W.D., Golany B. (1991), *Controlling Factor Weights in Data Envelopment Analysis*, IEEE Transactions, 23(1), 2-9.

Roll Y., Golany B. (1993), *Alternate Methods of Treating Factor Weights in DEA*, Omega, 21(1), 99-109.

Rosenthal E.C., Weiss H.J. (2017), *A Data Envelopment Analysis Approach for Ranking Journals*, Omega, 70, 135-147.

Ruiz J.L., Sirvent I. (2016), *Common Benchmarking and Ranking of Units with DEA*, Omega, 65, 1-9.

Thanassoulis E. (1993), *A Comparison of Regression Analysis and Data Envelopment Analysis as Alternative Methods for Performance Assessments*, J Operational Research Society, 44(11), 1129-1144.

Thompson R.G., Dharmapala P.S., Thrall R.M. (1993), *Importance for DEA of Zeros in Data, Multipliers, and Solutions*, Journal of Productivity Analysis, 4(4), 379-390.

Tofallis C. (2013), *An Automatic-democratic Approach to Weight Setting for the New Human Development Index*, Journal of Population Economics, 26(4), 1325-1345.

Wu J., Sun J., Liang L. (2020), *Methods and Applications of DEA Cross-efficiency: Review and Future Perspectives*, Front. Eng. Manag., 1-13.

Zohrehbandian M., Makui A., Alinezhad A. (2010), *A Compromise Solution Approach for Finding Common Weights in DEA: An Improvement to Kao and Hung's Approach*, Journal of the Operational Research Society, 61(4), 604-610.

Part II
Regular papers

MULTIPLE CRITERIA DECISION MAKING

Vol. 17 2022

Somdeb Lahiri[*]

AXIOMATIC CHARACTERIZATIONS OF PROBABILISTIC MAX-MIN EXTENDED CHOICE CORRESPONDENCE

DOI: 10.22367/mcdm.2022.17.05

Received: 24.11.2022 | Revised: 5.01.2023 | Accepted: 1.03.2023.

Abstract

In this paper we provide two axiomatic characterizations of the probabilistic max-min extended choice correspondence support, for a decision maker who has state-dependent preferences (represented by a linear order) over the set of alternatives and a (subjective) probability vector over states of nature, where both preferences and probability vectors are variable.

Keywords: state-dependent preferences, extended choice correspondence.

1 Introduction

A (fixed agenda) extended choice correspondence assigns to each profile of state-dependent strict rankings over the set of alternatives and probability vector over a non-empty finite set of states of nature, a non-empty (not necessarily proper) subset of alternatives, from a given non-empty finite and fixed set of alternatives. The genesis of this concept and a fairly detailed mathematical discussion of it can be found in Lahiri (2020/2021). With a different interpretation, Denicolo (1985) refers to a special case of the same mathematical entity as a social choice correspondence. The special case corresponds to equiprobable states of nature, but since the interpretation in the paper by Denicolo cited above is one of group decision-making under certainty, there is a point at which the analysis in our paper would remain incomplete, had we interpreted the framework differently. The different states of nature could be interpreted as different criteria and the probability vectors as weights, thus reducing it to a multi-criteria

[*] (Former Professor) PDE University, India and (Adjunct Professor) LJ University, India, e-mail: somdeb.lahiri@gmail.com, ORCID: 0000-0002-5247-3497.

decision making (MCDM) problem with a weight for each criterion (“MCDM with weights”). However, we don’t want to push that interpretation further and would rather root for interpreting “MCDM with weights” as decision-making under probabilistic uncertainty.

The framework introduced in Lahiri (2020/2021) is an extension of the seminal model of social choice theory developed by Kenneth J. Arrow. A decision maker is faced with making a choice under probabilistic uncertainty (risk) in which uncertainty is with regard to a future state of nature, which is realized after the decision has been made. The decision maker is provided with (or aware of) a data profile, which is a pair whose first component is a profile of state-dependent rankings over (the consequences) a non-empty finite set of alternatives and whose second component is a probability distribution over a non-empty finite set of states of nature. A decision support system (DSS) or decision aid is required to choose a non-empty “desirable” set of alternatives from which the final choice has to be made. The decision aid or DSS has no bias in favor of any one or more alternatives that it suggests. Such a decision support system is called an extended choice correspondence, i.e., a rule which associates with each data profile from a given set of data profiles a non-empty finite set of desirable alternatives.

The problem of choosing one or more alternatives from a given set of alternatives was raised and rigorously formulated for the first time in a seminal contribution on majority voting by Pattanaik (1970). For the classical theory of decision making under uncertainty in the state dependent case − which is the other and major motivation behind Lahiri (2020/2021) − one may refer to Karni (1985).

The initial concern that led to the frameworks discussed in this paper is that Arrowian voting theory framework does not have anything to say about the role of negotiations in group decision making and may therefore be very inadequate for our understanding of decision making in society. In view of this, slight extensions of voting models as models of choice under risk may serve a useful purpose.

The reasons for our interest in state-dependent preferences are precisely the same as the ones discussed in Karni (1985), i.e., it is so obviously true that it does not need justification beyond citing trivial day-to-day examples as Karni has done in his book. However, Karni focuses on state-dependent utility functions and it is our contention here that the informational requirements for (state-dependent?) utility of state-dependent monetary surplus derived by decision makers from consuming alternatives, may prove prohibitive and a significant reason for “bounded rationality”, thus leading to “useable” preferences being represented by rankings instead of utility functions. Knowledge of the exact state-dependent monetary surplus (and not necessarily the state-dependent utility functions) is not easy to obtain for the purpose of decision making, not only because the cost of obtaining such information is often exorbitant, but also because the knowledge of the monetary benefit from the chosen alternative may only be available on conditions prevailing at a future date that are neither accessible nor can be experienced at the time the decision has to be made.

Hence, the major justification for the framework and the investigation in Lahiri (2020/2021) is that the classical theory of decision making under uncertainty that rests on the assumption of maximization of expected utility (state-dependent or not) has an important limitation − i.e. the decision maker's preferences may not be "useable" in the form of cardinal utility functions, but only as rankings. That leads to a departure from the classical theory and opens up the possibility of decision makers using other algorithms (decision aids) for the purpose of decision making under risk. That is the line of investigation pursued in this paper. A full-fledged application using components of this framework to prove the existence of "preferred with probability at least half winners" has been provided in Lahiri (2020; 2021). This however, is not meant to be a denial of the worth of the huge literature based on utility functions that uses procedures other than expected utility maximization, to explain paradoxes that arise if the latter criterion is used to explain decision making under uncertainty. One such is the work of Gilboa (1988) which suggests that decision makers maximize a function that is increasing in both expected utility of an alternative and the worst utility of the alternative, in arriving at optimal choices. Such procedures would require information about the state-dependent utility of each alternative, and it is our contention here − as observed earlier − that such information may not be as easily available as expected utility theory presupposes. The ordinal equivalent of the procedure suggested by Gilboa (1988) would require maximizing a function of the Probabilistic Borda Score of an alternative (see Lahiri (2020/2021) and the worst rank that the alternative attains with positive probability, where the function is "increasing" in the first variable and "decreasing" in the second.

Here we begin by setting up the model for extended choice correspondences. In this framework we provide two axiomatic characterizations of the probabilistic max-min extended choice correspondence. This extended choice correspondence is based on the max-min choice correspondence due to Campbell, Kelly and Qi (2018). The max-min choice correspondence of Campbell, Kelly and Qi (2018) selects for each preference profile those alternatives which have the best "worst rank". In our framework, for a data profile − a pair comprising a strict preference profile and a probability vector (for the states of nature) − a "max-min alternative" is an alternative whose worst rank among states of nature that occur with probability is the best. The worst rank of a max-min alternative is said to be the "max-min rank". Our probabilistic max-min extended choice correspondence selects for each data profile those max-min alternatives which have the least positive probability of attaining the “max-min rank”. We ignore those states of nature that occur with probability zero, since if an alternative attains its worst rank with probability zero, it is improbable (though not impossible) that it will attain such a rank. Furthermore, if a max-min alternative attains the max--min rank with lowest probability, then it attains a superior rank with the highest probability among all max-min alternatives. It is very unlikely that a risk-averse individual, to whom the probabilistic max-min extended choice correspondence would be recommended, could wish for anything better. A related earlier paper is the one by Congar and Merlin (2012), where the main concern is with axiomatic characterization of max-min “social welfare function”. The domain of the probabilistic max-min extended choice correspondence, whose axiomatization we provide, is the set of all data profiles, such that for any non-empty subset of probability vectors, all strict preference profiles can be associated with any probability vector in the subset. The strict sub-domain where the data profiles are such that those states of nature that occur with positive probability have equal probability of occurrence is said to be one with equiprobable support. On the domain with equiprobable support, our solution concept is a refinement of the one discussed in Campbell, Kelly and Qi (2018), with a different interpretation. This would correspond to the sub-solution of the one in Campbell, Kelly and Qi (2018), where only those max-min alternatives with max-min ranks for the fewest number of agents are chosen.

Our study here concerns decision making under uncertainty and one of the earliest works dealing with axiomatic characterizations in such a scenario is that of Maskin (1979). However, the structure of the underlying set of alternatives from which choices are made in Maskin (1979) is completely different from what we assume here and hence our axiomatization, as well as the methodology we use to obtain our results, is completely different from the corresponding ones that are reported and used here. Another notable contribution in a related but different line of research is the work of Gilboa and Schmeidler (1989). A fairly comprehensive survey of research on decision-making uncertainty is the paper by Kelsey and Quiggin (1992). A paper that could have been an exact predecessor to our work here is the one due to Puppe and Schlag (2009), if they had used state dependent strict rankings (even rankings would do!) instead of state dependent pay-off functions. The fact that in their context, the set of alternatives from which choices in a state of nature can be made is state-dependent *may not* be a problem, if we take the given fixed set of alternatives to be the union of sets of alternatives available over the different states of nature and in each state of nature ranked those alternatives that are not available in that state of nature, strictly below those that are available in that state of nature. The paper by Congar and Merlin (2012) which is concerned with the max-min (Rawlsian) social welfare function, is related to the work that follows, but may fail to shed light on our results, since it uses a variable number of states (voters) argument via two axioms − duplication and weak separability − in the analysis reported there.

In a final section of this paper we provide an example to illustrate how "bounded rationality" arises in the context of probabilistic uncertainty due to absence of sufficient information about the state-dependent monetary surpluses of an alternative, thereby rendering expected utility maximization practically unusable. Under such circumstances, using a framework of analysis based on state-dependent rankings of alternatives may be unavoidable if not inevitable.

Proofs of results are available from the author on request.

2 The framework of analysis

The following framework is a fairly close adaptation of the ones available in Denicolo (1985) and section 2.2 of Endriss (2011) and discussed thoroughly in Lahiri (2020/2021).

Consider a decision maker (DM) faced with the problem of choosing one or more alternatives from a non-empty finite set of alternatives X, containing at least three elements. Let $\Psi(X)$ denote the set of all non-empty subsets of X. For a positive integer $n \geq 3$, let $N = \{1, 2, \ldots, n\}$. Contrary to convention we will refer to an element in N as a state of nature and to the set N as the set of states of nature.

A **strict preference relation/strict ranking** on X is a linear order (i.e. a reflexive, complete/connected/total, transitive and anti-symmetric binary relation) on X. Generally, a strict preference relation is denoted by R with P denoting its asymmetric part. If for $x, y \in X$, it is the case that $(x, y) \in R$, then we shall denote it by xRy and say that x is **at least as good as** y **for the strict preference relation** R. Similarly, xPy is interpreted as x is **strictly preferred** to y **for the strict preference relation** R.

Given a strict preference R and an alternative x, **the rank of** x **at** R, denoted by $rk(x, R) = |\{y \in X|yRx\}|$, i.e., 1 + cardinality of the set of alternatives strictly preferred to x for the strict preference relation R.

Let $\mathcal{L}$ denote the set of all strict preference relations on X.

A **strict preference profile**, denoted by R_N, is a function from N to $\mathcal{L}$. R_N is represented as the array $\langle R_i|i \in N\rangle$, where R_i is the strict preference relation/strict ranking in state of nature i. The set of all preference profiles is denoted by $\mathcal{L}^N$.

A **probability vector over** N is a vector $p \in \mathbb{R}_+^N$ satisfying $\sum_{i=1}^{N} p_i = 1$ where for $i \in N$, p_i is the probability that state of nature 'i' occurs.

The **set of probability vectors over** N is denoted by Δ.

Given a probability vector p, the set $\{j|p_j > 0\}$ is said to be **the support of** p and is denoted by support(p).

Since probabilities are associated with events, for each $i \in N$, the state of nature i represents a non-empty set and N is a finite partition of some underlying sample space.

Given $(R_N, p) \in \mathcal{L}^N \times \Delta$ and an alternative x (i.e. $x \in X$), a state of nature i (i.e., $i \in N$) will be said to be **a worst state of nature for** x at (R_N, p) if $i \in \operatorname{argmax}_{j \in \text{support}(p)} \text{rk}(x, R_j)$.

The above definition says that a state of nature is a worst state of nature for an alternative if the state of nature occurs with "positive probability" and the alternative does not attain any worse rank with "positive probability".

Given $(R_N, p) \in \mathcal{L}^N \times \Delta$ and an alternative x (i.e. $x \in X$), the set $WS(x, (R_N, p)) = \{i|$ i is a worst state of nature for x$\}$ is said to be **the set of worst states of nature** for x at (R_N, p), and for $i \in WS(x, (R_N, p))$, $\text{rk}(x, R_i)$, denoted worstrk$(x,(R_N, p))$, is said to be **the worst rank** of x at (R_N, p).

Clearly, worstrk$(x,(R_N, p)) = \max\{\text{rk}(x,R_i)|i \in \text{support}(p)\}$ for all $x \in X$.

For all $(R_N, p) \in \mathcal{L}^N \times \Delta$, let $Mm(R_N, p) = \operatorname{argmin}_{y \in X} \text{worstrk}(y, (R_N, p))$.

$Mm(R_N, p)$ is said to be **the set of max-min alternatives at** (R_N, p). The **max-min rank for** (R_N, p) is equal to the unique worstrk$(x,(R_N, p))$ for any $x \in Mm(R_N, p)$.

A **domain** is any non-empty subset of $\mathcal{L}^N \times \Delta$. We will denote a domain by $\mathcal{R}$.

An **extended choice correspondence** (ECC) on (domain) $\mathcal{R}$ is a function f from $\mathcal{R}$ to $\Psi(X)$.

The problem with $Mm(R_N, p)$ and any ECC that does not discriminate between states of nature which have positive probability is that they might overemphasize the "extremely unlikely" to absurd extents thereby denying the decision maker the right to exercise one's discretion within reasonable limits.

Example 1: $X = \{x,y\}$, $n = 2$, $p_1 = \frac{1}{100}$, $p_2 = \frac{99}{100}$.

$\text{rk}(x,R_1) = 1$, $\text{rk}(y,R_1) = 2$; $\text{rk}(x,R_2) = 2$, $\text{rk}(y,R_2) = 1$.

$Mm(R_N, p) = \{x,y\}$. But, does 'x' have any reason to be treated at par with 'y', when there is a 99% chance that 'y' is going to be preferred to 'x'?

Hence, we consider the following procedure.

The following notation will prove useful in what follows.

Given $(R_N, p) \in \mathcal{L}^N \times \Delta$ and $x \in X$, the **probability of the worst rank** of x at (R_N, p) denoted by $\Pr(WS(x, R_N, p)) = \sum_{i \in WS(x,R_N,p)} p_i$.

An ECC on $\mathcal{R}$ is said to be the **probabilistic max-min choice correspondence**, denoted by f^{PMm}, if for all $(R_N, p) \in \mathcal{R}$, $f^{PMm}(R_N, p) = \{x \in Mm(R_N, p) | Pr(WS(x, R_N, p)) \leq Pr(WS(y, R_N, p))$ for all $y \in Mm(R_N, p)\}$, i.e., $f^{PMm}(R_N, p)$ is the set of max-min alternatives with least total probability of securing the best worst rank at (R_N, p).

Thus, an ECC is f^{PMm} which at any (R_N, p) in the domain of the ECC, chooses those max-min alternatives whose max-min rank occur with least probability, i.e., the chosen alternatives are those max-min alternatives that each occurs at its worst rank with the least probability. In other words, f^{PMm} minimizes "the probability" with which a max-min rank occurs.

Clearly, $f^{PMm}(R_N, p)$ for Example 1 is {y}.

Choosing a "best ranked" alternative from among those which attain its worst rank with least probability may prove to be misleading as the following example reveals.

Example 2: X = {x,y,z}, n = 3, $p_1 = \frac{1}{100}$, $p_2 = \frac{98}{100}$, $p_3 = \frac{1}{100}$.

$rk(x,R_1) = 1$, $rk(y,R_1) = 2$, $rk(z,R_1) = 3$; $rk(x,R_2) = 2$, $rk(y,R_2) = 1$, $rk(z,R_2) = 3$; $rk(x,R_3) = 3$, $rk(y,R_3) = 2$, $rk(z,R_3) = 1$.

The probability with which x gets its worst rank, i.e. 3, is $\frac{1}{100}$. The probability with which y gets its worst rank, i.e. 2, is $\frac{2}{100} = \frac{1}{50}$. The probability with which z gets its worst rank, i.e. 3, is $\frac{99}{100}$. Hence (in this case) the unique alternative which attains its worst rank with least probability is x and the worst rank is equal to 3.

However, there is only one max-min ranker, i.e. y, and the first method selects 'y'. This seems quite reasonable for a risk-averse individual, since there is a 99% chance that 'y' will be preferred to 'x' and a 99% chance that 'y' will be preferred to 'z'.

In view of the fact that the domain $\mathcal{R}$ is a subset of $\mathcal{L}^N \times \Delta$, given any $(R_N, p) \in \in \mathcal{R}$ it is not possible for two different alternatives to have the same worst state of nature at (R_N, p).

In what follows we will be concerned only with those domains which satisfy the following property:

Domain Property: $\mathcal{R} = \mathcal{L}^N \times Q$, where Q is a non-empty subset of Δ.

3 Some axioms and a lemma that will be useful on the way

We begin this section with two very desirable axioms that few would wish to contest.

An ECC f on $\mathcal{R}$ is said to satisfy **Unanimity** if for $(R_N, p) \in \mathcal{R}$, $x \in X$: $[rk(x, R_i) = = 1$ for all $i \in N]$ implies $[f(R_N, p) = \{x\}]$.

An ECC f on $\mathcal{R}$ is said to satisfy **Independence of Irrelevant States** (to be **Independent of Irrelevant States**) (**IIS**) if for all (R_N, p), $(R'_N, p) \in \mathcal{R}$: $[\{j|p_j > 0\} \subset \subset \{j|R_j = R'_j\}]$ implies $[f(R'_N, p) = f(R_N, p)]$.

The next axiom is considerably more specific to our present context.

An ECC f on $\mathcal{R}$ is said to satisfy the **Worst-Rank Property** if for all $(R_N,p) \in \in \mathcal{R}$ and $x \in f(R_N,p)$: $[\text{worstrk}(x, (R_N,p)) > 1]$ implies [for no $y \in X$ is it the case that $\text{worstrk}(y, (R_N,p)) = \text{worstrk}(x, (R_N,p)) - 1$].

An ECC f on $\mathcal{R}$ is said to satisfy **Worst-Rank Positive Responsiveness (W-RPR)** if for all (R_N,p), $(R'_N,p) \in \mathcal{R}$, $x \in f(R_N,p)$ satisfies $\text{worstrk}(x, (R_N,p)) > 1$ and $i \in WS(x,(R_N,p))$: [(a) $R'_k = R_k$ for all $k \neq i$; (b) $rk(x,R'_i) = rk(x, R_i) - 1$, $rk(z, R'_i) = rk(z, R_i)$ if $z \neq x$ and $rk(z, R_i) \neq rk(x, R_i) - 1$] implies $[f(R'_N,p) = \{x\}]$.

From the construction of (R'_N,p) it is clear that $rk(z, R_i) = rk(x, R_i) - 1$ implies $rk(z, R'_i) = rk(z, R_i) + 1 = rk(x, R_i)$.

Note that from the definition of W-RPR, we get either $WS(x,(R'_N,p)) = = WS(x,(R_N,p))\backslash\{i\}$ in which case $\text{worstrk}\ (x,(R'_N,p)) = \text{worstrk}(x,(R_N,p))$ and $\sum_{j \in WS(x,(R'_N,p))} p_j = \sum_{j \in WS(x,(R_N,p))} p_j - p_i < \sum_{j \in WS(x,(R_N,p))} p_j$ or $WS(x,(R'_N,p)) = = WS(x,(R_N,p))$ in which case $\text{worstrk}\ (x,(R'_N,p)) = \text{worstrk}\ (x,(R_N,p)) - 1$.

W-RPR says that if i is a worst state of nature for some chosen alternative with worst rank greater than 1 and if in state of nature i this alternative exchanges its position with the alternative immediately above it at ‘i’, then after such a change this alternative becomes the uniquely chosen alternative and the unique max-min alternative.

As an immediate consequence of Unanimity, IIS, Worst Rank Property and W-RPR is the fact that chosen alternatives must be max-min alternatives.

Lemma 1: If an ECC f on a domain $\mathcal{R}$ satisfies Unanimity, IIS and W-RPR then for all $(R_N,p) \in \mathcal{R}$, it must be the case that $f(R_N, p) \subset Mm(R_N, p)$.

4 The main result

An ECC f on $\mathcal{R}$ is said to satisfy **Not Chosen After Worser Rank (NCWR)** if for (R_N, p), $(R'_N, p) \in \mathcal{R}$, $x \notin f(R_N, p)$ and $i \in N$: [(a) $R'_k = R_k$ for all $k \neq i$; (b) $rk(x, R'_i) = rk(x, R_i) + 1$, $rk(z, R'_i) = rk(z, R_i)$ if $z \neq x$ and $rk(z, R_i) \neq rk(x, R_i) + 1$] implies $[x \notin (R'_N, p)]$.

NCWR says that if initially an alternative is not chosen, then it remains unchosen if in any state of nature, it exchanges places with an alternative ranked immediately below it. It is easily verified that f^{PMm} satisfies NCWR.

An ECC f on $\mathcal{R}$ is said to satisfy **Not More Probable Worse Rank (NMPWR)** if for all (R_N,p), $(R'_N, p) \in \mathcal{R}$, $x \in X$ and $i \in I$: $[x \in f(R_N, e^I) \cap f(R'_N, e^I)]$ implies $[P(WS(x, (R'_N, p))) \leq P(WS(y, (R'_N, p)))$ for all $y \in f(R'_N, p)]$, where: [(a) $R'_k = R_k$ for all $k \neq i$; (b) $rk(x,R'_i) = rk(x, R_i) + 1$, $rk(z, R'_i) = rk(z, R_i)$ if $z \neq x$ and $rk(z, R_i) \neq rk(x, R_i) + 1$].

From the construction of (R'_N,p) it is clear that $rk(z, R_i) = rk(x, R_i) + 1$ implies $rk(z, R'_i) = rk(z, R_i) - 1 = rk(x, R_i)$.

NMPWR says that if a chosen alternative is chosen after it exchanges its position with the alternative immediately below it at a state of nature occurring with positive probability, then at the lower rank it occurs with positive probability not more often than any other chosen alternative does.

An ECC f on $\mathcal{R}$ is said to satisfy **Greater Probability if Exclusion After Worsening (GPEAW)** if for all (R_N,p), $(R'_N,p) \in \mathcal{R}$, $x \in X$ and $i \in N$: $[\{x\} = Mm(R_N,p)$ and $x \notin f(R'_N,p)]$ <u>implies</u> $[P(WS(x, (R'_N,p))) > P(WS(y, (R'_N,p)))$ for some $y \in f(R'_N,p)]$, where: [(a) $R'_k = R_k$ for all $k \neq i$; (b) $rk(x,R'_i) = rk(x, R_i) + 1$, $rk(z, R'_i) = rk(z, R_i)$ if $z \neq x$ and $rk(z, R_i) \neq rk(x, R_i) + 1$].

From the construction of (R'_N,p) it is clear that $rk(z, R_i) = rk(x, R_i) + 1$ implies $rk(z, R'_i) = rk(z, R_i) - 1 = rk(x, R_i)$.

Given Lemma 1, GPEAW is the converse of NMPWR. Along with Lemma 1, what NMPWR and GPEAW together say is the following:

A chosen alternative is chosen after it exchanges its position with the alternative immediately below it at a state of nature occurring with positive probability if and only if at the lower rank it occurs with positive probability not more often than any other chosen alternative does.

Note: By IIS, the three properties NCWR, NMPWR and GPEAW hold non-vacuously only when the state of nature 'i' in their definitions belong to support(p).

Proposition 1: If an ECC f on $\mathcal{R}$ satisfies Unanimity, IIS, Worst Rank Property, W-RPR, NCWR NMPWR and GPEAW then for all $(R_N,p) \in \mathcal{R}$, $f(R_N,p) = f^{PMm}(R_N,p)$.

It is easy to see that on the domain with equiprobable support f^{PMm} satisfies Unanimity, IIS, Worst Rank Property, NCWR, W-RPR, NMPWR and GPEAW.

Thus we arrive at the following theorem.

Theorem 1: An ECC f on $\mathcal{R}$ satisfies Unanimity, IIS, Worst Rank Property, W-RPR, NCWR, NMPWR and GPEAW <u>if and only if</u> $f = f^{PMm}$ on $\mathcal{R}$.

An alternative axiomatic characterization with a shorter proof can be obtained by replacing NMPWR by the following property.

An ECC f on $\mathcal{R}$ is said to satisfy **Non-Domination of Worst Rank (ND-WR)** if there does not exist $(R_N, p) \in \mathcal{R}$ and $x,y \in f(R_N, p)$ satisfying $\text{worstrk}(x, R_N, p) \leq \text{worstrk}(y, R_N, p)$, $\text{Pr.}(\text{WS}(x, R_N, p)) \leq \text{Pr.}(\text{WS}(y, R_N, p))$ with at least one strict inequality.

ND-WR says that given two chosen alternatives if one has a "better" worst rank than the other, then the probability of the first alternative securing its worst probable rank must be greater than the corresponding probability of the second.

With ND-WR replacing NMPWR, we have the following.

Proposition 2: If an ECC f on $\mathcal{R}$ satisfies Unanimity, IIS, Worst Rank Property, W-RPR, NCWR ND-WR and GPEAW then for all $(R_N,p) \in \mathcal{R}$, $f(R_N,p) = f^{PMm}(R_N,p)$.

Since f^{PMm} satisfies ND-WR, we get the following result.

Theorem 2: An ECC f on $\mathcal{R}$ satisfies Unanimity, IIS, W-RPR, NCWR, ND-WR and LPEAW if and only if $f = f^{PMm}$ on $\mathcal{R}$.

5 Bounded rationality in decision aiding − an example

This section was included at the behest of Professor Tadeusz Trzaskalik and I thank him for the suggestion.

Consider an individual who has to book a room in a hotel for an overnight stay that is supposed to take place a week later. There are three types of rooms in the hotel: Rooms with air conditioners "x", Rooms with air coolers "y" and Rooms with just a ceiling fan "z". The tariff for a room of type x is INR 3500 per night, for a room of type y it is INR 3000 per night, and for a room of type z it is INR 2500. The weather during the night of the proposed stay at the hotel could be either "1" hot and dry, "2" hot and humid, or "3" just pleasant, with equal probability of occurrence of each of the three types of weather.

The individual's satisfaction from each of the three types of rooms is reflected in a reservation price which depends not only on the weather but on other amenities (such as the availability of air freshener, room service etc.) and *in particular*, "the intensity of the weather condition", about which information is not available to the individual at the time of booking the room. This is a situation that may be referred to as "bounded rationality due to lack of sufficient information". However, on the basis of the information available − which includes room tariffs − the individual's weather-dependent preferences are as follows:

If the weather is as in 1, y is ranked first, x is ranked second and z is ranked third.

If the weather is as in 2, x is ranked first, z is ranked second and y is ranked third.

If the weather is as in 3, z is ranked first, y is ranked second and x is ranked third.

On the basis of the above information the individual prefers x to z with probability $\frac{2}{3}$, y to x with probability $\frac{2}{3}$ and z to y with probability $\frac{2}{3}$. This situation is referred to as Condorcet Paradox, where the individual is indecisive due to lack of sufficient information. Almost all reasonable ECC would recommend {x,y,z} under such circumstances.

However, if the "reservation price" for a type of room w ∊ {x,y,z} under the weather condition j ∊ {1,2,3} is denoted by WTP(w, j) (i.e. willingness to pay for w if the weather is as in "j") then the individual's expected surpluses for a type of room x is given by $\frac{1}{3}$[WTP(x, 1) + WTP(x, 2) + WTP(x, 3)] − 3500, for a type of room y is given by $\frac{1}{3}$[WTP(y, 1) + WTP(y, 2) + WTP(y, 3)] − 3000 and for a type of room z is given by $\frac{1}{3}$[WTP(z, 1) + WTP(z, 2) + WTP(z, 3)] − 2500.

It is not unreasonable to assume that WTP(x, 1) = 4000, WTP(y, 1) = 3600, WTP(z, 1) = −1000; WTP(x, 2) = 4500, WTP(y, 2) = 3000, WTP(z, 2) = 3000; and WTP(x, 3) = 3600, WTP(y, 3) = 3500, WTP(z, 3) = 3500. However, such information will be available only after arrival at the hotel and not at the time of booking a room.

During a night that is not humid, an air cooler can be made to serve exactly the same purpose as a ceiling fan, simply by turning off the water pump of the air cooler. On such nights the surplus (i.e., reservation price minus room tariff) is clearly greater for a room with a ceiling fan than for a room with an air cooler. It is only on a hot and dry night that the ceiling fan has the same effect as a "blast furnace".

Thus, on a hot and dry night the individual's surplus from a room of type x is 500, from a room of type y is 600 and from a room of type z is −3500.

On a hot and humid night, the individual's surplus from a room of type x is 1000, from a room type of room y is 0 and from a room of type z is 500.

On a pleasant night, the individual's surplus from a room of type x is 100, from a room of type y it is 500 and from a room of type z is 1000.

It is easy to see that the weather-dependent surpluses are consistent with the weather-dependent rankings.

The expected surplus from a room of type x is $533\frac{1}{3}$, the expected surplus from a room of type y is $366\frac{2}{3}$ and the expected surplus from a type of room z is $-666\frac{2}{3}$.

Suppose the individual is “risk neutral”.

Thus, had this information been available to the individual at the time of booking, the individual would have chosen x.

Hence the indecisiveness noticed earlier is an instance of “bounded rationality due to lack of sufficient information”.

In a private contribution, Professor Prasanta Pattanaik suggested that bounded rationality could arise out of a much greater informational deficit, i.e., lack of information about the probabilities of the states of nature. Clearly, this would be a very general starting point for investigating the consequences of “bounded rationality due to lack of sufficient information”.

In a different context, Professor Pattanaik mentioned his work in Pattanaik (1968), which seems to be related to what we are discussing here. I wish to thank him for showing me the way to Pattanaik (1968). In Pattanaik (1968), the problem faced by an individual is to choose one from a non-empty finite set X of societies that the individual could migrate to. Let #X denote the cardinality (i.e., the number of societies) in set X. For each society in X, there are ‘n’ possible positions that the individual may end up being placed in, resulting in a state of nature $s \in S = \{1, \dots, n\} \times X$. A typical state of nature, (j, x), represents the event “the individual chooses society ‘x’ and is assigned position j”. Pattanaik (1968) assumes the “first best” situation, where the individual’s preferences are represented by a **state-dependent utility function** $u: X \times S \to \mathbb{R}$ satisfying the property that for all $x \in X$ and $(j, x) \in S$, $u(x, (j, y)) = 0$ if $x \neq y$. Since, from the perspective of “nature” − the hypothetical entity that chooses or assigns the position to the individual − a priori, the probability of each society being chosen is the same as that of any other, the admissible set of subjective probability distributions over the states of nature S is a function of the form $p: \{1,\dots,n\} \times X \to [0,1]$ such that for each $x \in X$, $p(j, x) = \frac{q(j|x)}{\#X}$, where for each $x \in X$ and $j \in \{1, \dots, n\}$, $q(j|x) \in [0,1]$ and $\sum_{j=1}^{n} q(j|x) = 1$. Here q(j|x), may be interpreted as the probability of the event of being assigned the j^{th} position conditional on migrating to society ‘a’. The probability distribution ‘p’ is the individual’s assessment of the randomized strategy chosen by nature.

Note that $\sum_{s \in S} p(s) = \sum_{x \in X}(\sum_{j=1}^{n} \frac{q(j|x)}{\#X}) = \sum_{x \in X}(\frac{1}{\#X}\sum_{j=1}^{n} q(j|x)) = \sum_{x \in X} \frac{1}{\#X} = 1$.

Given such a ‘p’, the individual’s problem is to choose an $x \in X$ that maximizes $\sum_{s \in S} u(x, s)p(s)$ which is equivalent to choosing an $x \in X$ that maximizes $\sum_{j=1}^{n} u(x, (j, x))q(j|x)$.

The purpose of our example in this section is to point out that the information required to formulate individual preferences in terms of state-dependent utility functions may be difficult to access and at best one may have state-dependent preferences represented by a "partial preference relation", leading to "bounded rationality" that may be consistent with optimization and yet lead to sub-optimal outcomes, simply due to insufficient information.

I am also very grateful to Itzhak Gilboa for his informed observation which includes the following observations (which he honestly claims to be his personal views on "bounded rationality"):

(a) State-dependent preferences or utilities are not a reflection of bounded rationality.

(b) "(…) the term «bounded rationality» was coined by Simon, who had something much more dramatic in mind than what most people refer to by the term since he wanted to reject the entire optimization paradigm, replacing it by satisficing. While satisficing can also be embedded in the optimization framework, at least formally, it seems to me more of a deviation from the classical paradigm than, say, bounded memory, non-material payoffs and other models that explicitly are about optimization of something".

Given that expected surplus maximization is simply expected utility maximization by a risk-neutral individual, we have no reason to disagree with his claim that representation of preferences by state-dependent utility functions do not imply "bounded rationality". However, the example in this note clearly shows that state-dependent preferences may not always be able to perform the same "decision aiding" tasks that state-dependent utility functions are able to perform and so we would hesitate to treat the two concepts at par, at least in the context of decision aiding/making under uncertainty. Furthermore, the observation in (b) is a statement of fact, very correctly and succinctly expressed for the benefit of those like us, who may have limited knowledge of "mainstream bounded rationality" − theory and applications. Note that unlike the received theory of bounded rationality originating in the work of Herbert Simon, we focus on "lack of sufficient information" and not on "computational complexity" as our major concern for not being able to perform optimal decision making. Not being able to perform optimal decision making need not necessarily imply that the decision maker is not solving an optimization problem, as is mentioned in the last sentence of (b) above. The existing literature on bounded rationality focuses on behavioral issues related to computational constraints and complexity, which prevents individuals from solving optimization problems. We focus on the problems arising in "decision aiding" − the kind that technically qualified consultants may face − due to absence of sufficient information. Hence, although the decision aiding process involves optimization, the outcome of the process may turn out to be suboptimal, simply due to lack of available information.

Acknowledgment

Work on this paper began when I was with the School of Petroleum Management, PD Energy University (till June 5, 2022). I would like to thank Karl Schlag for comments on an earlier paper of which this paper is a considerably revised extension. The earlier paper was presented at the "Virtual Conference on Social Choice Theory and Organizations" on February 6, 2021, organized by Jac Heckelman (Wake Forest University, US), where I befitted immensely from comments of both Alexander Karpov and Elizabeth Maggie Penn. Those comments were an important reason for the revised extension discussed here. Thanks a lot, to all concerned, for their valuable inputs leading to this paper. I would also like to gratefully acknowledge comments on this paper and information about related research received from Itzhak Gilboa. His views about the solution concept available in this paper proved very useful in updating it. I would also like to thank Subhadip Chakraborty for valuable comments and suggesting corrections in the paper. As always, I wish to thank Professor Tadeusz Trzaskalik for his helpful comments and in particular for endorsing my attempts to re-orient "MCDM with weights". Finally, I wish to extend a very warm "thank you" to two very knowledgeable anonymous referees of this paper, particularly the one who forced me to take a fresh look at what is now a nonexistent and erroneous third theorem in a currently nonexistent section of the paper. Each and every comment from them was to say the least "very enlightening".

References

Campbell D., Kelly J., Qi S. (2018), *A Stability Property in Social Choice Theory*, International Journal of Economic Theory, 14, 85-95.

Congar R., Merlin V. (2012), *A Characterization of the Max-min Rule in the Context of Voting*, Theory and Decision, 72(1), 131-147.

Denicolo V. (1985), *Independent Social Choice Correspondences Are Dictatorial*, Economics Letters, 19(1), 9-12.

Endriss U. (2011), *Logic and Social Choice Theory* [in:] A. Gupta, J. van Benthem (eds.), *Logic and Philosophy Today*, College Publications.

Gilboa I. (1988), *A Combination of Expected Utility and Max-min Decision Criteria*, Journal of Mathematical Psychology, 32(4), 405-420.

Gilboa I., Schmeidler D. (1989), *Max-min Expected Utility with Non-Unique Prior*, Journal of Mathematical Economics, 18, 141-153.

Karni E. (1985), *Decision Making under Uncertainty: The Case of State-Dependent Preference*, Harvard University Press.

Kelsey D., Quiggin J. (1992), *Theories of Choice under Ignorance and Uncertainty*, Journal of Economic Surveys, 6(2), 133-153.

Lahiri S. (2020), *Generalized Sen-Coherence and Existence of Preferred with Probability at Least Half Winners*", International Journal of Operations Research, 17(3), 93-99.

Lahiri S. (2020/2021), *Extended Choice Correspondences & an Axiomatic Characterization of the Probabilistic Borda Rule*, Mathematical Methods in Economics and Finance − m^2ef, 16/17(1), 15-34.

Lahiri S. (2021), *Pattanaik's Axioms & the Existence of Winners Preferred with Probability at Least Half*, Operations Research & Decisions, 31(2), 109-122.

Maskin E. (1979), *Decision-Making under Ignorance with Implications for Social Choice*, Theory and Decision, 11, 319-337.

Pattanaik P.K. (1970), *Sufficient Conditions for the Existence of a Choice Set under Majority Voting*, Econometrica, 38(1), 165-170.

Pattanaik P.K. (1968), *Risk, Impersonality and Social Welfare Functions*, Journal of Political Economy, 76(6), 1152-1169.

Puppe C., Schlag K. (2009), *Choice under Complete Uncertainty when Outcome Spaces Are State Dependent*, Theory and Decision, 66(1), 1-16.