International Series in Operations Research & Management Science

Y. Ilker Topcu · Şule Önsel Ekici · Özgür Kabak · Emel Aktas · Özay Özaydın *Editors*

New Perspectives in Operations Research and Management Science

Essays in Honor of Fusun Ulengin

MOREMEDIA

Springer

International Series in Operations Research & Management Science

Volume 326

The book series **International Series in Operations Research and Management Science** encompasses the various areas of operations research and management science. Both theoretical and applied books are included. It describes current advances anywhere in the world that are at the cutting edge of the field. The series is aimed especially at researchers, advanced graduate students, and sophisticated practitioners.

The series features three types of books:

- Advanced expository books that extend and unify our understanding of particular areas.
- Research monographs that make substantial contributions to knowledge.
- Handbooks that define the new state of the art in particular areas. Each handbook will be edited by a leading authority in the area who will organize a team of experts on various aspects of the topic to write individual chapters. A handbook may emphasize expository surveys or completely new advances (either research or applications) or a combination of both.

The series emphasizes the following four areas:

Mathematical Programming: Including linear programming, integer programming, nonlinear programming, interior point methods, game theory, network optimization models, combinatorics, equilibrium programming, complementarity theory, multiobjective optimization, dynamic programming, stochastic programming, complexity theory, etc.

Applied Probability: Including queuing theory, simulation, renewal theory, Brownian motion and diffusion processes, decision analysis, Markov decision processes, reliability theory, forecasting, other stochastic processes motivated by applications, etc.

Production and Operations Management: Including inventory theory, production scheduling, capacity planning, facility location, supply chain management, distribution systems, materials requirements planning, just-in-time systems, flexible manufacturing systems, design of production lines, logistical planning, strategic issues, etc.

Applications of Operations Research and Management Science: Including telecommunications, health care, capital budgeting and finance, economics, marketing, public policy, military operations research, humanitarian relief and disaster mitigation, service operations, transportation systems, etc.

This book series is indexed in Scopus.

Y. Ilker Topcu • Şule Önsel Ekici • Özgür Kabak •
Emel Aktas • Özay Özaydın
Editors

New Perspectives in Operations Research and Management Science

Essays in Honor of Fusun Ulengin

Springer

Editors
Y. Ilker Topcu
Department of Industrial Engineering
Istanbul Technical University
Macka, Turkey

Özgür Kabak
Department of Industrial Engineering
Istanbul Technical University
Macka, Turkey

Özay Özaydın
Department of Industrial Engineering
Doğuş University
Dudullu, Turkey

Şule Önsel Ekici
Department of Industrial Engineering
İstanbul University-Cerrahpaşa
Avcılar, Turkey

Emel Aktas
School of Management
Cranfield University
Cranfield, United Kingdom

ISSN 0884-8289 ISSN 2214-7934 (electronic)
International Series in Operations Research & Management Science
ISBN 978-3-030-91850-7 ISBN 978-3-030-91851-4 (eBook)
https://doi.org/10.1007/978-3-030-91851-4

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

The first elective Prof. Dr. Fusun Ulengin proposed and started at Istanbul Technical University, Industrial Engineering Department, where she began to work as an Assistant Professor in the Fall of 1991, was "New Paradigms in Operations Research." That is why this book you are reading is inspired by this first elective, which has a special meaning for all of us. Some of us has been a teaching assistant to the New Paradigms in Operations Research elective, some of us studied with excitement during our graduate studies and expanded our horizons. The name of her first elective also reflects Dr. Ulengin's attitude towards her profession: Never being content with the existing achievements, always seeking innovations, and extending the state of the art. This search for new knowledge is not on her own but together with her colleagues. She seeks new knowledge with such hunger and curiosity; you also join in with her if you are one of the lucky people around her. She is a super-hardworking, extremely productive, always guiding and directing teacher, and a *friend who is always there with you on your good and bad days.*

Dr. Ulengin has touched the lives of many students and directed and motivated them to perform their profession to the best of their ability since we first met her in the first years of her career. She has continuously and efficiently contributed to knowledge while never losing her touch with the practice, developing solutions to real-life problems, being a role model for making a difference. We are so fortunate to be part of her team, and we are grateful to be her first students, then colleagues, and permanent members of her "extended family."

This book is an expression of our gratitude . . .

Y. Ilker Topcu
Şule Önsel Ekici
Özgür Kabak
Emel Aktas
Özay Özaydın

Preface

The roots of operations research and management science (OR/MS) go back to World War II when the British army benefitted from scientific methods to make decisions. Since World War II, OR/MS tools have been applied to the problems of many industries, including manufacturing, logistics, transportation, retail, finance, banking, energy, and others. Beginning with basic mathematical programming and statistics, OR/MS tools have evolved over the decades into many different techniques with new developments in science and fresh perspectives provided by scientists and practitioners. In this book, we aimed to present some contemporary OR/MS techniques and their recent applications. We are happy to see that these studies also fall in the research interests of Prof. Füsun Ulengin, for whom this book is dedicated.

The book starts with an application of a mathematical programming method to a fair assignment of debt cases to legal debt collection centers. Similar applications of mathematical programming models in real-life cases have increased in recent years mainly because of the significant software improvements that enable solving these models efficiently.

The second chapter presents a theoretical contribution to optimization theory. It introduces Lagrangian duality in an accessible way for less mathematically oriented readers and explains its main ideas. The chapter also provides an overview of the applications of the Lagrangian duality to problems in transportation and location.

The following three chapters use Data Envelopment Analysis (DEA) in different areas. This is a natural result of the increased interest in DEA in recent years. In the third chapter, DEA and Malmquist Index are applied to measure foreign trade-logistics efficiency. The fourth chapter presents an application of DEA in transportation. The authors develop a performance index for green transportation corridors using a DEA approach. Changing the context in the fifth chapter, the efficiency of nations is measured using DEA and random forest classification, integrating DEA and a machine learning algorithm.

Over the past decade, we have seen a prolific increase of machine learning approaches in the OR/MS field. Therefore, the fifth chapter is followed by three more studies that use machine learning tools. The sixth chapter explores the effectiveness of online marketing communication for healthcare services in Turkey using random forest, gradient boosting, and decision tree approaches. The seventh chapter investigates the third wave of feminism through hierarchical clustering and sentiment analysis. The eighth chapter compares forecasting methods such as Holt-Winters, ARIMA, and ARIMAX, to predict the evolution of COVID-19 figures.

Statistical methods have an important place in operations research. Not surprisingly, four chapters using several statistical methods are included in this book. Two of them are related to structural equation modeling (SEM). The ninth chapter explores the impact of dynamic innovation capabilities on firm performance. Four innovation capabilities and their effect on profitability, performance, and growth of companies in the food industry are examined using the PLS-SEM method. The tenth chapter investigates the relationships between e-service quality, information quality, e-satisfaction, and e-repurchase intention. The research model is tested by using SEM.

In the eleventh chapter, the results of a European survey on OR/MS Education are presented. Statistical analysis is performed to understand the relationship between respondents' perceptions ("Positive" and "Negative") and the types of OR/MS modules that were lectured at their institutions. The twelfth chapter analyzes the effect of an increase in the workforces share of the skilled labor on the skill wage premium in Turkey using multivariate regression analysis.

Multiple attribute decision-making (MADM) is an integral part of the OR/MS research. We included seven chapters using MADM methods in this book. The thirteenth chapter identifies the status value ascribed to organizations within socially constructed systems of norms and values. The Best-Worst Method, along with regression analysis, is used for this purpose. In the fourteenth chapter, the Analytic Hierarchy Process (AHP) and TOPSIS methods are integrated to select hazardous materials transportation routes. Alternative routes are generated by using the ArcGIS network analysis tool. Dealing with a current problem, in the fifteenth chapter, MADM methods based on Picture and Spherical Fuzzy Sets are developed to support strategic decisions of states during pandemics. The sixteenth chapter develops a methodology to evaluate suitable locations for a new hospital. The authors applied a hybrid MADM model that consists of the Analytic Network Process (ANP) for finding the importance of the attributes and PROMETHEE for evaluating the alternatives.

The seventeenth chapter uses Stepwise Weight Assessment Ratio Analysis (SWARA) method to prioritize the activities of technology transfer offices. A group decision-making approach is used to consider the views of multiple stakeholders. The eighteenth chapter develops the picture fuzzy set extension of the DEMATEL (DEcision-MAking Trial and Evaluation Laboratory) method to incorporate experts' hesitancy and refusal degrees into the analysis. The new method is applied to evaluate education quality.

Unlike other MADM chapters, the nineteenth chapter addresses the contingent dimensions (content and context) in multi-criteria decision-making. It explores the extent to which contingent factors are addressed in the literature. A sample of

46 MADM group decision-making papers from a single year of publication is examined regarding the “group decision-making” dimension.

The final three chapters touch on different OR/MS topics. The twentieth chapter analyzes the interaction between renewable energy penetration and the wealth of nations. For this purpose, the data provided by The World Bank and The International Renewable Energy Agency (IRENA) are analyzed by using Bayesian Nets.

The twenty-first chapter discusses the impact of logistics management in food supply chains based on two case studies: Soybean and Beef production. Social Network Analysis is used to understand the relationship among the countries involved in the related supply chains.

The book concludes with a philosophical discussion. The twenty-second chapter aims to enlighten the future of the global economy in general and transportation in particular, underpinning the revolutionary transformation we are going through.

This book is a collection of 22 studies related to recent developments and applications in OR/MS. As mentioned above, many OR/MS tools, from mathematical programming to DEA, machine learning, and MADM methods to social network analysis, are applied to various areas such as debt collection, transportation, location, education, energy, human development, pandemics, and so on. This shows the richness of OR/MS domain in terms of tools and application areas. It gives great hope for future developments to those of us who dedicate our lives to application and advancement of OR/MS. In this respect, we would like to thank our teacher, supervisor, and role model, Prof. Füsun Ülengin once again, who kindled the OR/MS fire in us.

There are also other special people to whom we owe appreciation. We would like to thank all the authors for their contributions to making this book possible. They all worked hard to ensure that their chapters were of good quality and delivered submissions on time. We also express our sincere gratitude to the referees, nameless heroes, for their invaluable comments that improved the quality of the chapters.

May the force of OR/MS be with you!

Macka, Turkey — Y. İlker Topcu
Avcılar, Turkey — Şule Önsel Ekici
Macka, Turkey — Özgür Kabak
Cranfield, UK — Emel Aktas
Dudullu, Turkey — Özay Özaydın

Contents

Fair Assignment of Debt Cases to Legal Debt Collection Centers

Elmira Farrokhizadeh, Sara Ghazanfari Khameneh, Nilüfer Altinok, Ahmet Tezcan Tekin, Basar Oztaysi, Özgür Kabak, Sezi Çevik Onar, Aslıhan Ceren Tari, Özge Gencer Aydemir, and Ersan Öztürk

Abstract Nowadays, unpaid invoices are major concerns in many industries or service companies that seriously affect firms' cash flows, credit ratings, and firms' image. To overcome the financial problems caused by unpaid debts, managing and controlling these debts is a vital task for companies' financial survival. Unpaid receivables are collected by the company or debt collection agency as a third party. The debt collection process consists of the collection phase and the legal phase. When the debt collectors are unsuccessful in obtaining the long-term unpaid receivables in the collection phase; the debts will be transferred to the legal debt collection agencies as the legal phase and carried out by experienced lawyers. In order to initiate the legal phase, the unpaid receivables, the so-called cases, are assigned legal collection agencies based on the features of both cases and agencies. In this study, a fair assignment model is built to allocate the customer cases to legal debt collection agencies to prevent company from increasing unpaid debts and manage the legal debt collection process effectively. This model maximizes the advantage of assigning in the same city or region considering the cost, capacity, and equality of each legal debt collection agency in getting the outstanding debts. Budget of the agencies and the number of the cases in the same city are balanced to have a fair assignment. The model is validated by an application in a big service company in Turkey.

Keywords Optimization problem · Debt collection · Legal debt collection · Assignment model · Fair assignment model

E. Farrokhizadeh · S. Ghazanfari Khameneh · N. Altinok · B. Oztaysi · Ö. Kabak (✉) · S. Çevik Onar · A. C. Tari · Ö. Gencer Aydemir · E. Öztürk
Industrial Engineering Department, Istanbul Technical University, Istanbul, Turkey
e-mail: oztaysib@itu.edu.tr; kabak@itu.edu.tr; cevikse@itu.edu.tr

A. T. Tekin
Management Engineering Department, Istanbul Technical University, Istanbul, Turkey

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_1

1 Introduction

Debt collection is one of the major problems of all organizations that significantly impacts their economic situation and survival in the competitive world. Debt collection is a set of deliberate attempts to compensate a debtor's overdue receivables in case of unpaid invoices or credit. Due to the current worldwide recession and the weakening people's economic power, these types of unpaid debts increase and significantly impact cash flows, turnover, credit ratings, and even organizations' credibility (S. C. Onar et al., 2018a, b). Accordingly, determining the most critical factors of succeeding in debt collection is very important in unpaid debts. Taneta- Skwiercz (2018) defines the debt collection process as a multidimensional phenomenon that is presented by the legal, economical, and psych sociological perspectives that the most significant changes should be made in the legal field. In practice, debts are collected by either the debt owner itself (using its name and own employees) or a third party as the debt collection agency that works for the creditor firm. In many countries, collection fees are added to the original debt as debt recovery cost and the debtor must pay it. In the debt collection process, the collector tries to collect both the debtor's debt and additional costs.

Generally, the debt collection process has two phases: The collection phase and the legal phase. After logging in to the debt collection system, the debtor enters the collection phase where collectors either own company or a third party tries to persuade debtors to repay the recovery cost consist of original debt plus the collection fee via letters, emails, and phone calls. At this stage, the collectors negotiate with the debtors and persuade them to collect their debts by giving a repayment plan. To successfully collect the debt at the maximum level, the human resources who provide communication with the customer should be used efficiently. It is necessary to communicate with the customer through the most appropriate channel with minimum cost. Achieving these goals requires fast and accurate actions to be taken in the light of the evaluation results after the correct processing, analysis, and evaluation of customer data. This stage can take from a few days to several months. Suppose the debt collectors are unsuccessful in obtaining the long-term unpaid debts in the collection phase, the receivables will be written off (i.e., the debtor has died or gone bankrupt) or transferred to the legal debt collection agencies as the legal stage that carried out by experienced lawyers. Due to the uncertain and expensive nature (this step requires a legal evaluation by a specialist) of the legal phase, both the debtor and collector prefer that the debt is repaid before this phase. Usually, large enterprises work with contractually legal debt collection agencies rather than expand internal resources. According to the impact of that legal debt collection agency's performance and experience on the closure or non-closure of debt cases, the agency's evaluation and optimal allocation of relevant cases to appropriate agencies are crucial and debatable issues (S. C. Onar et al., 2020).

Assignment Problem (AP) is a special type of linear programming problem (Kuhn, 1955). AP is a type of transportation problem to assign n tasks to m agents with minimum cost. Each task is assigned to precisely one agent with a set of capabilities for each agent and special requirements for each task (Series, 2019). Over the last decades, different types of classical APs are developed, such as Stochastic AP, Fuzzy AP, Bottleneck AP, Generalized AP, quadratic AP, etc. (Kumar & Gupta, 2011). Stochastic AP is based on the binary vector to determine agent m to assign to task n. Fuzzy AP is a linear problem that objective function, constrains coefficient, right-hand-side values or all of them can be fuzzy set instead of the crisp number to handle the nature of uncertainty in problems. In literature, many papers on APs have been reported in different aspects and applications. Related to the current study, Boah et al. (2015) applied AP to solve a problem for a legal firm in Kumasi, which had difficulty in assigning nine different cases to its nine junior lawyers.

In this chapter, a big service company that has millions of customers will be considered. Due to the confidentiality agreement, the name and the sector of the company cannot be pronounced. Every month this company faces thousands of customers who have to pay their debts. The legal debt collection agencies have emerged as a new business line to track these bills. In the administrative follow-up process, companies remind the debtor and warn the customer to pay their debt, but if clients do not pay their debts during this period, the legal process begins for these clients. This process involves a court process that must be done by experienced lawyers. When the legal process begins, the client must either pay the entire debt or go to court, which can penalize the debtors that have not paid their debts. Legal action is a process that specialized lawyers must carry out and it is essential to consider which case is assigned to which legal debt collection agency.

The company has contracts with a plenty of legal debt collection agencies that must be assigned thousands of new debt cases every month. All cases are assigned with a particular algorithm with some specific rules in the current working system of the company. The goal of this study is to allocate unpaid debts to legal debt collection agencies that maximize the company's objectives by using data stored in the company's databases. We build a fair assignment model to allocate the customers' debt cases to legal debt collection agencies by considering each agency's place, capacity, and equality. In this study, the unpaid case assignment priority to a legal debt collection agency is ordered. The agencies must be placed; 1—in the same city and 2—in the same region and otherwise can be assigned to other cities in Turkey. Therefore, to apply the fair assignment model's equality, two permitted deviation parameters are defined to prevent the unbalanced distribution of the cases to the legal debt collection agencies. This study aims to ensure that the files are assigned to the correct debt collection agencies, increases the firm's total collection, improves the agencies' performance in the medium term, and finally, decreases the courts' burden by preventing consumer victimization.

The rest of this chapter is structured as follows. Section 2 reviews previous studies related to assignment problems and legal debt collection. Section 3 develops the fair assignment model to assign debt cases to legal debt collection agencies. Section 4 presents the application to represent the applicability of the model. In this section, the case study is defined, parameters of the model are calculated, and finally, the results are explained. The last section concludes the model and gives further suggestions for future studies.

2 Literature Review

The aim of the study is to build a fair assignment model to allocate the customer debt files to legal debt collection agencies. For this, initially, we investigate the literature in two sub-sections: assignment problems and legal debt collection agencies.

2.1 Literature Review of Assignment Problems

The assignment problem is one of the main topics in the decision-making area. There are some different aspects to this problem. The classic assignment problem is designed to assign n agents to m task. There is a plenty of ways to solve this problem such as genetic algorithm (Younas et al., 2011). Wilson (1997) suggested a genetic algorithm for the assignment problem that provides good results with an accuracy less than 0.01% from the optimal solution. Liu et al. (2015) suggested an artificial bee colony algorithm that has been applied to assign the dynamic task to multi-agents. In this algorithm, each bee is connected to the number of agents and tasks, and the employed bee number is equal to onlooker bees. The result shows the similarity to the genetic algorithm as well.

Some algorithms solve assignment problems by assigning the tasks to the agents step by step such as the Hungarian algorithm by considering their cost to achieve the lowest cost. For instance, Mills-tettey and Stentz (2007) applied this algorithm based on the proposed cost and change it step by step to achieve the new optimal solution with lower cost. The transportation problem has been applied to illustrate this algorithm. Al-saeedi and Shiker (2020) have modified the Hungarian algorithm and they claim that this new algorithm proceeds based on a more efficient application with less time consuming and easier to use. Pala (2020) proposed the Hungarian algorithm in different aspects. They consider a more efficient way to achieve the optimal solution even if there are big n or m.

As reviewed by Pentico (2007), there are different types of the assignment problem. In the classical assignment problem, one task is assigned to each agent. In the k-cardinality assignment problem, only k tasks will be assigned to each agent. Another type is a bottleneck which is based on Min-Max. In the balance assignment problem, the difference between the minimum and maximum cost will be balanced. Assignment problems can be applied in multi-criteria or fractional nonlinear programming or one side constraint in linear programming.

Munkres (1957) created the algorithm of generalized assignment problem which has been introduced based on transportation problem. Some discussion and proof of algorithm proposed as well to make a better understanding of this algorithm. In some assignment problems more than one task are assigned to an agent (Gavish & Pirkul, 1991). In this article some relaxation of the problem considered and a heuristic solution is applied as an efficient procedure. The generalized assignment problem is an NP-hard optimization problem.

Wilson (1997) considered an assignment problem based on a linear program for polynomial-time algorithm when each coefficient increases linearly. It is aimed to decrease possible processing time when the total cost is increasing. Furthermore, Biggerstaff et al. (1993) concluded that if assignment problems would be solved with a plausible reasoning component. Basirzadeh (2012) proposed two mathematical models for assigning the task to labor with different skills. The model shows how it affects cost optimization when we assign the right task to the right labor. Also, a sensitivity analysis was applied. Some other literature concentrates on different linear, multi-dimension, and quadratic problems based on assignment problems (Krokhmal & Pardalos, 2009).

Pentico (2007) had reviewed mathematical programs to solve the assignment problem in the literature. Salehi (2014) proposed a MIP formulation with interval parameters. In order solve the problem efficiently, it is converted to a crisp linear program by applying the substitution variables approach. Min-max approach is applied as well to transform the multi-objective problem into a single objective program. Mills-tettey and Stentz (2007) proposed a simple and flexible algorithm that finds the results in competitive times. In the proposed algorithm Afroz and Hossen (2017) conclude some easy calculations to make some changes in the assignment problem which helped to propose easier and more efficient results. This method has been compared with the common methods to prove its efficiency.

There are different applications for assignment problem based on the literature. One of them is airport assignment problem. Bouras et al. (2014) built different solutions to achieve the optimal solution for airport assignment problem such as exact, heuristic, and metaheuristic algorithm. One of the assignment problems is the locomotive assignment problem (Piu & Speranza, 2014). This is about the set of trains which is needed to be satisfied by considering the budget constraint. Ho and Vaughan (2012) applied a task assignment problem for crowdsourcing markets such as Amazon. The online task assignment problem is formulated to maximize the benefit when each worker will be allocated to one of the tasks upon arrival. This algorithm is different from offline allocation hence there is an unknown skill level of each worker. It is proposed that this algorithm works better than random and greedy algorithms. Küçükmatçı & Acar (2019) worked on assignment problem based on commercial general proposed and Tabu search based on bottleneck assignment problem. It is based on minimizing assignment of the costly tasks to the agents. Tasks may be assign separately based on subgroup with less complexity. They merge these separated tasks to achieve more feasible solution. Tabu search method is one of the best facilities to work on multi-objective assignment problems. Munaam and Hammadi (2017) worked on some problems to measure the algorithm effectiveness in three objective sizes with applying annealing algorithm and genetic algorithm.

2.2 Literature Review of Legal Debt Collection Agencies

More than half a century ago, researchers began to optimize the debt collection process. Mitchner and Peterson (1957) proposed an optimal stopping problem to the debt collection model for Bank of America to optimize the cost and length of collector follow-up. A few years later, a simple Markov decision process and optimized credit control policy is proposed (Liebman, 1972). Abe et al. (2010) constructed a constrained Markov Decision Process for the debt collection process that considers legal, business, and resource constraints. Miller et al. (2012) used Markov Decision Process to optimize New York State's tax collection with the goal of maximizing long-term returns. A dynamic programming approach to optimization of consumer lending debt collection is proposed that the monthly decision depends on the action take in the month to come (De Almeida Filho et al., 2010). Geer et al. (2018) developed a Markov decision model for data-driven scheduling of outbound calls made by both own debt collectors and third-party debt collectors. They optimized the legal phase of debt collection procedure by maximizing the recovered debt during the collection phase. Duman et al. (2017) presented a novel optimization model to debt collection with considering the incremental churn effects of actions, capacity constraints, and path optimization for the banking sector. Antoine and Abdallah (2006) used the Markovian method to find the optimal collection dates and improve the debt collection procedure due to the financial and lost risks. Artificial Neural Networks to determine each debt collection center's score by considering resource limitation is presented (Georgopoulos & Giannaropoulos, 2007).

Before deciding on a debt collection agency, it is required to assess the agencies. Many companies are focused on money and are only interested in whether the debt is collected or not. However, there are many different levels of performance for a debt collection agency. Factors such as relationship, service, collection rate, collection time, collection fees, value-added data, flexibility, transparency, and brand representation are used to evaluate debt collection institutions (Stewart, 2017). In the literature, performance is defined according to a predefined definition that occurs after a completed activity and depends on time and place, vision and strategy, and forecasting future performance. The purpose of determining agency's performance is to measure objectively and periodically to what extent the agency achieve their goals before assigning the debt cases to these offices. An organization's performance measurement can be used to evaluate inputs, internal processes, procedures, outputs, and results. Therefore, by analyzing indicators, the right decision can be made, limited resources that can create continuous improvement can be effectively developed and targets can be given to individuals. In the literature, multi-criteria decision-making techniques such as Cognitive Maps, Regression Analysis and Artificial Neural Networks, and Analytical Hierarchy processes are used in performance evaluation (S. Ç. Onar et al., 2018a, b). Several studies in the literature focus on the performance of legal debt collection agencies. Among these studies, the AHP method is applied to assess legal debt collection agencies' performance (Brock et al., 2006). In another study, researchers determined the technical performance of major agencies in the US with regression analysis and investigated the correlation of the number of lawyers, income, support staff and partners, with legal debt collection agency performance (W. Wang, 2000). Boah et al. (2015) applied the assignment problem to model the optimal assignment of lawyer's cases with the goal of minimizing the total time elapsed for the case. They applied the assignment problem concept to solve a problem that was difficult to allocate nine different cases to nine junior lawyers of a law firm in Kumasi. S. Ç. Onar et al. (2018a, b) found the optimal assignment in legal debt collection agencies by evaluated their performances by Pythagorean fuzzy AHP. In other work, Hesitant Pythagorean (Intuitionistic type 2) fuzzy AHP is proposed to determine the performance of legal debt collection agencies under hesitancy and vagueness environment (S. C. Onar et al., 2020).

Han and Jang (2013) proposed the Loss Given Default (LGD) model, which assesses the necessary steps to repay the debt in legal and internal debt collection. The success of debt collection largely depends on two factors: the customer's willingness to pay and the agent's negotiation and communication skills. Determining the appropriate data sets to provide a valid answer to the research question is one of the popular aspects of the debt collection problem investigated in the literature. Ho Ha and Krishnan (2012) developed Cox proportional hazard analysis to recognize the credit estimation repayment pattern. Several studies in the literature have a hybrid model that combines the predictive analysis and assignment problem. Kim and Kang (2016) developed a predictive model for fair allocation of customer contact lists to call center agents in the debt collection procedure. They evaluated customers' data using five machine-learning-based payment prediction models under ten customer scoring rules and assigned the customer to call centers by snake draft method. Chen et al. (2013) combined the model with association rules, clustering, and decision tree methods to predict the customer's payment behavior before assignment. L. Wang et al. (2019) create the decision tree for a behavioral customer scoring model to minimize the bad debt in Taiwan's telecommunications market. A behavioral scoring model to evaluate the credit cardholders' repayment behavior is proposed in the literature. First, he classified factors by Chi-square automatic interaction detector and artificial neural networks and then applies these factors in the second stage to construct a DEA model (I-Fei, 2010). Takahashi and Tsuda (2013) found that location and monetary value of the transaction are two important characteristics of bad debtor by random forest approach.

Vecchio et al. (2006) determined the correct address of bad debtors with different addresses by the Hierarchical Clustering method Levenshtein Distance in MS Excel Macro. Rule-Based Decision engines were used to optimize the debt collection procedure (Chin & Kotak, 2006). In the other study, three different debts (loan debts, credit card debts, and invoices) are investigated by the Fuzzy Inference System method with using the amount of loan, the wealth of debtor's historical data, amount of other debts, be an active customer or not, credit limit, and criticality as inputs (S. C. Onar et al., 2018a, b).

According to the articles reviewed in this section, some studies considered optimizing the debt collection process or predicting the performance of customers or debt collection agencies by machine learning method, but in this article, we try to build a fair assignment model to assign the debt cases to legal debt collection agencies.

3 Problem Definition

The problem of optimizing the legal phase of the debt collection process is formulated as a fair assignment problem in this study. Our model is suited for legal debt collection problems in which the unpaid cases enter to legal phase with the contracted expert legal debt collection agencies. The proposed company is the big service company in Turkey that wants to optimize the legal debt collection to maximize the total benefit of assigning the cases in the same city and establish justice and equality between the contracted legal debt collection agencies. The model has settled equality by minimizing each city's deviation of the average budget and assigning rate. This company has several contracted legal debt collection agencies. The unpaid case assignment priority to an agency is ordered as the case and agency must be placed; 1—in the same city and 2—in the same region and otherwise can be assigned to other cities in Turkey. Thousands of cases are assigned to these legal debt collection agencies from all over Turkey every month. Assumptions of the problem are listed as follows:

- Each debt case can be assigned to only one legal debt collection agency.
- Each legal debt collection agency has limited capacity.
- To achieve a fair assignment, the average budget of assigned cases to legal debt collection agencies must be close to each other.
- For legal debt collection agencies in the same city, the city's percentage of debt cases to all assigned debt cases should be close to each other.

Parameters and variables of the fair assignment model that optimizes the legal debt collection under these assumptions and limitations are represented as follows:

Indices:	
i	Set of debt cases $i \in \{1, 2, \ldots, I\}$
j	Set of legal debt collection agencies $j \in \{1, 2, \ldots, J\}$
c	*Set of cities* $c \in \{1, 2, \ldots, C\}$
r	*Set of regions* $r \in \{1, 2, \ldots, R\}$
Parameters (input):	
fc_{ic}	If case i is placed in city c, equal to 1; otherwise, 0
b_i	Budget value of case i
oc_{jc}	If legal debt collection agency j is placed in city c, equal to 1; otherwise, 0
cap_j	The capacity of legal debt collection agency j
re_{cr}	If city c is placed in region r, equal to 1; otherwise, 0
ba_c	If cases of city c are required to distributed evenly, equal to 1; otherwise, 0
fr_{ir}	If case i is placed in region r, equal to 1; otherwise, 0
or_{jr}	If office j is placed in region r, equal to 1; otherwise, 0
sd	The permitted deviation of assigning case to offices that placed in the same city
sd'	Permitted deviation from the average budget
Parameters (Calculated based on input parameters or assumed for the IP formulation)	
A_{ij}	Objective function coefficient for assigning case i to office j
$\overline{b}$	Average of budgets of the cases
no_c	Total number of offices in city c
cr_c	Ratio of cases in city c to capacity of legal debt collection agencies in city c
u_j	Upper bound of assigning cases to same city of legal debt collection agency j
l_j	Lower bound of assigning cases to same city of legal debt collection agency j
na_{ij}	If case i cannot be assigned to office j, equal to 1; otherwise 0
d_c	Rate of cases in city c, where don't have contracted legal debt collection agency
ul_{jc}	Upper capacity limit of legal debt collection agency j in city c that must be balanced assigned
Decision Variables:	
x_{ij}	If debt case i is assigned to legal debt collection agency j equal to 1, otherwise 0

Based on these parameters and limitations, our fair assignment model in legal debt collection is represented as follows:

$$\text{Max} \sum_{i \in \{I \mid na_{ij}=0\}} \sum_{j \in \{J \mid na_{ij}=0\}} A_{ij} \times x_{ij} \tag{1}$$

Subject to:

$$\sum_{j \in \{J \mid na_{ij}=0\}} x_{ij} = 1 \qquad \forall i \tag{2}$$

$$\sum_{i \in \{I \mid na_{ij}=0\}} x_{ij} \le cap_j \qquad \forall j \tag{3}$$

$$\sum_{i\in\{I|na_{ij}=0\}}\left(x_{ij}\times b_i\right)-\overline{b}\times\sum_{i\in\{I|na_{ij}=0\}}x_{ij}\ \le sd'\times\overline{b}\times\sum_{i\in\{I|na_{ij}=0\}}x_{ij}\qquad\forall j \tag{4}$$

$$-\sum_{i\in\{I|na_{ij}=0\}}\left(x_{ij}\times b_i\right)+\overline{b}\times\sum_{i\in\{I|na_{ij}=0\}}x_{ij}\ \le sd'\times\overline{b}\times\sum_{i\in\{I|na_{ij}=0\}}x_{ij}\qquad\forall j \tag{5}$$

$$\sum_{i\in\{I|na_{ij}=0\}}\sum_{c}\left(x_{ij}\times fc_{ic}\times oc_{jc}\right)\le u_j\qquad\forall j \tag{6}$$

$$\sum_{i\in\{I|ay_{ij}=0\}}\sum_{c}\left(x_{ij}\times fc_{ic}\times oc_{jc}\right)\ge l_j\qquad\forall j \tag{7}$$

$$\sum_{i\in\{I|na_{ij}=0,fc_{ic}=1\}}x_{ij}\le ul_{jc}\qquad\forall j,c\in\left\{J,C|ba_c=1,oc_{jc}\ne 1\right\} \tag{8}$$

$$x_{ij}\in\{0,1\}\qquad\forall i,j \tag{9}$$

The objective of the proposed assignment model is to maximize the advantage of assigning the debt cases to legal debt collection agencies placed in the same city and in the same region. Constraint (2) guarantees that each customer's debt case must assign to only one legal debt collection agency. Constraint (3) ensures that the total of assigned customer's debt cases to each legal debt collection agency cannot exceed its capacity. Constraints (4–8) state equality and fair assignment constraints. Limitations (4 and 5) ensure that the total assigned debt cases' budget can be as much or less than the average budget's permitted deviation. Constraints (6 and 7) show that the total assigned debt cases to each legal debt collection agency must be less or equal than the upper bound and greater or equal than the lower bound of each office. Also, limitation (8) guarantees the equality rate of assigned debt cases placed in cities that haven't legal debt collection agencies. And finally, limitation (9) represents that x_{ij} be a binary variable that takes 0 or 1 values.

4 A Case Study

An illustrative example of the company is presented in this section to validate the fair assignment model's applicability. As mentioned previously, the company is one of the biggest companies in the service sector in Turkey with millions of customers. In this case study, monthly information about the company's debt cases and contracted legal debt collection agencies are available. Due to the confidentiality agreement, some private information has been masked and manipulated, provided that it does not disturb the general structure of the data. In this section, we first give information related to the case data and then explain how the given data is processed for the application. Finally, the application results are provided.

4.1 The Case Data

The data provided here are related to a specific month in 2020. The company has 10,000 debt cases all over Turkey that are required to be distributed to 80 contracted legal debt collection agencies. These cases must be assigned to legal debt collection agencies with our fair assignment model and under the company's limitations. Debt cases come from the whole of Turkey, consisting of 81 cities and 7 regions represented in Fig. 1. It shows the total number of cities, legal debt collection agencies, and debt cases of each region too. Therefore, in the model, the c and r indexes are defined as $c \in \{1, 2, \ldots, 81\}$ and $r \in \{1, 2, \ldots, 7\}$ to represent the set of debt cities and regions of Turkey, respectively. The total budget of the company's debt cases from the whole of Turkey is calculated as 8,304,468 TL (Turkish lira).

Information about each legal debt collection agency's places, capacity, and total debts in each city is represented in Table 1. As you can see in the table, there are 7 regions and 81 cities in the whole of Turkey some of which have legal debt collection agencies. Information about the dispersal of legal debt collection agencies and debt cases is shown in Figs. 2 and 3. Figure 2 represents ten cities with the most debt cases and explains that approximately half of the debt cases belong to Istanbul as the most populous city in Turkey, with over 15 million people constituting 18.4% of the total population. The other two cities that have a large number of debt cases are Ankara and Izmir as second and third most populous provinces. The total number of contracted legal debt collection agencies for each region is defined in Fig. 1. It explains that Marmara and Central Anatolia have the most contracted legal debt collection agencies. In Fig. 3 total numbers of debt cases are compared with the total capacity of these agencies in each region to show the lack or excess of legal debt collection agencies in contrast to debt cases frequency. It explains that the highest number of offices and highest number of case capacity of the offices are in the Marmara region. This is because of Istanbul that alone has 32 legal debt collection agencies and 5071 case capacity. The capacity of Marmara is much more than the cases in Marmara. Therefore, the legal debt collection agencies in

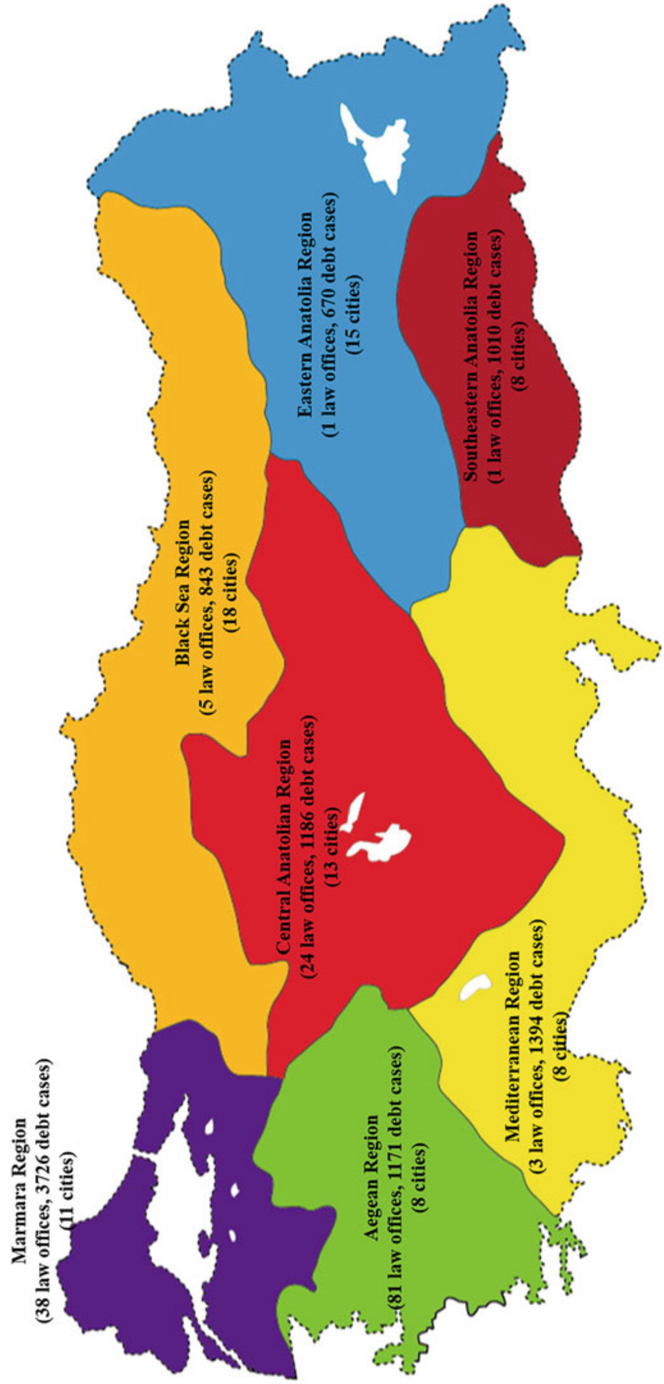


Fig. 1 Regions of Turkey

Table 1 Information about each legal debt collection agencies and debt cases in cities

Region	City	Number of debt cases	Total debt budget	Number of legal debt collection agency (no_c)	Total legal debt collection agency's capacity	Region	City	Number of debt cases	Total debt budget	Number of legal debt collection agency (no_c)	Total legal debt collection agency's capacity
1	1	288	207,793	0	0	5	6	488	408,835	21	2081
	7	338	328,018	1	189		18	20	10,890	0	0
	15	23	18,389	0	0		26	95	86,933	1	230
	31	228	184,329	1	57		38	109	81,892	1	91
	32	54	42,721	0	0		40	15	12,218	0	0
	33	270	214,250	0	0		42	189	152,860	1	22
	46	137	112,688	0	0		50	29	20,624	0	0
	80	56	46,942	1	88		51	33	20,671	0	0
2	4	66	45,706	0	0		58	82	57,453	0	0
	12	16	9494	0	0		66	41	21,593	0	0
	13	25	17,341	0	0		68	39	24,512	0	0
	23	50	47,768	0	0		70	21	15,532	0	0
	24	29	24,269	0	0		71	25	15,713	0	0
	25	85	59,777	1	145	6	5	28	20,148	0	0
	30	29	22,452	0	0		8	23	23,441	0	0
	36	40	24,808	0	0		14	32	25,194	0	0
	44	103	78,683	0	0		19	52	42,756	1	32
	49	34	23,631	0	0		28	52	40,368	0	0
	62	9	5662	0	0		29	12	7164	0	0
	65	113	91,135	0	0		37	31	26,312	1	22
	73	22	27,132	0	0		52	78	60,622	0	0

(continued)

Table 1 (continued)

Region	City	Number of debt cases	Total debt budget	Number of legal debt collection agency (no_c)	Total legal debt collection agency's capacity	Region	City	Number of debt cases	Total debt budget	Number of legal debt collection agency (no_c)	Total legal debt collection agency's capacity
	75	23	15,949	0	0		53	39	30,682	0	0
	76	26	23,448	0	0		55	144	98,845	1	101
3	3	65	49,451	0	0		57	19	16,966	0	0
	9	119	107,983	1	69		60	57	51,609	0	0
	20	74	60,500	0	0		61	81	56,375	2	196
	35	562	458,782	6	659		67	69	52,248	0	0
	43	57	49,283	0	0		69	4	2829	0	0
	45	139	108,098	0	0		74	28	19,724	0	0
	48	120	107,948	0	0		78	18	12,050	0	0
	64	35	37,502	1	22		81	76	72,518	0	0
4	2	48	29,238	0	0	7	10	138	102,589	1	22
	21	285	195,595	0	0		11	18	13,953	0	0
	27	319	244,279	1	85		16	388	298,661	3	574
	47	52	45,753	0	0		17	58	50,843	0	0
	56	19	13,923	0	0		22	57	45,781	1	22
	63	231	176,325	0	0		34	2438	2,237,789	32	5071
	72	38	36,796	0	0		39	47	32,148	0	0
	79	18	9673	0	0		41	222	231,777	1	378
							54	149	127,508	0	0
							59	163	115,768	0	0
							77	48	54,560	0	0

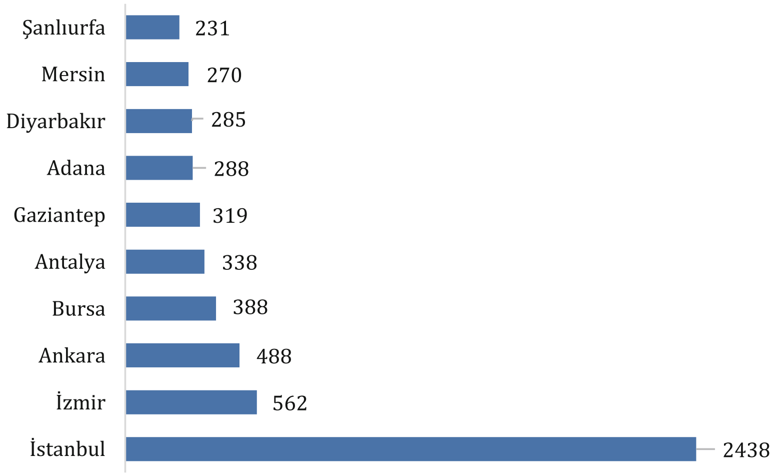


Fig. 2 Top 10 cities with large number of debt cases

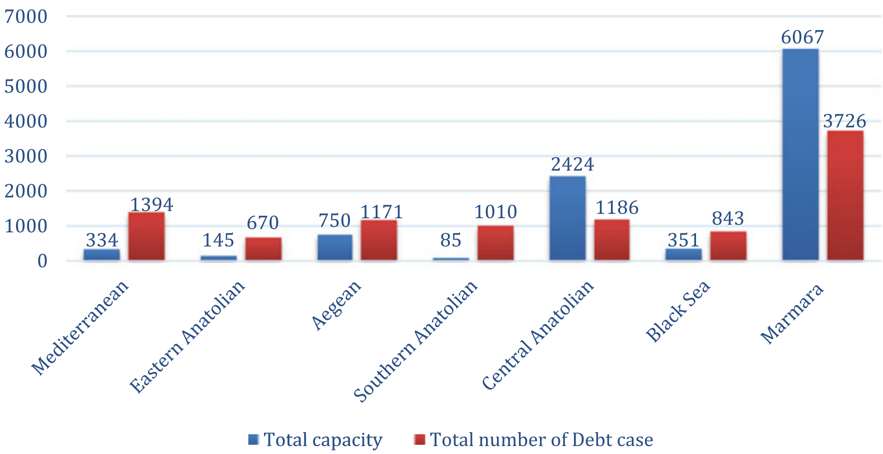


Fig. 3 Comparison of total number of debt cases and total capacity of legal debt collection agencies in regions

Marmara will have cases from all over Turkey. This is also true for these agencies in Central Anatolia Region. However, other regions have less capacity than the cases; therefore, the legal debt collection agencies in these regions will take cases from the same area.

4.2 Data Processing for the Case Study

Due to the model's objective, we want to assign debt cases to legal debt collection agencies in the same cities as possible with equality in distribution rate determined by company managers. Therefore, calculations of some of the parameters are given in this section.

The objective function coefficients are calculated as follows:

$$A_{ij} = \left\{ \begin{array}{ll} 1 & \forall i \in \left\{ I | fr_{ir} = 1 \right\}, \quad j \in \left\{ J | or_{jr} = 1 \right\} \\ 10 & \forall i \in \left\{ I | fc_{ic} = 1 \right\}, \quad j \in \left\{ J | oc_{jc} = 1 \right\} \\ 0 & otherwise \end{array} \right\} \tag{10}$$

According to Eq. (10), if a case is assigned to a legal debt collection agency in the case city, it gets an objective function coefficient of 10. If it is assigned to a legal debt collection agency in the same region (but not the case city), the related objective function coefficient will be 1. For other assignments, the objective function coefficient would be 0. In Eq. (10), fr_{ir} and or_{jr} are calculated as follows:

$$fr_{ir} = \sum_{c} fc_{ic} \times re_{cr} \qquad \forall i, r \tag{11}$$

$$or_{jr} = \sum_{c} oc_{jc} \times re_{cr} \qquad \forall j, r \tag{12}$$

We define some equality parameters to use as input in the model due to the fair model's equality objective. These parameters are found as follows:

$$cr_c = \mathrm{Min} \left\{ 1, \frac{\sum_i fc_{ic}}{\sum_j \left(oc_{jc} \times cap_j \right)} \right\} \qquad \forall c \in \{ C | no_c > 0 \} \tag{13}$$

$$u_j = \left[\sum\nolimits_{c \in \{ C | oc_{jc} = 1 \}} cap_j \times (cr_c + sd) \right]^+ \qquad \forall j \tag{14}$$

$$l_j = \left[\sum\nolimits_{c \in \{ C | oc_{jc} = 1 \}} cap_j \times (cr_c - sd) \right]^- \qquad \forall j \tag{15}$$

One of the important equality parameters is the rate of cases in city c to the capacity of legal debt collection agencies cr_c which is defined for cities that have legal debt collection agencies with Eq. (13). Based on this formula, the rate of assigning cases to each city is equal to 1 or total debt cases of a city divided by

Table 2 Rate of assigning debt case to each city

City_No	cr_c	City_No	cr_c	City_No	cr_c
6	0.23	25	0.59	38	1
7	1	26	0.41	41	0.59
9	1	27	1	42	1
10	1	31	1	55	1
16	0.68	34	0.48	61	0.41
19	1	35	0.85	64	1
22	1	37	1	80	0.64

the total capacity of contracted agency in that city and this value must between 0 and 1. Table 2 shows the assigning rate for each city based on their legal debt collection agencies' capacity. The upper and lower bounds of assigning cases to each agency based on its cities' rate are determined by Eqs. (14) and (15). Suppose that the permitted deviation of assigning cases to offices that placed in the same city sd is equal to 0.02 so u_j and l_j is represented in Table 3. Based on Eqs. (14) and (15), the upper and lower bounds of legal debt collection agencies increase or decrease by the permitted deviation of assigning case determined by the company manager and prevents over-dispersion of data and balance in an assignment.

In our application, we set $sd = 0.02$ and $sd' = 0.05$ based on the suggestions of the company.

na_{ij} is a parameter to eliminate impossible assignments before executing the model. By this way, it is aimed to decrease the number of the decision variables in the model and decrease the solution time.

If a city's assign rate (cr_c) is lower than 1 (i.e., the capacity of legal debt collection agencies are more than cases in the city), we can be sure that the cases of this city will be assigned to that city. The cases in these cities cannot be assigned to the other cities. Therefore na_{ij} is defined as follows:

$$na_{ij} = \left\{ \begin{array}{ll} 1 & \forall i, j, c \in \{I, J, C | fc_{ic} = 1, 0 < cr_c < 1, \;\; oc_{jc} \neq 1\} \\ 0 & \qquad\qquad\qquad\qquad\qquad Otherwise \end{array} \right\} \tag{16}$$

According to Eq. (16), case i cannot be assigned to office j if debt case i placed in city c that have assigned rate between 0 and 1 and for legal debt collection agencies j that not placed in that cities. One of the important issues in this fair model is the equal assignment of debt cases from cities that have not an agency. Therefore, to apply this equality, we determine the rate of cases in city c, where don't have a contracted legal debt collection agencies d_c for cities must have balanced assignment ($ba_c = 1$) with Eq. (17). It represents the rate of not assigned debt cases in a city relative to the total debt cases. This parameter is equal to the subtraction of total debt case of city c and total legal debt collection agencies' capacity of the same city divided on total debt case in whole Turkey, which is calculated in Table 4. Finally, calculate the upper capacity limit of each office in each city that must be

Table 3 Upper and lower bounds of assigning cases from same city to each legal debt collection agency

Agency_ID	u_j	l_j	Agency_ID	u_j	l_j	Agency_ID	u_j	l_j
1	103	94	28	135	123	55	193	185
2	17	14	29	25	20	56	206	188
3	27	22	30	72	64	57	32	29
4	44	41	31	23	20	58	93	89
5	71	66	32	36	29	59	20	16
6	58	54	33	9	6	60	174	159
7	17	13	34	136	124	61	37	33
8	17	14	35	88	82	62	12	9
9	174	159	36	53	47	63	42	35
10	64	58	37	26	21	64	26	21
11	76	69	38	230	214	65	33	27
12	286	268	39	103	94	66	12	10
13	71	67	40	21	18	67	42	37
14	110	104	41	14	12	68	17	13
15	105	99	42	9	6	69	47	39
16	105	95	43	17	13	70	100	90
17	111	101	44	24	21	71	33	31
18	55	52	45	69	57	72	31	25
19	143	130	46	92	84	73	108	102
20	41	34	47	179	170	74	9	6
21	92	84	48	9	6	75	21	17
22	39	35	49	104	98	76	23	21
23	158	145	50	87	83	77	23	21
24	10	8	51	32	29	78	20	18
25	206	188	52	17	14	79	23	21
26	59	55	53	42	37	80	23	21
27	23	21	54	64	58			

Table 4 Case's rate in cities where don't have contracted legal debt collection agency

ba_c	1	2	4	21	33	44	46	63	65
d_c	0.029	0.005	0.007	0.029	0.027	0.010	0.014	0.023	0.011

balanced assigned ul_{jc} with Eq. (18). Therefore, Table 5 represents each legal debt collection agencies' upper capacity to accept the debt cases from each city that don't have any agencies and must have a balanced assignment. It means that legal debt collection agencies with excess capacity must devote a percentage of their capacity to accepting cities' debt cases without contracted agencies. And to establish equality and prevent over-accumulation of cases in a particular city and office, we defined a

Table 5 Upper capacity limit of each legal debt collection agency for assigning the case from cities that have not agencies

ba_c	1	2	4	21	33	44	46	63	65	ba_c	1	2	4	21	33	44	46	63	65
1	7	3	3	7	7	4	4	6	4	41	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	42	2	2	2	2	2	2	2	2	2
3	5	2	2	5	4	3	3	4	3	43	3	2	2	3	3	2	2	3	2
4	3	2	2	3	3	2	2	3	2	44	3	2	2	3	3	2	2	3	2
5	4	2	2	4	4	3	3	4	3	45	9	3	3	9	9	4	5	8	5
6	4	2	2	4	4	2	3	4	3	46	7	2	3	7	6	3	4	6	4
7	3	2	2	3	3	2	2	3	2	47	7	3	3	7	7	4	4	6	4
8	2	2	2	2	2	2	2	2	2	48	2	2	2	2	2	2	2	2	2
9	12	3	4	11	11	5	6	10	5	49	4	2	2	4	4	3	3	4	3
10	5	2	2	5	5	3	3	4	3	50	4	2	2	4	4	2	3	4	3
11	6	2	3	6	6	3	4	5	3	51	3	2	2	3	3	2	2	3	2
12	13	4	4	13	13	6	7	11	6	52	2	2	2	2	2	2	2	2	2
13	4	2	2	4	3	2	2	3	2	53	4	2	2	4	4	2	3	3	2
14	5	2	2	5	5	3	3	4	3	54	5	2	2	5	5	3	3	4	3
15	5	2	2	5	5	3	3	4	3	55	7	2	3	7	7	3	4	6	4
16	8	3	3	7	7	4	4	6	4	56	13	4	4	13	13	6	7	11	6
17	8	3	3	8	8	4	5	7	4	57	3	2	2	3	3	2	2	3	2
18	3	2	2	3	3	2	2	3	2	58	4	2	2	4	4	2	3	4	3
19	10	3	3	10	9	4	5	8	5	59	4	2	2	4	4	2	3	3	2
20	6	2	3	6	6	3	4	5	3	60	12	3	4	11	11	5	6	10	5
21	7	2	3	7	6	3	4	6	4	61	4	2	2	4	3	2	3	3	2
22	4	2	2	4	4	2	3	3	2	62	3	2	2	3	3	2	2	3	2
23	11	3	4	11	10	5	6	9	5	63	6	2	3	6	6	3	4	5	3
24	2	2	2	2	2	2	2	2	2	64	4	2	2	4	4	3	3	4	3
25	13	4	4	13	13	6	7	11	6	65	5	2	2	5	5	3	3	4	3
26	3	2	2	3	3	2	2	3	2	66	2	2	2	2	2	2	2	2	2
27	2	2	2	2	2	2	2	2	2	67	4	2	2	4	4	2	3	3	2
28	9	3	3	9	9	4	5	8	5	68	3	2	2	3	3	2	2	3	2
29	4	2	2	4	4	3	3	4	3	69	7	2	3	7	6	3	4	6	4
30	6	2	3	6	6	3	4	5	3	70	8	3	3	8	8	4	5	7	4
31	3	2	2	3	3	2	2	3	2	71	2	2	2	2	2	2	2	2	2
32	6	2	2	6	5	3	3	5	3	72	5	2	2	5	5	3	3	4	3
33	2	2	2	2	2	2	2	2	2	73	5	2	2	5	5	3	3	4	3
34	9	3	3	9	9	4	5	8	5	74	2	2	2	2	2	2	2	2	2
35	6	2	3	6	5	3	4	5	3	75	4	2	2	4	4	2	3	3	2
36	5	2	2	5	4	3	3	4	3	76	2	2	2	2	2	2	2	2	2
37	4	2	2	4	4	3	3	4	3	77	2	2	2	2	2	2	2	2	2
38	12	3	4	12	12	5	7	10	6	78	2	2	2	2	2	2	2	2	2
39	7	3	3	7	7	4	4	6	4	79	2	2	2	2	2	2	2	2	2
40	3	2	2	3	3	2	2	2	2	80	2	2	2	2	2	2	2	2	2

variable called sd' that determine in this example as $sd' = 0.05$. So the results are represented in Table 5.

$$d_c = \frac{\left(\sum_i fc_{ic} - \sum_{j\in\{J|oc_{jc}=1\}} cap_j\right)}{|I|} \qquad \forall c \in \{C|ba_c = 1\} \tag{17}$$

$$ul_{jc} = \left(d_c \times cap_j\right) + \left(1 + sd'\right) \qquad \forall j, s \in \left\{J, C|ba_c = 1, fc_{jc} \neq 1\right\} \tag{18}$$

4.3 Results of the Case Study

In this section, we examine our mathematical model with described input parameters and analyze the output results. To achieve this, we solve the proposed fair assignment model for assigning the debt cases to contracted legal debt collection agencies with the General Algebraic Modeling System (GAMS)—version 27.1 using CPLEX solver. We solved models with described parameters in this example and found the maximum advantage of assign in the same city subject to the company's limitations and equality constraints. In this example, the maximum value of assigning advantage is calculated as 53,265. Results of the fair assignment model to each contracted legal debt collection agency are represented in Table 6. Information about the number of assigned debt cases, total budget, average budget, and deviation of the average budget from the overall average for each legal debt collection agency are listed in Table 6. Results represent that all debt cases are assigned to the contracted agencies considering the equality measure. As you can see in Fig. 4 average budget of assigned debt cases to each legal debt collection agency is scattered around the overall mean axis and a significant difference between the allocated cases' average is not seen. This shows the equality in the proposed fair assignment model. Also, each agency's average budget deviation is between its upper and lower bounds of permission deviation, which is assumed in this example as 0.05.

Based on the objective of the fair assignment model, which is to maximize the advantage of assigning the debt case to the same city as a first priority, transferring to the same region as a second priority, and otherwise assigning in the whole of Turkey, enriched results are presented in Tables 7, 8, and 9. The assignment matrix of debt cases to legal debt collection agencies based on the region is represented in Table 7. Results explain the number of assigned debt cases from each region to each region of legal debt collection agencies. For example, the Mediterranean region with code 1 has 1394 debt cases assigned 334 to the own region, 385 to the Central Anatolian region, and 675 to the Marmara region because its capacity is 334 and it couldn't be assigned all its debt cases. But Marmara and Central Anatolian are two regions whose debt cases are assigned only to their own regions' legal debt

Table 6 Results of proposed fair assignment model according to contracted legal debt collection agencies

Agencies ID	Total number of assigned cases	Total budget of assigned cases	Average budget of assigned cases	Deviation of assigned cases budget	Agencies ID	Total number of assigned cases	Total budget of assigned cases	Average budget of assigned cases	Deviation of assigned cases budget
1	205	175, 072	854.01	23.56	41	32	27, 140	848.13	17.68
2	31	26, 827	865.39	34.94	42	32	27, 228	850.88	20.43
3	104	88, 190	847.98	17.53	43	63	49, 746	789.62	−40.83
4	63	52, 395	831.67	1.22	44	47	37, 193	791.34	−39.11
5	101	81, 801	809.91	−20.54	45	212	168, 556	795.08	−35.37
6	88	72, 845	827.78	−2.66	46	183	145, 815	796.80	−33.64
7	63	53, 785	853.73	23.28	47	205	167, 366	816.42	−14.03
8	32	26, 300	821.88	−8.57	48	32	27, 767	867.72	37.27
9	347	294, 485	848.66	18.21	49	101	79, 985	791.93	−38.52
10	126	104, 694	830.90	0.46	50	85	70, 806	833.01	2.56
11	151	129, 778	859.46	29.01	51	63	50, 535	802.14	−28.30
12	409	322, 887	789.45	−40.99	52	32	25, 269	789.66	−40.79
13	69	57, 078	827.22	−3.23	53	82	68, 406	834.22	3.77
14	126	108, 886	864.17	33.73	54	126	109, 050	865.48	35.03
15	120	95, 456	795.47	−34.98	55	189	161, 537	854.69	24.25
16	208	180, 492	867.75	37.30	56	373	321, 779	862.68	32.23
17	221	190, 703	862.91	32.46	57	63	53, 054	842.13	11.68
18	63	54, 815	870.08	39.63	58	91	71, 881	789.90	−40.55
19	284	236, 561	832.96	2.51	59	76	61, 258	806.03	−24.42
20	161	127, 590	792.48	−37.96	60	347	278, 002	801.16	−29.29
21	183	144, 985	792.27	−38.18	61	72	59, 480	826.11	−4.34
22	76	66, 004	868.47	38.03	62	43	34, 397	799.93	−30.52
23	315	272, 507	865.10	34.65	63	164	131, 214	800.09	−30.36

(continued)

Table 6 (continued)

Agencies ID	Total number of assigned cases	Total budget of assigned cases	Average budget of assigned cases	Deviation of assigned cases budget	Agencies ID	Total number of assigned cases	Total budget of assigned cases	Average budget of assigned cases	Deviation of assigned cases budget
24	18	15, 074	837.44	7.00	64	101	83, 567	827.40	−3.05
25	360	305, 859	849.61	19.16	65	126	102, 639	814.60	−15.85
26	57	49, 291	864.75	34.31	66	22	19, 114	868.82	38.37
27	22	18, 937	860.77	30.33	67	82	64, 719	789.26	−41.19
28	268	233, 512	871.31	40.87	68	63	49, 900	792.06	−38.38
29	95	80, 771	850.22	19.77	69	176	139, 792	794.27	−36.17
30	164	129, 641	790.49	−39.95	70	229	184, 029	803.62	−26.83
31	44	36, 266	824.23	−6.22	71	32	25, 929	810.28	−20.17
32	139	111, 511	802.24	−28.21	72	120	95, 953	799.61	−30.84
33	32	26, 095	815.47	−14.98	73	123	105, 720	859.51	29.07
34	271	228, 111	841.74	11.29	74	32	26, 764	836.38	5.93
35	145	114, 518	789.78	−40.67	75	82	66, 878	815.59	−14.86
36	104	84, 216	809.77	−20.68	76	22	18, 729	851.32	20.87
37	100	82, 408	824.08	−6.37	77	22	17, 892	813.27	−17.17
38	378	327, 725	867.00	36.55	78	22	17, 432	792.36	−38.08
39	205	177, 606	866.37	35.92	79	22	18, 575	844.32	13.87
40	41	34, 986	853.32	22.87	80	22	18, 709	850.41	19.96

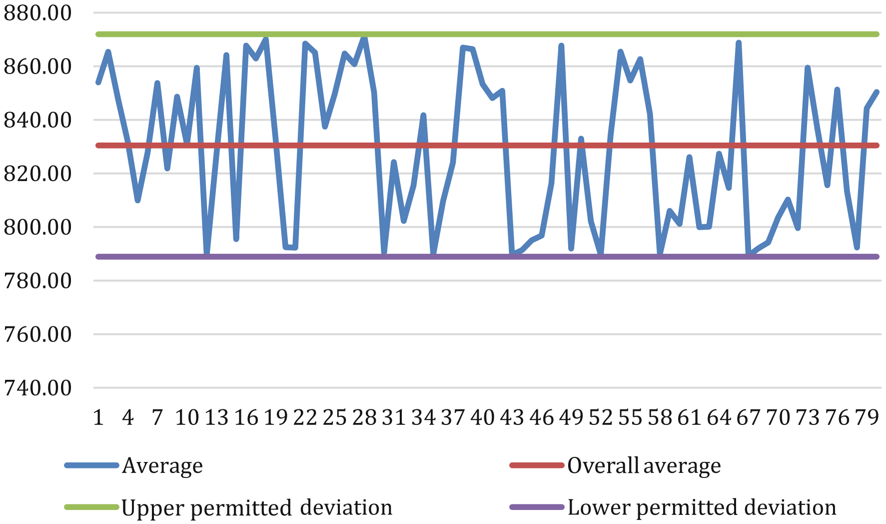


Fig. 4 Information about average budget of assigned cases for each legal debt collection agencies

Table 7 Assignment matrix based on regions

		1	2	3	4	5	6	7	Number of cases assigned to the same region	Rate of cases assigned to the same region
Mediterranean	1	334	0	0	0	385	0	675	334	100%
Eastern Anatolian	2	0	145	0	0	220	0	305	145	100%
Aegean	3	0	0	750	0	183	0	238	750	100%
Southern Anatolian	4	0	0	0	85	321	0	604	85	100%
Central Anatolian	5	0	0	0	0	1186	0	0	1186	51%
Black Sea	6	0	0	0	0	63	698	82	698	100%
Marmara	7	0	0	0	0	0	0	3726	3726	61%

collection agencies. In Table 7, the cumulative results of the number of the assigned debt cases to the same regions are defined too. It means how many debt cases are assigned to their own region and how many of the cases are assigned to the other regions. For example, as you can see in Table 7, Mediterranean, East Anatolian, Aegean, Southern Anatolian, and Black Sea regions' legal debt collection agencies only accept their own region debt cases and other regions' cases are not assigned to these regions, so their rate of assigned case to same region is 100% but region of Central Anatolian and Marmara assignment rate are 51% and 61% which means that these regions can fill the extra capacity by the other region's debt case. Central Anatolian and Marmara debt cases are assigned to their own region's legal debt

Table 8 Information about assigning in the same cities

City_Code	Number of cases assigned to same city	Rate of cases assigned to same city	City Code	Number of cases assigned to same city	Rate of cases assigned to same city
6	488	24%	34	2438	53%
7	189	100%	35	562	85%
9	69	100%	37	22	100%
10	22	100%	38	91	100%
16	388	68%	41	222	59%
19	32	100%	42	22	100%
22	22	100%	55	101	100%
25	85	59%	61	81	15%
26	95	41%	64	22	100%
27	85	100%	80	56	64%
31	57	100%			

Table 9 Number of cases assigned to each legal debt collection agencies from own city

Legal debt collection agencies_ID	Number of cases from same city	u_j	l_j	Legal debt collection agencies_ID	Number of cases from same city	u_j	l_j
6	488	440	533	34	2438	2333	2538
7	189	185	193	35	562	547	577
9	69	67	71	37	22	21	23
10	22	21	23	38	91	89	93
16	388	376	401	41	222	214	230
19	32	31	33	42	22	21	23
22	22	21	23	55	101	99	103
25	85	82	88	61	81	70	93
26	95	90	100	64	22	21	23
27	85	83	87	80	56	54	58
31	57	56	58				

collection agencies and these agencies extra capacity in these two regions assigned to the debt case from the other region.

Table 8 proposes the number of the assigned debt cases to the same city, which is the first priority of the assignment model. As far as mentioned before, there is no legal debt collection agency in some cities, so those debt cases are assigned to the agencies in the other cities which are in the same region or other regions. There are 21 cities that have legal debt collection agencies. Some cities' agencies assign to the debt case from their own cities. Hence, they do not have more capacity for the other cities' debt case, such as Central Anatolian (code number 7) or Samsun (code number 55), but Istanbul (code number 34) has been assigned 53% of its capacity to debt case form Istanbul and the rest of the extra capacity has been assigned to the other cities. Based on this model, extra capacity is distributed among variety of debt cases from different cities.

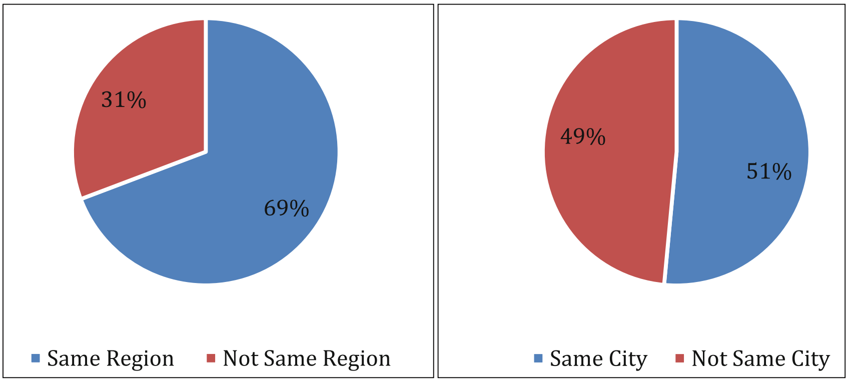


Fig. 5 Rate of debt case to the same region and same city

In Fig. 5 the rate of the debt cases assigned to the same region and same city is defined to compare with the other debt case assigned to the other regions besides their own region. It is proposed that about 50% of the debt case has been assigned to the same cities' legal debt collection agencies.

In the mathematical model, we defined two parameters u_j and l_j which are limited each legal debt collection agency to accept the equal number of debt cases from own city with permitted deviation $sd = 0.02$ and based on the city's rate cr_c. The model's results based on the upper and lower bounds of each agency to assign to the same city debt cases are shown in Table 9. For example, legal debt collection agencies with code 10 that placed in Istanbul by $cr_c = 0.48$ can accept 126 cases but based on the model results, 22 debt cases of Istanbul are assigned to this agency and the rest of the capacity is assigned by other cities' cases. As you can see in Table 9 this number is in the accepted range of the problem, which was calculated to establish equality between the Istanbul city's legal debt collection agencies in accepting cases related to their own city.

5 Conclusions

This chapter proposes an optimization model for the fair assignment of debt cases of a big service company to its contracted legal debt collection agencies. Regarding to past research studies as mentioned in literature review, optimization models and machine learning models have been applied to improve the debt collection process. But none of them had studied the assignment of debt cases to legal debt collection agencies. Therefore, in this study, a mathematical programming model is formulated and solved for fair assignment of debt cases to legal debt collection agencies. The objective is set as to assign the cases to the agencies in the same city or in the same region. The fair assignment is achieved by balancing the average budget of the agencies and assigning similar percentage of cases from the same city. A variety of parameters has been applied to reach the best model. This model is solved on GAMS platform using CPLEX solver. Results of the model represented the validity of the proposed fair assignment model. Model constraints limited this assignment based on the capacity of agencies or equality limitations that minimize the permitted deviation in the number of assignment cases and average budget of cases.

The results of the case study show that the fair assignment can be reached by the use of the proposed model. We analyzed the statistics of assigning the cases to the agency in the same city or the same region. Based on the fair assignment model outputs, approximately 51% of the cases are assigned to their own cities and about 69% of them assigned to their regions. And we can see in results that the city such as Istanbul with high capacity in legal debt collection centers assigned all their debt cases and allocated their extra capacity to cities of its region and other regions. On the other hand, a region such as the Mediterranean only assigns part of its debt cases due to its capacity and the rest of its debt cases are assigned to other regions. Another objective of the proposed model was equality which was achieved by permitted deviation parameters in the number of same city assignment and average of assigned cases' budget. For analyzing the model, a sensitivity analysis was applied for sd and $sd^{'}$ parameters. However, the cities' legal debt collection agencies allowing capacity in assigning own city's cases were changed but the total number of assignments in each city and region did not change because of the capacity limitation.

For the future study, a fuzzy linear assignment model can be proposed to handle the model's uncertainty environment. Alternatively, a bi-objective model can be suggested that maximize the advantage of assigning in the same city or region and minimize the permitted deviation parameters sd and $sd^{'}$ at the same time. In the assignment of cases to debt collection agencies, the past performance of the agencies can be an significant input. For this, a machine learning algorithm can be used to find the expected performance of agencies for each of the cases. This input can be integrated to the proposed model as new objective function or a constraint. By this way, the efficiency of the debt collection process can be improved.

Acknowledgment This work is supported by TUBITAK (The Scientific and Technological Research Council of Turkey, Project Id: 5200012).

References

Abe, N., Jensen, D. L., Thomas, V. P., Bennett, J. J., Anderson, G. F., Cooley, B. R., & Kowalczyk, M. (2010). Optimizing debt collections using constrained reinforcement learning. *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC*, 25–28. https://doi.org/10.1145/1835804.1835817.

Afroz, H. D., & Hossen, D. M. A. (2017). New proposed method for solving assignment problem and comparative study with the existing methods. *IOSR Journal of Mathematics, 13*(02), 84–88. https://doi.org/10.9790/5728-1302048488

Al-Saeedi, H. A. H. A.-D. & Shiker, M. A. K. (2020). Two new effective methods to find the optimal solution for the assignment problems. *Journal of Advanced Research in Dynamical and Control Systems, 12*(7), 49–54. https://doi.org/10.5373/JARDCS/V12I7/20201983

Antoine, D., & Rabie, N. A. (2006). Modeling of the debts collection process for service companies. *IFAC Proceedings Volumes, 39*(3), 349–354. https://doi.org/10.3182/20060517-3-FR-2903.00188

Basirzadeh, H. (2012). Ones assignment method for solving mathematical formulation of assignment problems. *Applied Soft Computing, 6*(47), 2345–2355.

Biggerstaff, T. J., Mitbander, B. G., & Webster, D. (1993). The concept assignment problem in program understanding. *Proceedings Working Conference on Reverse Engineering*, 27–43. https://doi.org/10.1109/WCRE.1993.287781.

Boah, D. K., Adu, I. K., & Gyebil, F. J. (2015). Assignment problem of a legal firm in Kumasi, Ghana. *International Journal of Computing and Optimization, 2*(1), 1–5. https://doi.org/10.12988/ijco.2015.41226

Bouras, A., Ghaleb, M. A., Suryahatmaja, U. S., & Salem, A. M. (2014). The airport gate assignment problem: A Survey. *The Scientific World Journal, 2014*, 1–27. https://doi.org/10.1155/2014/923859

Brock, D. M., Yaffe, T., & Dembovsky, M. (2006). International diversification and performance: A study of global law firms. *Journal of International Management, 12*(4), 473–489. https://doi.org/10.1016/j.intman.2006.08.003

Chen, C.-H., Chiang, R.-D., Wu, T.-F., & Chu, H.-C. (2013). A combined mining-based framework for predicting telecommunications customer payment behaviors. *Expert Systems with Applications, 40*(16), 6561–6569. https://doi.org/10.1016/j.eswa.2013.06.001

Chin, A. G., & Kotak, H. (2006). Improving debt collection processes using rule-based decision engines: A case study of Capital One. *International Journal of Information Management, 26*(1), 81–88. https://doi.org/10.1016/j.ijinfomgt.2005.10.002

De Almeida Filho, A. T., Mues, C., & Thomas, L. C. (2010). Optimizing the collections process in consumer credit. *Production and Operations Management, 19*(6), 698–708. https://doi.org/10.1111/j.1937-5956.2010.01152.x

Duman, E., Ecevit, F., Çakır, Ç., & Altan, O. (2017). A novel collection optimisation solution maximising long-term profits: A case study in an international bank. *Journal of Decision Systems, 26*(4), 328–340. https://doi.org/10.1080/12460125.2017.1422318

Gavish, B., & Pirkul, H. (1991). Algorithms for the multi-resource generalized assignment problem. *Management Science, 37*(6), 695–713. https://doi.org/10.1287/mnsc.37.6.695

Georgopoulos, E. F., & Giannaropoulos, S. M. (2007). Solving resource management optimization problems in contact centers with artificial neural networks. In *19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007)* (pp. 405–412). https://doi.org/10.1109/ICTAI.2007.128

Han, C., & Jang, Y. (2013). Effects of debt collection practices on loss given default. *Journal of Banking and Finance, 37*(1), 21–31. https://doi.org/10.1016/j.jbankfin.2012.08.009

Ho, C. J., & Vaughan, J. W. (2012). Online task assignment in crowdsourcing markets. *Proceedings of the National Conference on Artificial Intelligence, 1*, 45–51.

Ho Ha, S., & Krishnan, R. (2012). Predicting repayment of the credit card debt. *Computers and Operations Research, 39*(4), 765–773. https://doi.org/10.1016/j.cor.2010.10.032

I-Fei, C. (2010). Evaluate the performance of cardholders' repayment behaviors using artificial neural networks and data envelopment analysis. *The 6th International Conference on Networked Computing and Advanced Information Management, Seoul*, 478–483. https://ieeexplore.ieee.org/document/5572085?arnumber=5572085

Kim, J., & Kang, P. (2016). Late payment prediction models for fair allocation of customer contact lists to call center agents. *Decision Support Systems, 85*, 84–101. https://doi.org/10.1016/j.dss.2016.03.002

Krokhmal, P. A., & Pardalos, P. M. (2009). Random assignment problems. *European Journal of Operational Research, 194*(1), 1–17. https://doi.org/10.1016/j.ejor.2007.11.062

Küçük Matci, D., & Acar, I. (2019). Modelling and optimization of the exam invigilator assignment problem based on preferences. *Turkish Journal of Engineering*. https://doi.org/10.31127/tuje.467003

Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly, 2*(1–2), 83–97. https://doi.org/10.1002/nav.3800020109

Kumar, A., & Gupta, A. (2011). Methods for solving fuzzy assignment problems and fuzzy travelling salesman problems with different membership functions. *Fuzzy Information and Engineering, 3*(1), 3–21. https://doi.org/10.1007/s12543-011-0062-0

Liebman, L. H. (1972). A Markov decision model for selecting optimal credit control policies. *Management Science, 18*(10), B-519–B-525. https://doi.org/10.1287/mnsc.18.10.B519

Liu, H., Zhang, P., Hu, B., & Moore, P. (2015). A novel approach to task assignment in a cooperative multi-agent design system. *Applied Intelligence, 43*(1), 162–175. https://doi.org/10.1007/s10489-014-0640-z

Miller, G., Weatherwax, M., Gardinier, T., Abe, N., Melville, P., Pendus, C., Jensen, D., Reddy, C. K., Thomas, V., Bennett, J., Anderson, G., & Cooley, B. (2012). Tax collections optimization for New York state. *Interfaces, 42*(1), 74–84. https://doi.org/10.1287/inte.1110.0618

Mills-tettey, G. A., & Stentz, A. (2007, July). *The dynamic Hungarian algorithm for the assignment problem with changing costs*.

Mitchner, M., & Peterson, R. P. (1957). An operations research study of the collection of defaulted loans. *Operations Research, 5*(4), 522–545. https://doi.org/10.1287/opre.5.4.522

Munaam, A., & Hammadi, K. (2017). Solving multi objective assignment problem using Tabu search algorithm. *Global Journal of Pure and Applied Mathematics, 13*(9), 4747–4764. http://www.ripublication.com.

Munkres, J. (1957). Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics, 5*(1), 32–38. https://doi.org/10.1137/0105003

Onar, S. C., Oztaysi, B., & Kahraman, C. (2018a). A fuzzy rule based inference system for early debt collection. *Technological and Economic Development of Economy, 24*(5), 1845–1865. https://doi.org/10.3846/20294913.2016.1266409

Onar, S. C., Oztaysi, B., Kahraman, C., & Ozturk, E. (2020). Evaluation of legal debt collection services by using Hesitant Pythagorean (Intuitionistic Type 2) fuzzy AHP. *Journal of Intelligent & Fuzzy Systems, 38*(1), 883–894. https://doi.org/10.3233/JIFS-179456

Onar, S. Ç., Öztürk, E., Öztayşi, B., Yüksel, M., & Kahraman, C. (2018b). Pisagor bulanik akilli çok ölçütlü yasal takip avukatlik ofisi performans değerlendirme modeli. In *Mühendislik ve Teknoloji Yönetimi Zirvesi* (pp. 88–97). İstanbul Teknik Üniversitesi & Bahçeşehir Üniversitesi.

Pala, O. (2020). Kuadratik atama problemine yeni bir melez karinca kolonisi optimizasyon algoritmasi önerisi. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 29*(4), 21–32. https://doi.org/10.35379/cusosbil.658693

Pentico, D. W. (2007). Assignment problems: A golden anniversary survey. *European Journal of Operational Research, 176*(2), 774–793. https://doi.org/10.1016/j.ejor.2005.09.014

Piu, F., & Speranza, M. G. (2014). The locomotive assignment problem: A survey on optimization models. *International Transactions in Operational Research, 21*(3), 327–352. https://doi.org/10.1111/itor.12062

Salehi, K. (2014). An approach for solving multi-objective assignment problem with interval parameters. *Management Science Letters, 4*(9), 2155–2160. https://doi.org/10.5267/j.msl.2014.7.031

Series, C. (2019). A new approach solution for fuzzy assignment problem using the development Zimmermann method. *IOSR Journal of Engineering, 8*(5), 50–58. https://doi.org/10.1088/1742-6596/1280/2/022045

Stewart, A. (2017). How to evaluate your debt collection agency. *Debt Recoveries Australia*. https://debtrecoveries.com.au/how-to-evaluate-your-debt-collection-agency/

Takahashi, M., & Tsuda, K. (2013). Towards early detections of the bad debt customers among the mail order industry. *Studies in Computational Intelligence, 484*, 167–176. https://doi.org/10.1007/978-3-642-37932-1_12

Taneta- Skwiercz, D. (2018). Wskaźniki pomiaru zrównoważonego rozwoju – Polska na tle krajów Unii Europejskie. *Prace Naukowe Uniwersytetu Ekonomicznego We Wrocławiu, 516*, 121–132. https://doi.org/10.15611/pn.2018.515.11

van de Geer, R., Wang, Q., & Bhulai, S. (2018). Data-Driven consumer debt collection via machine learning and approximate dynamic programming. *SSRN Electronic Journal, 1–32*. https://doi.org/10.2139/ssrn.3250755

Vecchio, M., Jin, S., Mistretta, A., Rolando, H., & Tuck, H. (2006). Designing a search mechanism for debt collection. *2006 IEEE Systems and Information Engineering Design Symposium*, 168–173. https://doi.org/10.1109/SIEDS.2006.278733

Wang, L., Wang, H., Xu, Z., & Ren, Z. (2019). The interval-valued hesitant Pythagorean fuzzy set and its applications with extended TOPSIS and Choquet integral-based method. *International Journal of Intelligent Systems, 34*(6), 1063–1085. https://doi.org/10.1002/int.22086

Wang, W. (2000). Evaluating the technical efficiency of large US law firms. *Applied Economics, 32*(6), 689–695. https://doi.org/10.1080/000368400322309

Wilson, J. M. (1997). A genetic algorithm for the generalised assignment problem. *Journal of the Operational Research Society, 48*(8), 804–809. https://doi.org/10.1057/palgrave.jors.2600431

Younas, I., Kamrani, F., Schulte, C., & Ayani, R. (2011, April). Optimization of task assignment to collaborating agents. *2011 IEEE Symposium on Computational Intelligence in Scheduling (SCIS)*, 17–24. https://doi.org/10.1109/SCIS.2011.5976547

On the Principle of Lagrange in Optimization Theory and Its Application in Transportation and Location Problems

J. B. G. Frenk and Sonya Javadi

Abstract In mathematical optimization, the Lagrangian approach is a general method to find an optimal solution of a finite (infinite) dimensional constrained continuous optimization problem. This method has been introduced by the Italian mathematician Joseph-Louis Lagrange in 1755 in a series of letters to Euler. This approach became known under the name *The Principle of Lagrange* and was also applied much later to integer optimization problems. The basic idea behind this method is to replace a constrained optimization problem by a sequence of easier solvable optimization problems having fewer constraints and penalizing the deletion of some of the original constraints by replacing the original objective function. To select the best penalization, the so-called Lagrangian dual function needs to be optimized and a possible algorithm to do so is given by the so-called subgradient method. This method is discussed in detail at the end of this chapter. The Lagrangian approach led to the introduction of dual optimization problems and penalization methods in nonlinear programming and recently to the development of interior point methods and the identification of polynomial solvable classes of continuous optimization problems. Also it had its impact on how to construct algorithms to generate approximate solutions of integer optimization problems. In this chapter, we discuss in the first part the main ideas behind this approach for any type of finite dimensional optimization problem. In the remaining parts of this chapter we focus in more detail on how this approach is used in continuous optimization problems and show its full impact on the so-called K-convex continuous optimization problems. Also we consider its application within linear integer programming problems and show how it is used to solve these type of problems. To illustrate its application to the well-known integer programming problems, we consider in the final section

J. B. G. Frenk
Faculty of Engineering and Natural Science, Sabancı University, Istanbul, Turkey
e-mail: frenk@sabanciuniv.edu

S. Javadi (✉)
Department of Industrial Engineering, Doğuş University, Istanbul, Turkey
e-mail: sjavadi@dogus.edu.tr

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_2

its application to some classical vehicle routing and location models. As such this chapter should be regarded as an introduction to duality theory and the Lagrangian approach for less mathematically oriented readers proving at the same time most of the results using the simplest possible proofs.

Keywords Constrained optimization problems · Lagrangian relaxation technique · Dual problems · Transportation and location problems

1 Introduction

In this chapter we give a unified overview of the Lagrangian relaxation technique used in finite dimensional optimization problems with either continuous and/or discrete decision variables and show as an example of how to apply these techniques by solving some classical transportation and location models. The main purpose and goal of this chapter is to give an introduction to Lagrangian duality accessible for less mathematically oriented readers and explain its main ideas. At the same time we try to be rigorous proving the most well-known results about strong duality for the class of K-convex optimization problems by means of the simplest possible proofs. All the proofs related to strong duality will be given with one exception. We will only mention the most basic separation result in convex analysis separating a nonempty convex set from a point outside this set. In general, contrary to continuous K-convex optimization problems, no strong duality holds for integer linear programming problems. By means of elementary proofs, we show how to apply the Lagrangian relaxation technique to these problems. Another more abstract approach to Lagrangian duality given by the perturbation function approach (Frenk and Kassay, 2005; Rockafellar, 1970) requires the knowledge of conjugate and biconjugate functions well-known in convex analysis.

To show strong duality results for K-convex optimization problems, one uses in this approach the Fenchel–Moreau theorem. This theorem states that under some mild topological conditions a convex function equals its biconjugate function. Although this approach is more general and has the advantage of shortening proofs, applying directly the Fenchel–Moreau theorem the perturbation function approach hides the intuitive ideas that are shown more clearly in the Lagrangian relaxation approach. For this reason, we have chosen this approach to explain the main results in duality theory. This approach is closely related to the minmax approach used in a two-person noncooperative game with the decision maker playing against nature. In this framework strong duality is equivalent to the existence of a Nash equilibrium point (Frenk and Kassay, 2006, 2008).

In the first section we introduce the Lagrangian relaxation approach and its application to finite dimensional optimization problems. In the second section we apply this approach to (integer) linear programming problems and discuss the main relations between such an integer linear program and some of the most well-known relaxations used within this field. In the third section we consider the subgradient method to find the optimal Lagrangian multiplier in a Lagrangian dual problem.

Finally in the last section we consider some classical location and transportation models and show how the Lagrangian relaxation technique results in heuristics by solving these problems.

Although no new results are discussed, this chapter introduces the reader to the main ideas of the Lagrangian relaxation technique originally introduced by Joseph-Louis Lagrange in his treatise “Leçons sur le calcul des functions” (nouvelle edition) published in 1806 (Lagrange, 1806) and revealed in a series of letters to Euler between 1754 and 1756. Joseph-Louis Lagrange applied his approach to infinite dimensional optimization problems, and his approach eventually lead to the development of optimal control theory. To prove the claims of Lagrange, the theory of convex sets and functions was later developed. Without any difficulty, one can also extend the finite dimensional approach discussed in this chapter to infinite dimensional optimization problems using the well-known Hahn–Banach theorem in functional analysis (Choquet, 1976) or the related minmax approach (Frenk Kas and Kassay, 2007).

This chapter also tries to be complete in discussing in a hopefully transparent way some of the mathematical technicalities found in the books on convex analysis. As such the rationale behind the Lagrangian relaxation approach can be seen as one of the main ideas in optimization theory. It relates a difficult constrained optimization problem to an easier solvable constrained optimization problem with fewer constraints and another objective function. In this newly selected objective function the deleted constraints of the original problem are penalized using a linear penalty function.

Developments occurring 150 years later within the field of nonlinear and linear programming like the primal–dual relations in linear programming and the well-known Karush–Kuhn–Tucker conditions in nonlinear programming are special instances of this Lagrangian relaxation approach. Also the developments of penalty-based optimization methods in nonlinear programming (Fiacco and McCormick, 1969) can be seen as an easy extension of the ideas behind this relaxation approach. These penalty methods eventually led to the development of interior point polynomial algorithms in convex programming (see page 2 of Nesterov and Nemirovski (2001)). By these observations, it seems clear that the ideas of Lagrange play a fundamental role in the development of the theory of optimization known nowadays.

2 On the Principle of Lagrange

Before discussing the principle of Lagrange for general optimization problems with continuous and/or discrete decision variables, we first introduce some definitions. If the sets $A, B \subseteq \mathbb{R}^n$ are nonempty sets, then for every $\alpha, \beta \in \mathbb{R}$ we define the so-called Minkowski sum:

$$\alpha A + \beta B := \{\alpha \mathbf{x} + \beta \mathbf{y} : \mathbf{x} \in A, \mathbf{y} \in B\}. \tag{1}$$

Also we denote by $\mathbb{Z}_+^n$ ($\mathbb{R}_+^n$) the set of nonnegative integer valued (real valued) n-dimensional vectors and by $\mathbb{Z}_-^n := -\mathbb{Z}_+^n$ and $\mathbb{R}_-^n := -\mathbb{R}_+^n$.

Definition 2.1 A nonempty set $L \subseteq \mathbb{R}^n$ is called a *linear* space if $\alpha L + \beta L \subseteq L$ for every $\alpha, \beta \in \mathbb{R}$. A nonempty set $M \subseteq \mathbb{R}^n$ is called *affine* if $\alpha M + (1-\alpha)M \subseteq M$ for every $\alpha \in \mathbb{R}$. A nonempty set $C \subseteq \mathbb{R}^n$ is called *convex* if $\alpha C + (1-\alpha)C \subseteq C$ for every $0 < \alpha < 1$. A nonempty set $K \subseteq \mathbb{R}^n$ is called a *cone* if $\alpha K \subseteq K$ for every $\alpha > 0$ and it is called a *pointed cone* if K is a cone satisfying $K \cap (-K) = \{\mathbf{0}\}$. A nonempty set $A \subseteq \mathbb{R}^n$ is called *proper* if the set A is strictly contained in $\mathbb{R}^n$.

It is easy to verify that a pointed cone is proper and a cone K is convex if and only if $K + K \subseteq K$. To represent feasible sets within optimization theory, we need to introduce an ordering on a set.

Definition 2.2 A binary relation $\preceq$ on a nonempty set A is called a transitive ordering on the set A if for every $\mathbf{x}_i \in A$, $i = 1, 2, 3$ satisfying $\mathbf{x}_1 \preceq \mathbf{x}_2$ and $\mathbf{x}_2 \preceq \mathbf{x}_3$ it follows that $\mathbf{x}_1 \preceq \mathbf{x}_3$. A binary relation $\preceq$ on a nonempty set A is called a partial ordering on the set A if it is a transitive ordering and for every $\mathbf{x} \in A$ it satisfies $\mathbf{x} \preceq \mathbf{x}$ (reflexive property) and for every $\mathbf{x}_i \in A$,$i = 1, 2$, satisfying $\mathbf{x}_1 \preceq \mathbf{x}_2$ and $\mathbf{x}_2 \preceq \mathbf{x}_1$ it holds that $\mathbf{x}_1 = \mathbf{x}_2$ (antisymmetry property)

For any convex cone $K \subseteq \mathbb{R}^n$, we introduce the transitive ordering $\preceq_K$ on the set $\mathbb{R}^n$ (Boyd and Vandenberghe , 2004; Frenk and Kassay, 1999) given by

$$\mathbf{y} \preceq_K \mathbf{x} \Leftrightarrow \mathbf{x} - \mathbf{y} \in K. \tag{2}$$

It follows that the transitive ordering $\preceq_K$ on $\mathbb{R}^n$ for any convex pointed cone K is a partial ordering on $\mathbb{R}^n$. In most cases, unless otherwise specified, we will only consider orderings $\preceq_K$ on $\mathbb{R}^n$ with K a convex pointed cone. Consider now a nonempty set $X \subseteq \mathbb{R}^n$ and $K \subseteq \mathbb{R}^m$ a nonempty proper convex cone. Moreover, let $f : \mathbb{R}^n \to \mathbb{R}$ be some finite valued function and $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$ a finite valued vector function given by $\mathbf{g}(\mathbf{x}) := (g_1(\mathbf{x}), \ldots, g_m(\mathbf{x}))$ with $g_i : \mathbb{R}^n \to \mathbb{R}$ and introduce the minimization problem:

$$\upsilon(P) := \inf\{f(\mathbf{x}) : \mathbf{x} \in \mathcal{F}\} \tag{P}$$

with

$$\mathcal{F} := \{\mathbf{x} \in X : \mathbf{g}(\mathbf{x}) \preceq_K \mathbf{0}\} \tag{3}$$

the so-called feasible region. The optimization problem (P) is called a *primal problem*, and using the ordering $\preceq_K$ with K a proper convex cone yields the possibility to model a lot of different optimization problems. The value $\upsilon(P)$ is by definition equal to ∞ if the feasible region $\mathcal{F}$ of optimization problem (P) is empty. In this case optimization problem (P) is called *infeasible*. If the feasible region $\mathcal{F}$ is nonempty, then optimization problem (P) is called *feasible* and $-\infty \leq \upsilon(P) < \infty$. It is not assumed that an *optimal feasible solution* exists and so we cannot replace inf by min. Since

$$-\sup\{-f(\mathbf{x}) : \mathbf{x} \in \mathcal{F}\} = \inf\{f(\mathbf{x}) : \mathbf{x} \in \mathcal{F}\},$$

we also cover in this framework maximization problems. Important instances of optimization problem (P) are given by nonlinear optimization problems with feasible region $\mathcal{F}$ consisting of p inequality and $m-p$ equality constraints (Bazaraa, Sherali and Shetty, 1993; Nocedal and Wright, 2006). If this holds, the convex cone K is given by $K = \mathbb{R}_+^p \times \{\mathbf{0}\}, \mathbf{0} \in \mathbb{R}^{m-p}, 0 \leq p \leq m$. For $p = 0$, this reduces to a feasible region $\mathcal{F}$ only consisting of m equality constraints and we obtain $K = \{\mathbf{0}\}$. Special cases are linear programming problems (Bazaraa, Jarvis, and Sherali, 1990; Chvatal, 1983) and for X a mixed discrete–continuous set mixed integer linear programming problems (Nemhauser and Wolsey, 1988; Wolsey, 1998). Finally we also mention conic convex programming problems (Ben-Tal and Nemirovski, 2001) given by

$$\inf\{\mathbf{c}^\top \mathbf{x} : \mathbf{x} - \mathbf{b} \in L, \mathbf{x} \in K\} \tag{4}$$

with K a proper convex cone and L a linear space. In most cases it is difficult to find a feasible solution of optimization problem (P) and even more difficult (if it exists!) to find an optimal solution. A possible way to overcome this problem is to replace optimization problem (P) by an easier optimization problem selected in such a way that it resembles the original problem. Having this idea in mind we introduce the next definition.

Definition 2.3 An optimization problem given by

$$\upsilon(R) := \inf\{f_R(\mathbf{x}) : \mathbf{x} \in \mathcal{F}_R\} \tag{R}$$

is called a relaxation of the primal problem (P) if the feasible region $\mathcal{F}_R$ satisfies $\mathcal{F} \subseteq \mathcal{F}_R$ and $f_R(\mathbf{x}) \leq f(\mathbf{x})$ for every $\mathbf{x}$ belonging to $\mathcal{F}$.

An immediate consequence of the definition of a relaxation is given by

$$-\infty \leq \upsilon(R) \leq \upsilon(P) \leq \infty. \tag{5}$$

An easy way to obtain a relaxation is to delete some of the difficult constraints within the feasible set $\mathcal{F}$ given by relation (3). However, to achieve that the new constructed optimization problem resembles the original problem (P) it is in general not sufficient to delete constraints. A more refined way to achieve this goal is to delete constraints and *at the same time* incorporate within the objective function some penalty cost to compensate for the loss of these constraints and/or give some reward for still satisfying feasibility. To formalize this idea, we first introduce the standard inner product on $\mathbb{R}^n$ given by

$$\mathbf{x}^\mathsf{T}\mathbf{y} := \sum\nolimits_{i=1}^{n} x_i y_i. \tag{6}$$

The most simplest way to achieve penalization of the loss of constraints is to consider a linear penalty function and introduce for some penalty parameter $\boldsymbol{\lambda} \in \mathbb{R}^m$ the problem

$$\inf\{f(\mathbf{x}) + \boldsymbol{\lambda}^\mathsf{T}\mathbf{g}(\mathbf{x}) : \mathbf{x} \in X\}. \tag{7}$$

Clearly it follows that $\mathcal{F} \subseteq X$ and to guarantee that

$$f(\mathbf{x}) + \boldsymbol{\lambda}^\mathsf{T}\mathbf{g}(\mathbf{x}) \leq f(\mathbf{x}) \tag{8}$$

for every $\mathbf{x}$ belonging to $\mathcal{F}$ it is sufficient and necessary to assume that $\boldsymbol{\lambda}^\mathsf{T}\mathbf{g}(\mathbf{x}) \leq 0$ for every $\mathbf{x}$ belonging to $\mathcal{F}$. Since we are dealing with minimization problems, a reward means that $\boldsymbol{\lambda}^\mathsf{T}\mathbf{g}(\mathbf{x}) \leq 0$ for every feasible $\mathbf{x}$. Due to $\mathbf{g}(\mathbf{x}) \preceq_K \mathbf{0}$ for every $\mathbf{x} \in \mathcal{F}$ and introducing the so-called dual cone K^* of a set $K \subseteq \mathbb{R}^m$ given by

$$K^* := \{\mathbf{y} \in \mathbb{R}^m : \mathbf{y}^\mathsf{T}\mathbf{x} \geq 0 \ \text{ for every } \mathbf{x} \in K\}, \tag{9}$$

this inequality certainly holds for every $\boldsymbol{\lambda}$ belonging to K^*. Also by the definition of a dual cone, the set K^* is a nonempty closed convex cone. A similar operation for the special case of a linear space $L \subseteq \mathbb{R}^m$ is to introduce its *orthogonal complement* $L^\perp = \{\mathbf{y} \in \mathbb{R}^m : \mathbf{y}^\mathsf{T}\mathbf{x} = 0 \text{ for every } \mathbf{x} \in L\}$. It is easy to show for a linear space L that

$$L^* = L^\perp. \tag{10}$$

At the moment, except for finiteness of the functions f and $\mathbf{g}$ on their domain we do not impose any other conditions on these functions and formally introduce the minimization problem:

$$\theta(\boldsymbol{\lambda}) := \inf\{L(\mathbf{x}, \boldsymbol{\lambda}) : \mathbf{x} \in X\} \tag{$D(\boldsymbol{\lambda})$}$$

with the function $L : X \times K^* \to [-\infty, \infty)$ the so-called Lagrangian *function* given by

$$L(\mathbf{x}, \boldsymbol{\lambda}) := f(\mathbf{x}) + \boldsymbol{\lambda}^\mathsf{T}\mathbf{g}(\mathbf{x}). \tag{11}$$

Since the set X is nonempty, it follows that $-\infty \leq \theta(\boldsymbol{\lambda}) < \infty$. The objective function $\theta : K^* \to [-\infty, \infty)$ is called the *Lagrangian dual function* and the vector $\boldsymbol{\lambda}$ the vector of *Lagrangian multipliers* or *dual variables*. To extend the definition of the function θ to $\mathbb{R}^m$, we set $\theta(\boldsymbol{\lambda}) = -\infty$ for every $\boldsymbol{\lambda}$ not belonging to K^*. By the construction of the Lagrangian dual function, the following result is easy to verify. Implicitly it is always assumed that solving optimization problem $(D(\boldsymbol{\lambda}))$ is much more easier for any $\boldsymbol{\lambda}$ than solving the original problem (P).

Lemma 2.1 *The function* $\theta : \mathbb{R}^m \to [-\infty, \infty)$ *is concave satisfying* $\theta(\boldsymbol{\lambda}) \leq \upsilon(P)$ *for every* $\boldsymbol{\lambda} \in \mathbb{R}^m$.

However, even for a primal problem having an optimal solution it might happen that the optimization problem $(D(\boldsymbol{\lambda}))$ does not have an optimal solution for any $\boldsymbol{\lambda} \in K^*$ and so $\theta(\boldsymbol{\lambda}) = -\infty$ for every $\boldsymbol{\lambda} \in \mathbb{R}^m$. This is shown by the following example. If this happens, the convex set $dom(\theta) = \{\boldsymbol{\lambda} \in \mathbb{R}^m : \theta(\boldsymbol{\lambda}) > -\infty\}$ is an empty set and solving optimization problem $(D(\boldsymbol{\lambda}))$ for some given $\boldsymbol{\lambda}$ does not give any information about the value $\upsilon(P)$.

Example 2.1 Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = -x^2$ if $x < 0$ and $f(x) = x$ if $x \geq 0$ and consider the optimization problem:

$$\inf\{f(x) : 0 \leq x < \infty\} = \inf\{f(x) : x \preceq_K 0\}$$

with $K = \mathbb{R}_-$ and $X = \mathbb{R}$. This problem has optimal solution $x = 0$, and the dual cone K^* is given by $K^* = \mathbb{R}_-$. Constructing now the Lagrangian function, we obtain for $\lambda \leq 0$ the Lagrangian dual function:

$$\theta(\lambda) = \inf\{f(x) + \lambda x : x \in \mathbb{R}\}.$$

Since $\lim_{x \downarrow -\infty} f(x) + \lambda x = -\infty$ for every $\lambda \leq 0$, it follows that $\theta(\lambda) = -\infty$ for every $\lambda \leq 0$.

We now discuss how much the Lagrangian relaxation optimization problem $(D(\boldsymbol{\lambda}))$ resembles the original problem (P). A similar result for an optimization problem consisting only of equality constraints or equivalently $K = \{\mathbf{0}\}$ is discussed by Everett (1963).

Lemma 2.2 *If* $\mathbf{x}_{opt}(\boldsymbol{\lambda})$ *is an optimal solution of optimization problem:*

$$\theta(\boldsymbol{\lambda}) = \inf\{f(\mathbf{x}) + \boldsymbol{\lambda}^\mathsf{T}\mathbf{g}(\mathbf{x}) : \mathbf{x} \in X\}$$

for some $\boldsymbol{\lambda} \in K^*$ *with* K *a convex pointed cone, then* $\mathbf{x}_{opt}(\boldsymbol{\lambda})$ *is an optimal solution of the perturbed primal problem:*

$$\inf\{f(\mathbf{x}) : \mathbf{g}(\mathbf{x}) \preceq_K \mathbf{g}(\mathbf{x}_{opt}(\boldsymbol{\lambda})), \mathbf{x} \in X\}.$$

Proof Let $\mathbf{x}_{opt}(\boldsymbol{\lambda})$ be an optimal solution of $\inf\{f(\mathbf{x}) + \boldsymbol{\lambda}^\mathsf{T}\mathbf{g}(\mathbf{x}) : \mathbf{x} \in X\}$ for some $\boldsymbol{\lambda} \in K^*$ and consider an arbitrary $\mathbf{x} \in X$ satisfying $\mathbf{g}(\mathbf{x}) \preceq_K \mathbf{g}(\mathbf{x}_{opt}(\boldsymbol{\lambda}))$. We show that $\mathbf{x}_{opt}(\boldsymbol{\lambda})$ is feasible for the perturbed primal problem and $f(\mathbf{x}_{opt}(\boldsymbol{\lambda})) \leq f(\mathbf{x})$. Since the cone is pointed and hence $\mathbf{0} \in K$, the feasibility is obvious. To show $f(\mathbf{x}_{opt}(\boldsymbol{\lambda})) \leq f(\mathbf{x})$ for every $\mathbf{x} \in X$ and $\mathbf{g}(\mathbf{x}) \preceq_K \mathbf{g}(\mathbf{x}_{opt}(\boldsymbol{\lambda}))$, we observe for $\boldsymbol{\lambda} \in K^*$ that $\boldsymbol{\lambda}^\mathsf{T}(\mathbf{g}(\mathbf{x}_{opt}(\boldsymbol{\lambda}) - \mathbf{g}(\mathbf{x})) \geq 0$. This yields for $\boldsymbol{\lambda} \in K^*$ and $\mathbf{x}_{opt}(\boldsymbol{\lambda})$ an optimal solution of the minimization problem $(D(\boldsymbol{\lambda}))$ that

$$\begin{aligned} f(\mathbf{x}_{opt}(\boldsymbol{\lambda})) &= L(\mathbf{x}_{opt}(\boldsymbol{\lambda}), \boldsymbol{\lambda}) - \boldsymbol{\lambda}^\intercal \mathbf{g}(\mathbf{x}_{opt}(\boldsymbol{\lambda})) \leq L(\mathbf{x}, \boldsymbol{\lambda}) - \boldsymbol{\lambda}^\intercal \mathbf{g}(\mathbf{x}_{opt}(\boldsymbol{\lambda})) \\ &= f(\mathbf{x}) + \boldsymbol{\lambda}^\intercal (\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{x}_{opt}(\boldsymbol{\lambda})) \ \leq \ f(\mathbf{x}) \end{aligned}$$

and we have verified the result. □

To approximate the objective value $\upsilon(P)$ of a primal problem (P) as good as possible by means of the *relaxation approach*, it is therefore by Lemma 2.1 necessary (*maybe not sufficient!*) to maximize the dual function $\theta(\boldsymbol{\lambda})$ over the set K^*. This motivates the following definition.

Definition 2.4 For a primal problem (P), the optimization problem (D) given by

$$\upsilon(D) = \sup\{\theta(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in K^*\} \tag{D}$$

is called the Lagrangian dual of the primal problem (P).

An immediate consequence of Lemma 2.1 is given by the next result and this result is known as *weak duality*.

Lemma 2.3 *It follows that* $-\infty \leq \upsilon(D) \leq \upsilon(P) \leq \infty$.

In the next example we will use the principle of Lagrange to construct the Lagrangian dual problem of a canonical linear programming problem as proposed on page 237 in Bazaraa, Sherali and Shetty (1993). In different books canonical linear programming problems are defined differently, and one can derive similar results for these linear programs.

Example 2.2 Consider the canonical linear programming problem (Bazaraa, Sherali and Shetty, 1993):

$$\upsilon(P) := \inf\{\mathbf{c}^\intercal \mathbf{x} : A\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$$

with A some $m \times n$ matrix, $\mathbf{c} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$. It follows that

$$\upsilon(P) = \inf\{\mathbf{c}^\intercal \mathbf{x} : \mathbf{g}(\mathbf{x}) \preceq_K \mathbf{0}, \mathbf{x} \geq \mathbf{0}\}$$

with $\mathbf{g}(\mathbf{x}) = \mathbf{b} - A\mathbf{x}$ and $K = \mathbb{R}^m_+$. Hence by penalizing the constraints $\mathbf{b} - A\mathbf{x}$, we obtain for every $\boldsymbol{\lambda} \in K^* = \mathbb{R}^m_+$ that the Lagrangian dual function $\theta : \mathbb{R}^m_+ \to [-\infty, \infty)$ equals

$$\theta(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^\intercal \mathbf{b} + \inf\{(\mathbf{c} - A^\intercal \boldsymbol{\lambda})^\intercal \mathbf{x} : \mathbf{x} \geq \mathbf{0}\} = \begin{cases} \boldsymbol{\lambda}^\intercal \mathbf{b} & \text{if } \mathbf{c} - A^\intercal \boldsymbol{\lambda} \geq \mathbf{0} \\ -\infty & \text{otherwise} \end{cases}$$

This yields that the Lagrangian dual problem (D) is given by

$$\sup\{\theta(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in \mathbb{R}^m\} = \sup\{\boldsymbol{\lambda}^\intercal \mathbf{b} : A^\intercal \boldsymbol{\lambda} \leq \mathbf{c}, \boldsymbol{\lambda} \geq \mathbf{0}\},$$

and this coincides with the so-called dual linear programming of the above canonical linear programming problem mostly introduced without any explanation. To generalize this construction to conic convex problems, we first observe that for $X = K$

$$\inf\{(\mathbf{c} - \boldsymbol{\lambda})^\top \mathbf{x} : \mathbf{x} \in X\} = \begin{cases} 0 & \text{if } \mathbf{c} - \boldsymbol{\lambda} \in K^* \\ -\infty & \text{otherwise} \end{cases} \tag{12}$$

Introducing the vector function $\mathbf{g}(\mathbf{x}) = \mathbf{b} - \mathbf{x}$ and using that any linear space is a convex cone and $L^* = L^\top$, it follows that

$$\inf\{\mathbf{c}^\top \mathbf{x} : \mathbf{x} - \mathbf{b} \in L, \mathbf{x} \in K\} = \inf\{\mathbf{c}^\top \mathbf{x} : \mathbf{g}(\mathbf{x}) \preceq_L \mathbf{0}, \mathbf{x} \in K\}.$$

By a similar reasoning as in the first part of this example, using relation (12) we obtain that the Lagrangian dual of the above conic convex programming problem is given by $\sup\{\boldsymbol{\lambda}^\top \mathbf{b} : \mathbf{c} - \boldsymbol{\lambda} \in K^*, \boldsymbol{\lambda} \in L^\perp\}$.

In Example 2.2 it is shown that the well-known LP-dual of a linear programming problem (Bazaraa, Sherali and Shetty, 1993) mostly introduced without any explanation can be constructed using the principle of Lagrange. A direct consequence of Lemma 2.3 for standard linear programming problems is given by the observation that $\upsilon(P) = -\infty$ implies that the dual problem is infeasible. If in Lemma 2.3 we have $\upsilon(D) < \upsilon(P)$, a so-called duality gap exists. As shown by the following example, such a duality gap can even occur for linear programming problems (see also page 60 of Chvatal (1983)).

Example 2.3 Consider the linear programming problem:

$$\inf\{-x_1 - x_2 : x_1 - x_2 \geq 1, -x_1 + x_2 \geq 1, \mathbf{x} \in \mathbb{R}^2_+\}.$$

Clearly this optimization problem is infeasible and so $\upsilon(P) = \infty$. Penalizing the constraints $x_1 - x_2 - 1 \geq 0$ and $-x_1 + x_2 - 1 \geq 0$ using the nonpositive Lagrangian multipliers λ_1 and λ_2, we obtain that the Lagrangian dual function $\theta : R^2_- \to [-\infty, \infty)$ is given by

$$\theta(\boldsymbol{\lambda}) = \inf\{x_1(\lambda_1 - \lambda_2 - 1) + x_2(\lambda_2 - \lambda_1 - 1) : \mathbf{x} \in \mathbb{R}^2_+\}.$$

Observe now for every $\boldsymbol{\lambda} \in \mathbb{R}^2_-$ that $\lambda_1 - \lambda_2 - 1 \geq 0 \Rightarrow \lambda_2 - \lambda_1 \leq -1$ and $\lambda_2 - \lambda_1 - 1 \geq 0 \Rightarrow \lambda_1 - \lambda_2 \leq -1$ and by this observation it follows that $\theta(\boldsymbol{\lambda}) = -\infty$ for every $\boldsymbol{\lambda} \in \mathbb{R}^2_-$ or equivalently $\upsilon(D) = -\infty$.

By Example 2.3, it is clear that one should be careful in applying the principle of Lagrange to a primal problem with an empty feasible region. Fortunately in a lot of practical problems one can decide beforehand that the feasible region $\mathcal{F}$ is nonempty and so we only consider primal problems satisfying $-\infty \leq \upsilon(P) < \infty$. By Lemma 2.3, we obtain that

$$\upsilon(P) = -\infty \Rightarrow \upsilon(P) = \upsilon(D) \tag{13}$$

and so in this case no duality gap exists and every $\boldsymbol{\lambda} \in K^*$ is an optimal dual solution. By relation (13) and Example 2.3, we therefore only need to check for $\upsilon(P)$ finite whether a duality gap exists. For the class of feasible primal problems satisfying $\upsilon(P) > -\infty$, it is now important to decide under which conditions

$$\upsilon(D) = \upsilon(P). \tag{14}$$

An important issue in optimization theory is to determine the largest possible class satisfying the above equality. By identifying such a class, we are able in theory to solve an optimization problem belonging to this class using the dual approach or a combination of the dual and primal approach. The most well-known example of such an approach is the primal dual simplex method in linear programming (Chvatal, 1983). We return to this *identification problem* in the next section. Remember in Example 2.1 we constructed a primal optimization problem not having the property listed in relation (14). In this example an optimal solution of the primal optimization problem exists and $\upsilon(P) > \upsilon(D) = -\infty$ and the Lagrangian dual problem does not have any optimal solution. As shown in the next section, this never happens for linear programming problems or K-convex optimization problems satisfying a so-called regularity condition.

Despite the negative result as shown by Example 2.1, the Lagrangian relaxation approach is often used. This applies in particular to discrete optimization problems since for these problems mostly no strong duality holds. During the search for an upper bound on the objective value $\upsilon(P)$, a number of times the optimization problem $(D(\boldsymbol{\lambda}))$ for different values $\boldsymbol{\lambda}_i \in K^*, i = 1, \ldots, k$ is solved and we obtain the dual feasible optimal solutions $\mathbf{x}_{opt}(\boldsymbol{\lambda}_i), 1 \leq i \leq k$. Sometimes the dual feasible optimal solution $\mathbf{x}_{opt}(\boldsymbol{\lambda}_i)$ might also be primal feasible or equivalently $\mathbf{g}(\mathbf{x}_{opt}(\boldsymbol{\lambda}_i)) \preceq_K \mathbf{0}$. However, in most cases this does not happen and so $\mathbf{x}_{opt}(\boldsymbol{\lambda}_i)$ is not primal feasible. If $\mathbf{x}_{opt}(\boldsymbol{\lambda}_i)$ is not primal feasible, it is sometimes possible to construct using the primal nonfeasible $\mathbf{x}_{opt}(\boldsymbol{\lambda}_i), i = 1 \ldots, k$ a primal feasible $\mathbf{x}_i$ and we obtain a feasible candidate of our primal problem. The heuristic procedure that converts a primal nonfeasible $\mathbf{x}_{opt}(\boldsymbol{\lambda})$ into a primal feasible $\mathbf{x}$ is called a *Lagrangian heuristic*. On the other hand, it is also possible to generate for most problems a feasible solution by means of a so-called *primal heuristic*. Such a primal heuristic is mostly an ad hoc procedure particularly suited for the specific problem and it does not use dual information as done by a Lagrangian heuristic. We consider some examples of this approach in location and transportation problems. One might now wonder whether it is possible to say something about the quality of the generated primal feasible solution by this Lagrangian heuristic with respect to the optimal objective value of the primal problem. This motivates the next definition.

Definition 2.5 For any $\epsilon > 0$, a vector $\mathbf{x} \in \mathbb{R}^n$ is called an ϵ-optimal solution of the optimization problem $\upsilon(P) = \inf\{f(\mathbf{x}) : \mathbf{x} \in \mathcal{F}\}$ if $f(\mathbf{x}) \leq \upsilon(P) + \epsilon$ and $\mathbf{x} \in \mathcal{F}$.

The next result discusses the quality of a primal feasible ϵ-optimal solution of the optimization problem $(D(\boldsymbol{\lambda}))$. The assumption of the next lemma automatically implies that $\upsilon(P)$ is finite.

Lemma 2.4 *If for some $\boldsymbol{\lambda} \in K^*$ there exists a primal feasible $\mathbf{x}_0$ with $\mathbf{x}_0$ an ϵ-optimal solution of optimization problem:*

$$\theta(\boldsymbol{\lambda}) = \inf\{f(\mathbf{x}) + \boldsymbol{\lambda}^\top \mathbf{g}(\mathbf{x}) : \mathbf{x} \in X\},$$

then

$$\upsilon(P) \leq f(\mathbf{x}_0) \leq \upsilon(P) + \epsilon - \boldsymbol{\lambda}^\top \mathbf{g}(\mathbf{x}_0) < \infty.$$

Proof Since $\mathbf{x}_0$ is primal feasible, we obtain by the weak duality result given by Lemma 2.3 that

$$\begin{aligned}\upsilon(P) &\leq f(\mathbf{x}_0) = L(\mathbf{x}_0, \boldsymbol{\lambda}) - \boldsymbol{\lambda}^\top \mathbf{g}(\mathbf{x}_0)\\ &\leq \theta(\boldsymbol{\lambda}) + \epsilon - \boldsymbol{\lambda}^\top \mathbf{g}(\mathbf{x}_0) \leq \upsilon(D) + \epsilon - \boldsymbol{\lambda}^\top \mathbf{g}(\mathbf{x}_0)\\ &\leq \upsilon(P) + \epsilon - \boldsymbol{\lambda}^\top \mathbf{g}(\mathbf{x}_0)\end{aligned}$$

and this shows the result. □

The next result is partly an easy implication of Lemma 2.4.

Lemma 2.5 *The following conditions are equivalent:*

1. The vector $\mathbf{x}_0$ is primal feasible ($\mathbf{g}(\mathbf{x}_0) \preceq_K \mathbf{0}$), $\boldsymbol{\lambda}_0$ is dual feasible ($\boldsymbol{\lambda}_0 \in K^*$), and $f(\mathbf{x}_0) = \theta(\boldsymbol{\lambda}_0)$.
2. The vector $\mathbf{x}_0$ is an optimal solution of optimization problem (P), the Lagrangian multiplier $\boldsymbol{\lambda}_0 \in K^*$ is an optimal solution of the Lagrangian dual problem (D), and $\upsilon(P) = \upsilon(D)$.
3. The primal feasible solution $\mathbf{x}_0$ is an optimal solution of $(D(\boldsymbol{\lambda}_0))$ with $\boldsymbol{\lambda}_0 \in K^*$ and $\theta(\boldsymbol{\lambda}_0)$ finite and $\boldsymbol{\lambda}_0^\top \mathbf{g}(\mathbf{x}_0) = 0$.

Proof To prove the implication $1 \to 3$, we observe

$$f(\mathbf{x}_0) = \theta(\boldsymbol{\lambda}_0) \leq f(\mathbf{x}_0) + \boldsymbol{\lambda}_0^\top \mathbf{g}(\mathbf{x}_0) \tag{15}$$

and this shows $\boldsymbol{\lambda}_0^\top \mathbf{g}(\mathbf{x}_0) \geq 0$. Since $\mathbf{g}(\mathbf{x}_0) \preceq_K \mathbf{0}$ and $\boldsymbol{\lambda}_0 \in K^*$, this yields $\boldsymbol{\lambda}_0^\top \mathbf{g}(\mathbf{x}_0) = 0$. By relation (15), we obtain $f(\mathbf{x}_0) + \boldsymbol{\lambda}_0^\top \mathbf{g}(\mathbf{x}_0) = \theta(\boldsymbol{\lambda}_0)$ and so $\mathbf{x}_0$ is an optimal solution of $\theta(\boldsymbol{\lambda}_0)$ showing part 3. The implication $3 \to 2$ is a direct consequence of Lemma 2.4, while the implication $2 \to 1$ is trivial.

The condition $\boldsymbol{\lambda}_0^{\mathsf{T}}\mathbf{g}(\mathbf{x}_0) = 0$ mentioned in part 3 of Lemma 2.5 is called the *complementary slackness condition*. Observe if $K = \{\mathbf{0}\}$ this complementary slackness condition is automatically satisfies for $\mathbf{x}_0$ primal feasible. Unfortunately in most integer linear programming problems, there exists a duality gap unless the matrix describing the feasible region has a special structure like total unimodularity (Nemhauser and Wolsey, 1988) and so in general condition 3 in Lemma 2.5 does not hold. Observe that we are interested in condition 3 since our procedure is to solve optimization problems $(D(\boldsymbol{\lambda}))$ for different values of $\boldsymbol{\lambda} \in K^*$. If during this procedure we find a vector $\mathbf{x}_0$ satisfying condition 3 of Lemma 2.5, then this vector $\mathbf{x}_0$ is an optimal solution of our problem (P). In Lemma 2.5 we assume that no duality gap exists and so it is important to know for which optimization problems such a result holds. This is discussed in the next section.

2.1 K-Convex Minimization Problems and Strong Duality

In this section we consider a class of optimization problems (P) for which under a so-called regularity condition on the feasible region no duality gap exists. To identify these classes of optimization problems, we introduce the following class of vector valued functions (Giannessi , 1984; Wolkowicz, 1981). Observe in this definition we do not assume that K is a convex pointed cone.

Definition 2.6 Let $X \subseteq \mathbb{R}^n$ be a nonempty convex set and $K \subseteq \mathbb{R}^m$ is a nonempty convex cone. The vector valued function $\mathbf{g} : X \to \mathbb{R}^m$ is called K-convex if and only if

$$\mathbf{g}(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \preceq_K \alpha\mathbf{g}(\mathbf{x}_1) + (1-\alpha)\mathbf{g}(\mathbf{x}_2) \tag{16}$$

for every $0 < \alpha < 1, \mathbf{x}_1, \mathbf{x}_2 \in X$ and $\preceq_K$ the transitive ordering introduced in relation (2).

If $K = \mathbb{R}^m_+$ (and so a convex pointed cone), then K-convexity reduces to the classical definition of convexity. Also it is easy to check that any affine function $\mathbf{g}(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ is K-convex for any nonempty convex pointed cone K. To relate the above definition of a K-convex function for $K \subseteq \mathbb{R}^m$ a convex pointed cone to the modern definition of a convex function (Rockafellar, 1970), we introduce for any convex pointed cone $K \subseteq \mathbb{R}^m$ and any vector function $\mathbf{g} : \mathbb{R}^n \to (\mathbb{R} \cup \{\infty\})^m$ the so-called epigraph $epi_K(\mathbf{g})$ given by

$$epi_K(\mathbf{g}) := \{(\mathbf{x}, \mathbf{r}) : \mathbf{g}(\mathbf{x}) \preceq_K \mathbf{r}\} \subseteq \mathbb{R}^{n+m}.$$

Since K is a convex pointed cone, we know that the ordering $\preceq_K$ is a partial ordering. It is shown in (Frenk and Kassay, 2007) for K a convex pointed cone and $\mathbf{h} : X \to \mathbb{R}^m$ some given vector function on a nonempty domain $X \subseteq \mathbb{R}^n$ that the set $epi_K(\mathbf{h})$ is a convex set if and only if X is a convex set and $\mathbf{h} : X \to \mathbb{R}^m$ is K-convex and so for K a convex pointed cone the old definition of convexity coincides with the modern definition of convexity.

Definition 2.7 The optimization problem $\inf\{f(\mathbf{x}) : \mathbf{g}(\mathbf{x}) \preceq_K \mathbf{0}, \mathbf{x} \in X\}$ with K a proper convex cone and nonempty feasible region $\mathcal{F}$ is called a K-convex minimization problem if the set $X \subseteq \mathbb{R}^n$ is convex, the function $f : X \to \mathbb{R}$ is convex, and the function $\mathbf{g} : X \to \mathbb{R}^m$ is K-convex.

In the above definition we could also have assumed that K is a convex cone. However, using this alternative definition we only added the case $K = \mathbb{R}^n$, and since in this case $K^* = \{\mathbf{0}\}$, the Lagrangian dual function reduces to the original problem. At the same time, we could have imposed that the cone K is pointed (mostly assumed for K-convex optimization problems) but this condition only implies that the ordering $\preceq_K$ is a partial ordering and this property is not necessary for achieving strong duality for K-convex optimization problems. Since by relation (16) the function $\mathbf{x} \mapsto \boldsymbol{\lambda}^\top \mathbf{g}(\mathbf{x})$ is convex for $\boldsymbol{\lambda} \in K^*$ and any K-convex function $\mathbf{g} : X \to \mathbb{R}^m$, it follows for any K-convex minimization problem that the optimization problem $(D(\boldsymbol{\lambda}))$ is a convex minimization problem for every $\boldsymbol{\lambda} \in K^*$. To show that for the class of K-convex minimization problems satisfying some additional constraint qualification the strong duality property holds, we need the following definition.

Definition 2.8

1. If $S \subseteq \mathbb{R}^n$ is nonempty, then the set $\mathrm{aff}(S)$ is the smallest affine set containing S. This set is called the *affine hull* of S.
2. If $S \subseteq \mathbb{R}^n$ and $\mathbf{y} \in S$, then the point $\mathbf{y}$ is called a *relative interior point* of the set S if there exists some $\epsilon > 0$ such that the intersection of $\mathrm{aff}(S)$ and the set $\mathbf{y} + \epsilon\mathcal{B}$ with $\mathcal{B}$ the closed Euclidean ball given by

$$\mathcal{B} := \{\mathbf{x} \in \mathbb{R}^n : \| \mathbf{x} \| \leq 1\}$$

   and $\| \, . \, \|$ the Euclidean norm is contained in S.
3. The *relative interior* $\mathrm{ri}(S)$ of a set $S \subseteq R^n$ is the collection of all relative interior points of the set S.

By the generalization of inequalities to cones, we needed the concept of K-convex functions. However, as for classical convex programming problems with equalities the property of K-convexity is stronger than is needed for the proof of the no-duality gap for K-convex optimization problems. What we need for this proof is the following implication of K-convex functions, which is weaker than K-convexity.

Lemma 2.6 *If $K \subseteq \mathbb{R}^m$ is a nonempty convex cone and the function $\mathbf{g} : X \to R^m$ is K-convex, then the set $\mathbf{g}(X) + ri(K)$ is nonempty and convex.*

Proof Since the set $K \subseteq \mathbb{R}^m$ is a nonempty convex cone, it is well-known (Frenk and Kassay, 2005; Rockafellar, 1970) that $ri(K)$ is also a nonempty convex cone and so the set $\mathbf{g}(X) + ri(K)$ is nonempty. To check that the set $\mathbf{g}(X) + ri(K)$ is convex, we need to verify for any $\mathbf{x}_1, \mathbf{x}_2$ belonging to X and $0 < \alpha < 1$ that

$$\alpha(\mathbf{g}(\mathbf{x}_1) + ri(K)) + (1-\alpha)(\mathbf{g}(\mathbf{x}_2) + ri(K)) \subseteq \mathbf{g}(X) + ri(K).$$

By the convexity of the set $ri(K)$, it follows that $\alpha ri(K) + (1-a)ri(K) \subseteq ri(K)$ and this implies that

$$\alpha(\mathbf{g}(\mathbf{x}_1) + ri(K)) + (1-\alpha)(\mathbf{g}(\mathbf{x}_2) + ri(K)) \subseteq \alpha\mathbf{g}(\mathbf{x}_1) + (1-\alpha)\mathbf{g}(\mathbf{x}_2) + ri(K).$$

By the K-convexity of the function $\mathbf{g}$ and relation (2), we know that

$$\alpha\mathbf{g}(\mathbf{x}_1) + (1-\alpha)\mathbf{g}(\mathbf{x}_2) \in \mathbf{g}(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) + K. \tag{17}$$

Since it is well-known that $K + ri(K) \subseteq ri(K)$ (Frenk and Kassay, 2005; Rockafellar, 1970) and the set X is convex, we obtain by relation (17) that

$$\begin{aligned}\alpha(\mathbf{g}(\mathbf{x}_1) + ri(K)) + (1-\alpha)(\mathbf{g}(\mathbf{x}_2) + ri(K)) &\subseteq \mathbf{g}(X) + K + ri(K)\\ &\subseteq \mathbf{g}(X) + ri(K)\end{aligned}$$

and we have verified the desired result. □

We now need a separation result to prove that under some regularity condition no duality gap exists for the class of K-convex minimization problems. However, we did not discuss what we mean by separation and so we list the following definition (Frenk and Kassay, 2005; Rockafellar, 1970).

Definition 2.9 If $C \subseteq \mathbb{R}^n$ is a nonempty set and $\mathbf{0} \in R^n$ does not belong to C, then the sets C and $\{\mathbf{0}\}$ are said to be properly separated if there exists some vector $\boldsymbol{\sigma}$ belonging to $\mathbb{R}^n$ satisfying

$$\inf\{\boldsymbol{\sigma}^\top\mathbf{x} : \mathbf{x} \in C\} \geq 0$$

and for some $\mathbf{x}$ belonging to C it follows that $\boldsymbol{\sigma}^\top\mathbf{x} > 0$. The vector $\boldsymbol{\sigma}$ is called the normal vector of the separating hyperplane between the sets C and $\{\mathbf{0}\}$.

By Definition 2.9, it follows that $\boldsymbol{\sigma}$ is not equal to the zero vector. The next theorem mentions an important separation result for an arbitrary proper convex set (Frenk and Kassay, 2005; Hiriart-Urruty and Lemaréchal, 2013; Rockafellar, 1970).

Theorem 2.1 *If $C \subseteq \mathbb{R}^n$ is a nonempty convex set and $\mathbf{0} \in \mathbb{R}^n$ does not belong to C, then the sets C and $\{\mathbf{0}\}$ can be properly separated. Moreover, the normal vector $\boldsymbol{\sigma}$ of the separating hyperplane can be chosen to belong to the affine hull $aff(C)$ of the set C.*

By the above separation result, it is possible to show that any K-convex minimization problem satisfying an additional constraint qualification has no duality gap.

Theorem 2.2 *If the K-convex minimization problem satisfies $-\infty < \upsilon(P) < \infty$ and the vector $\mathbf{0}$ belongs to $ri(\mathbf{g}(X) + K)$, then the Lagrangian dual (D) of the K-convex optimization problem given by* $\sup\{\theta(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in K^*\}$ *has an optimal solution and the duality gap equals zero.*

Proof Without loss of generality, we may assume that $\upsilon(P) = 0$. Since optimization problem (P) is a K-convex optimization problem, we obtain that the function $\mathbf{h} : X \to R^{m+1}$ given by

$$\mathbf{h}(\mathbf{x}) = (f(\mathbf{x}), \mathbf{g}(\mathbf{x}))$$

is K_e-convex with $K_e = [0, \infty) \times K$. This implies by Lemma 2.6 that the set $\mathbf{h}(X) + ri(K_e)$ is convex. If $\mathbf{0}$ belongs to the convex set $\mathbf{h}(X) + ri(K_e)$, then there exists some $\mathbf{x}_0 \in X$ satisfying $\mathbf{h}(\mathbf{x}_0) \in -ri(K_e)$. Since $ri(K_e) = (0, \infty) \times ri(K)$, we obtain $f(\mathbf{x}_0) < 0$ and $\mathbf{g}(\mathbf{x}_0) \preceq_K \mathbf{0}$. This contradicts our assumption $\upsilon(P) = 0$ and so $\mathbf{0} \notin \mathbf{h}(X) + ri(K_e)$. Applying now Theorem 2.1, there exists some nonzero vector $\boldsymbol{\sigma} = (\mu, \boldsymbol{\lambda}_0)$ satisfying for every $\mathbf{x} \in X$ and $\mathbf{k}_e = (r, \mathbf{k})$ with $r > 0$ and $\mathbf{k} \in ri(K)$

$$\boldsymbol{\sigma}^\top(\mathbf{h}(\mathbf{x}) + \mathbf{k}_e) = \mu(f(\mathbf{x}) + r) + \boldsymbol{\lambda}_0^\intercal(\mathbf{g}(\mathbf{x}) + \mathbf{k}) \geq 0 \tag{18}$$

with a strict inequality for at least one $\mathbf{x}_0 \in X, r_0 > 0$ and $\mathbf{k}_0 \in ri(K)$. Also by Theorem 2.1, the component $\boldsymbol{\lambda}_0$ of this vector belongs to the set $aff(\mathbf{g}(X)+K)$. We now show that $\mu > 0$. Since $ri(K_e)$ is a convex cone, this implies for every $\alpha > 0$ and $\mathbf{k}_e = (r, \mathbf{k}) \in ri(K_e)$ that $\alpha\mathbf{k}_e \in ri(K_e)$ and so we obtain by relation (18) for $\mathbf{x} \in X$

$$\boldsymbol{\sigma}^\top(\mathbf{h}(\mathbf{x}) + \alpha\mathbf{k}_e) \geq 0. \tag{19}$$

This shows by contradiction for $\mathbf{x} \in X$ fixed and letting α tend to infinity in relation (19) that $\boldsymbol{\sigma}^\top\mathbf{k}_e \geq 0$ for every $\mathbf{k}_e \in ri(K_e)$ and so $\boldsymbol{\sigma}$ belongs to $(ri(K_e))^* = K_e^* = [0, \infty) \times K^*$. Hence it follows that $\mu \geq 0$ and if we assume that $\mu = 0$ or $\boldsymbol{\sigma} = (0, \boldsymbol{\lambda}_0)$ it follows again by relation (18) that

$$\boldsymbol{\lambda}_0^\intercal(\mathbf{g}(\mathbf{x}) + \mathbf{k}) \geq 0 \tag{20}$$

for every $\mathbf{x} \in X$ and $\mathbf{k} \in ri(K)$. Since by assumption $\mathbf{0}$ belongs to $ri(\mathbf{g}(X) + K)$ and $\boldsymbol{\lambda}_0 \in aff(\mathbf{g}(X) + K)$, one can find some $\epsilon > 0$ such that $-\epsilon^\intercal\boldsymbol{\lambda}_0 \in \mathbf{g}(X) + K$ and this implies by relation (20) that

$$-\epsilon\boldsymbol{\lambda}_0^\intercal\boldsymbol{\lambda}_0 = -\epsilon \parallel \boldsymbol{\lambda}_0 \parallel^2 \geq 0$$

or equivalently $\boldsymbol{\lambda}_0 = \mathbf{0}$. Hence the normal vector $\boldsymbol{\sigma}$ equals the null vector and this contradicts Theorem 2.1. Therefore $\mu > 0$ and applying now relation (19), we obtain for any $\alpha > 0$, $r \in ri(K)$, $\mathbf{x} \in X$ and $\mathbf{k} \in K$ that

$$(f(\mathbf{x}) + \alpha r) + \mu^{-1}\boldsymbol{\lambda}_0^{\mathsf{T}}(\mathbf{g}(\mathbf{x}) + \mathbf{k}) \geq 0. \tag{21}$$

Since for $\mathbf{k} \in K$ it follows that $\alpha\mathbf{k} \in K$ for any $\alpha > 0$, we can replace $\mathbf{k}$ in relation (21) by $\alpha\mathbf{k}$ and letting α tend to zero we conclude from the same relation that for every $\mathbf{x} \in X$

$$f(\mathbf{x}) + \mu^{-1}\boldsymbol{\lambda}_0^{\mathsf{T}}\mathbf{g}(\mathbf{x}) \geq 0$$

with $\mu^{-1}\boldsymbol{\lambda}_0 \in K^*$. This shows $\theta(\mu^{-1}\boldsymbol{\lambda}_0) \geq 0$ and applying the weak duality result yields $0 = \upsilon(D)$ and $\mu^{-1}\boldsymbol{\lambda}_0$ is an optimal dual solution. $\square$

The additional constraint qualification in Theorem 2.2 is called a *Slater-type condition* and as shown in the proof of Theorem 2.2 this condition is sufficient to guarantee that an optimal solution of the Lagrangian dual problem (D) of a general K-convex minimization problem exists. The constraint qualification is slightly stronger than the feasibility condition since

$$\mathcal{F} = \{\mathbf{x} \in X : \mathbf{g}(\mathbf{x}) \preceq_K \mathbf{0}\} \text{ nonempty } \Leftrightarrow \mathbf{0} \in \mathbf{g}(X) + K.$$

and it means that the feasible region has a nonempty relative interior. At the same time, it might happen that the primal problem does not have an optimal solution. Actually the *Slater-type condition* is a sufficient condition to guarantee the existence of an optimal solution of the dual problem and this condition guarantees that an optimal solution of the dual problem is contained in a compact subset of K^*. Due to the convexity-type assumptions on the bifunction $L : X \times K^*$ given by $L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^{\mathsf{T}}\mathbf{g}(\mathbf{x})$ as listed in relation (11), the strong duality follows immediately as a corollary of Sion's minmax theorem already proven in 1928 by Neumann (1928) using Brouwer's fixed point theorem. This result was later rediscovered by Sion (1958) and proved using the Knaster–Kuratowski–Mazurkiewicz (KKM) lemma (George Xian-Zhi Yuan, 1999). This lemma is equivalent to Brouwer's fixed point theorem. This means that the Lagrangian relaxation approach can be seen as a particular instance of a noncooperative two-person game in which nature plays again the decision maker. The action set of nature is given by K^*, and the action set of the decision maker is given by $X \subseteq \mathbb{R}^n$. The strong duality result states now that the value of this particular two-person game exists, and Sion's minmax theorem shows under which sufficient conditions on the bifunction L this game has a value. In fact the strong duality result is an immediate consequence of the property that a convex set is connected and so it is a topological result. For more details on how to prove a generalization of Sion's minmax theorem using a more elementary approach by contradiction and Joos method and the obvious fact that a convex set is connected, the reader should consult (Frenk and Kassay, 2006). Connectedness of a set means that any two points in such a set can be connected by a continuous curve contained within this set.

For linear programming problems given by $\inf\{\mathbf{c}^\top \mathbf{x} : A\mathbf{x} \le \mathbf{b}, \mathbf{x} \in R^n\}$, one can verify by inspection of the last lemma in the simplex method (Chvatal, 1983; Saigal, 1997) used to solve the above linear programming problem that the Lagrangian dual of this linear programming problem has an optimal solution and no duality gap exists. *This means for linear programming problems that we only need to assume υ (P) is finite and so no Slater-type condition is needed.* Next to linear programming another class of important optimization problems is represented by the function $d_{cl(K)} : \mathbb{R}^n \to \mathbb{R}$ with

$$\begin{aligned} d_{cl(K)}(\mathbf{x}) &:= \inf\{\tfrac{1}{2}\|\mathbf{x}-\mathbf{y}\|^2 : \mathbf{y} \in cl(K)\} \\ &= \inf\{\tfrac{1}{2}\|\mathbf{x}-\mathbf{y}\|^2 : -\mathbf{y} \preceq_{cl(K)} \mathbf{0}\} \end{aligned}$$

and $K \subseteq \mathbb{R}^n$ a nonempty proper convex cone, $\|.\|$ denoting the Euclidean norm, and $cl(K)$ the closure of this convex cone. This optimization problem denotes the orthogonal projection of a vector on a given convex cone, and such a problem shows up in the subgradient method. It is easy to see that

$$d_{cl(K)}(\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} \in cl(K) \tag{22}$$

and for any fixed $\mathbf{x}$ the above optimization problem is clearly a K-convex programming problem (even convex!) with $f(\mathbf{y}) = \frac{1}{2}\|\mathbf{x}-\mathbf{y}\|^2$ and $\mathbf{g}(\mathbf{y}) = -\mathbf{y}$ and $X = \mathbb{R}^n$. It follows trivially that $\mathbf{g}(\mathbb{R}^n) = \mathbb{R}^n$ and so the Slater-type condition reduces to $\mathbf{0} \in ri(\mathbb{R}^n + cl(K)) = ri(\mathbb{R}^n) = \mathbb{R}^n$, which is automatically satisfied. Since $(cl(K))^* = K^*$ and applying Theorem 2.2 we obtain that

$$d_{cl(K)}(\mathbf{x}) = \sup\{\theta(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in K^*\}$$

with

$$\theta(\boldsymbol{\lambda}) = \inf\left\{\frac{1}{2}\|\mathbf{x}-\mathbf{y}\|^2 - \boldsymbol{\lambda}^\top \mathbf{y} : \mathbf{y} \in \mathbb{R}^n\right\}.$$

Since the function $\mathbf{y} \to \frac{1}{2}\|\mathbf{x}-\mathbf{y}\|^2 - \boldsymbol{\lambda}^\top \mathbf{y}$ is strictly convex and differentiable on $\mathbb{R}^n$, an optimal solution for the optimization problem $(D(\boldsymbol{\lambda})$ is obtained by solving the necessary and sufficient first-order conditions. This implies with $\mathbf{y}_{opt}(\boldsymbol{\lambda})$ denoting the optimal feasible solution of the optimization problem $(D(\boldsymbol{\lambda}))$ that

$$-(\mathbf{x} - \mathbf{y}_{opt}(\boldsymbol{\lambda})) - \boldsymbol{\lambda} = \mathbf{0} \Leftrightarrow \mathbf{y}_{opt}(\boldsymbol{\lambda}) = \mathbf{x} + \boldsymbol{\lambda}$$

and so $\theta(\boldsymbol{\lambda}) = \frac{1}{2}\|\boldsymbol{\lambda}\|^2 - \boldsymbol{\lambda}^\top(\mathbf{x} + \boldsymbol{\lambda}) = -\frac{1}{2}\|\boldsymbol{\lambda}\|^2 - \boldsymbol{\lambda}^\top\mathbf{x}$. This finally shows

$$d_{cl(K)}(\mathbf{x}) = \max\left\{-\frac{1}{2}\|\boldsymbol{\lambda}\|^2 - \boldsymbol{\lambda}^\top\mathbf{x} : \boldsymbol{\lambda} \in K^*\right\}. \tag{23}$$

and using relation (23) the next so-called important bipolar theorem can be verified.

Theorem 2.3 *If $K \subseteq \mathbb{R}^n$ is a nonempty convex cone, then it follows that $cl(K) = K^{**}$ with $K^{**} := (K^*)^*$.*

Proof If $\mathbf{x}$ belongs to K, then by the definition of K^* it follows for every $\boldsymbol{\lambda} \in K^*$ that $\boldsymbol{\lambda}^\intercal\mathbf{x} \geq 0$ and this shows that $\mathbf{x}$ belongs to K^{**}. Hence $K \subseteq K^{**}$ and since K^{**} is closed we obtain $cl(K) \subseteq K^{**}$. To prove the reverse inclusion, let $\mathbf{x}$ belong to K^{**}. By the definition of K^{**}, it follows that $\mathbf{x}^\top\boldsymbol{\lambda} \geq 0$ for every $\boldsymbol{\lambda} \in K^*$ and this shows $-\frac{1}{2}\|\boldsymbol{\lambda}\|^2 - \boldsymbol{\lambda}^\top\mathbf{x} \leq -\frac{1}{2}\|\boldsymbol{\lambda}\|^2$ for every $\boldsymbol{\lambda} \in K^*$. Since $\mathbf{0}$ belongs to K^*, this implies by relation (23) that

$$0 \leq d_{cl(K)}(\mathbf{x}) \leq \max\left\{-\frac{1}{2}\|\boldsymbol{\lambda}\|^2 : \boldsymbol{\lambda} \in K^*\right\} = 0$$

and hence by relation (22) we obtain $\mathbf{x} \in cl(K)$. □

A very special case of the bipolar theorem is the so-called lemma of Farkas known in linear programming, and this result can be used to show (Saigal, 1997) that any linear programming problem with a finite optimal objective function value satisfies the strong duality property. By the above bipolar theorem, it is easy to give a univariate characterization of K-convex functions in case K is a closed convex cone.

Lemma 2.7 *For a closed convex cone K, it follows that the function $\mathbf{x} \to \boldsymbol{\lambda}^\intercal\mathbf{h}(\mathbf{x})$ is convex on the convex set X for every $\boldsymbol{\lambda} \in K^*$ if and only if $\mathbf{h}$ is K-convex.*

An important implication of Theorem 2.2 and Lemma 2.7 is given by the following result. Observe a strictly convex function is a convex function with the inequalities in the definition replaced by strict inequalities for any strict convex combination of any two vectors.

Lemma 2.8 *If the objective function f in the K-convex optimization problem*

$$\inf\{f(\mathbf{x}) : \mathbf{g}(\mathbf{x}) \preceq_K \mathbf{0}, \mathbf{x} \in X\}$$

is strictly convex and $\mathbf{0} \in ri(\mathbf{g}(X) + K)$ and this optimization problem has an optimal solution, then the Lagrangian dual problem has an optimal solution $\boldsymbol{\lambda}_{opt}$ and the optimal solution $\mathbf{x}_{opt}(\boldsymbol{\lambda}_{opt})$ of the optimization problem $D(\boldsymbol{\lambda}_{opt})$ exists and is unique and coincides with the unique optimal solution of the primal problem.

Proof Since it is assumed that the primal K-convex optimization problem (P) has an optimal solution and the objective function f is strictly convex, the optimal solution of the problem (P) denoted by $\mathbf{x}_{opt}$ is unique. By Theorem 2.2, we know that an optimal solution $\boldsymbol{\lambda}_{opt} \in K^*$ of the Lagrangian dual problem exists and it satisfies

$$f(\mathbf{x}_{opt}) = \upsilon(P) = \upsilon(D) = \theta(\boldsymbol{\lambda}_{opt}).$$

This shows that condition 1 of Lemma 2.5 is satisfied and again by Lemma 2.5 the primal feasible solution $\mathbf{x}_{opt}$ is an optimal solution of $D(\boldsymbol{\lambda}_0)$. Since by Lemma 2.7 we obtain that the function $\mathbf{x} \to \boldsymbol{\lambda}_{opt}^{\intercal}\mathbf{g}(\mathbf{x})$ is convex on X and hence the function $\mathbf{x} \to f(\mathbf{x}) + \boldsymbol{\lambda}^{\intercal}\mathbf{g}(\mathbf{x})$ is strictly convex, the optimization problem $D(\boldsymbol{\lambda}_{opt})$ has a unique optimal solution and since $\mathbf{x}_{opt}$ is such an optimal solution we obtain the desired result.

From a computational point of view, the following theorem for K-convex optimization problems is important and is a direct consequence of Lemma 2.5 and Theorem 2.2.

Theorem 2.4 *If the optimization problem (P) is a K-convex minimization problem and $\mathbf{0}$ belongs to $ri(\mathbf{g}(X) + K)$, then the vector $\mathbf{x}_0$ is an optimal solution of (P) if and only if for some optimal dual variable $\boldsymbol{\lambda}_0 \in K^*$ it follows that*

$$\mathbf{x}_0 = argmin\{f(\mathbf{x}) + \boldsymbol{\lambda}_0^{\intercal}\mathbf{g}(\mathbf{x}) : \mathbf{x} \in X\}, \boldsymbol{\lambda}_0^{\intercal}\mathbf{g}(\mathbf{x}_0) = 0 \textit{ and } -\mathbf{g}(\mathbf{x}_0) \in K.$$

Proof Since (P) is a K-convex optimization problem and $\mathbf{0} \in ri(\mathbf{g}(X) + K)$, it follows by Theorem 2.2 that there exists an optimal dual solution $\boldsymbol{\lambda}_0$ and $\upsilon(P) = \upsilon(D)$. Hence condition 2 of Lemma 2.5 is satisfied and this yields by $2 \to 3$ of Lemma 2.5 that $\mathbf{x}_0$ satisfies the above system of equations. The reverse implication is also a direct consequence of Lemma 2.5.

In the next result the so-called Karush–Kuhn–Tucker conditions for K-convex programs with differentiable functions are given and this result is an immediate consequence of Theorem 2.4. It also shows that the *Karush–Kuhn–Tucker vector* is an optimal solution of the Lagrangian dual problem.

Theorem 2.5 *If the optimization problem (P) is a K-convex optimization problem consisting of differentiable functions and the convex set X equals $\mathbb{R}^n$, then under the regularity condition $\mathbf{0} \in ri(\mathbf{g}(X) + K)$ the vector $\mathbf{x}_0$ is an optimal solution of (P) if and only if for some optimal dual variable $\boldsymbol{\lambda} \in K^*$ it follows that*

$$\nabla f(\mathbf{x}_0) + \sum_{i=1}^{n} \lambda_i \nabla g_i(\mathbf{x}_0) = \mathbf{0}, \boldsymbol{\lambda}^{\intercal}\mathbf{g}(\mathbf{x}_0) = 0 \textit{ and } -\mathbf{g}(\mathbf{x}_0) \in K.$$

Proof By Theorem 2.4, the vector $\mathbf{x}_0$ is an optimal solution of (P) if and only if for some optimal dual variable $\boldsymbol{\lambda} \in K^*$ we obtain

$$\mathbf{x}_0 = \text{argmin}\{f(\mathbf{x}) + \boldsymbol{\lambda}^{\intercal}\mathbf{g}(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\}, \boldsymbol{\lambda}^{\intercal}\mathbf{g}(\mathbf{x}_0) = 0, -\mathbf{g}(\mathbf{x}_0) \in K.$$

Since the optimization problem (P) is a differentiable K-convex optimization problem, it follows that the function $\mathbf{x} \to f(\mathbf{x}) + \boldsymbol{\lambda}^{\mathsf{T}}\mathbf{g}(\mathbf{x})$ is convex and differentiable on $\mathbb{R}^n$ and so the result follows by the sufficient and necessary first-order conditions for an unconstrained convex optimization problem.

If we consider a K-convex optimization problem satisfying Slater's constraint qualification and having an optimal solution and we identified an optimal dual solution, then by a similar proof as in the first part of Lemma 2.8 we can only show that the optimal solution set

$$\arg\min\{f(\mathbf{x}) + \boldsymbol{\lambda}^{\mathsf{T}}\mathbf{g}(\mathbf{x}) : \mathbf{x} \in X\}$$

contains a primal feasible optimal solution satisfying the complementary slackness condition. However, *the optimal solution set of the optimization problem* $(D(\lambda))$ *can contain more than one element and so it might be computationally difficult to identify this primal feasible element satisfying the complementary slackness condition.* The implication of Theorem 2.5 is that for differentiable K-convex minimization problems satisfying the Slater-type condition the Karush–Kuhn–Tucker vector $\boldsymbol{\lambda}$ completely coincide with the optimal dual variables and so the Karush–Kuhn–Tucker conditions for K-convex programs are actually duality results. These conditions are also known in linear programming as *primal–dual relations*. We show in the next section the relation between Lagrangian relaxation and linear programming relaxation applied to an integer linear programming problem.

2.2 On Integer Linear Programming and Lagrangian and LP Relaxations

In this section we consider for D some $p \times n$ matrix and A some $m \times n$ matrix the integer linear programming problem:

$$\inf\{\mathbf{c}^{\top}\mathbf{x} : \mathbf{x} \in \mathcal{F}_{INT}\} \qquad (INT)$$

with

$$\mathcal{F}_{INT} := \{\mathbf{x} \in \mathbb{Z}_{+}^{n} : D\mathbf{x} \leq \mathbf{d}, A\mathbf{x} = \mathbf{b}\}$$

a nonempty feasible region and apply the Lagrangian relaxation approach to this problem. In the above representation it is assumed that the optimization problem $\inf\{\mathbf{h}^{\top}\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathbb{Z}_{+}^{n}\}$ is relatively easy to solve for any $\mathbf{h} \in \mathbb{R}^n$ by means of a polynomially bounded algorithm, while the original problem (INT) is extremely difficult to solve. Due to this assumption, we penalize in optimization problem (INT) the constraint $D\mathbf{x} - \mathbf{d} \leq \mathbf{0}$ and so the Lagrangian dual problem (sometimes called the *partial dual* problem) is given by

$$\sup\{\theta(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \geq \mathbf{0}\} \tag{D}$$

with

$$\theta(\boldsymbol{\lambda}) = -\boldsymbol{\lambda}^\top \mathbf{d} + \inf\{(\mathbf{c} + D^\top \boldsymbol{\lambda})^\top \mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathbb{Z}_+^n\}. \tag{24}$$

Also, a relaxation of optimization problem (INT) is given by the optimization problem:

$$\inf\{\mathbf{c}^\top \mathbf{x} : D\mathbf{x} \leq \mathbf{d}, \mathbf{x} \in conv(X)\} \tag{LPR}$$

with $conv(X)$ denoting the smallest convex set containing $X = \{\mathbf{x} \in \mathbb{Z}_+^n : A\mathbf{x} = \mathbf{b}\}$. The set conv$(X)$ (also called the *convex hull* of the set X) can be completely characterized by linear inequalities and so the optimization problem (LPR) is actually a linear programming problem. Unfortunately it is not possible to generate the linear inequality description of the set conv(X) and so optimization problem (LPR) cannot be solved by a linear programming solver. Remember we only used a relaxation of the original problem if this relaxation can be easily solved and for problem (LPR) this is certainly not true. A relaxation of problem (INT) that can be solved by a linear programming solver is given by the standard LP relaxation:

$$\inf\{\mathbf{c}^\top \mathbf{x} : D\mathbf{x} \leq \mathbf{d}, A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\} \tag{SLPR}$$

of optimization problem (INT). In the next result we relate the optimal objective function values of the above optimization problems.

Lemma 2.9 *If $\mathcal{F}_{INT}$ is nonempty, then*

$$\upsilon(SLPR) \leq \upsilon(D) = \upsilon(LPR) \leq \upsilon(INT).$$

Proof The inequalities $\upsilon(SLPR) \leq \upsilon(LPR) \leq \upsilon(INT)$ are obvious and so we only need to verify that $\upsilon(D) = \upsilon(LPR)$. As observed, the optimization problem (LPR) is a linear programming problem and if $\upsilon(LPR)$ is finite, we obtain by strong duality and the observation after Theorem 2.2 that $\upsilon(LPR) = \sup_{\boldsymbol{\lambda} \geq \mathbf{0}}\{\theta(\boldsymbol{\lambda})\}$ with

$$\theta(\boldsymbol{\lambda}) = -\boldsymbol{\lambda}^\intercal \mathbf{d} + \inf\{(\mathbf{c}+D^\top \boldsymbol{\lambda})^\top \mathbf{x} : \mathbf{x} \in conv(X)\}.$$

Since the function $\mathbf{x} \to (\mathbf{c} + D^\top \boldsymbol{\lambda})^\top \mathbf{x}$ is linear on $conv(X)$, it follows that

$$\begin{aligned}\inf_{\mathbf{x} \in conv(X)}\{(\mathbf{c} + D^\top \boldsymbol{\lambda})^\top \mathbf{x}\} &= \inf_{\mathbf{x} \in X}\{(\mathbf{c} + D^\top \boldsymbol{\lambda})^\top \mathbf{x}\} \\ &= \inf\{(\mathbf{c} + D^\top \boldsymbol{\lambda})^\top \mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathbb{Z}_+^n\}\end{aligned}$$

and this shows by relation (24) the desired result. □

In Lemma 2.9 it might happen that the optimal objective function value $\upsilon(D)$ is strictly above $\upsilon(SLPR)$ and in this case we obtain a stronger lower bound on $\upsilon(INT)$. Hence it might be computationally more efficient to compute $\upsilon(D)$ and use this approach in a branch and bound procedure. In general it is computationally easier to compute $\upsilon(SLPR)$ instead of $\upsilon(D)$ but due to the lower bound difference computing $\upsilon(D)$ might be more efficient in branching in a classical branch and bound procedure. However, under some conditions the values $\upsilon(SLPR)$ and $\upsilon(D)$ are equal, and this happens if the optimization problem satisfies the so-called integrability property. Observe in this case all the lower bounds discussed in Lemma 2.9 are the same.

Definition 2.10 The optimization problem $\upsilon(R) = \inf\{\mathbf{h}^\top \mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathbb{Z}_+^n\}$ satisfies the so-called integrability property if $\upsilon(R) = \inf\{\mathbf{h}^\top \mathbf{x} : A\mathbf{x} = b, \mathbf{x} \geq \mathbf{0}\}$ for every vector $\mathbf{h}$.

A sufficient condition for the integrability property is that the matrix A is totally unimodular (Schrijver (1998)), and these matrices are extremely important in polyhedral combinatorics. For sufficient conditions to check whether a matrix is totally unimodular, consult Schrijver (1998)). These matrices appear a lot in network flow problems. By a similar argument as used in Lemma 2.9, one can now show the following result.

Lemma 2.10 *If $\mathcal{F}_{INT}$ is nonempty and the optimization problem* $\inf\{\mathbf{h}^\top \mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathbb{Z}_+^n\}$ *satisfies the integrability property, then* $\upsilon(SLPR) = \upsilon(D) = \upsilon(LPR)$.

Proof By penalizing the constraints $\mathbf{D}x \leq \mathbf{d}$ in problem ($SLPR$) and using the integrability property, we obtain by the strong duality result that $\upsilon(SLPR) = \sup_{\boldsymbol{\lambda} \geq \mathbf{0}}\{\theta(\boldsymbol{\lambda})\}$ with

$$\begin{aligned}\theta(\boldsymbol{\lambda}) &= -\boldsymbol{\lambda}^\top \mathbf{d} + \inf\{(\mathbf{c} + D^\top \boldsymbol{\lambda})^\top \mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\} \\ &= -\boldsymbol{\lambda}^\top \mathbf{d} + \inf\{(\mathbf{c} + D^\top \boldsymbol{\lambda})^\top \mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathbb{Z}_+^n\}\end{aligned}$$

and the desired result follows. □

By the above observation, it follows that the Lagrangian dual has the same optimal objective function value as the standard linear programming relaxation of (INT) and so it seems useless to solve the Lagrangian dual. However, from a computational point of view it is sometimes more efficient to solve the Lagrangian dual than to use a standard linear programming package to solve the standard linear programming relaxation of (INT). Hence to solve the Lagrangian dual or the standard linear programming relaxation depends on the problem under consideration. This concludes our discussion of the principle of Lagrange for finite dimensional optimization problems. Until now we only dealt in general with the principle of Lagrange and we did not discuss how to optimize the Lagrangian dual function. This is the topic of the next section.

2.3 On Subgradient Optimization and the Dual Problem

Until now we only dealt with the general structure of a Lagrangian dual problem and we did not discuss how to solve the Lagrangian dual problem $\sup\{\theta(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in K^*\}$ with K^* a closed convex cone and the Lagrangian dual function $\theta : K^* \to [-\infty, \infty)$ given by

$$\theta(\boldsymbol{\lambda}) = \inf\{f(\mathbf{x}) + \boldsymbol{\lambda}^\top \mathbf{g}(\mathbf{x}) : \mathbf{x} \in X\}. \tag{25}$$

Since the set $X \subseteq R^n$ is nonempty, it follows $-\infty \leq \theta(\boldsymbol{\lambda}) < \infty$. Moreover, by Lemma 2.1 the function $-\theta$ is convex and

$$-\sup\{\theta(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in K^*\} = \inf\{-\theta(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in K^*\}. \tag{$-D$}$$

The optimization problem ($-D$) has the same optimal solution as optimization problem (D) if an optimal solution exists and without loss of generality we always assume that the set $\text{dom}(-\theta) := \{\theta \in K^* : \theta(\boldsymbol{\lambda}) > -\infty\}$ is nonempty. In most practical cases the nonemptyness of the set $\text{dom}(-\theta)$ can be easily verified (look at location problems!). It is now easy to check that

$$-\infty \leq \upsilon(-D) < \infty \Leftrightarrow \text{dom}(-\theta) \text{ is nonempty.}$$

Observe that the function $-\theta : K^* \to (-\infty, \infty]$ is in general not differentiable. This can be easily checked by means of a picture if we assume for simplicity that the set X consists of two elements. Also the optimization problem ($-D$) is an *unconstrained convex programming problem* in case $K = \{\mathbf{0}\}$ and a *constrained convex programming problem* otherwise. Hence in principle we can apply to this problem the subgradient optimization procedure (Hiriart-Urruty and Lemaréchal, 2013) if it is possible to compute a ϵ-*subgradient* with $\epsilon \geq 0$ of the function $-\theta : K^* \to (-\infty, \infty]$ at every point $\boldsymbol{\lambda} \in K^*$. Observe that the set of ϵ-subgradients of the function $-\theta$ at the point $\boldsymbol{\lambda} \in K^0$ is called the ϵ-*subgradient set* and this set is denoted by $\partial_\epsilon(-\theta)(\boldsymbol{\lambda})$. In case ϵ equals 0 a 0-subgradient can be seen as a generalization of a gradient for a differentiable function.

Definition 2.11 A vector $\boldsymbol{\lambda}_0^* \in \mathbb{R}^n$ is called a ϵ-subgradient for some $\epsilon \geq 0$ of the function $f : K^* \to (-\infty, \infty]$ at the point $\boldsymbol{\lambda}_0 \in K^*$ if for every $\boldsymbol{\lambda} \in K^*$ it follows that

$$f(\boldsymbol{\lambda}) \geq f(\boldsymbol{\lambda}_0) + (\boldsymbol{\lambda} - \boldsymbol{\lambda}_0)^\top \boldsymbol{\lambda}_0^* - \epsilon.$$

It is called a subgradient if $\epsilon = 0$.

In general it is difficult to *compute* a ϵ-subgradient for an arbitrary finite valued convex function. However, for the function $-\theta$ a ϵ-subgradient for this function at a given point $\boldsymbol{\lambda} \in K^*$ is easy to compute by solving optimization problem $(D(\boldsymbol{\lambda}))$. This is shown in the next lemma and can be easily proved.

Lemma 2.11 *If for some $\boldsymbol{\lambda}_0 \in K^*$ the optimization problem:*

$$\theta(\boldsymbol{\lambda}_0) := \inf\{f(\mathbf{x}) + \boldsymbol{\lambda}_0^\top \mathbf{g}(\mathbf{x}) : \mathbf{x} \in X\}$$

has an optimal solution $\mathbf{x}_{opt}(\boldsymbol{\lambda}_0)$, then it follows that $-\mathbf{g}(\mathbf{x}_{opt}(\boldsymbol{\lambda}_0))$ is a subgradient of the function $-\theta$ at $\boldsymbol{\lambda}_0$. Moreover, if the feasible solution $\mathbf{x}_\epsilon(\boldsymbol{\lambda}_0)$ is an ϵ-optimal solution of the optimization problem $(D(\boldsymbol{\lambda}_0))$, then the vector $-\mathbf{g}(\mathbf{x}_\epsilon(\boldsymbol{\lambda}_0))$ is a ϵ-subgradient of the function $-\theta$ at $\boldsymbol{\lambda}_0$.

Observe that we introduce ϵ-subgradients since in general one cannot perform exact calculations on a computer and so in most cases we actually are calculating a ϵ-optimal solution. Solving the optimization problem $(D(\boldsymbol{\lambda}))$ with $\boldsymbol{\lambda} = \boldsymbol{\lambda}_0$ and obtaining the optimal solution $\mathbf{x}(\boldsymbol{\lambda}_0)$ yield both the value $-\theta(\boldsymbol{\lambda}_0)$ and the affine function $\boldsymbol{\lambda} \to \theta(\boldsymbol{\lambda}_0) - (\boldsymbol{\lambda} - \boldsymbol{\lambda}_0)^\top \mathbf{g}(\mathbf{x}_{opt}(\boldsymbol{\lambda}_0))$. By Lemma 2.11, the above affine function serves as a lower bound on the convex function $-\theta$ on K^*.

We now discuss in detail the subgradient method applied to the optimization problem:

$$\inf\{f(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in D\} \qquad (C)$$

with $D \subseteq \mathbb{R}^n$ a closed convex set and $f : \mathbb{R}^n \to (-\infty, \infty]$ with $f(\boldsymbol{\lambda}) < \infty$ for some $\boldsymbol{\lambda} \in D$. By our assumption, it follows that $-\infty \leq \upsilon(C) < \infty$. Moreover, for this problem we assume that for every $\boldsymbol{\lambda} \in D$ a ϵ-subgradient of f at $\boldsymbol{\lambda}$ can be computed and by Lemma 2.11 this condition is clearly satisfied by our optimization problem $\inf\{-\theta(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in K^*\}$ selecting an optimal Lagrangian multiplier. To introduce the subgradient method, we observe the following. If $\boldsymbol{\lambda}_{opt}$ is the unknown optimal solution of optimization problem (C) and we apply an algorithm to find this optimal solution, this algorithm generates a sequence $\{\boldsymbol{\lambda}_n : n \in N\} \subseteq D$. It is *desirable* that at each next step of this algorithm the newly generated point $\boldsymbol{\lambda}_{n+1}$ is closer to $\boldsymbol{\lambda}_{opt}$, then the presently generated point $\boldsymbol{\lambda}_n$. This means that the inequality:

$$\| \boldsymbol{\lambda}_{n+1} - \boldsymbol{\lambda}_{opt} \|^2 < \| \boldsymbol{\lambda}_n - \boldsymbol{\lambda}_{opt} \|^2 \qquad (26)$$

with

$$\| \boldsymbol{\lambda} \|^2 := \sum_{i=1}^{n} \lambda_i^2$$

the squared Euclidean norm should hold for our algorithm. To achieve this, suppose that we are after n steps of our iteration procedure at the vector $\boldsymbol{\lambda}_n$ and some oracle supplies us with a search direction $\mathbf{d}_n$. If the new point $\boldsymbol{\lambda}_{n+1}$ is given by

$$\boldsymbol{\lambda}_{n+1} := \boldsymbol{\lambda}_n - t_n \mathbf{d}_n$$

for some $t_n > 0$, then it follows that

$$\| \boldsymbol{\lambda}_{n+1} - \boldsymbol{\lambda}_{opt} \|^2 = \| \boldsymbol{\lambda}_n - \boldsymbol{\lambda}_{opt} \|^2 + t_n^2 \| \boldsymbol{\lambda}_{opt} \|^2 + 2 t_n \mathbf{d}_n^{\intercal} (\boldsymbol{\lambda}_{\text{opt}} - \boldsymbol{\lambda}_n). \tag{27}$$

To guarantee that relation (26) holds, it is *necessary* (maybe not sufficient!) to choose the direction $\mathbf{d}_n$ in such a way that

$$\mathbf{d}_n^{\intercal} (\boldsymbol{\lambda}_{opt} - \boldsymbol{\lambda}_n) \leq 0. \tag{28}$$

If $\mathbf{d}_n$ is an ϵ-subgradient of the function f at $\boldsymbol{\lambda}_n$ for some $\epsilon \geq 0$ (remember for $\epsilon = 0$ the ϵ-subgradient is actually a subgradient!), this implies

$$\mathbf{d}_n^{\intercal} (\boldsymbol{\lambda}_{opt} - \boldsymbol{\lambda}_n) \leq f(\boldsymbol{\lambda}_{opt}) - f(\boldsymbol{\lambda}_n) + \epsilon \leq \epsilon \tag{29}$$

and since it is *necessary* that relation (28) holds it follows by relation (29) that a hopefully good choice of the direction $\mathbf{d}_n$ is selecting a ϵ-subgradient of f at $\boldsymbol{\lambda}_n$. By this observation, we obtain the usual *subgradient iteration scheme* given by the following algorithm.

Algorithm 2.1 Subgradient optimization scheme:

$$\boldsymbol{\lambda}_{n+1} := \boldsymbol{\lambda}_n - t_n \boldsymbol{\lambda}_n^* , t_n > 0 \text{ and } \boldsymbol{\lambda}_n^* \text{ a } \epsilon\text{-subgradient of } f \text{ at } \boldsymbol{\lambda}_n .$$

However, using the above formula might create problems for constrained optimization problems. It can happen that the point $\boldsymbol{\lambda}_n$ belongs to D and $\boldsymbol{\lambda}_{n+1}$ does not belong to D and so the point $\boldsymbol{\lambda}_{n+1}$ becomes infeasible. A way to solve this is to use the orthogonal projection of $\boldsymbol{\lambda}_{n+1}$ onto the closed convex set D. This means that the *modified subgradient iteration scheme* has the form:

Algorithm 2.2 Modified subgradient iteration scheme:

$$\boldsymbol{\lambda}_{n+1} = P_D(\boldsymbol{\lambda}_n - t_n \boldsymbol{\lambda}_n^*), t_n > 0 \text{ and } \boldsymbol{\lambda}_n^* \text{ a } \epsilon\text{-subgradient of } f \text{ at } \boldsymbol{\lambda}_n \tag{30}$$

with

$$P_D(\mathbf{x}_0) := \arg\min\{\parallel \mathbf{x}_0 - \boldsymbol{\lambda} \parallel^2 : \boldsymbol{\lambda} \in D\}, \mathbf{x}_0 \text{ fixed vector}$$

and argmin$\{Q\}$ denoting the optimal solution of optimization problem (Q).

Since the function $\boldsymbol{\lambda} \to \|\mathbf{x}_0 - \boldsymbol{\lambda} \|^2$ for $\mathbf{x}_0$ fixed is strictly convex and D is a closed nonempty convex set, there exists a *unique optimal solution* and so the vector $P_D(\mathbf{x}_0)$ is well-defined or equivalently $\mathbf{x} \to P_D(\mathbf{x})$ represents a function. Moreover, we mention without proof (check first-order conditions of the above constrained optimization problem!) that the function $\mathbf{x} \to P_D(\mathbf{x})$ is a so-called contraction and this means that

$$\parallel P_D(\mathbf{x}_0) - P_D(\mathbf{x}_1) \parallel \leq \parallel \mathbf{x}_0 - \mathbf{x}_1 \parallel \tag{31}$$

for every $\mathbf{x}_0$ and $\mathbf{x}_1$ belonging to $\mathbb{R}^n$. For our primal optimization problem, it almost always follows that $K = \mathbb{R}^m_+ \times \mathbb{R}^{n-m}$ and so $K^* = \mathbb{R}^m_+ \times \{\mathbf{0}\}$. This means that we can solve the projection by hand (make a picture!). It is now possible to show the following fundamental inequality (Correa and Lemaréchal , 1993).

Lemma 2.12 *If* $\boldsymbol{\lambda} \in D$ *and* $\boldsymbol{\lambda}_n^*$ *is an* ϵ_n*-subgradient of the function* f *at the present iteration point* $\boldsymbol{\lambda}_n$ *for some* $\epsilon_n \geq 0$ *and* $\boldsymbol{\lambda}_{n+1} = P_D(\boldsymbol{\lambda}_n - t_n\boldsymbol{\lambda}_n^*)$*, then it follows that*

$$\parallel \boldsymbol{\lambda}_{n+1} - \boldsymbol{\lambda} \parallel^2 \leq \parallel \boldsymbol{\lambda}_n - \boldsymbol{\lambda} \parallel^2 + t_n^2 \parallel \boldsymbol{\lambda}_n^* \parallel^2 + 2t_n(f(\boldsymbol{\lambda}) - f(\boldsymbol{\lambda}_n) + \epsilon_n).$$

Proof Since obviously $P_D(\boldsymbol{\lambda}) = \boldsymbol{\lambda}$ for $\boldsymbol{\lambda} \in D$, it follows by relation (30) that

$$\parallel \boldsymbol{\lambda}_{n+1} - \boldsymbol{\lambda} \parallel^2 = \parallel P_D(\boldsymbol{\lambda}_n - t_n\boldsymbol{\lambda}_n^*) - P_D(\boldsymbol{\lambda}) \parallel^2$$

and this implies by relation (31) that

$$\parallel \boldsymbol{\lambda}_{n+1} - \boldsymbol{\lambda} \parallel^2 \leq \|\boldsymbol{\lambda}_n - \boldsymbol{\lambda} \parallel^2 + t_n^2 \parallel \boldsymbol{\lambda}_n^* \parallel^2 + 2t_n(\boldsymbol{\lambda} - \boldsymbol{\lambda}_n)^\intercal \boldsymbol{\lambda}_n^*. \tag{32}$$

Since $\boldsymbol{\lambda}_n^*$ is an ϵ_n-subgradient of the function f at $\boldsymbol{\lambda}_n$, we obtain that $f(\boldsymbol{\lambda}) - f(\boldsymbol{\lambda}_n) \geq (\boldsymbol{\lambda} - \boldsymbol{\lambda}_n)^\intercal \boldsymbol{\lambda}_n^* - \epsilon_n$ and this implies by relation (32) that

$$\parallel \boldsymbol{\lambda}_{n+1} - \boldsymbol{\lambda} \parallel^2 \leq \parallel \boldsymbol{\lambda}_n - \boldsymbol{\lambda} \parallel^2 + t_n^2 \parallel \boldsymbol{\lambda}_n^* \parallel^2 + 2t_n(f(\boldsymbol{\lambda}) - f(\boldsymbol{\lambda}_n) + \epsilon_n)$$

showing the desired inequality. □

By the above inequality, the convergence of our subgradient optimization scheme is easily proved and this is shown by the following result.

Theorem 2.6 *If* $\lim_{n\uparrow\infty} \epsilon_n = 0$ *and the positive sequence* $\{t_n : n \in \mathbb{N}\}$ *satisfies*

$$\sum_{n=1}^{\infty} t_n = \infty \text{ and } \lim_{n\uparrow\infty} t_n \parallel \boldsymbol{\lambda}_n^* \parallel^2 = 0,$$

then it follows that

$$\lim_{n\uparrow\infty} m_n = \upsilon(C) \geq -\infty$$

with

$$m_n := \min\{f(\boldsymbol{\lambda}_k) : k \leq n\}.$$

Proof Clearly the sequence $m_n, n \in N$ is decreasing and so it has a limit c. If the limit c equals $-\infty$, the result is proved and so we assume that the limit c is finite. If $c > \upsilon(C)$, one can find by the definition of an infimum some $\delta > 0$ and $\boldsymbol{\lambda}_e \in D$ satisfying

$$\upsilon(C) \leq f(\boldsymbol{\lambda}_e) \leq c - \delta \leq f(\boldsymbol{\lambda}_n) - \delta \tag{33}$$

for every $n \in \mathbb{N}$. Hence by Lemma 2.12 with $\boldsymbol{\lambda} = \boldsymbol{\lambda}_e$, we obtain that

$$\| \boldsymbol{\lambda}_{n+1} - \boldsymbol{\lambda}_e \|^2 \leq \| \boldsymbol{\lambda}_n - \boldsymbol{\lambda}_e \|^2 + t_n^2 \| \boldsymbol{\lambda}_n^* \|^2 + 2t_n(f(\boldsymbol{\lambda}_e) - f(\boldsymbol{\lambda}_n) + \epsilon_n).$$

This implies by relation (33) that

$$\| \boldsymbol{\lambda}_{n+1} - \boldsymbol{\lambda}_e \|^2 \leq \| \boldsymbol{\lambda}_n - \boldsymbol{\lambda}_e \|^2 + t_n(t_n \| \boldsymbol{\lambda}_n^* \|^2 + 2(\epsilon_n - \delta)).$$

Since we assume that

$$\lim_{n\uparrow\infty} t_n \| \boldsymbol{\lambda}_n^* \|^2 = 0 \text{ and } \lim_{n\uparrow\infty} \epsilon_n = 0,$$

one can find some $n_0 \in N$ such that for every $n \geq n_0$ it follows that

$$t_n \| \boldsymbol{\lambda}_n^* \|^2 + 2(\epsilon_n - \delta) \leq -\delta$$

and this yields for every $n \geq n_0$ that

$$\| \boldsymbol{\lambda}_{n+1} - \boldsymbol{\lambda}_e \|^2 \leq \| \boldsymbol{\lambda}_n - \boldsymbol{\lambda}_e \|^2 - \delta t_n. \tag{34}$$

Hence we obtain by relation (34) that

$$0 \leq \| \boldsymbol{\lambda}_m - \boldsymbol{\lambda}_e \|^2 \leq \| \boldsymbol{\lambda}_{n_0} - \boldsymbol{\lambda}_e \|^2 - \delta \sum_{k=n_0}^{m-1} t_k$$

for every $m \geq n_0$ and this implies by the assumption $\sum_{k=n_0}^{\infty} t_k = \infty$ that

$$0 \leq \lim_{m\uparrow\infty} \| \boldsymbol{\lambda}_m - \boldsymbol{\lambda}_e \|^2 = -\infty.$$

Hence we obtain a contradiction and so it must follow that $c = \upsilon(C)$. $\square$

Since it can happen that an optimal solution of optimization problem (C) does not exist, we cannot show that the sequence $\boldsymbol{\lambda}_n$ generated by the subgradient method converges to the optimal solution of (C). In general it also does not hold that the sequence $f(\boldsymbol{\lambda}_n)$ is decreasing. Although we have shown under fairly general conditions that the sequence m_n converges to $v(C)$, it is very difficult to give a *proper stopping rule* for the subgradient method. Also in case we apply the subgradient method to optimization problem ($-D$), the selection of the step sizes $t_n, n \in N$ is a matter of trial and error (Nemhauser and Wolsey, 1988) and so devising a proper scheme is an art in itself. For more details on these practical issues, the reader is referred to Beasley et al. (1993). We finally consider one theoretical case in which one knows in advance that the current iteration point is indeed optimal. This result is presented in the next lemma.

Lemma 2.13 *If at the* $(n + 1)th$ *step it follows that*

$$\boldsymbol{\lambda}_{n+1} := P_D(\boldsymbol{\lambda}_n - t_n\boldsymbol{\lambda}_n^*)$$

equals $\boldsymbol{\lambda}_n$ *with* $\boldsymbol{\lambda}_n^*$ *a* ϵ_n*-subgradient of the function* f *at* $\boldsymbol{\lambda}_n$*, then the vector* $\boldsymbol{\lambda}_n$ *is an* ϵ_n*-optimal solution of optimization problem (*C*).*

Proof We only give a proof of this result in case $\boldsymbol{\lambda}_n - t_n\boldsymbol{\lambda}_n^*$ belongs to D. If this holds, it follows since $t_n > 0$ that

$$\boldsymbol{\lambda}_{n+1} := P_D(\boldsymbol{\lambda}_n - t_n\boldsymbol{\lambda}_n^*) = \boldsymbol{\lambda}_n - t_n\boldsymbol{\lambda}_n^*$$

and so $\boldsymbol{\lambda}_n^* = \mathbf{0}$. Since $\boldsymbol{\lambda}_n^*$ is an ϵ_n-subgradient, this implies that

$$f(\boldsymbol{\lambda}) \geq f(\boldsymbol{\lambda}_n) - \epsilon_n$$

for every $\boldsymbol{\lambda} \in D$ and the desired result follows. □

For more theoretical results regarding the subgradient method, the reader is referred to Shor (2012) and Correa and Lemaréchal (1993). In the next section we apply the above results to derive Lagrangian heuristics for location and transportation problems.

3 Application of the Principle of Lagrange to Problems in Transportation and Location Theory

In this section we consider two classical problems in location and transportation theory and apply to these problems the Lagrangian relaxation approach. In the first subsection we discuss the well-known set covering problem and in the second subsection several versions of the capacitated facility location problem.

3.1 The Set Covering Problem

In this subsection we formulate the set covering problem. Observe that this problem shows up in a lot of different applications within computer and management science. To give a particular example in transportation, consider a company that needs to serve every day m different customers given by the set $\mathcal{M} = \{1, \ldots, m\}$. To serve these customers, the company selects from a given set $\mathcal{N} = \{1, \ldots, n\}$ of known routes a subset of these routes covering all present day customers. Each route $j \in \mathcal{N}$, has a known cost $c_j > 0$ and so the company wants to select the subset that serves all customers on a given day and has minimal cost. Selecting a route $j \in \mathcal{N}$ for today is a yes\no decision and so we introduce for each route $j \in \mathcal{N}$ the decision variable

$$x_j = \begin{cases} 1 & \text{if route } j \text{ is selected} \\ 0 & \text{otherwise} \end{cases}.$$

To model that every customer who needs to be served today is at least covered by one selected route, we construct the $m \times n$ matrix $A = (a_{ij})$, given by

$$a_{ij} = \begin{cases} 1 & \text{if customer } i \text{ is on route } j \\ 0 & \text{otherwise} \end{cases}.$$

Clearly the value $\sum_{j \in \mathcal{N}} a_{ij} x_j$, $i \in \mathcal{M}$ yields the number of times customer $i \in \mathcal{M}$ is on the subset of selected routes and we need to solve the optimization problem:

$$\inf\left\{\sum_{j \in \mathcal{N}} c_j x_j : \sum_{j \in \mathcal{N}} a_{ij} x_j \geq 1, i \in \mathcal{M}, x_j \in \mathbb{B}, j \in \mathcal{N}\right\} \tag{SP}$$

with $\mathbb{B} = \{0, 1\}$. Hence the set covering problem (SP) is given by

$$\inf\{\sum_{j \in \mathcal{N}} c_j x_j : \mathbf{g}(\mathbf{x}) \preceq_K \mathbf{0}, \mathbf{x} \in \mathbb{B}^n\}$$

with $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \ldots, g_m(\mathbf{x}))$, $g_i(\mathbf{x}) = 1 - \sum_{j \in \mathcal{N}} a_{ij} x_j$ and $K = \mathbb{R}^m_+$. The feasible set contains only a finite number of elements bounded above by 2^n and so the set covering problem has an optimal solution. It is well-known that the set covering problem is $\mathcal{NP}$-hard (Garey and Johnson, 1979) and so by present day belief it seems very unlikely that there exists a polynomially bounded algorithm to solve this problem. To analyze this problem by means of the Lagrangian relaxation approach, we penalize the constraints $\sum_{j \in \mathcal{N}} a_{ij} x_j \geq 1$, $i \in \mathcal{M}$ and construct the Lagrangian dual function $\theta : \mathbb{R}^m_+ \to \mathbb{R}$ given by

$$\theta(\boldsymbol{\lambda}) = \min\left\{\sum_{j \in \mathcal{N}} c_j x_j + \sum_{i \in \mathcal{M}} \lambda_i (1 - \sum_{j \in \mathcal{N}} a_{ij} x_j) : \mathbf{x} \in \mathbb{B}^n\right\}.$$

After some easy calculations, it follows that

$$\theta(\boldsymbol{\lambda}) = \sum_{i \in \mathcal{M}} \lambda_i + \sum_{j \in \mathcal{N}} \min\left\{c_j - \sum_{i \in \mathcal{M}} a_{ij}\lambda_i, 0\right\} \tag{35}$$

and the optimal solution $\mathbf{x}(\boldsymbol{\lambda}) = (x_1(\boldsymbol{\lambda}), \ldots, x_n(\boldsymbol{\lambda}))$ is given by

$$x_j(\boldsymbol{\lambda}) = 1_{S_j}(\boldsymbol{\lambda}) \tag{36}$$

with $S_j = \{\boldsymbol{\lambda} \in \mathbb{R}^m_+ : c_j \leq \sum_{i \in \mathcal{M}} a_{ij}\lambda_i\}$, $j \in \mathcal{N}$. We now apply the subgradient method to solve the Lagrangian dual problem

$$\upsilon(D) = \min_{\lambda \geq \mathbf{0}} \left\{\sum_{i \in \mathcal{M}} \lambda_i + \sum_{j \in \mathcal{N}} \min\{c_j - \sum_{i \in \mathcal{M}} a_{ij}\lambda_i, 0\}\right\}$$

and obtain as an optimal dual solution a vector $\boldsymbol{\lambda}_{opt}$ and construct the solution $\mathbf{x}(\boldsymbol{\lambda}_{opt})$ given in relation (36). By Lemma 2.9, this dual problem satisfies

$$\upsilon(D) = \min\left\{\sum_{j \in \mathcal{N}} c_j x_j : \sum_{j \in \mathcal{N}} a_{ij} x_j \geq 1, i \in \mathcal{M}, 0 \leq x_j \leq 1, j \in \mathcal{N}\right\}.$$

If the solution $\mathbf{x}(\boldsymbol{\lambda}_{opt})$ is primal feasible and it satisfies the complementary slackness conditions

$$\lambda_i(1 - \sum_{j \in \mathcal{N}} a_{ij} x_j(\boldsymbol{\lambda}_{opt})) = 0, i \in \mathcal{M},$$

then it follows by Lemma 2.5 that this solution is an optimal primal solution. If the solution $\mathbf{x}(\boldsymbol{\lambda}_{opt})$ is only primal feasible, we have found a feasible solution for our problem and so $\theta(\boldsymbol{\lambda}_{opt}) \leq \upsilon(SP) \leq \sum_{j \in \mathcal{N}} c_j x_j(\boldsymbol{\lambda}_{opt})$. If the solution is not primal feasible, then we use the following Lagrangian heuristic to convert the solution $\mathbf{x}(\boldsymbol{\lambda}_{opt})$ into a primal feasible solution. Consider the set of routes selected by our solution $\mathbf{x}(\boldsymbol{\lambda}_{opt})$ given by

$$R(\boldsymbol{\lambda}_{opt}) = \{1 \leq j \leq n : x_j(\boldsymbol{\lambda}_{opt}) = 1\}.$$

Starting with the initial set $R(\boldsymbol{\lambda}_{opt})$, we check for each customer $i \in \mathcal{M}$ whether this customer is visited by one of the routes in the presently selected set. If not, we select the cheapest possible route covering this customer and add this route to this presently selected set of routes. After evaluating all the customers, we have constructed a set of routes that covers all customers. Some customers might be assigned to different routes and we delete them from all the routes except one. Observe that this heuristic is applied in a branch and bound procedure and at each branch we solve a special case of the set covering problem with some routes already selected. This concludes our discussion of the set covering problem. In the next section we consider the (un)capacitated facility location problem.

3.2 Fixed Charged Location Models on Discrete Spaces

To introduce the so-called fixed charged location problems in discrete location, we denote as before the set of customers by $\mathcal{M} = \{1, \ldots, m\}$ and the set of possible sites of facilities by $\mathcal{N} = \{1, \ldots, n\}$. We denote by $f_j, j \in \mathcal{N}$ the fixed setup cost of site j, while c_{ij} denotes the transportation costs of supplying the demand of customer $i \in \mathcal{M}$ by site $j \in \mathcal{N}$. The *uncapacitated facility location problem* tries to determine a subset of the set $\mathcal{N}$ of possible sites in such a way that the total transportation and setup costs are minimized. In this problem the demand of all customers should be satisfied. Moreover, each open facility has unlimited capacity. A more difficult and related problem is given by the capacitated version of this problem. In this problem each site j has a fixed capacity q_j. For the capacitated facility location problem, there actually exist two versions. In the first version the total demand of each customer can only be supplied by one facility and this problem is called the *single source capacitated facility location* problem, while in the other version the total demand of each customer can be supplied by more than one facility. This problem is called the *multiple source capacitated facility location* problem. To formulate an optimization problem for the uncapacitated version, we introduce the binary decision vector $\mathbf{y} = (y_j)_{j \in \mathcal{N}}$ given by

$$y_j = 1 \Leftrightarrow \text{facility } j \text{ is opened at site } j.$$

Since in the uncapacitated facility location problem the capacity of each facility is unlimited, we assign each customer to the “cheapest ” open facility and this implies that the model can be represented by the optimization problem:

$$\inf\left\{\sum\nolimits_{i \in \mathcal{M}} \min\{c_{ij} : y_j = 1\} + \sum\nolimits_{j \in \mathcal{N}} f_j y_j : \mathbf{y} \in \mathbb{B}^n\right\}..$$

However, in this optimization problem the objective function is not linear and to linearize the objective function we introduce the decision variables $\mathbf{x} = (x_{ij})_{i \in \mathcal{M}, j \in \mathcal{N}}$ given by

$$x_{ij} = \text{fraction of demand of customer } i \in \mathcal{M} \text{ supplied by site } j \in \mathcal{N}.$$

Using these decision variables, the objective function is given by

$$\sum\nolimits_{i \in \mathcal{M}} \sum\nolimits_{j \in \mathcal{N}} c_{ij} x_{ij} + \sum\nolimits_{j \in \mathcal{N}} f_j y_j.$$

Since all demand should be delivered, it is clear that these decision variables need to satisfy the so-called *assignment constraints* $\sum_{j \in \mathcal{N}} x_{ij} = 1$ for every $i \in \mathcal{M}$ and so the feasible region $\mathcal{F}_{UFL}$ of the uncapacitated facility location problem is given by

$$\mathcal{F}_{UFL} = \{(\mathbf{x}, \mathbf{y}) : 0 \leq x_{ij} \leq y_j, \sum_{j \in \mathcal{N}} x_{ij} = 1, y_j \in \mathbb{B}, i \in \mathcal{M}, j \in \mathcal{N}\}. \tag{37}$$

Hence the optimization problem for the uncapacitated facility location problem is given by

$$\inf \left\{ \sum_{i \in \mathcal{M}} \sum_{j \in \mathcal{N}} c_{ij} x_{ij} + \sum_{j \in \mathcal{N}} f_j y_j : (\mathbf{x}, \mathbf{y}) \in \mathcal{F}_{UFL} \right\}. \tag{UFL}$$

If we consider the multiple-source capacitated facility location problem, we need to add the capacity constraints $\sum_{i \in \mathcal{M}} d_i x_{ij} \leq q_j y_j$ for every $j \in \mathcal{N}$ with d_i denoting the total demand of customer $i \in \mathcal{M}$. Hence the feasible region $\mathcal{F}_{MSCFL}$ of the multiple-source capacitated facility location problem is given by

$$\mathcal{F}_{MSCFL} = \mathcal{F}_{UFL} \cap \left\{ (\mathbf{x}, \mathbf{y}) : \sum_{i \in \mathcal{M}} d_i x_{ij} \leq q_j y_j, j \in \mathcal{N} \right\}$$

and we need to solve the optimization problem:

$$\inf \left\{ \sum_{i \in \mathcal{M}} \sum_{j \in \mathcal{N}} c_{ij} x_{ij} + \sum_{j \in \mathcal{N}} f_j y_j : (\mathbf{x}, \mathbf{y}) \in \mathcal{F}_{MSCFL} \right\}. \tag{$MSCFL$}$$

If we consider the single source capacitated facility location problem, each customer can only be supplied by exactly one open facility and this implies that the decision variables $x_{ij} \in \mathbb{B}$. Hence the feasible region $\mathcal{F}_{\text{SSCFL}}$ is given by

$$\mathcal{F}_{SSCFL} = \mathcal{F}_{UFL} \cap \left\{ (\mathbf{x}, \mathbf{y}) : \sum_{i \in \mathcal{M}} d_i x_{ij} \leq q_j y_j, x_{ij} \in \mathbb{B}, i \in \mathcal{N}, j \in \mathcal{M} \right\}$$

and the optimization problem is given by

$$\inf \left\{ \sum_{i \in \mathcal{M}} \sum_{j \in \mathcal{N}} c_{ij} x_{ij} + \sum_{j \in \mathcal{N}} f_j y_j : (\mathbf{x}, \mathbf{y}) \in \mathcal{F}_{SSCFL} \right\}. \tag{$SSCFL$}$$

For an extensive overview of transportation-location models on discrete spaces, the reader should consult Daskin (2011) or Francis, McGinnis and White (1992). It is well-known that the uncapacitated facility location problem (UFL) is $\mathcal{NP}$-hard (Garey and Johnson, 1979) and as already observed represented by

$$\inf \left\{ \sum_{i \in \mathcal{M}} \sum_{j \in \mathcal{N}} c_{ij} x_{ij} + \sum_{j \in \mathcal{N}} f_j y_j : \mathbf{g}(\mathbf{x}) \preceq_K \mathbf{0}, 0 \leq x_{ij} \leq y_j, \right.$$
$$\left. i \in \mathcal{N}, j \in \mathcal{M}, \mathbf{y} \in \mathbb{B}^n \right\}$$

with

$$\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \ldots, g_m(\mathbf{x})), g_i(\mathbf{x}) = \sum\nolimits_{j \in \mathcal{N}} x_{ij} - 1$$

and $K = \{\mathbf{0}\} \subset \mathbb{R}^m$. To construct the Lagrangian dual function of the uncapacitated facility location model, we first observe for $f_j = 0$ that we always open a facility at site j and so without loss of generality we may assume that $f_j > 0$. In the uncapacitated facility location model we penalize the assignment constraints $\sum_{j \in \mathcal{N}} x_{ij} - 1 = 0$ for every $i \in \mathcal{M}$. Since $K^* = \{\mathbf{0}\}^* = \mathbb{R}^m$, the Lagrangian dual function $\theta : \mathbb{R}^m \to [-\infty, \infty)$ is defined on $\mathbb{R}^m$ and it is given by

$$\theta(\boldsymbol{\lambda}) = -\sum\nolimits_{i \in \mathcal{M}} \lambda_i + \sum\nolimits_{j \in \mathcal{N}} \inf \left\{ \sum\nolimits_{i \in \mathcal{M}} (c_{ij} + \lambda_i) x_{ij} + f_j y_j : (\mathbf{x}, \mathbf{y}) \in X_j \right\} \tag{38}$$

with $X_j := \{(\mathbf{x}, \mathbf{y}) : 0 \leq x_{ij} \leq y_j, i \in \mathcal{M}, y_j \in \mathbb{B}\}$, $j \in \mathcal{N}$. To solve the minimization problem:

$$\inf \left\{ \sum\nolimits_{i \in \mathcal{M}} (c_{ij} + \lambda_i) x_{ij} + f_j y_j : (\mathbf{x}, \mathbf{y}) \in X_j \right\} \tag{H_j}$$

for a given $j \in \mathcal{N}$, we take x_{ij} as large (small) as possible in case $c_{ij} + \lambda_i$ is negative (positive). This implies that an optimal solution $(\mathbf{x}(\boldsymbol{\lambda}), \mathbf{y}(\boldsymbol{\lambda}))$ of optimization problem (H_j) is given by

$$x_{ij}(\boldsymbol{\lambda}) = \begin{cases} y_j(\boldsymbol{\lambda}) & \text{if } c_{ij} + \lambda_i < 0 \\ 0 & \text{if } c_{ij} + \lambda_i \geq 0 \end{cases}. \tag{39}$$

By relation (39), it follows that the decision variables x_{ij}, $i \in \mathcal{M}$, $j \in \mathcal{N}$ can be eliminated and we obtain

$$\inf \left\{ \sum\nolimits_{i \in \mathcal{M}} (c_{ij} + \lambda_i) x_{ij} + f_j y_j : (\mathbf{x}, \mathbf{y}) \in X_j \right\} = \inf\{(f_j + b_j(\boldsymbol{\lambda})) y_j : y_j \in \mathbb{B}\} \tag{40}$$

with $b_j : \mathbb{R}^m \to (-\infty, \infty)$ given by

$$b_j(\boldsymbol{\lambda}) := \sum\nolimits_{i \in \mathcal{M}} \min\{c_{ij} + \lambda_i, 0\}. \tag{41}$$

To determine the optimal objective value of problem (H_j), we finally observe that

$$\inf\{(f_j + b_j(\boldsymbol{\lambda})) y_j : y_j \in \mathbb{B}\} = \min\{f_j + b_j(\boldsymbol{\lambda}), 0\} \tag{42}$$

with optimal solution $\mathbf{y}(\boldsymbol{\lambda}) = (y_j(\boldsymbol{\lambda}))_{j \in \mathcal{N}}$ given by

$$y_j(\boldsymbol{\lambda}) = \begin{cases} 1 & \text{if } f_j + b_j(\boldsymbol{\lambda}) \leq 0 \\ 0 & \text{if } f_j + b_j(\boldsymbol{\lambda}) > 0 \end{cases}. \tag{43}$$

Hence the Lagrangian dual function $\theta : \mathbb{R}^m \to \mathbb{R}$ of the uncapacitated facility location problem is given by

$$\theta(\boldsymbol{\lambda}) = -\sum\nolimits_{i \in \mathcal{N}} \lambda_i + \sum\nolimits_{j \in \mathcal{M}} \min\{f_j + b_j(\boldsymbol{\lambda}), 0\}$$

and its Lagrangian dual by

$$\sup\{\theta(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in \mathbb{R}^n\}.$$

Applying the subgradient method to $\inf\{-\theta(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in \mathbb{R}^m\}$, we determine the optimal Lagrangian multiplier $\boldsymbol{\lambda}_{opt}$. In case the optimal solution $(\mathbf{x}(\boldsymbol{\lambda}_{opt}), \mathbf{y}(\boldsymbol{\lambda}_{opt}))$ given by relations (39) and (43) is also primal feasible, we obtain since $K = \{\mathbf{0}\}$ that the complementary slackness conditions are automatically satisfied. This implies by Lemma 2.5 that this solution is an optimal solution of the uncapacitated facility location problem. However, in most cases this solution is not primal feasible and this means that either no facilities are opened or there exists a customer whose demand is not satisfied. The following Lagrangian heuristic now transforms the solution $(\mathbf{x}(\boldsymbol{\lambda}_{opt}), \mathbf{y}(\boldsymbol{\lambda}_{opt}))$ into a primal feasible solution.

Algorithm 3.1 Lagrangian heuristic:

1. Consider the possibly empty set $S_1(\boldsymbol{\lambda}_{opt})$ of open facilities given by

   $$S_1(\boldsymbol{\lambda}_{opt}) := \{j \in \mathcal{N} : y_j(\boldsymbol{\lambda}_{opt}) = 1\}$$

   and go to step 2.
2. If this set is empty, open a facility and assign all customers to this facility. If this set is nonempty, consider all customers "overassigned" or not assigned and make them nonassigned. Assign now these nonassigned customers to the cheapest facility belonging to the set $S_1(\boldsymbol{\lambda}_{opt})$.

As for the set covering problem, we need to solve in a branch and bound procedure at each branch a special instance of the uncapacitated facility location problem with some facilities already opened and or some customers already assigned to some open facilities. A more efficient and special purpose procedure to solve the uncapacitated facility location problem starts with the observation that for positive setup costs we may replace the decision variables $y_j \in \mathbb{B}$ in problem (UFL) by $y_j \in \mathbb{Z}_+$ without changing the set of optimal solutions and the optimal objective value. This means that we replace the feasible region $\mathcal{F}_{UFL}$ by the bigger set $\mathcal{F}_{RUFL}$ given by

$$\mathcal{F}_{RUFL} := \{(\mathbf{x}, \mathbf{y}) : 0 \leq x_{ij} \leq y_j, \sum\nolimits_{j \in \mathcal{N}} x_{ij} = 1, y_j \in \mathbb{Z}_+\},$$

and so we obtain the relaxation:

$$\inf \left\{ \sum\nolimits_{i \in \mathcal{M}} \sum\nolimits_{j \in \mathcal{N}} c_{ij} x_{ij} + \sum\nolimits_{j \in \mathcal{M}} f_j y_j : (\mathbf{x}, \mathbf{y}) \in \mathcal{F}_{RUFL} \right\}. \qquad (RUFL)$$

As already observed, it follows that $\upsilon(UFL) = \upsilon(RUFL)$ and the optimal solutions of both optimization are the same. Penalizing now the assignment constraints in the above optimization problem ($RUFL$), it follows by the same procedure as used for the uncapacitated facility location problem that the Lagrangian dual problem is given by

$$\upsilon(D) = \sup \left\{ -\sum\nolimits_{i \in \mathcal{M}} \lambda_i : f_j + b_j(\boldsymbol{\lambda}) \geq 0, j \in \mathcal{M}, \boldsymbol{\lambda} \in \mathbb{R}^m \right\}.$$

This Lagrangian dual problem is also called the *condensed dual*. There exists now a special purpose heuristic called DUALOC (Erlenkotter, 1978) to solve the above dual problem and this is surprisingly effective. Finally we look at the multisource capacitated version of the facility location problem. Again penalizing the assignment constraints, we obtain by the same approach as for the uncapacitated version that the Lagrangian dual function $\theta : \mathbb{R}^m \to \mathbb{R}$ is given by

$$\theta(\boldsymbol{\lambda}) = -\sum\nolimits_{i \in \mathcal{N}} \lambda_i + \sum\nolimits_{j \in \mathcal{M}} \min\{0, f_j + h_j(\boldsymbol{\lambda})\}$$

with

$$\begin{aligned} h_j(\boldsymbol{\lambda}) &= \inf \left\{ \sum\nolimits_{i \in \mathcal{M}} (c_{ij} + \lambda_j) x_{ij} : 0 \leq x_{ij} \leq 1, \sum\nolimits_{i \in \mathcal{M}} d_i x_{ij} \leq q_j, j \in \mathcal{N} \right\} \\ &= -\sup\{\sum\nolimits_{i \in \mathcal{M}} (-c_{ij} - \lambda_j) x_{ij} : 0 \leq x_{ij} \leq 1, \sum\nolimits_{i \in \mathcal{M}} d_i x_{ij} \leq q_j, j \in \mathcal{N}\}. \end{aligned}$$

The last problem is a LP-relaxation of the knapsack problem. Without loss of generality, we may now assume that $-c_{ij} - \lambda_j > 0$ for every $i \in \mathcal{M}$ (if not set $x_{ij} = 0$ and reduce problem!). To solve this problem, it is well-known that we need to order the ratios $\frac{-c_{ij} - \lambda_j}{d_i}$ in decreasing order given by

$$\frac{-c_{\pi(1)j} - \lambda_j}{d_{\pi(1)}} \geq \frac{-c_{\pi(2)j} - \lambda_j}{d_{\pi(2)}} \geq \ldots . \geq \frac{-c_{\pi(m)j} - \lambda_j}{d_{\pi(m)}}$$

for some permutation π and start assigning the demand of customers $\pi(1), \pi(2), \ldots, \pi(n)$ to location j until the capacity of this location is satisfied. The last customer assigned might only get a part of his demand delivered from that facility j. Again we can easily construct as for the uncapacitated version a

Lagrangian heuristic to convert the optimal solution of $\theta(\boldsymbol{\lambda})$ to a primal feasible solution. Observe that for the single source capacitated facility location problem we obtain the classical knapsack problem. Although this problem is $\mathcal{NP}$-hard, there are fast heuristics solving this problem approximately (Martello, 1990). This ends our discussion of fixed charged location problems on discrete spaces.

4 Conclusion

In this study, the Lagrangian relaxation approach and the main ideas behind this approach are discussed in detail for finite dimensional optimization problems with continuous and/or discrete decision variables. Contrary to many courses taught in graduate programs distinguishing between continuous and discrete optimization problems, we start in this chapter with a generic description of an optimization problem incorporating both type of problems and apply to that optimization problem the Lagrangian relaxation approach. By introducing the main ideas behind this approach, we identify the so-called Lagrangian dual problem and a class of continuous optimization problems for which the optimal objective value of the primal problem equals the optimal objective value of the Lagrangian dual problem. At the same time, we show that many important results much later developed in the field of Operational Research are special instances of the Lagrangian relaxation approach. The most important ones are the primal–dual relations in linear programming and the dual linear program and the Karush–Kuhn–Tucker conditions in nonlinear programming. As such this chapter does not contain any new results but tries to give an easy introduction to duality theory and its use in finite dimensional optimization problems for less mathematically oriented readers using the simplest possible proofs. Since many models in Operational Research can be formulated as optimization problems and the Lagrangian relaxation technique tries to find solutions of these models, a basic understanding of this technique is of importance to more application oriented researchers in this field. Also it is important to understand that seemingly different algorithms for solving certain problems are actually based on the same ideas. As an example (although not discussed in this chapter), we mention the Dantzig–Wolfe decomposition technique in linear programming and the dual version of it called Benders method. To illustrate its use to some specific examples, we showed in the last part of this chapter how these techniques can be applied in generating approximative solutions for the classical set covering and fixed charged facility location models.

Acknowledgments The authors like to thank the comments of anonymous referees, which greatly improved the previous version of this chapter.

References

Bazaraa, M. S., Jarvis, J. J., & Sherali, H. D. (1990). *Linear programming and network flows*. New York: John Wiley and Sons.

Bazaraa, M. S., Sherali, H. D., & Shetty, C. M. (1993). *Nonlinear programming: theory and algorithms*. New York: John Wiley and Sons.

Beasley, J., Dowsland, K., Glover, F., Laguna, M., Peterson, C., Reeves, C. R., & Soderberg, B. (1993). *Modern heuristic techniques for Combinatorial Problems*. New York: John Wiley and Sons.

Ben-Tal, A., & Nemirovski, A. (2001), *Lectures on modern convex optimization: analysis, algorithms and engineering applications*. Philadelphia: SIAM.

Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge: Cambridge University Press.

Choquet, G. (1976). *Lectures on analysis volume II: representation theory*. London: W.A. Benjamin.

Chvatal, V. (1983). *Linear programming*. New York: Macmillan Company.

Correa, R., & Lemaréchal, C. (1993). Convergence of some algorithms for convex minimization. *Mathematical Programming, 62*(1), 261–275.

Daskin, M. S. (2011). *Network and discrete location: models, algorithms, and applications*. New York: John Wiley and Sons.

Erlenkotter, D. (1978). A dual-based procedure for uncapacitated facility location. *Operations Research, 26*(6), 992–1009.

Everett III, H. (1963). Generalized Lagrange multiplier method for solving problems of optimum allocation of resources. *Operations Research, 11*(3), 399–417.

Fiacco, A. V., & McCormick, G. P. (1969). *Nonlinear programming: sequential unconstrained minimization techniques*. New York: Wiley.

Francis, R. L., McGinnis, L. F., & White, J. A. (1992). *Facility Layout and Location: an analytical approach*. Pearson College Division.

Frenk, J. B. G., & Kassay, G. (1999). On classes of generalized convex functions, Gordan–Farkas type theorems, and Lagrangian duality. *Journal of Optimization Theory and Applications, 102*(2), 315–343.

Frenk, J. B. G., & Kassay, G. (2005). Introduction to convex and quasiconvex analysis. In *Handbook of generalized convexity and generalized monotonicity* (pp. 3–87). New York: Springer.

Frenk, J. B. G., & Kassay, G. (2006). The level set method of Joo and its use in minimax theory. *Mathematical Programming, 105*(1), 145–155.

Frenk, J. B. G., & Kassay, G. (2007). Lagrangian duality and cone convexlike functions. *Journal of Optimization Theory and Applications, 134*(2), 207–222.

Frenk, J. B. G., & Kassay, G. (2008). On noncooperative games, minimax theorems and equilibrium problems. In A. Chinchuluum, P. M. Pardalos, A. Migdalas & L. Pitsoulis (Eds.) *On Noncooperative Games, Minimax Theorems and Equilibrium Problems.* (pp. 211–250). New York: Springer.

Frenk, J. B. G., Kas, P., & Kassay, G. (2007). On linear programming duality and necessary and sufficient conditions in minimax theory. *Journal of Optimization Theory and Applications, 132*, 423–439.

Garey, M. R., & Johnson, D. S. (1979). *Computers and intractability, a guide to the theory of NP completeness*. San Francisco: W.F Freeman.

Giannessi, F. (1984). Theorems of the alternative and optimality conditions. *Journal of Optimization Theory and Applications, 42*(3), 331–365.

Hiriart-Urruty, J. B., & Lemaréchal, C. (2013). *Convex analysis and minimization algorithms I: fundamentals*. New York: Springer Verlag.

Lagrange, J. L. (1806). *Leçons sur le calcul des fonctions*. Paris: Courcier.

Martello, S. (1990). *Knapsack problems: algorithms and computer implementations*. New York: John Wiley and Sons.

Nemhauser, G. L., & Wolsey, L. A. (1988). *Integer and combinatorial optimization*. New York: John Wiley and Sons.

Neumann, J. V. (1928). Zur theorie der gesellschaftsspiele. *Mathematische Annalen, 100*(1), 295–320.

Nesterov, Y., & Nenirovski, A. (2001). *Interior point polynomial algorithms in convex programming*. Philadelphia: SİAM.

Nocedal, J., & Wright, S. (2006). *Numerical optimization*. New York: Springer Verlag.

Rockafellar, R. (1970). *Convex analysis*. Princeton: Princeton University Press.

Saigal, R. (1997). Linear programming: a modern integrated analysis. *Journal of the Operational Research Society, 48*(7), 762–762.

Schrijver, A. (1998). *Theory of Linear and Integer Programming*. New York: John Wiley and Sons.

Shor, N. Z. (2012). *Minimization methods for non-differentiable functions*. New York: Springer Verlag.

Sion. (1958). On general minimax theorems. *Pacific Journal of Mathematics* (8), 171–178.

Wolkowicz, H. (1981). Some applications of optimization in matrix theory. *Linear Algebra and Its Applications, 40*, 101–118.

Wolsey, L. A. (1998). *Integer programming*. New York: John Wiley and Sons.

Yuan, G. X. (1999). *KKM theory and applications to nonlinear analysis*. New York: Marcel Dekker.

Measuring Foreign Trade-Logistics Efficiency: A DEA Approach and the Malmquist Index

Mehmet Fatih Acar and Alev Özer Torgalöz

Abstract With the effect of globalization, the importance of logistics activities has increased for both countries and companies. In particular, governments have aimed to improve their logistics performance in recent years. The World Bank has started to publish the Logistics Performance Index (LPI) since 2007, evaluating countries regarding their logistics activities. LPI has six sub-dimensions: "customs," "infrastructure," "international shipment," "quality of logistics services," "tracking/tracing," and "timeliness." This study aims to investigate the logistics performance of the member countries of the Organization for Economic Cooperation and Development (OECD) in respect to foreign trade. For this, LPI sub-dimensions, foreign direct investment, and export volumes are considered to measure the efficiencies. In this research, LPI and foreign trade values belongs to OECD member countries for 2007–2018 period are analyzed with using Data Envelopment Analysis (DEA) and Malmquist Index (MI). The main motivation of this study is to examine the changes in the foreign trade-logistics efficiencies of the OECD member countries from 2007 to 2018. This research can be regarded as being among the first studies in this context and provides a perspective for countries to improve their logistics efficiencies. Furthermore, the case of Turkey is investigated in detail to point out how developing countries can evaluate themselves compared to the developed ones. In conclusion, according to the results, some countries appear performing better in terms of logistics efficiencies related to foreign trade values. Moreover, the comparison analyses reveal that Turkey's efficiency is at a low level, and it needs to make necessary technical and legal arrangements for all dimensions

M. F. Acar (✉)
İzmir Katip Çelebi Üniversitesi, İİBF, Uluslararası Ticaret ve İşletmecilik Bölümü, Çiğli, İzmir, Turkey
e-mail: mehmet.fatih.acar@ikcu.edu.tr

A. Özer Torgalöz
İzmir Demokrasi Üniversitesi, İİBF, İşletme Bölümü, Karabağlar, İzmir, Turkey
e-mail: alev.ozer@idu.edu.tr

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_3

of LPIs. Required steps which can be rearranged for Turkey are also discussed at the end of the study.

Keywords Logistics Performance Index (LPI) · Data envelopment analysis · Turkey · Organization for Economic Cooperation and Development (OECD)

1 Introduction

Foreign trade transactions are critical for the economic development of countries (Malkowska & Malkowski, 2021). With the increase of global trade, logistics and supply chain operations have become very important factors for exporters and importers (Chu et al., 2015). Therefore, the logistics sector is gradually growing (Akdoğan & Durak, 2016). In today's competitive business environment, governments and companies should give importance to logistics operations to respond quickly for the demands of the market. Moreover, companies and governments need to pay attention to logistics processes to be more efficient in terms of time and cost.

Logistics is basically the management and controlling of transfer of goods, services, and information from producers to consumers (De Souza et al., 2007). Effectively managed and well-organized supply chains enable customers to receive products or services in a safe, fast, and cost-effective way. Therefore, firms should be aware of the importance of logistics to increase both their financial profits and reputation in their markets. According to Christopher (1992), in today's business world competition is not among firms but among supply chains. Therefore, not only transfer activities but also various supply chain management operations such as production planning and inventory management have become primary functions for enterprises. In this regard, transportation infrastructure, laws, the quality of personnel, the capacities of the warehouses, and using of different transport channels determine the supply chain performance for states and companies.

With the Logistics Performance Index (LPI) published by the World Bank since 2007, it has become easier to compare and analyze the logistics efficiencies of countries. This index provides information and overview on customs regulations, logistics costs, and transport infrastructure for different countries (Arvis et al., 2018). LPI basically measures the logistics performance of countries with six different dimensions via a survey. These dimensions are customs, infrastructure, international shipments, quality of logistics services, tracking/tracing, and timeliness. Customs is related with the efficiency levels of customs procedures of logistics activities. Infrastructure refers to the quality level of transport infrastructure. International shipments indicate competitive prices for international transfers within the market. Logistics quality and competence show the adequacy and quality of logistics services in a country. Tracking/tracing refers to the controllability of the shipment in terms of time and place. Finally, timeliness is an indicator of whether the delivery is on planned and expected time or not (Arvis et al., 2018).

In this study, foreign trade-logistics efficiencies based on LPI scores are analyzed using Data Envelopment Analysis (DEA) for OECD member countries. Since these countries play an important role in the global economy, investigating their foreign trade-logistics efficiencies provides valuable insight and information. Meanwhile, only data of OECD member countries are included in this study to meet the homogeneity for the data set. For this study, homogeneity is important as DEA determines the efficiency scores of decision-making units (DMUs) relatively. In other words when the efficiency of any DMU is calculated by using DEA methodology, other DMUs' performances play an important role. In addition, DEA calculates the efficiency as the ratio of weighted outputs to weighted inputs. If there is no homogeneity, it can be problematic to calculate efficiencies, because DMUs with low inputs and low outputs may appear more efficient. Therefore, if all countries are considered, some underdeveloped or developing countries may rank upper in efficiency rankings.

The purpose of this research is to point out how inefficient countries can improve themselves compared to the efficient ones. In this study, Turkey is determined as a case study to provide advice for developing countries and discuss their needs. There are some reasons for why we specifically choose Turkey. First, Turkey is a developing country and is commonly expected to be inefficient compared to developed OECD members. Second, it is a member of the OECD, and its European Union (EU) participation process still continues. In short, Turkey is a developing country, and its economy has faced significant changes.

The Malmquist Index (MI) provides information about the change in logistics efficiencies over different periods. Therefore, in this study, Malmquist Index (MI) is calculated for all OECD member countries to investigate the change of foreign trade-logistics efficiency for different years. There are many studies about the logistics efficiencies within the related literature, however, most of them consider specific LPI provided only for a particular year. The main motivation of this study is to examine the changes in the foreign trade-logistics efficiencies of OECD member countries by considering all LPI values from 2007 to 2018. Therefore, the study can be regarded as being among the first ones in this context and it contributes to the literature by showing how foreign trade-logistics efficiencies of countries have changed over the time with giving MI values. In addition, this study points out how a developing country can compare its foreign trade-logistics efficiencies with those of developed countries.

The research questions of this study are as follows: (1) Which OECD countries perform better in terms of foreign trade-logistics efficiency?, (2) Which countries have a very low foreign trade-logistics efficiency?, (3) Which countries have high and low levels of improvement in terms of foreign trade-logistics efficiency over the years?, (4) What is the situation of Turkey about foreign trade-logistics efficiency compared to other OECD countries?, and (5) Which types of precautions should be taken by Turkish policy makers with regard to the logistics efficiency?

The rest of the paper is organized as follows. Section 2 provides an extensive literature review about the studies on logistics performance. Sections 3 and 4 present the background information and application of the proposed model, respectively. The obtained results are illustrated and discussed in Sects. 5 and 6. Finally, Sect. 7 provides concluding remarks and future research directions.

2 Literature Review

Logistics efficiency is a critical issue within the operations management literature. Many studies have evaluated logistics operations at the micro-level. However, only a limited number of studies have discussed the performance of logistics activities at the macro-level (Rashidi & Cullinane, 2019).

This study uses the competitive advantage theory to explain the importance of logistics for trade and economies in a broader perspective. Porter's (1990) national competitive advantage theory basically argues that the main pillar of a country's development depends on whether it has competitive advantage within international markets or not (Ren & Ma, 2018). The reason why we focus on logistics is that it provides countries with a competitive advantage within the global trade. In fact, having a well-established logistics infrastructure and logistics system can be considered as supporting factors for a favorable or superior economic position. Bhatnagar and Teo (2009) regarded logistics management as criteria that enhances competitive positioning as a part of the value chain proposed by Porter (1985). In this context, the foreign trade-logistics efficiencies of countries can be seen as a competitive advantage for all countries, especially for the developing ones. Similarly, Puertas et al. (2014) stated that logistics is very important in exporting nations.

As mentioned above, LPI has been published by the World Bank since 2007 and these data have been used in various studies in the literature (Martí et al., 2017; Onsel Ekici et al., 2019; Rashidi & Cullinane, 2019). This index offers a wide range of different perspectives regarding the logistics operations of countries (Arvis et al., 2018). Moreover, different techniques have been applied for the data of LPI in the literature. For example, Quariguasi et al. (2009) examined logistics networks and efficiencies in Germany by taking environmental and economic conditions into consideration. In addition, Kabak et al. (2019) investigated the relationship between countries' logistics performances and competitiveness and their methodology included Bayesian Net (BN), Partial Least Square (PLS), and Importance-Performance Map Analysis (IPMA). This study points out the importance of "Business Sophistication," "Financial Market Development," "Infrastructure," "Good Market Efficiency," and "Higher Education and Training" which mostly affect the logistics performance. Myers et al. (2004) demonstrated that job skills have a positive effect on performances of logistics managers, however, they were unable to provide similar evidence for both experience and education. Fechner (2010) stated that the infrastructure is very critical for the improvement of logistics operations in countries. Furthermore, Martí et al. (2014) investigated the effects of LPI sub-scales on trade performance for developing countries and the research highlighted the importance of all LPI sub-dimensions on international trade. In addition, D'Aleo and Sergi (2017) examined the relationship between LPI and the Global Competitiveness Index (GCI) using panel data analysis. In a study by Sternad et al. (2018) DEA was used to analyze the logistics efficiency of EU countries. Similarly, Mesjasz-Lech (2019) demonstrated the importance of LPI for entrepreneurship within the fields of transportation and storage in EU. Liu et al. (2018) investigated the effects of logistics performance on environmental damage. In the study, it was shown that LPI sizes are significantly associated with CO_2 emissions. In addition, Luttermann et al. (2020) pointed out a positive relationship between logistics performance and FDI as a result of a panel data analysis for 20 Asian countries.

Min and Kim (2010) constructed a mixed logistics and environmental sustainability index using DEA method. Similarly, Lu et al. (2019) evaluated green transportation and logistics practices with an Environmental Logistics Performance Index (ELPI) for 112 selected countries. According to their research, there is a strong correlation between ELPI and LPI. Moreover, Yu and Hsiao (2016) presented an alternative way for countries to evaluate the LPI. In their studies, meta-boundary data envelopment analysis (Meta-DEA-AR) model was applied to evaluate the LPI scores, and recommendations were made for certain countries. In addition, Coto-Millán et al. (2013) analyzed the relationship between the logistics performance and economic growth of countries. Ekici et al. (2019) evaluated the effects of the Global Competitiveness Index (GCI) on the LPI and provided a general framework for policy makers on how they can improve their countries' logistics performance and showed that digitalization and supply chain analytics are crucial for countries to increase their logistics efficiencies. In addition, Richey et al. (2007) emphasized the impact of technology on logistics performance in their research. Likewise, Erkan (2014) examined the effect of technology with Global Competitiveness Index (GCI) and LPI and revealed that the technological infrastructure and market size affect LPI.

There are also different studies about LPI and Turkey within the related literature. In recent years, various statistical methods have been used for LPI-based comparative analyzes between Turkey and other countries. Ekici et al. (2016) investigated the relationship between LPI and Global Competitiveness Index (GCI) with using the Artificial Neural Network (ANN) for Turkey. In that research, fixed broadband Internet was found to be the most important topic for Turkey, as the fixed broadband Internet had a significant positive impact on the growth of logistics. With their methodology, Ekici et al. (2016) showed how resources of countries could be used to improve logistics competitiveness. Iris and Tanyas (2011) evaluated the LPI values of Turkey and discussed the necessary points to improve the performance in related fields. Danaci and Nacar (2017) compared Turkey and EU countries with considering the LPI and found a statistically significant relationship between the logistics performance and Gross Domestic Product (GDP). Guner and Coskun (2012) examined the relationship between economic/social factors and LPI. They found a significant correlation between them. Unalan and Yaprakli (2016) examined the Turkey's ranking in terms of LPI. They stated that Turkey's performance has improved but there is no significant change has been observed in the rankings over the years. Uca et al. (2015) studied the relationship between GDP and LPI and they revealed a significant relationship among customs, infrastructure, and GDP. In addition, Erdogan (2015) evaluated the situation of Turkey according to the LPI and discussed how to improve Turkey's logistics performance. Orhan (2019) compared Turkey and EU countries based on LPI results by using Entropy-weighted Edas methodology. Gungor et al. (2019) analyzed the LPI scores for Mediterranean economies and the research highlighted the importance of infrastructure and customs to have high level of GDP. Moreover, in their bibliometric analysis, Pekmezci and Mutlu (2018) provided a detailed and comprehensive literature review about Turkey and LPI.

In many studies, DEA has been frequently used to measure logistics efficiency. Since DEA allows using many inputs and outputs, it is used in this study to assess the foreign trade-logistics performances of selected countries. Although there are various studies about the LPI within the literature (e.g., Rashidi & Cullinane, 2019) many of them have investigated the logistics efficiencies for a specific time-period. However, this current research measures the changes in the foreign trade-logistics efficiencies of the OECD member countries for a period between 2007 and 2018. Therefore, we can argue that this research contributes to the logistics literature by calculating the Malmquist Index to observe the change of foreign trade-logistics efficiencies. Furthermore, this study suggests how a developing country can evaluate itself in comparison to the developed countries with respect to foreign trade-logistics efficiency based on the example of Turkey.

3 Methodology

In this study, the Data Envelopment Analysis (DEA) measures the logistics efficiencies of countries. This method bases on a mathematical programming, and it has been developed by Charnes et al. (1978) to find the relative efficiency of decision-making units (DMUs) with respect to each other. Units are often referred to as decision-making units (DMU) in DEA terminology. DEA is a popular tool to evaluate the relative efficiencies of different DMUs and is often used in many academic studies due to some advantages. The first advantage is to offer a comparable efficiency score for each DMU. In particular, it allows assessment between DMUs by defining a frontier to define efficient and inefficient units. Many inputs and outputs with different units can be analyzed in DEA together. In addition to these, there is no any prerequisite the nature of the production function for this method. Finally, DEA reveals comparative efficiency performance that permits to evaluate all input and output levels (Sevkli et al., 2007).

3.1 The CCR-DEA Model

The most popular DEA models are CCR (Charnes et al., 1978) and BCC (Banker et al., 1984) models. In this research, the CCR model is applied because overall productivity is considered and constant returns to scale are presumed. Nevertheless, BCC model considers pure technical efficiency and assumes variable returns to scale.

DEA model can be input- or output-oriented. Input-oriented models investigate the efficiency levels by deciding how much input should be utilized to obtain a specific output level for any DMU to become efficient. However, an output-oriented model aims to determine the suitable levels of outputs for a DMU to become efficient according to inputs. Briefly, if making changes on inputs is more proper than changes on outputs, input-oriented analysis is better to apply, however, if making changes on the outputs is more suitable, then output-oriented analysis should be chosen. In this research, the input-oriented DEA model is considered to determine how logistic performance can be improved by changing the level of current inputs.

To illustrate the basic DEA-CCR model mathematically, let us assume that each decision-making unit (DMUs) uses m inputs to produce n outputs at a given technology level. While X_{ij} represents the ith ($i = 1, 2, \ldots, m$) input of the m inputs of the jth DMU ($j = 1, 2, \ldots, k$), Y_{sj} represents the sth ($s = 1, 2, \ldots n$) output of the n outputs produced by jth DMU. The variables U_r ($r = 1, 2, \ldots, n$) and W_i ($i = 1, 2, \ldots, m$) are the weights of each output and input, respectively. The efficiency of DMU(o) can be written as follows: the efficiency of any decision unit is the ratio of the weighted output to the weighted input quantity. The representation of efficiency as a mathematical model is given below.

$$\text{Max} = \frac{\sum_{r=1}^{n} u_r Y_{r0}}{\sum_{i=1}^{m} w_i X_{i0}} = 1 \tag{1}$$

$$\frac{\sum_{r=1}^{n} u_r Y_{rj}}{\sum_{i=1}^{m} w_i X_{ij}} \leq 1 \quad j = 1, 2 \ldots k \tag{2}$$

$$u_r \text{ and } w_i \geq 0 \ \ (r = 1, 2, \ldots.., n) \ \text{ and } \ (i = 1, 2, \ldots\ldots, m) \tag{3}$$

The CCR model is shown below. In this algorithm, the score of the most efficient DMU is 1.

$$\text{Max} = \sum\nolimits_{r=1}^{0} \mu_r Y_{r0} \tag{4}$$

Constraints;

$$\sum\nolimits_{i=1}^{m} w_i X_{i0} = 1 \tag{5}$$

$$\sum_{r=1}^{n} \mu_r Y_{rj} - \sum_{i=1}^{m} w_i X_{ij} \leq 0 \tag{6}$$

$$\mu_r \text{and } w_i \geq 0 \ \ (r = 1, 2, \ldots.., n) \ \text{ and } \ (i = 1, 2, \ldots\ldots, m) \tag{7}$$

3.2 Malmquist Index

Normally, DEA methods finds the efficiency of DMUs for a sole time. However, in this study, DEA-based Malmquist productivity index developed by Färe et al. (1994) is operated to measure the efficiency change of a DMU among time stages. The Malmquist index is equal to the product of "Catch-up" and "Frontier-shift." The catch-up term is the ratio of the observed DMU efficiency score in different time periods. It shows the relationship between periods $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$. Moreover, the frontier-shift term implies the change in the efficient frontiers for specific DMU between the time periods $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$. The frontier-shift is calculated as the geometric mean of frontier-shift effects for periods $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$. The Malmquist Index (MI) is generated by Färe et al. (1994) between periods $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$ is given as follows:

$$\mathrm{MI} = \left[\frac{\delta^{t_1}\left((x_o, y_o)^{t_2}\right)}{\delta^{t_1}\left((x_o, y_o)^{t_1}\right)} \times \frac{\delta^{t_2}\left((x_o, y_o)^{t_2}\right)}{\delta^{t_2}\left((x_o, y_o)^{t_1}\right)} \right]^{1/2} \tag{8}$$

where $\delta^{t_1}\left((x_o, y_o)^{t_2}\right)$ shows the distance between the period t_2 observation to period t_1 technology. MI > 1 displays a rise in total product productivity of the observed DMU from the period t_1 to t_2. Likewise, MI = 1 and MI < 1 imply no change and a decrease in total product productivity, respectively. Further discussion on the calculation of Malmquist index is given in Tone (2004).

4 The Analysis of the Model

In this section, details of DEA model used in this research are explained. Moreover, DEA results are given for OECD member countries.

4.1 Construction of DEA Model

OECD is one of the most important organizations within the global economy and therefore it was selected as target set in this study to satisfy homogeneity. Inputs of this analysis are sub-dimensions of LPI, these are "customs," "infrastructure," "international shipment," "quality of logistics services," "tracking/tracing," and "delivery on time." In addition to this, outputs of DEA are the inward foreign direct investment (FDI) and export volumes. The value of outputs is chosen in terms of percentage of national Gross Domestic Product (GDP). The sources of data are the official website of World Bank and OECD. The inward FDI stock is the investment amount in the related country by foreign investors. As suggested in the literature (Martí et al., 2017), a monotone decreasing transformation (five minus the original values) is applied, because DEA computes the efficiency scores with considering the ratio weighted outputs to weighted inputs. If not, in the DEA results, efficient countries' score may be so low due to high scores of inputs. The Malmquist Index (MI) provides information about the change in efficiency at different times. For this reason, in this study, Malmquist Index (MI) values are shown for all OECD member countries to examine the change in foreign trade-logistics efficiency in different periods.

4.2 DEA Results for OECD

In this research, for data set of The Organization for Economic Co-operation and Development (OECD) countries, the CCR-input based DEA are used with DEA-Solver 13 software. Data of six different time periods (for 2007, 2010, 2012, 2014, 2016, and 2018) are evaluated in this study. Table 1 provides information about the highest foreign trade-logistics efficiency values among OECD countries for the relevant years. For instance, Belgium, Ireland, Netherlands, and Switzerland appear as the most efficient countries in 2007. Meanwhile, Chile (0.97) follows these countries. Looking at the 2007–2018 period, some countries outshine in terms of foreign trade-logistics efficiency. In this point, Ireland, Netherlands, Austria, and Belgium are frequently listed among the five most efficient ones. Furthermore, Netherlands and Ireland are found as the most efficient places for each year.

Table 2 displays the lowest foreign trade-logistics efficiency values for related countries during the 2007–2018 period. For instance, in 2007, Colombia (0.12) has the lowest efficiency score, followed by Greece (0.22), Mexico (0.22), Turkey (0.23), and the USA (0.25), respectively. Looking at the 2007–2018 period in a general, results show that some countries have frequently listed among the last five countries in terms of foreign trade-logistics efficiency such as Colombia, Turkey, Japan, and the USA. In particular, Colombia has the lowest foreign trade-logistics efficiency scores in many times.

As a result of the analysis, the USA and Japan are among the inefficient countries. The main reason is that these two countries are the main sources of global FDI investments. In short, the outward FDI values of these two countries are very high. Since inward FDI values are considered as output in the analyses, it is acceptable for these two countries to be among the inefficient countries. In fact, these countries are important actors within the global trade.

Table 1 Top five countries with the highest logistics efficiency values by years for OECD

2007		2010		2012		2014		2016		2018	
Countries	Eff. score	Countries	Eff. score	Countries	Eff. score	Countries	Eff. score	Countries	Eff. score	Countries	Eff. score
Belgium	1.00	Belgium	1.00	Belgium	1.00	Austria	1.00	Austria	1.00	Belgium	1.00
Ireland	1.00	Ireland	1.00	Czech Republic	1.00	Ireland	1.00	Belgium	1.00	Ireland	1.00
Netherlands	1.00	Netherlands	1.00	Estonia	1.00	Netherlands	1.00	Ireland	1.00	Netherlands	1.00
Switzerland	1.00	Switzerland	1.00	Ireland	1.00	Hungary	0.94	Netherlands	1.00	Switzerland	0.89
Chile	0.97	Chile	0.87	Netherlands	1.00	Belgium	0.93	Czech Republic	0.96	Denmark	0.77

Table 2 The five countries with the lowest logistic efficiency value by years for OECD

2007		2010		2012		2014		2016		2018	
Countries	Eff. score	Countries	Eff. score	Countries	Eff. score	Countries	Eff. score	Countries	Eff. score	Countries	Eff. score
USA	0.25	Greece	0.23	New Zealand	0.28	Italy	0.28	New Zealand	0.19	Japan	0.23
Turkey	0.23	Japan	0.22	Turkey	0.23	USA	0.24	Japan	0.17	Israel	0.22
Mexico	0.22	USA	0.19	USA	0.22	Japan	0.22	USA	0.15	Turkey	0.20
Greece	0.22	Turkey	0.14	Japan	0.18	Turkey	0.18	Turkey	0.15	USA	0.19
Colombia	0.12	Colombia	0.12	Colombia	0.14	Colombia	0.15	Colombia	0.12	Colombia	0.18

Table 3 The best and worst MI values

Countries	(MI)
Top three countries with highest improvement in efficiency according to MI2	
Colombia	1.50
Latvia	1.22
Lithuania	1.21
Bottom three countries with lowest reduction in efficiency according to MI2	
Israel	0.35
Chile	0.34
Canada	0.34

Table 4 DEA results of Turkey for OECD data set

	2007	2010	2012	2014	2016	2018
Efficiency score	0.23	0.14	0.23	0.18	0.15	0.20
Ranking	33	35	33	35	35	34
Catch-up effect		0.71	1.69	0.86	0.82	1.41
Frontier effect		1.17	0.81	1.28	1.01	0.87
MI1		0.83	1.37	1.10	0.82	1.24
MI2						1.30

At the end of Malmquist analysis, three different scores are obtained for each country based on the data. These are catch-up effect, frontier-shift, and Malmquist index. As mentioned above, catch-up reflects the change in efficiency between two different periods. The frontier-shift term measures the change in the efficient frontiers for the any decision-making unit between two different periods. Malmquist index is calculated as the product of catch-up and frontier-shift effect.

The MI value in Table 3 shows the way of change for the foreign trade-logistics efficiency of countries from 2007 to 2018. The countries with the highest and lowest MI values for OECD countries are given below. Colombia has the highest MI value (1.50). It is followed by Latvia (1.22) and Lithuania (1.21), respectively. Among OECD countries, these has been the top three countries that have increased their foreign trade-logistics efficiency from 2007 to 2018. Moreover, the countries with the biggest reduction in foreign trade-logistics efficiency are the Israel (0.35), Chile (0.34), and Canada (0.34).

Table 4 includes the foreign trade-logistics efficiency value, ranking related to efficiency scores, catch-up effect, frontier effect, and Malmquist Index (MI) of Turkey for 2007–2018 period. MI1 reveals the Malmquist Index value for two consecutive years, however, MI2 gives the direct Malmquist value from 2007 to 2018.

Turkey's foreign trade-logistics efficiency value has decreased over the years, for example, it is 0.23 in 2007 and 0.29 in 2018. In addition, Turkey's foreign trade-logistics efficiency score is quite low as 0.14 in 2010. There is no serious fluctuation for Turkey's ranking among OECD countries, for instance; Turkey's

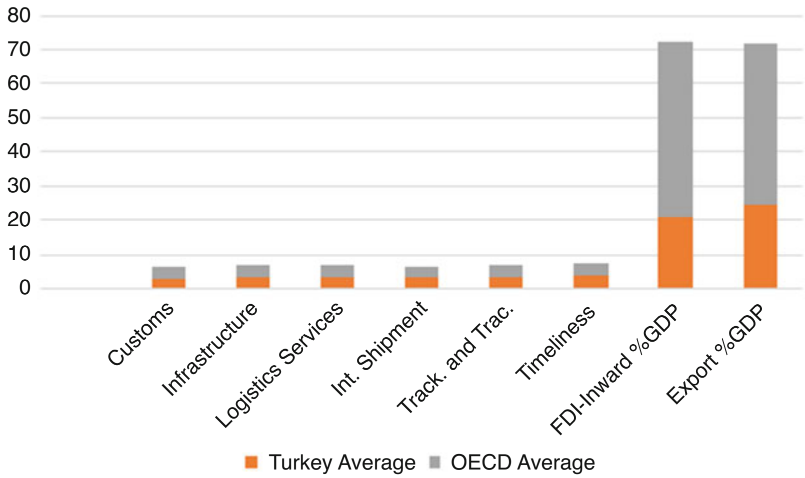


Fig. 1 Average LPI and foreign trade values for Turkey and OECD during 2007–2018

ranking changes between 33 and 35 over the years. Catch-up Effect, Frontier Effect, and MI1 and MI2 values are also given in Table 4. MI1 values are bigger than 1 in 2012 and 2014. It means that, for 2012 and 2014, Turkey's foreign trade-logistics efficiencies have improved compared to the previous year. Although the catch-up effect of Turkey (0.86) is below 1 in 2014, foreign trade-logistics efficiency of Turkey has risen compared to 2012 with the contribution of frontier effect. However, in many years, MI1 value has remained below 1 except 2012 and 2014. Moreover, MI2 value is 1.30 which is higher than 1. This implies that Turkey's foreign trade-logistics efficiency has increased from 2007 to 2018.

During 2007–2018 period, the average values for each of the sub-index of LPI of Turkey and OECD members are shown in Fig. 1 The results show that for all sub-indices, Turkey's average scores are lower than the average scores of OECD countries. Furthermore, Turkey's average percentage values of export and foreign direct investment to GDP are lower than those of OECD. This situation also points to the reason for Turkey's inefficiency compared to OECD countries.

Table 5 shows the projection of input values of Turkey for the period 2007–2018. This value means how many percentages should be increased or decreased for Turkey's input values to be efficient in terms of foreign trade-logistics efficiency. As mentioned above, a monotone decreasing transformation (five minus the original values) is applied for the inputs (LPI sub-dimensions) of analysis, as recommended by the literature. Therefore, the input values in this research are not the same as the LPI sub-dimension scores, in contrast, they have been inverted. In this case, there is a need to interpret the values shown in Table 5. Table 5 emphasizes that all input values should be reduced by a certain percentage. In this context, Turkey's input values for each year should be lower and this situation implies that the score of LPI sub-dimensions should be higher. Because the input values in the analysis are found

Table 5 Projection values for inputs of Turkey during 2007–2018

Year	Customs	Infrastructure	Logistics quality and competence	International shipments	Tracking and tracing	Timeliness
	Diff.(%)	Diff.(%)	Diff.(%)	Diff.(%)	Diff.(%)	Diff.(%)
2007	−77.38	−77.38	−86.97	−79.36	−81.73	−82.69
2010	−86.38	−86.70	−86.33	−86.33	−87.79	−92.95
2012	−80.39	−77.17	−77.17	−78.16	−78.17	−77.17
2014	−84.44	−81.88	−82.19	−81.67	−86.56	−83.05
2016	−84.54	−84.97	−86.33	−84.54	−85.90	−86.53
2018	−79.30	−79.30	−83.00	−79.30	−79.82	−82.31

with the formula of five minus the original LPI sub-dimension scores. In conclusion, projection values says that Turkey should improve its LPI sub-dimensions values in each year.

5 Discussion

In recent years, logistics has become a very important topic for both countries and companies as well as their supply chains. In today's competitive environment, supply chains compete with other supply chains (Christopher, 1992). In this sense, as stated previously, the Logistics Performance Index (LPI) published by the World Bank is a good source of data for countries to evaluate themselves.

Logistics efficiencies can positively affect a country's domestic and foreign trade. International shipment and customs sub-factors of LPI are directly related to import and export transactions. Therefore, the impact of qualifications of the LPI on any country's foreign trade is inevitable. Moreover, LPI sub-criterion such as infrastructure and quality of logistics services also can determine the speed and cost of logistics operations. For this reason, higher logistics efficiencies lead to lower costs. When considering that one of the factors affecting inflation the most is logistics costs, efficiencies in transportation operations are crucial for the national and household economies. In addition, LPI is also critical for the image of countries. Sub-factors such as timeliness and tracing/tracking are directly related to logistics quality of countries. These advantages contribute to an increase in investments like Foreign Direct Investment.

In the research, by considering the OECD member countries and using DEA and Malmquist method, it is attempted to measure the foreign trade-logistics efficiencies and the changes over the time. Analyses show that Ireland and the Netherlands appear among the most efficient countries in many years. Countries aiming to improve in terms of logistics should carefully examine the transportation operations implemented in these most efficient countries and imitate similar practices according to their own needs. Turkey and Colombia appear among the countries with lowest logistics efficiency values. These countries need to develop seriously in terms of transportation activities.

According to the Malmquist Index, the countries improving the most are Colombia, Latvia, and Lithuania. Therefore, countries that want to improve their logistics operations should also take the logistics applications in these countries into consideration. Countries that have a low logistics efficiency can determine a roadmap for themselves by examining the law systems, legal regulations, infrastructure, and tax/incentive systems in these three countries. In addition, the frequently observed countries that have not improved over time in the analysis reveal as Israel, Chile, and Canada. These countries should also study their logistics competencies. Moreover, all countries should investigate these three countries to determine why they are falling behind in order not to make the same mistakes.

6 Policy Implications for Turkey

In this study, using the LPI data, logistics efficiencies of Turkey have been compared with the OECD countries. When the DEA and MI results are evaluated, Turkey has had a low performance. For this reason, Turkey should develop its performance with respect to the six sub-LPI dimensions. First, for the customs sub-dimension, the relevant laws and regulations should be revised to make customs operations more efficient. Moreover, the bureaucratic procedures in customs should be made easier and shorter. In addition, customs tax rates should be reviewed, and the customs gates should be increased. For the infrastructure sub-dimension, related investments should be continued rapidly, and unfinished projects should be completed in a short time. For example, ongoing constructions in Turkey such as Galataport in Istanbul and Çandarli Port in Izmir projects are quite crucial for the international trade. Moreover, new projects such as logistics villages should be developed, and legal arrangements should continue to facilitate logistics activities. Regarding the quality of logistics services and the ease of delivery arrangement sub-dimensions, authorized institutions should find solutions for market needs, especially in rail and maritime transport. Furthermore, Turkish government should ensure the usage of various alternatives for effective logistics operations. As an example, the Cabotage Law (seaway transportation law) can be updated to benefit more from maritime transportation. In addition, authorized corporations should take legal steps to make railways operated by private companies. For the tracking and tracing sub-dimension, Turkey should also improve its situation. First, necessary infrastructures such as satellite communication, telecommunication, radio frequency systems should be completed flawlessly by the state. The private sector should be encouraged to follow-up and monitor logistics operations and new sanctions should be introduced for some sectors such as food and medicine. Especially in cold chain activities, companies should be audited more carefully. Furthermore, various credit and incentive systems should be implemented by the government to encourage private organizations to make new investments on their logistics operations. Finally, in terms of the timeliness sub-dimension, the state and the private sector have duties as for other fields. Firstly, within the scope of public activities, infrastructure needs such as roads and bridges should be met quickly. Moreover, to accelerate the processes, legal steps that slow down logistics operations such as customs and tax should be reviewed in accordance with today's needs and conditions. In addition, companies should pay more attention to business ethics in this regard and provide training to their employees on ethics. Besides, companies should impose sanctions on their employees if necessary and should be responsive in terms of on time delivery.

7 Conclusion

In earlier times, logistics was only concerned with the transportation activities. With the increasingly competitive environment, it has gained a wider perspective. Normally, buyers and sellers engage in logistics activities to transport products. Therefore, logistics can be considered as an essential part of commercial life. Studies by Langley et al. (2008), Mangan et al. (2008), and Rushton et al. (2009) have provided a definition for logistics from a broader perspective. These studies suggest that logistics is an integration of information, packaging, storage, and transportation system.

Today, the correct and timely transportation of products is more important than the past due to the international trade. Especially with the increasing internet facilities, demands of companies and consumers have largely increased. Therefore, governments are rapidly making infrastructure investments such as ports, terminals, and highways to provide efficiency in international trade. Moreover, different projects such as intermodal transportation and logistics villages are frequently organized nowadays. All these developments facilitate the trade of goods and significantly reduce supply chain costs. Trade logistics is significant to improve regional and international trade (Pavcnik, 2002). Today's competitive market conditions have made logistics more important than ever, as developing countries increasingly compete for access to the international markets in efficient ways (Devlin & Yee, 2005). As stated by the World Bank's report (2010), countries' policies and logistics procedures directly affect foreign trade performance. In this context, the Logistics Performance Index (LPI) published by the World Bank is a good data source for research. This index allows to make comparisons for countries and can provide guidance for politicians and academics to evaluate trade performances.

In this research, it was aimed to analyze foreign trade values of OECD countries according to the six different LPIs and provide an understanding of how developing countries can improve their positions by comparing their own logistics performance with the developed ones'. Foreign trade-logistics efficiency was calculated by using DEA. Sub-dimensions of LPI are considered as inputs, while percentage values of foreign direct investment and export to GDP are assumed as outputs for OECD countries. The results revealed that Ireland and the Netherlands have high performances in terms of foreign trade-logistics efficiency. In addition, it is observed that Colombia and Turkey had very low levels of foreign trade-logistics efficiency. According to the Malmquist Index results, it was found that Colombia, Latvia, and Lithuania made very good improvement. Conversely, Israel, Chile, and Canada could not improve their transportation capabilities over the time. Therefore, the practices of these countries should be carefully examined by the field experts. These inferences can provide a roadmap for the developing countries regarding their logistics systems. Moreover, with the case of Turkey, it is emphasized how developing countries' self-evaluations could be made for relevant improvements. Turkey's logistics efficiency is lower than many OECD countries and this study reveals which regulations are necessary to improve Turkey's logistics performance.

As with all studies, this research has some limitations. One limitation is about using non-parametric method and only considering the developed and developing countries when calculating the logistics efficiencies. In future research, more countries can be included to generate foreign trade-logistics efficiency values and parametric efficiency methods like Stochastic Frontier can be used to compare countries' efficiencies.

References

Akdoğan, M. Ş., & Durak, A. (2016). Logistic and marketing performances of logistics companies: A comparison between Germany and Turkey. *Procedia-Social and Behavioral Sciences, 235*, 576–586. https://doi.org/10.1016/j.sbspro.2016.11.084

Arvis, J.-F., Ojala, L., Wiederer, C., Shepherd, B., Raj, A., Dairabayeva, K., & Kiiski, T. (2018). *Connecting to compete 2018: Trade logistics in the global economy*. World Bank.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30*, 1078–1092. https://doi.org/10.1287/mnsc.30.9.1078

Bhatnagar, R., & Teo, C. C. (2009). Role of logistics in enhancing competitive advantage. *International Journal of Physical Distribution and Logistics Management, 39*(3), 202–226. https://doi.org/10.1108/09600030910951700

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research, 2*, 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Christopher, M. L. (1992). *Logistics and supply chain management*. Pitman Publishing.

Chu, D., Li, C., Xu, X., & Zhang, X. (2015). A graph-based framework for route optimization in sea-trade logistics. *Mathematical Problems in Engineering*, 2015.

Coto-Millán, P., Agüeros, M., Casares-Hontañón, P., & Pesquera, M. Á. (2013). Impact of logistics performance on world economic growth (2007–2012). *World Review of Intermodal Transportation Research, 4*, 300–310.

Danaci, T., & Nacar, R. (2017). *Comparing the foreign trade and logistic performance of turkey and EU members with cluster analysis* [Conference presentation]. Global Business Research Congress (GBRC), Istanbul, May 24–25, 2017.

D'Aleo, V., & Sergi, B. S. (2017). Does logistics influence economic growth? The European experience. *Management Decision, 55*(8), 1613–1628.

De Souza, R., Goh, M., Gupta, S., & Lei, L. (2007). An investigation into the measures affecting the integration of ASEAN's priority sectors (Phase 2): The case of logistics. *REPSF Project, 6*(001).

Devlin, J., & Yee, P. (2005). Trade logistics in developing countries: The case of the Middle East and North Africa. *World Economy, 28*(3), 435–456.

Ekici, S. O., Kabak, O., & Ulengin, F. (2016). Linking to compete: Logistics and global competitiveness interaction. *Transport Policy, 48*, 117–128.

Erdogan, A. (2015). Performance of Turkey in logistics and suggestions for solutions to improve performance. *Social Sciences Research Journal, 4*(1), 1–7.

Erkan, B. (2014). Türkiye'de lojistik sektörü ve rekabet gücü. *Assam Uluslararası Hakemli Dergi, 1*(1), 44–65.

Färe, R., Grosskopf, S., Lindgren, B., & Roos, P. (1994). Productivity developments in Swedish hospitals: A Malmquist output index approach. In *Data envelopment analysis: Theory, methodology, and applications* (pp. 253–272). Springer.

Fechner, I. (2010). Role of logistics centers in national logistics system. *Electronic Scientific Journal of Logistics, 6*(2).

Guner, S., & Coskun, E. (2012). Comparison of impacts of economic and social factors on countries' logistics performance: a study with 26 OECD countries. *Research in Logistics and Production, 2*(4), 329–343.

Gungor, Ş., Dursun, E., & Karaoglan, A. (2019). Lojistik faaliyetlerin Akdeniz ülke ekonomileri ile ilişkisi. *İktisat Politikası Araştırmaları Dergisi, 6*(2), 91–106.

Iris, C., & Tanyas, M. (2011). Analysis of Turkish logistics sector and solutions selection to emerging problems regarding criteria listed in logistics performance index (LPI). *International Journal of Business and Management Studies, 3*(1), 93–102.

Kabak, O., Onsel Ekici, S., & Ulengin, F. (2019). Analyzing two-way interaction between the competitiveness and logistics performance of countries. *Transport Policy*. https://doi.org/10.1016/j.tranpol.2019.10.007

Langley, C., Coyle, J., Gibson, B., Novak, R., & Bardi, E. (2008). *Managing supply chain: A logistics approach*. South–Western College Publishing.

Liu, J., Yuan, C., Hafeez, M., & Yuan, Q. (2018). The relationship between environment and logistics performance: Evidence from Asian countries. *Journal of Cleaner Production, 204*, 282–291. https://doi.org/10.1016/j.jclepro.2018.08.310

Lu, M., Xie, R., Chen, P., Zou, Y., & Tang, J. (2019). Green transportation and logistics performance: An improved composite index. *Sustainability, 11*, 2976.

Luttermann, S., Kotzab, H., & Halaszovich, T. (2020). The impact of logistics performance on exports, imports and foreign direct investment. *World Review of Intermodal Transportation Research, 9*, 27–46.

Malkowska, A., & Malkowski, A. (2021). International trade in transport services between Poland and the European Union. *Sustainability, 13*(1), 424. https://doi.org/10.3390/su13010424

Mangan, J., Lalwani, C., & Butcher, C. (2008). *Global logistics and supply chain management*. Wiley.

Martí, L., Martín, J. C., & Puertas, R. (2017). A DEA-logistics performance index. *Journal of Applied Econometrics, 20*, 169–192. https://doi.org/10.1016/S1514-0326(17)30008-9

Martí, L., Puertas, R., & García, L. (2014). The importance of the Logistics Performance Index in international trade. *Applied Economics, 46*, 2982–2992. https://doi.org/10.1080/00036846.2014.916394

Mesjasz-Lech, A. (2019). Logistics performance of European Union markets: Towards the development of entrepreneurship in the transport and storage sector. *Global Journal of Environmental Science and Management, 5*, 122–130. https://doi.org/10.22034/gjesm.2019.05.SI.14

Min, H., & Kim, I. (2010). *Measuring the effectiveness of the country's green supply chain from a macro perspective*. Proceedings of the First Annual Kent State International Symposium on Green Supply Chains Canton, Ohio. 29–30.

Myers, M. B., Griffith, D. A., Daugherty, P. J., & Lusch, R. F. (2004). Maximizing the human capital equation in logistics: Education, experience, and skills. *Journal of Business Logistics, 25*(1), 211–232.

Orhan, M. (2019). Türkiye ile Avrupa Birliği ülkelerinin lojistik performanslarının Entropi ağırlıklı EDAS yöntemiyle karşılaştırılması. *Avrupa Bilim ve Teknoloji Dergisi, 17*, 1222–1238.

Onsel Ekici, S., Kabak, O., & Ulengin, F. (2019). Improving logistics performance by reforming the pillars of Global Competitiveness Index. *Transport Policy, 81*, 197–207. https://doi.org/10.1016/j.tranpol.2019.06.014

Quariguasi Frota Neto, J., Walther, G., Bloemhof, J., van Nunen, J. A. E. E., & Spengler, T. (2009). A methodology for assessing eco-efficiency in logistics networks. *European Journal of Operational Research, 193*, 670–682. https://doi.org/10.1016/j.ejor.2007.06.056

Pavcnik, N. (2002). Trade liberalization, exit, and productivity improvement: Evidence from Chilean plants. *Review of Economic Studies, 69*(1), 245–276.

Pekmezci, M., & Mutlu, M. (2018). *Türkiye*'*yi konu alan 'lojistik performans indeksi' çalışmalarına yönelik bibliyometrik bir analiz* [Conference presentation]. IV. International Caucasus-Central Asia Foreign Trade and Logistics Congress, Didim/Aydin.

Porter, M. E. (1985). Technology and competitive advantage. *Journal of Business Strategy*.

Porter, M. E. (1990). The competitive advantage of nations. *Competitive Intelligence Review, 1*(1), 14–14.

Puertas, R., Martí, L., & García, L. (2014). Logistics performance and export competitiveness: European experience. *Empirica, 41*(3), 467–480.

Rashidi, K., & Cullinane, K. (2019). Evaluating the sustainability of national logistics performance using data envelopment analysis. *Transport Policy, 74*, 35–46. https://doi.org/10.1016/j.tranpol.2018.11.014

Ren, Z., & Ma, Y. (2018). *The significance of comparative advantage theory and competitive advantage theory to the development of China's foreign trade* [Conference presentation]. 2018 International Conference on Social Science and Education Reform (ICSSER 2018) (10–13). Atlantis Press.

Richey, R. G., Daugherty, P. J., & Roath, A. S. (2007). Firm technological readiness and complementarity: Capabilities impacting logistics service competency and performance. *Journal of Business Logistics, 28*(1), 195–228.

Rushton, A., Oxley, J., & Croucher, P. (2009). *The handbook of logistics and distribution management*. Kogan Page.

Sevkli, M., Lenny Koh, S. C., Zaim, S., Demirbag, M., & Tatoglu, E. (2007). An application of data envelopment analytic hierarchy process for supplier selection: A case study of BEKO in Turkey. *International Journal of Production Research, 45*, 1973–2003. https://doi.org/10.1080/00207540600957399

Sternad, M., Skrucany, T., & Jereb, B. (2018). International logistics performance based on the DEA analysis. *Communications – Scientific Letters of the University of Zilina*. 20.

Tone, K. (2004). Malmquist productivity index. In W. W. Cooper, L. M. Seiford, & J. Zhu (Eds.), *Handbook on data envelopment analysis* (pp. 203–227). Springer.

Uca, N., Civelek, M. E., & Çemberci, M. (2015). The effect of the components of logistics performance index on gross domestic product: Conceptual model proposal. *Eurasian Business & Economics Journal, Eurasian Academy of Sciences, 1*, 86–93.

Unalan, M. S., & Yaprakli, T. S. (2016). *Küresel lojistik performans endeksi ve Türkiye'nin son 10 yıllık lojistik performansının analizi*. http://hdl.handle.net/11508/8945

World Bank. (2010). *Trade and transport facilitation assessment: a practical toolkit for country implementation*. World Bank.

Yu, M.-M., & Hsiao, B. (2016). Measuring the technology gap and logistics performance of individual countries by using a meta-DEA–AR model. *Maritime Policy & Management, 43*, 98–120. https://doi.org/10.1080/03088839.2015.1037372

Further Reading

Chow, G., Heaver, T. D., & Henriksson, L. E. (1994). Logistics performance: Definition and measurement. *International Journal of Physical Distribution and Logistics Management, 24*(1), 17–28.

Jane, C.-C. (2011). Performance evaluation of logistic systems under cost and reliability considerations. *Transportation Research Part E: Logistics and Transportation Review, 47*, 130–137. https://doi.org/10.1016/j.tre.2010.09.012

Markovits-Somogyi, R., & Bokor, Z. (2014). Assessing the logistics efficiency of European countries by using the DEA-PC methodology. *Transport, 29*(2), 137–145.

Partially Non-discretionary Measures for Green Transportation Corridors Performance Index: A DEA Approach

Isotilia Costa Melo, Paulo Nocera Alves Junior, Thiago Guilherme Péra, Daisy Aparecida Do Nascimento Rebelatto, and José Vicente Caixeta-Filho

Abstract Freight transportation is vital to a nation's long-term development and its performance needs to be carefully evaluated to ensure the efficiency of haulage infrastructure decisions. Frequently, real-world physical barriers pose transportation constraints that are impossible to be completely overpassed or ignored. Previous studies on benchmarking Green Transport Corridors (GTCs) through routes efficiency have not considered the possibility of partially non-discretionary (pND) measures (only a certain percentage of the measure is controllable). The present paper creates a long-distance cargo haulage performance index that will be deemed as Logistic Composite Index (LCI) integrating pND measures using a Data Envelopment Analysis (DEA) methodology. Since infrastructure aspects can be assumed to be a Variable Returns to Scale (VRS), huge investments may be necessary for the possibility of just partially reducing the length of a route in a certain percentage by private and public investment strategies. This characteristic was incorporated, for the first time, with pND measures in a Double-Frontier of a

I. Costa Melo (✉)
São Carlos Engineering School (EESC), University of São Paulo (USP), São Carlos, SP, Brazil

Faculty of Engineering and Science (FES), Universidad Adolfo Ibáñez (UAI), Viña del Mar, Chile

P. Nocera Alves Junior
Luiz de Queiroz College of Agriculture (ESALQ), University of São Paulo (USP), Piracicaba, SP, Brazil

Universidad Católica del Norte (UCN), Escuela de Ingeniería de Coquimbo (EIC), Larrondo, Coquimbo, Chil
e-mail: epaulo.alves@ce.ucn.cl

T. G. Péra · J. Vicente Caixeta-Filho
Luiz de Queiroz College of Agriculture (ESALQ), University of São Paulo (USP), Piracicaba, SP, Brazil
e-mail: thiago.pera@usp.br; jose.caixeta@usp.br

D. Aparecida Do Nascimento Rebelatto
São Carlos Engineering School (EESC), University of São Paulo (USP), São Carlos, SP, Brazil
e-mail: daisy@sc.usp.br

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_4

Slack-Based Measure (SBM), and under VRS assumptions (pND-DF-SBM-VRS). Therefore, the present chapter integrates a novelty in DEA literature with practical implications for public investments. The method is applied to the context of soybean transportation, one of the relevant Brazilian exporting products, during the harvest of 2018/2019, from the main mid-sized producing regions to the key exporting ports. The proposed approach and findings provide insights into the public and private long-term investment strategies and infrastructure policies, especially in Brazil and developing countries.

Keywords Data envelopment analysis (DEA) · Partially non-discretionary slack-based measure (pND-SBM) · Construction of composite index · Freight transportation · Brazilian soybean

1 Introduction

In 2007, the European Commission's Freight Transport Logistics Action Plan introduced the concept of Green Transport Corridors (GTC) for freight transportation between major hubs as integrated multimodal used to reduce environmental impact via road, rail, waterways, and intelligent technologies (European Commission, 2007).

Green Transport Corridors (GTCs) promote environmental-friendly freight transportation through the efficient management of investments, operations, integration of transportation routes, and transportation modes (land, water, and air) (Panagakos, 2015). For achieving and managing the maximum efficiency, it is necessary to measure the GTCs performance, through Logistics Performance Index (LPI), especially considering economically relevant hubs and long-distance routes (Panagakos, 2015). It is important to note that the LPI from the World Bank (2018) was developed using Principal Component Analysis (PCA). Also, the LPI is applied at a country-level, without considering regional specifications or in-country transportation routes. Hence, the LPI from the World Bank (2018) differs from the LPIs developed using Data Envelopment Analysis (DEA) to evaluate routes and corridors. Though, in practical terms, the proposal of a wide accepted LPI-development methodology faced different obstacles, depending on if the LPI is supposed to be applied at a country level or at a regional level (Alves Junior et al., 2021; Melo et al., 2018, 2020; Rentizelas et al., 2019).

Specifically, previously proposed LPIs which rely on DEA models to assess and compare regions have faced the common challenge of factoring "route length" into their models. Firstly, some papers (Alves Junior et al., 2021; Rentizelas et al., 2019) excluded the transported length of the model, considering only the indirect variables (dependent on the length) such as transportation costs, fuel consumption, emissions, etc. This approach has the back draw of ignoring one of the most affecting logistic characteristics.

For example, a DEA Slack-Based Measure (SBM) model, with variable returns to scale (VRS), was applied for choosing alternatives in the international biomass supply chain (Rentizelas et al., 2019). Three variables were considered: costs and energy input (as inputs) and emissions (as undesirable output). It can be stated that the number of chosen variables (three) is very limited to incorporate the complexity of the system. Alves Junior et al. (2021) proposed a single multi-criteria Logistics Composite Index (LCI) for GTCs. The authors applied it to Brazilian agricultural bulk transport export corridors, considering the existing and planned infrastructure in the harvest year of 2018/2019. They used seven variables (classified as desired and undesired inputs as well as desired and undesired outputs) but did not use the length in the DEA model.

Secondly, on the other hand, it is also possible to consider the length as a totally non-discretionary (tND) measure, assuming, i.e., decision- and policy-makers cannot change the length of the route, independently of their amount of investments or efforts. This was proposed by Melo et al. (2018) for investigating 102 soybean haulage routes in Brazil and the USA. The authors considered nine variables, classified them into inputs, outputs, undesirable outputs, and length as a tND measure.

Thirdly, it is also possible to consider the length as a DEA input, i.e., a measure, which the minimization is aimed (Cook et al., 2014). The classification of the length as an input implies the assumption that the transported length is fully under the control of decision- and policy-makers, depending exclusively on interests and effort focus.

We argue that, in real-world applications, decision- and policy-makers can change the transported length. Though they are usually limited by external constraints at a certain level. Hence, we investigate the possibility of integrating the length to the model as a partially Non-Discretionary (pND) measure, i.e., a measure that can be reduced until a certain percentage.

Along these lines, we aim to propose a long-distance cargo haulage performance index (LPI) methodology, integrating pND measures. For the first time in an LPI application, the pND characteristic was incorporated in a Double-Frontier of a Slack-Based Measure (SBM) under Variable Return to Scale (VRS) assumptions. The application is in 12 GTCs (encompassing 254 routes), considering the soybean transportation in Brazil, during the harvest of 2018/2019.

Hence, the LCI proposed here, incorporating pND measures and applying a Double-Frontier Data Envelopment Analysis (DEA), Slack-Based Measure (SBM) under Variable Return to Scale (VRS) assumption (pND-DF-SBM-VRS) to evaluate GTCs and their multimodal routes is a novelty, resulting in innovative methodology with practical implications for public investments.

Subsequently, the results of the proposed methodology were compared to the results considering the length as a tND measure and as an input. The pND efficiency results were similar to the efficiency results considering the length as an input. Though the pND assumption proved to be useful for constructing efficiency-improvement goals. Goals constructed based on the input assumption can be physically unachievable (such as proposing 18% of the length reduction for reaching efficiency, passing through a natural reserve area).

The long-distance cargo haulage performance index integrating pND measures may be used to guide future investments in infrastructure. And the methodology can be a useful tool in different contexts of application (such as other countries and other transported cargos).

2 Literature Review

Based on the multi-attribute utility theory (MAUT) and the decision theory, Dyckhoff and Souren (2020) proposed the multi-criteria production theory (MCPT) for applying methods to multi-criteria decision making (MCDM) problems—such as Data Envelopment Analysis (DEA) for decision-making in production systems. Many previous authors tried to formulate special DEA-MCDM models (Belton & Vickers, 1993; Doyle & Green, 1993; Joro et al., 1998) with some specific characteristics from Multi-Objective Linear Programming (MOLP). However, in general, DEA is a method to measure the efficiency of DMUs (Charnes et al., 1978), but its concept also relies on decision theory, even though this aspect has been ignored by part of the DEA literature, as well as it relies on the production theory (Charnes et al., 1985).

For example, Li and Reeves (1999) presented a Multiple Criteria DEA which can be used to improve discrimination power. Sarkis (1997) and Dvorakova and Klicnarova (2017) also applied DEA as an MCDM tool. Besides, it is argued that assigning arbitrary weights lead to the subjectivity problems in some MCDM approaches, as this limitation can be seen in AHP, TOPSIS, VIKOR, etc. (Hu et al., 2017; Noryani et al., 2018; Shen et al., 2018) because it requires subjective assessments of the decision-maker to prioritize performance attributes (Alinezhad et al., 2011). According to Jahedi and Méndez (2014), although subjectivity can be useful in some situation, for example, mainly when objective data is difficult to obtain, subjectivity suffer from systematic biases, it can be uncorrelated or negatively correlated to the objective data or it can be difficult to interpret. DEA is less subjective, because it does not rely on the decision-makers' preference, so it is more suitable in the present context (Greco et al., 2018).

Among DEA models, Dyckhoff and Souren (2020) highlighted the adequacy and relevance of non-oriented additive DEA models for MCDM, especially, the slack-based measure (SBM), created by Tone (2001). These models take all slacks into account for efficiency measurement. Consequently, they directly identify strongly efficient solutions without the additional calculations necessary in radial models. In addition, as it is often hard to justify an orientation of a DEA model, the absent orientation of SBM represents yet another advantage.

One of the seminal assumptions of DEA is the homogeneity among DMUs. The acceptable limits of heterogeneity remain under discussion. Li et al. (2016) proposed the adoption of a non-homogeneous DEA model for solving non-homogeneity problems. Among DEA pitfalls, Cook et al. (2014) pointed out the misjudgment of efficiency when inputs and outputs simultaneously deal with ratio and raw data.

However, under certain circumstances, the authorsstated that i the dealing with different types of data in the same DEA model is acceptable. The present paper did not assume the restriction of data type as a condition for this index construction.

The discrimination power in DEA is affected by the ratio between the number of DMUs and variables. Banker et al. (1989) stated that the DMUs may be, at least, three times more than variables. Notwithstanding, it is not an imperative rule, just accepted by convenience (Cook et al., 2014). It was assumed as a desirable target here.

Besides outputs and inputs, DEA also may have variables classified as undesirable outputs, e.g., pollutions. An interested reader about this variable type may consult (Hua & Bian, 2007; Liu et al., 2010; Seiford & Zhu, 2002). Among the possible treatments, this paper chose to insert inverted emissions as inputs (for minimization), based on the judgment of specialists.

There are also variables (inputs and outputs) classified as partially non-discretionary (pND). Melo et al. (2018) incorporated the concepts of non-discretion of Saen (2005) to the SBM, assuming no control of the variable (i.e., totally non-discretionary, tND). This paper goes a step further, incorporating a pND (e.g., assuming up to 5% of control of the variable) under VRS in a Double-Frontier-SBM applied to the context of Green Transport Corridors. This incorporation came from the assumption of the possibility of reducing the length of the route in a certain percentage by public investment strategies. The value of 5% was assumed because the percentage of yearly changes in the road (from 2001 to 2017) was up to 4.21% (DNIT, 2020).

3 Methods

The current investigation involved: (*1*) defining DMUs (routes from GTCs) and collecting data, (*2*) analyzing available variables and classifying them into DEA measures, (*3*) applying the pND-SBM model, and, finally, (*4*) applying the tiebreaking tool.

3.1 DMUs Definition and Data Collection

We considered a total of 245 DMUs (routes from 12 GTCs) during the harvest year of 2018/2019. Since Alves Junior et al. (2021) have already studied several routes and Green Transport Corridors in Brazil, we are using the same databases described in their paper, so the results of the present chapter can be compared to the literature. We considered only the currently existing infrastructure and not the planned projects with estimated values. The DMUs originated from producing mid-sized regions in all Brazilian macro-sized regions (IBGE, 2019) and are destined for the 12 main exporting ports. Figure 1 shows the ports and multimodal infrastructure and transport network in the main soybean export corridors (Ministry of Infrastructure, 2021).

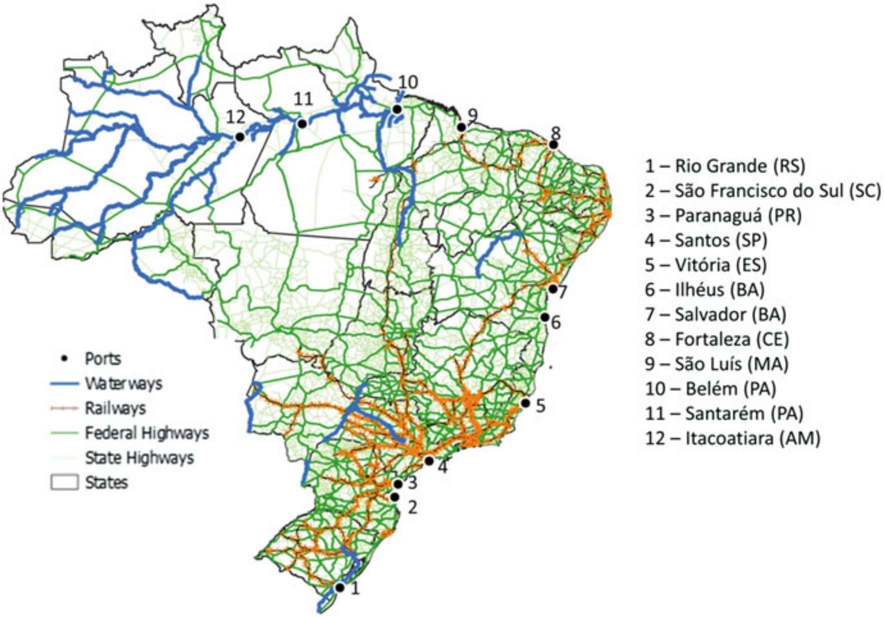


Fig. 1 Ports and multimodal infrastructure and transport network in the main soybean export corridors. Source: Ministry of Infrastructure (2021)

3.2 *Variable Classification and Analysis*

In DEA literature, the most usual variable classification is formulated considering desirable outputs (O) (measures to be maximized), and desirable inputs (I) (measures to be minimized) (Cook et al., 2014). Though, in real-life problems, there are also undesirable outputs (UO) (to be minimized) and undesirable inputs (UI) (to be maximized) (Liu et al., 2015). Finally, some variables can be classified as a partially non-discretionary (pND) measure, i.e., including a quasi-fixed factor that is almost not under control (Saen, 2005). The criterion for choosing a variable was the systematic judgment of specialists about the relevance of a variable for the model (Golany & Roll, 1989), considering the whole scenario, previous papers (Alves Junior et al., 2021; Melo et al., 2018), and the objective of the index. Table 1 shows the classification and the descriptive statistics of the eight used variables.

As can be noted in Table 1, EXPORTS (O) is the only measure that presents a greater standard deviation of the observed values than the mean. This is caused by the difference between productivity capacity in Brazil that lead to using the same ports for exporting (for example, the following ports: Santos, Santarém, Paranaguá, and Rio Grande). Similarly, PAVED (UI) is the measure with the smallest standard deviation in comparison to the mean. The unique pND measure (LENGTH) presents a standard deviation related to the mean of 68.42%, similarly what happens to the

Table 1 Classification and descriptive statistics of selected variables

Variable	Class.	Obs.	Mean	Std. Dev.	Min.	Max.
EXPORT (10^3 tonne)	O	245	435.846	982.697	0.053	9577.108
DEPTH (meters)	UI	245	15.669	7.141	8.000	45.000
STORAGE (10^3 tonnes)	UI	245	1033.745	718.880	3.200	2296.945
PAVED (100 km/km^2)	UI	245	1.162	0.699	0.218	3.517
LENGTH (km)	pND	245	655.666	448.628	53.215	2250.136
COST ($/tonne)	I	245	102.854	39.667	38.311	236.730
CO_2 (kg/tonnes)	UO	245	19.733	10.925	1.964	58.935
ACCIDENTS (per 100 km)	UO	245	33.120	26.865	0.533	68.800

STORAGE (69.54%). But the fact that LENGTH is established as a pND measure may restrict more the impact of the dispersion of the observed values on the final DEA rank results. In other words, the dispersion of STORAGE is expected to have more impact on results. Further in the *Findings*.

Similarly to Alves Junior et al. (2021), because we are using the same databases (available through the same GitHub link informed in their paper, please see the data for DEA application to evaluate Brazilian GTCs at GITHUB (2021)), EXPORTS is the amount of exported soybeans and corn by each port (10^3 tons) and is classified as an Output (O). DEPTH is the highest draft depth of each port (meters), STORAGE is the grain storage capacity in the catchment area (10^3 ton), and PAVED is the paved road density (10^2 km of road/km^2 of the area). DEPTH, STORAGE, and PAVED are classified as Undesirable Inputs (UI). COST is the weighted average freight cost of the flows arriving in each export port (BRL/ton) and is classified as an Input (I). CO_2 is the weighted average CO_2 emission (kg of CO_2/ton) and ACCIDENT is the number of accidents per kilometer estimated in the transportation corridor (accidents/km). CO_2 and ACCIDENT are classified as undesirable outputs (UO).

LENGTH is the length of the route from the origin to the final destination (km). Here we propose to classify LENGTH as a non-discretionary measure (pND). In most cases, we assumed that it is not physically possible to meaningly shorten the transportation distance, by moving the position of the most productive areas, the position of the main infrastructure poles and destinations (ports), planting in similar areas with shorter length of the routes or investing in the construction of a straighter route.

3.3 *Slack-Based Measure Model with Partially Non-discretionary Measures (pND–SMB)*

The equating of the SBM model (Tone, 2001) with incorporated non-discretionary measures (Saen, 2005) follows the objective function in Eq. (1), and it is constrained by Eqs. (2)–(4), and (7) (SBM constraints), and Eqs. (5) and (6) (non-discretionary constraints) (Saen, 2005):

$$\text{Minimize } \tau = t - \left({}^{1}\!\middle/_{m}\right) \sum_{i=1}^{m} {}^{S_i^-}\!\middle/_{x_{i0}} \tag{1}$$

Subject to:

$$t + \left({}^{1}\!\middle/_{s}\right) \sum_{r=1}^{s} {}^{S_r^+}\!\middle/_{y_{r0}} = 1 \tag{2}$$

$$\sum_{k=1}^{z} \Lambda_k x_{ik} + S_i^- - t x_{i0} = 0 \quad i = 1, 2, \ldots, m \tag{3}$$

$$\sum_{k=1}^{z} \Lambda_k y_{rk} - S_r^+ - t y_{r0} = 0 \quad r = 1, 2, \ldots, s \tag{4}$$

$$S_i^- \leq \beta_i x_{i0} \quad i = 1, 2, \ldots, m \tag{5}$$

$$S_r^+ \leq \gamma_r y_{r0} \quad r = 1, 2, \ldots, s \tag{6}$$

$$\Lambda_k \geq 0, S_i^- \geq 0, S_r^+ \geq 0 \text{ and } t > 0 \tag{7}$$

where τ is the efficiency, t is the model linearization variable, S_i^- is the slack of the ith input, S_r^+ is the slack of the rth output, Λ_k is the contribution of the kth DMU to the analyzed DMU, x_{i0} is the ith input of the DMU under analysis, y_{r0} is the rth output of the DMU under analysis, x_{ik} is the ith input of the kth DMU, y_{rk} is the rth output of the kth DMU, m is the number of inputs, s is the number of outputs, z is the number of DMUs, and β_i and γ_r are constants of discretion, respectively, for inputs and outputs (when assuming a value equal to 0, they represent a tND measure and infinite or excluding the constraint represents a totally discretionary input, i.e., a standard SBM model).

As explained in the *Literature Review*, it was assumed that the length of the route could be 5% controllable due to slight changes on the routes. For example, even in a microregion, there are differences in the length of the route depending on how distant from the center of the origin it is or it can be changed by public investments in transportation infrastructure and land use (DNIT, 2020). In other words, we assumed $\beta_i = 0.05$ in Eq. (5).

According to Cook et al. (2014), mixing raw data with ratios is permissible in DEA, but the Variable Return to Scale (VRS) assumption is preferable, mainly if the ratio data is in percentages because considering Constant Return to Scale (CRS) assumption not always maintain the projection between 0% and 100%. As the present application requires the VRS assumption, it was necessary to add a constraint, according to Eq. (8).

$$\sum_{k=1}^{z} \Lambda_k = t \tag{8}$$

The optimum solution $\left(\tau^*, t^*, \Lambda_k^*, S_i^{-*}, S_r^{+*}\right)$ is described by the conditions in Eq. (9):

$$\tau_{\text{optimal}} = \tau^*, \lambda_k^* = {\Lambda_k^*}/{t^*}, s_i^{-*} = {S_i^{-*}}/{t^*}, s_r^{+*} = {S_r^{+*}}/{t^*} \tag{9}$$

In this model, a DMU will be considered efficient when $\tau^* = 1$. Where λ_k^*, S_i^{-*}, and S_r^{+*} are the original optimal variables (before linearizing) solutions. In the model, we treated UO as a negative factor and UI as a positive factor. In other words, UO is mathematically treated as the opposite of an output, i.e., as an input, so, in a post-efficiency analysis, the goal is to decrease the UO. Similarly, UI is mathematically treated as the opposite of an input, i.e., as an output, so, in a post-efficiency analysis, the goal is to increase the UI. This approach was already adopted and discussed by previous papers (Alves Junior et al., 2021; Melo et al., 2018).

3.4 Tiebreaking Method: Double-Frontier Logistic Composite Index (LCI)

The tiebreaking method of the composite index (Leta et al., 2005), also named as Double-Frontier method, was applied, according to Eq. (10). It represents an arithmetic average between standard and inverted efficiencies standardized by the maximum composite index of the analyzed population.

$$\text{LCI} = {\left[E_k^{\text{standard}} + \left(1 - E_k^{\text{inverted}}\right)\right]/2} \Big/ {\max\left\{\left[E_k^{\text{standard}} + \left(1 - E_k^{\text{inverted}}\right)\right]/2\right\}} \quad k = 1, 2, \ldots, z \tag{10}$$

where $E_k^{standard}$ is the standard efficiency resulted from the application of the DEA model for the kth DMU, $E_k^{inverted}$ is the inverted efficiency of the kth DMU, i.e., the resulted efficiency when inputs are inserted in the SBM model as outputs and vice versa.

4 Findings

Table 6 in the *Appendix* presents the resulting LCI when considering LENGTH as pND, for each DMU, as well as the Rank position based on the LCI. For demonstrating and discussing the proposed approach, Table 6 also presents LCI and Rank results, when considering LENGTH as a Controllable measure (input) as well as considering LENGTH as a totally non-discretionary (tND) measure.

Observing the results (Table 6, in Appendix) and the data descriptive statistics (Table 1), it is possible to see that the five routes with the best performance in all configurations were not those with great EXPORTS. The routes with greatest STORAGE were related to Belém (PA), Itacoatiara (AM), Santarém (PA), Santos (SP), and São Luis (MA). As can be seen in Table 6, when LENGTH is treated as a controllable measure, the most efficient routes are those with the shortest length. When LENGTH is treated as tND, there is a relative performance improvement of those routes with other desired measures (for example, those with the greatest STORAGE). Finally, when LENGTH is treated as pND, there is a balance between routes with short LENGTH and other desired measures.

For a faster and easier visualization, Table 2 presents the same results of Table 6, but aggregated by GTC, through the arithmetic average of the results of the DMUs in the same GTC.

It is possible to observe in Table 2 that the main differences in the aggregate results regarding the models with pND, tND, and controllable measures are between the GTC from Santarém (PA) and Itacoatiara (AM). Santarém (PA) is in the second and Itacoatiara (AM) is in the fourth position in controllable and pND ranks, while Santarém (PA) is in the fourth and Itacoatiara (AM) is in the second position in the

Table 2 Aggregated GTC's LCI results considering LENGTH as a Controllable (control) measure (input), a totally non-discretionary (tND) measure, and a partially non-discretionary (pND) measure, followed by their respective rank positions

	Average LCI			Rank		
GTCs	Control	tND	pND	Control	tND	pND
Rio Grande (RS)	0.714	0.728	0.704	1	1	1
Santarém (PA)	0.630	0.485	0.601	2	4	2
Paranaguá (PR)	0.548	0.495	0.515	3	3	3
Itacoatiara (AM)	0.403	0.517	0.41	4	2	4
São Luís (MA)	0.384	0.324	0.332	6	5	5
São Francisco do Sul (SC)	0.397	0.273	0.301	5	6	6
Vitória (ES)	0.261	0.27	0.274	8	7	7
Santos (SP)	0.285	0.188	0.225	7	8	8
Belém (PA)	0.218	0.124	0.156	9	9	9
Ilhéus (BA)	0.121	0.094	0.122	10	10	10
Salvador (BA)	0.022	0.028	0.025	11	11	11
Fortaleza (CE)	0.001	0.001	0.001	12	12	12

tND rank. Despite a measure being not controllable, it happens, because the tND neutralize a measure in terms of source of inefficiency, and the average length from Santarém (PA) is 1360.98 km while the ones from Itacoatiara (AM) is 1880.92 km, so that huge difference in distance is totally ignored in a model with tND, but the model with pND allows it being almost no controllable and be a source of inefficiency yet.

On the other hand, São Luís (MA) was the fifth GTC under controllable assumptions and São Francisco do Sul (SC) was the sixth. Under both tND and pND assumptions, they inverted positions. In other words, São Luís (MA) is sixth and São Francisco do Sul (SC) fifth in both assumptions. The rank change from controllable (discritionary) assumption to a partially Non-Discretionary (pND) and a totally Non-Discretionary (tND) assumptions. This is explained by the fact that São Luís (MA) corridor presents desired observed values for other target measures (e.g., DEPTH of the port, multimodal infrastructure, and very low ACCIDENTS). Once São Luís (MA) and São Francisco do Sul (SC) present similar LENGHT of roads, but São Luís (MA) has more multimodal infrastructure, when it is constrained, its better-observed values in these three aspects improve its relative position. However, the aggregation through arithmetic average presents limitations. One of them is the dependency on the number of routes in a GTC. Although the aggregated values are useful for fast visualization and understanding, it is recommended to investigate routes' (DMUs') results (Table 6) for taking decisions and making policies. Also, other types of aggregations and models can be explored, as the network ones.

Even though with these results, someone could argue about the lack of big differences between the models with controllable and pND measures, but a deep investigation in the percentage of variation to achieve the goals to be in the efficient frontier, computed as a post-efficiency analysis and shown in Table 3.

As it can be seen in Table 3, the model with the LEGTH as a controllable measure shows changes (reductions) up to 18.74% in the length of the routes in a GTC. Considering long-distance haulage, it could be enough to move to another state, so

Table 3 GTCs' % of variation to achieve the goal to be in the frontier

GTCs	LENGTH (control)	LENGTH (tND)	LENGTH (pND)
Paranaguá (PR)	−18.74%	0.00%	−3.21%
São Francisco do Sul (SC)	−10.38%	0.00%	−0.81%
Vitória (ES)	−8.14%	0.00%	−1.73%
Santos (SP)	−7.56%	0.00%	−1.84%
Rio Grande (RS)	−6.19%	0.00%	−2.16%
Salvador (BA)	−3.13%	0.00%	−1.70%
São Luís (MA)	−2.56%	0.00%	−2.21%
Ilhéus (BA)	−2.49%	0.00%	−1.13%
Belém (PA)	−2.35%	0.00%	−1.98%
Itacoatiara (AM)	−1.28%	0.00%	−0.63%
Santarém (PA)	−0.50%	0.00%	−0.50%
Fortaleza (CE)	0.00%	0.00%	0.00%

sometimes it is not a viable outcome to be implementable in practice. While the model with the LENGTH as a pND measure shows similar final average ranks for the GTCs, but with changes in (reductions) up to 18.74% in the length of the routes in a GTC. It is a more viable outcome. And about the model with the LENGTH as a tND measure, it does not even allow changes in it, what sometimes it is not in accordance with the practice (e.g., when the farmer is far away from the center of an origin region).

5 Discussion

As stated in the *Findings*, the five routes with the best performance in all configurations were not those with great EXPORTS. Such as Alves Junior et al. (2021), the model configuration proposed here dealt well in avoiding bias due to productive inequalities. This represents one step further in methodological evolution, one it was not achieved by Melo et al. (2018, 2020), which presented, among the admitted limitations of the results, the greater producers also as part of the most efficient routes. The current paper, as well as Alves Junior et al. (2021), is focused on the destination (origin) instead of the origin (production).

The aggregated results in Table 2 shows the GTC of Rio Grande in the top position independently of the LENGTH treatment. Also, Paranaguá maintained the third position in the three treatments. Besides, the worst performers Belém, Ilhéus, Salvador, and Fortaleza did not change rank positions. These relative positions agree with the previous literature, which demonstrated that, in general, routes and corridors in the Southern of Brazil are more efficient than those from Northern and North-eastern (Alves Junior et al., 2021; Garcia et al., 2019; Branco et al., 2020; Melo et al., 2018, 2019; Rentizelas et al., 2019).

It is possible to observe in Table 2 that the last four corridors (average of routes) are from the North and Northeast regions from Brazil. With this in mind and comparing with Alves Junior et al. (2021), Branco et al. (2020), and the Brazilian Planning and Logistics Company (2021), it is possible to suggest public policies to improve the performance of the corridors from the North and Northeast regions. For example, investing in new railways. Also, investing in new waterways, multimodal routes, and GTCs enable connecting these regions to other productive areas. In this regard, the synergy of integration supports the mitigation of CO_2 emissions. It is possible to highlight the prioritization of four railways: *Ferrograo* (connecting the Center-West productive region to the PA state, and providing alternative access to the Port of Santarém (PA) using the Tapajós waterway). *Ferrovia Norte-Sul* (connecting the North and Northeast regions to the Southeast one). *Ferrovia de Integração Oeste-Leste* [connecting the west of BA state to the Port of Ilhéus (BA)]. *Ferrovia Nova Transnordestina* (connecting the North to the Northeast region).

Another point to discuss is related to the aggregated values that are useful for fast visualization and understanding, it is recommended to investigate routes' (DMUs') results (Table 6). These detailed results permit the decision-makers to understand where in the GTC (and how) is required to guide more efforts to improve local and aggregated performance.

For example, the GTC of Itacoatiara has 12 routes (DMUs 18–29) (Table 4). One of them is the fourth best ranked, considering the 245 DMUs under analysis (DMU24). Though other DMUs are among the worst-ranked (18, 20, 23, 26, and 27). Efforts directed to improve the efficiency of the worst-ranked routes will result in a GTC better performance as well as promote regional development.

In parallel, the GTC of Santarém has eight routes (DMUs 88–95) (Table 5). Following the proposed methodology, efforts should be guided to worst-ranked DMUs. The aggregation through the arithmetic average may have benefited the GTCs with fewer routes. In practical terms, it may be not possible to build more routes due to natural barriers such as mountains and forests. In this case, the DMUs' results point which existing route should be the focus of efforts. For example, in the case of Santarem (in Amazon Forest), they are DMUs 88, 93, and 92. Though, in cases where it is possible to build more routes, planned routes can also be incorporated into the analysis and their expected performance can be investigated.

Table 4 DMUs' results of the GTC of Itacoatiara (MA)

		LCI			Rank		
DMU	Destination	Control	tND	pND	Control	tND	pND
24	Itacoatiara (AM)	0.937	0.949	0.947	6	7	4
25	Itacoatiara (AM)	0.654	0.868	0.662	37	13	32
21	Itacoatiara (AM)	0.610	0.618	0.617	44	39	39
19	Itacoatiara (AM)	0.515	0.522	0.521	69	60	63
22	Itacoatiara (AM)	0.515	0.522	0.521	70	61	64
29	Itacoatiara (AM)	0.423	0.522	0.457	100	64	80
28	Itacoatiara (AM)	0.451	0.522	0.456	90	63	81
27	Itacoatiara (AM)	0.323	0.348	0.326	142	104	113
20	Itacoatiara (AM)	0.290	0.696	0.293	153	30	122
18	Itacoatiara (AM)	0.110	0.111	0.111	209	164	197
26	Itacoatiara (AM)	0.007	0.522	0.007	226	68	226
23	Itacoatiara (AM)	0.007	0.011	0.007	225	206	227

Table 5 DMUs' results of the GTC of Santarém (PA)

		LCI			Rank		
DMU	Destination	Control	tND	pND	Control	tND	pND
89	Santarém (PA)	0.888	0.522	0.886	11	57	8
91	Santarém (PA)	0.851	0.522	0.837	15	58	13
90	Santarém (PA)	0.817	0.776	0.808	19	23	17
95	Santarém (PA)	0.741	0.726	0.738	24	25	22
94	Santarém (PA)	0.731	0.716	0.725	27	27	23
92	Santarém (PA)	0.522	0.228	0.348	67	130	106
93	Santarém (PA)	0.285	0.282	0.286	157	120	127
88	Santarém (PA)	0.208	0.107	0.178	181	168	150

6 Conclusions

We presented a methodology for building a long-distance cargo-haulage performance index, named Logistic Composite Index (LCI). In this context, the LCI brings the novelty of incorporating partially Non-Discretionary (pND) measures in Double-Frontier Data Envelopment Analysis (DEA), Slack-Based Measure (SBM) under Variable Return to Scale (VRS) assumption to study Green Transport Corridors and its routes.

For deepening the discussion about the impact of the partial non-discretionarily treatment, we also ran the model considering two other possibilities: (*1*) route transport distances as controllable measures (inputs), i.e., assuming decision-makers and policy-makers have the possibility of shortening the physical transport distance between producers and exporting ports, without any external constraints (this was the most adopted assumption in previous studies); (*2*) route transport distances as totally non-discretionary (tND), i.e., assuming decision-makers and policy-makers have no possibility of shortening the physical transport distance between producers and exporting ports. They are completely limited by external constraints.

The three results were aligned to the previous literature, pointing routes and corridors in Southern Brazil more efficient than those in the Northern and Northeastern. But treating the length of the route as a partially Non-Discretionary (pND) measure proved to be more accurate, mainly when calculating the percentages of variation to achieve the goals to be in the frontier. Once the top-ranked DMUs under the pND assumption also presented better-ranked positions under controllable assumptions. They these DMUs presented worse-ranked positions under the tND-distance assumption. Also, both assumptions (tND and controllable) are not achievable in real life for the studied context.

For creating a Green Transport Corridor's (GTC) index and avoiding the lower number of GTCs, we considered the routes as DMUs, computed the LCIs, and aggregated the routes' LCIs of each GTC, through an arithmetic average. Although the GTC values are useful for fast visualization and understanding, DMU's results should be considered when planning efforts for improving GTC's efficiency as well as promoting regional development.

For future investigations, in terms of application, we recommend studies focused on the logistic operators, such as related to the availability of return freight. We also recommend the use of big data and real-time logistic data, when they are available. This application can improve the model developed through the incorporation of other techniques, such as hierarchical network models and deep learning. For example, we recommend the development of a model where the discretionary level of measure could be customized for each DMU. This way, the same model could assume (for the same measure) a higher discretionary level for those DMUs where the measure is less externally constrained. Once the data is available, the discretionary level of each DMU could be calculated through deep learning and other techniques. Similarly, it is also possible to improve the aggregation method from routes to corridors, such as proposing (dynamic) network-DEA and hierarchical-DEA models. Finally, we suggest for future studies to investigate the impact of the dispersion of the data (standard-deviation) on the efficiency results, and other aggregation methods or Network models applied to evaluate the GTCs.

Appendix

Table 6 DMU's (Route's) LCI results considering DISTANCE as a Controllable measure (input), a totally non-discretionary (tND) measure, and a partially non-discretionary (pND) measure, followed by their respective rank positions

DMU	Destination	LCI			Rank		
		Control	tND	pND	Control	tND	pND
79	Rio Grande (RS)	0.99	1	1	2	1	1
82	Rio Grande (RS)	0.99	1	1	3	2	2
83	Rio Grande (RS)	0.983	0.991	0.991	4	5	3
24	Itacoatiara (AM)	0.937	0.949	0.947	6	7	4
237	São Luís (MA)	0.933	0.945	0.943	7	8	5
231	São Luís (MA)	0.908	0.92	0.918	10	9	6
64	Rio Grande (RS)	1	0.899	0.906	1	12	7
89	Santarém (PA)	0.888	0.522	0.886	11	57	8
70	Rio Grande (RS)	0.923	0.841	0.843	9	15	9
235	São Luís (MA)	0.832	0.843	0.842	16	14	10
75	Rio Grande (RS)	0.673	0.906	0.838	35	10	11
175	Santos (SP)	0.829	0.84	0.838	17	16	12
91	Santarém (PA)	0.851	0.522	0.837	15	58	13
176	Santos (SP)	0.823	0.834	0.832	18	17	14
192	São Francisco do Sul (SC)	0.873	0.777	0.813	13	22	15
55	Paranaguá (PR)	0.803	0.812	0.811	21	19	16
90	Santarém (PA)	0.817	0.776	0.808	19	23	17
61	Rio Grande (RS)	0.877	0.777	0.792	12	21	18
183	Santos (SP)	0.778	0.788	0.786	22	20	19
73	Rio Grande (RS)	0.587	0.905	0.779	47	11	20
76	Rio Grande (RS)	0.809	0.745	0.75	20	24	21
95	Santarém (PA)	0.741	0.726	0.738	24	25	22
94	Santarém (PA)	0.731	0.716	0.725	27	27	23
54	Paranaguá (PR)	0.704	0.997	0.723	31	4	24
238	São Luís (MA)	0.694	0.722	0.701	33	26	25
62	Rio Grande (RS)	0.749	0.694	0.697	23	31	26
189	Santos (SP)	0.695	0.702	0.694	32	29	27
78	Rio Grande (RS)	0.735	0.691	0.694	25	32	28
67	Rio Grande (RS)	0.549	0.826	0.682	56	18	29
74	Rio Grande (RS)	0.718	0.645	0.667	29	35	30

(continued)

Table 6 (continued)

DMU	Destination	LCI			Rank		
		Control	tND	pND	Control	tND	pND
66	Rio Grande (RS)	0.725	0.28	0.667	28	121	31
25	Itacoatiara (AM)	0.654	0.868	0.662	37	13	32
205	São Francisco do Sul (SC)	0.872	0.652	0.662	14	34	33
49	Paranaguá (PR)	0.732	0.643	0.644	26	36	34
68	Rio Grande (RS)	0.649	0.685	0.634	39	33	35
57	Paranaguá (PR)	0.628	0.642	0.634	43	37	36
77	Rio Grande (RS)	0.554	0.711	0.632	54	28	37
33	Paranaguá (PR)	0.704	0.303	0.624	30	113	38
21	Itacoatiara (AM)	0.61	0.618	0.617	44	39	39
63	Rio Grande (RS)	0.685	0.601	0.61	34	42	40
169	Santos (SP)	0.567	0.618	0.606	50	40	41
52	Paranaguá (PR)	0.591	0.6	0.595	46	43	42
180	Santos (SP)	0.566	0.639	0.592	52	38	43
45	Paranaguá (PR)	0.654	0.582	0.585	38	44	44
214	São Luís (MA)	0.959	0.522	0.584	5	55	45
51	Paranaguá (PR)	0.552	0.61	0.582	55	41	46
36	Paranaguá (PR)	0.628	0.569	0.572	42	46	47
172	Santos (SP)	0.476	0.576	0.56	81	45	48
50	Paranaguá (PR)	0.54	0.566	0.558	58	47	49
39	Paranaguá (PR)	0.647	0.541	0.548	40	54	50
60	Paranaguá (PR)	0.543	0.56	0.546	57	48	51
41	Paranaguá (PR)	0.538	0.554	0.545	60	50	52
186	Santos (SP)	0.476	0.56	0.543	82	49	53
200	São Francisco do Sul (SC)	0.927	0.522	0.541	8	56	54
32	Paranaguá (PR)	0.531	0.549	0.54	62	52	55
174	Santos (SP)	0.489	0.552	0.534	75	51	56
69	Rio Grande (RS)	0.61	0.519	0.533	45	69	57
80	Rio Grande (RS)	0.525	0.99	0.531	65	6	58
46	Paranaguá (PR)	0.661	0.513	0.531	36	70	59
81	Rio Grande (RS)	0.522	1	0.528	66	3	60
171	Santos (SP)	0.479	0.548	0.526	78	53	61
9	Belém (PA)	0.515	0.522	0.521	68	59	62
19	Itacoatiara (AM)	0.515	0.522	0.521	69	60	63
22	Itacoatiara (AM)	0.515	0.522	0.521	70	61	64
166	Santos (SP)	0.515	0.522	0.521	71	62	65
71	Rio Grande (RS)	0.563	0.513	0.516	53	71	66
43	Paranaguá (PR)	0.566	0.509	0.514	51	72	67
42	Paranaguá (PR)	0.539	0.471	0.502	59	77	68
56	Paranaguá (PR)	0.443	0.502	0.487	93	73	69
53	Paranaguá (PR)	0.483	0.492	0.487	77	74	70
31	Paranaguá (PR)	0.526	0.46	0.478	64	83	71

(continued)

Table 6 (continued)

DMU	Destination	LCI			Rank		
		Control	tND	pND	Control	tND	pND
170	Santos (SP)	0.475	0.491	0.477	83	75	72
44	Paranaguá (PR)	0.532	0.469	0.476	61	78	73
206	São Francisco do Sul (SC)	0.488	0.46	0.468	76	84	74
48	Paranaguá (PR)	0.455	0.156	0.468	87	144	75
72	Rio Grande (RS)	0.528	0.079	0.466	63	177	76
167	Santos (SP)	0.311	0.522	0.465	146	65	77
207	São Francisco do Sul (SC)	0.475	0.464	0.462	84	81	78
242	Vitória (ES)	0.455	0.482	0.46	88	76	79
29	Itacoatiara (AM)	0.423	0.522	0.457	100	64	80
28	Itacoatiara (AM)	0.451	0.522	0.456	90	63	81
34	Paranaguá (PR)	0.575	0.437	0.455	49	87	82
59	Paranaguá (PR)	0.444	0.462	0.453	92	82	83
65	Rio Grande (RS)	0.465	0.442	0.446	86	86	84
179	Santos (SP)	0.41	0.469	0.443	106	79	85
40	Paranaguá (PR)	0.489	0.437	0.443	74	88	86
185	Santos (SP)	0.43	0.447	0.435	98	85	87
208	São Francisco do Sul (SC)	0.637	0.423	0.433	41	89	88
35	Paranaguá (PR)	0.489	0.408	0.432	73	92	89
245	Vitória (ES)	0.347	0.465	0.428	130	80	90
215	São Luís (MA)	0.493	0.413	0.421	72	91	91
168	Santos (SP)	0.407	0.423	0.411	108	90	92
30	Paranaguá (PR)	0.451	0.402	0.408	91	95	93
58	Paranaguá (PR)	0.389	0.404	0.396	111	94	94
182	Santos (SP)	0.388	0.401	0.393	112	96	95
188	Santos (SP)	0.354	0.408	0.388	126	93	96
196	São Francisco do Sul (SC)	0.469	0.377	0.384	85	98	97
38	Paranaguá (PR)	0.438	0.372	0.379	96	99	98
216	São Luís (MA)	0.412	0.371	0.377	105	100	99
195	São Francisco do Sul (SC)	0.453	0.362	0.374	89	101	100
165	Santos (SP)	0.362	0.382	0.366	121	97	101
241	Vitória (ES)	0.355	0.354	0.358	125	103	102
221	São Luís (MA)	0.376	0.34	0.357	116	105	103
224	São Luís (MA)	0.363	0.356	0.354	120	102	104
213	São Luís (MA)	0.358	0.058	0.352	123	187	105
92	Santarém (PA)	0.522	0.228	0.348	67	130	106
47	Paranaguá (PR)	0.34	0.335	0.342	133	107	107
219	São Luís (MA)	0.375	0.316	0.342	118	111	108
193	São Francisco do Sul (SC)	0.394	0.329	0.341	109	108	109
164	Santos (SP)	0.334	0.338	0.338	136	106	110
209	São Francisco do Sul (SC)	0.435	0.194	0.335	97	135	111
211	São Luís (MA)	0.351	0.318	0.327	129	110	112

(continued)

Table 6 (continued)

DMU	Destination	LCI			Rank		
		Control	tND	pND	Control	tND	pND
27	Itacoatiara (AM)	0.323	0.348	0.326	142	104	113
227	São Luís (MA)	0.441	0.239	0.326	95	127	114
222	São Luís (MA)	0.333	0.182	0.326	138	137	115
244	Vitória (ES)	0.319	0.325	0.323	143	109	116
218	São Luís (MA)	0.324	0.297	0.322	141	116	117
2	Belém (PA)	0.584	0.19	0.32	48	136	118
223	São Luís (MA)	0.479	0.043	0.319	79	190	119
191	São Francisco do Sul (SC)	0.357	0.297	0.317	124	115	120
1	Belém (PA)	0.338	0.301	0.311	134	114	121
20	Itacoatiara (AM)	0.29	0.696	0.293	153	30	122
126	Santos (SP)	0.407	0.288	0.293	107	118	123
14	Ilhéus (BA)	0.284	0.284	0.292	159	119	124
230	São Luís (MA)	0.287	0.313	0.29	156	112	125
229	São Luís (MA)	0.376	0.254	0.288	117	124	126
93	Santarém (PA)	0.285	0.282	0.286	157	120	127
181	Santos (SP)	0.233	0.165	0.284	172	142	128
112	Santos (SP)	0.414	0.247	0.273	101	126	129
217	São Luís (MA)	0.414	0.263	0.27	103	123	130
226	São Luís (MA)	0.303	0.073	0.27	148	180	131
190	Santos (SP)	0.265	0.29	0.268	163	117	132
243	Vitória (ES)	0.263	0.266	0.266	164	122	133
156	Santos (SP)	0.441	0.122	0.263	94	163	134
111	Santos (SP)	0.298	0.251	0.259	149	125	135
151	Santos (SP)	0.333	0.233	0.245	137	129	136
139	Santos (SP)	0.308	0.227	0.24	147	131	137
173	Santos (SP)	0.235	0.239	0.237	171	128	138
125	Santos (SP)	0.425	0.094	0.236	99	173	139
124	Santos (SP)	0.352	0.213	0.234	128	132	140
37	Paranaguá (PR)	0.38	0	0.219	114	236	141
202	São Francisco do Sul (SC)	0.338	0.076	0.218	135	179	142
159	Santos (SP)	0.414	0.195	0.208	102	134	143
5	Belém (PA)	0.413	0.063	0.206	104	183	144
160	Santos (SP)	0.296	0.182	0.205	151	138	145
141	Santos (SP)	0.245	0.06	0.202	168	185	146
7	Belém (PA)	0.189	0.207	0.194	187	133	147
148	Santos (SP)	0.222	0.179	0.191	177	139	148
96	Santos (SP)	0.261	0.132	0.183	166	159	149
88	Santarém (PA)	0.208	0.107	0.178	181	168	150
199	São Francisco do Sul (SC)	0.387	0.166	0.177	113	141	151
105	Santos (SP)	0.207	0.162	0.177	182	143	152
127	Santos (SP)	0.241	0.033	0.175	170	195	153

(continued)

Table 6 (continued)

DMU	Destination	LCI			Rank		
		Control	tND	pND	Control	tND	pND
148	Santos (SP)	0.222	0.179	0.191	177	139	148
96	Santos (SP)	0.261	0.132	0.183	166	159	149
88	Santarém (PA)	0.208	0.107	0.178	181	168	150
199	São Francisco do Sul (SC)	0.387	0.166	0.177	113	141	151
105	Santos (SP)	0.207	0.162	0.177	182	143	152
127	Santos (SP)	0.241	0.033	0.175	170	195	153
194	São Francisco do Sul (SC)	0.173	0.169	0.172	190	140	154
100	Santos (SP)	0.297	0.155	0.168	150	145	155
113	Santos (SP)	0.318	0.154	0.166	144	146	156
97	Santos (SP)	0.209	0.141	0.166	180	152	157
128	Santos (SP)	0.262	0.15	0.165	165	148	158
136	Santos (SP)	0.182	0.024	0.165	189	199	159
143	Santos (SP)	0.324	0.151	0.163	140	147	160
129	Santos (SP)	0.281	0.039	0.161	161	191	161
133	Santos (SP)	0.29	0.019	0.161	154	201	162
121	Santos (SP)	0.156	0.021	0.158	196	200	163
101	Santos (SP)	0.342	0.144	0.157	132	150	164
162	Santos (SP)	0.164	0.142	0.156	194	151	165
134	Santos (SP)	0.202	0.067	0.156	184	181	166
16	Ilhéus (BA)	0.16	0.025	0.154	195	198	167
153	Santos (SP)	0.259	0.018	0.154	167	202	168
115	Santos (SP)	0.228	0.017	0.154	174	203	169
103	Santos (SP)	0.214	0.013	0.154	179	205	170
131	Santos (SP)	0.28	0.14	0.152	162	153	171
146	Santos (SP)	0.379	0.137	0.151	115	155	172
17	Ilhéus (BA)	0.148	0.148	0.15	198	149	173
104	Santos (SP)	0.242	0.033	0.15	169	194	174
155	Santos (SP)	0.291	0.01	0.15	152	208	175
154	Santos (SP)	0.223	0.01	0.15	176	209	176
144	Santos (SP)	0.196	0.136	0.149	185	157	177
149	Santos (SP)	0.153	0.127	0.148	197	162	178
99	Santos (SP)	0.168	0.044	0.147	192	189	179
117	Santos (SP)	0.289	0.133	0.146	155	158	180
157	Santos (SP)	0.192	0.132	0.146	186	160	181
107	Santos (SP)	0.285	0.131	0.143	158	161	182
137	Santos (SP)	0.331	0.102	0.137	139	170	183
119	Santos (SP)	0.145	0.097	0.137	199	171	184
150	Santos (SP)	0.216	0.003	0.134	178	218	185
120	Santos (SP)	0.132	0.003	0.134	204	219	186
163	Santos (SP)	0.131	0.137	0.133	205	156	187
102	Santos (SP)	0.132	0.027	0.133	203	196	188

(continued)

Table 6 (continued)

DMU	Destination	LCI			Rank		
		Control	tND	pND	Control	tND	pND
110	Santos (SP)	0.13	0.107	0.132	206	169	189
132	Santos (SP)	0.189	0.088	0.129	188	174	190
122	Santos (SP)	0.14	0.085	0.129	202	176	191
138	Santos (SP)	0.167	0.002	0.129	193	222	192
232	São Luís (MA)	0.126	0.139	0.128	207	154	193
118	Santos (SP)	0.358	0.107	0.125	122	167	194
198	São Francisco do Sul (SC)	0.118	0.006	0.117	208	213	195
161	Santos (SP)	0.144	0.001	0.117	201	229	196
18	Itacoatiara (AM)	0.11	0.111	0.111	209	164	197
142	Santos (SP)	0.318	0.003	0.109	145	217	198
204	São Francisco do Sul (SC)	0.172	0.095	0.108	191	172	199
158	Santos (SP)	0.353	0.079	0.108	127	178	200
225	São Luís (MA)	0.478	0.066	0.103	80	182	201
108	Santos (SP)	0.367	0.006	0.099	119	212	202
87	Salvador (BA)	0.085	0.109	0.097	212	165	203
109	Santos (SP)	0.207	0.002	0.094	183	221	204
145	Santos (SP)	0.091	0.001	0.092	210	230	205
11	Belém (PA)	0.229	0.004	0.087	173	216	206
197	São Francisco do Sul (SC)	0.085	0.086	0.086	213	175	207
239	Vitória (ES)	0.085	0.001	0.086	214	231	208
140	Santos (SP)	0.144	0.001	0.084	200	228	209
184	Santos (SP)	0.065	0.108	0.066	215	166	210
236	São Luís (MA)	0.06	0.522	0.063	216	66	211
187	Santos (SP)	0.051	0.063	0.052	217	184	212
106	Santos (SP)	0.39	0.01	0.048	110	207	213
234	São Luís (MA)	0.041	0.522	0.044	219	67	214
177	Santos (SP)	0.044	0.06	0.044	218	186	215
4	Belém (PA)	0.089	0.026	0.042	211	197	216
178	Santos (SP)	0.03	0.05	0.031	220	188	217
123	Santos (SP)	0.344	0.002	0.028	131	220	218
233	São Luís (MA)	0.025	0.037	0.026	221	192	219
10	Belém (PA)	0.021	0.037	0.024	222	193	220
147	Santos (SP)	0.227	0.001	0.022	175	227	221
201	São Francisco do Sul (SC)	0.284	0	0.015	160	237	222
13	Ilhéus (BA)	0.012	0.014	0.013	223	204	223
3	Belém (PA)	0.01	0.01	0.01	224	210	224
135	Santos (SP)	0.007	0.008	0.008	227	211	225
26	Itacoatiara (AM)	0.007	0.522	0.007	226	68	226
23	Itacoatiara (AM)	0.007	0.011	0.007	225	206	227
8	Belém (PA)	0.006	0.006	0.006	228	214	228
98	Santos (SP)	0.006	0.006	0.006	229	215	229

(continued)

Table 6 (continued)

DMU	Destination	LCI Control	LCI tND	LCI pND	Rank Control	Rank tND	Rank pND
86	Salvador (BA)	0.002	0.002	0.002	230	223	230
210	São Francisco do Sul (SC)	0.002	0.002	0.002	231	224	231
212	São Luís (MA)	0.002	0.002	0.002	232	225	232
228	São Luís (MA)	0.002	0.002	0.002	233	226	233
6	Belém (PA)	0.001	0.001	0.001	234	232	234
12	Fortaleza (CE)	0.001	0.001	0.001	235	233	235
114	Santos (SP)	0.001	0.001	0.001	236	234	236
220	São Luís (MA)	0.001	0.001	0.001	237	235	237
15	Ilhéus (BA)	0	0	0	238	238	238
84	Salvador (BA)	0	0	0	239	239	239
85	Salvador (BA)	0	0	0	240	240	240
116	Santos (SP)	0	0	0	241	241	241
130	Santos (SP)	0	0	0	242	242	242
152	Santos (SP)	0	0	0	243	243	243
203	São Francisco do Sul (SC)	0	0	0	244	244	244
240	Vitória (ES)	0	0	0	245	245	245

References

Alinezhad, A., Makui, A., Mavi, R. K., & Zohrehbandian, M. (2011). An MCDM-DEA approach for technology selection. *Journal of Industrial Engineering International, 7*(12), 32–38.

Alves Junior, P. N., Melo, I. C., Branco, J. E. H., Bartholomeu, D. B., & Caixeta-Filho, J. V. (2021). Which green transport corridors (GTC) are efficient? A dual-step approach using network equilibrium model (NEM) and data envelopment analysis (DEA). *Journal of Marine Science and Engineering, 9*(3), 247. https://doi.org/10.3390/jmse–47

Banker, R., Charnes, A., Cooper, W., & Swarts, J. (1989). An introduction to data envelopment analysis with some of its models and their uses. *Research in Governmental and Non-Profit Accounting, 5*, 125–163.

Belton, V., & Vickers, S. P. (1993). Demystifying DEA-A visual interactive approach based on multiple criteria analysis. *The Journal of the Operational Research Society, 44*(9), 883. https://doi.org/10.2307/–81

Branco, J. E. H., Bartholomeu, D. B., Alves Junior, P. N., & Caixeta Filho, J. V. (2020). Evaluation of the economic and environmental impacts from the addition of new railways to the Brazilian's transportation network: An application of a network equilibrium model. *Transport Policy*. https://doi.org/10.1016/j.tranpol.2020.03.011

Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *Company European Journal of Operational Research, 2*.

Charnes, A., Cooper, W. W., Golany, B., Seiford, L., & Stutz, J. (1985). Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions. *Journal of Econometrics, 30*(1–2), 91–107. https://doi.org/10.1016/0304-4076(85)–-2

Cook, W. D., Tone, K., & Zhu, J. (2014). Data envelopment analysis: Prior to choosing a model. *Omega (United Kingdom), 44*, 1–4. https://doi.org/10.1016/j.omega.2013.09.004

DNIT. (2020). *PNV and SNV*. http://www.dnit.gov.br/sistema-nacional-de-viacao/pnv-e-snv

Doyle, J., & Green, R. (1993). Data envelopment analysis and multiple criteria decision making. *Omega, 21*(6), 713–715. https://doi.org/10.1016/0305-0483(93)–-B

Dvorakova, M., & Klicnarova, J. (2017). On the differences between DEA and selected MCDM methods. *The International Scientific Conference INPROFORUM, 2017*, 338–343.

Dyckhoff, H., & Souren, R. (2020). Data envelopment methodology of performance evaluation. In *Performance evaluation – foundations and challenges* (pp. 47–82). Springer International. https://doi.org/10.1007/978-3-030-—-7_3

European Commission. (2007). *Freight transport logistics action plan*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:tr0053

Garcia, B. T. d. G., Lopes, D. M. M., Leal Junior, I. C., Amorim, J. C. C., da Silva, M. A. V., & Guimarães, V. d. A. (2019). Analysis of the performance of transporting soybeans from Mato Grosso for export: A case study of the Tapajós-Teles Pires waterway. *Sustainability, 11*(21), 6124. https://doi.org/10.3390/su–124

GITHUB. (2021). *NEM-results-and-data-for-DEA—JMSE−5*. https://github.com/pjnocera/NEM-results-and-data-for-DEA%2D%2D-JMSE–5

Golany, B., & Roll, Y. (1989). An application procedure for DEA. *Omega, 17*(3), 237–250. https://doi.org/10.1016/0305-0483(89)–-7

Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2018). On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Social Indicators Research, 1–34*. https://doi.org/10.1007/s–-017-1832-9

Hu, C.-K., Liu, F.-B., & Hu, C.-F. (2017). A hybrid fuzzy DEA/AHP methodology for ranking units in a fuzzy environment. *Symmetry, 9*(11), 273. https://doi.org/10.3390/sym–73

Hua, Z., & Bian, Y. (2007). DEA with undesirable factors. In *Modeling data irregularities and structural complexities in data envelopment analysis* (pp. 103–121). Springer US. https://doi.org/10.1007/978-0-387-—-7_6

IBGE – Brazilian Institute of Geography and Statistics. (2019). *Area of districts*. https://www.ibge.gov.br/geociencias-novoportal/organizacao-do-territorio/estrutura-territorial/–-areas-dos-municipios.html?=&t=o-que-e

Jahedi, S., & Méndez, F. (2014). On the advantages and disadvantages of subjective measures. *Journal of Economic Behavior and Organization, 98*, 97–114. https://doi.org/10.1016/j.jebo.2013.12.016

Joro, T., Korhonen, P., & Wallenius, J. (1998). Structural comparison of data envelopment analysis and multiple objective linear programming. *Management Science, 44*(7), 962–970. https://doi.org/10.1287/mnsc.44.7.962

Leta, F. R., Mello, J. C. C. B. S. d., Gomes, E. G., & Meza, L. A. (2005). Métodos de melhora de ordenação em DEA aplicados à avaliação estática de tornos mecânicos. *Investigação Operacional, 25*(2), 229–242.

Li, W., Liang, L., Cook, W. D., & Zhu, J. (2016). DEA models for non-homogeneous DMUs with different input configurations. *European Journal of Operational Research, 254*, 946–956. https://doi.org/10.1016/j.ejor.2016.04.063

Li, X.-B., & Reeves, G. R. (1999). A multiple criteria approach to data envelopment analysis. *European Journal of Operational Research, 115*(3), 507–517. https://doi.org/10.1016/S0377-2217(98)–-1

Liu, W. B., Meng, W., Li, X. X., & Zhang, D. Q. (2010). DEA models with undesirable inputs and outputs. *Annals of Operations Research, 173*(1), 177–194. https://doi.org/10.1007/s–-009-0587-3

Liu, W., Zhou, Z., Ma, C., Liu, D., & Shen, W. (2015). Two-stage DEA models with undesirable input-intermediate-outputs. *Omega, 56*, 74–87. https://doi.org/10.1016/j.omega.2015.03.009

Melo, I. C., Alves Junior, P. N., Pera, T. G., Caixeta-Filho, J. V., & Rebelatto, D. A. d. N. (2019). Framework for logistics performance index construction using DEA: An application for soybean haulage in Brazil. *Proceedings of World Conference on Transportation Research*.

Melo, I. C., Alves Junior, P. N., Perico, A. E., Guzman, M. G. S., & Rebelatto, D. A. d. N. (2018). Benchmarking freight transportation corridors and routes with data envelopment analysis (DEA). *Benchmarking: An International Journal, 25*(2), 713–742. https://doi.org/10.1108/BIJ-11-2016-0175

Melo, I. C., Péra, T. G., Alves Júnior, P. N., Nascimento Rebelatto, D. A. d., & Caixeta-Filho, J. V. (2020). Framework for logistics performance index construction using DEA: An application for soybean haulage in Brazil. *Transportation Research Procedia, 48*, 3090–3106. https://doi.org/10.1016/j.trpro.2020.08.179

Ministry of Infrastructure. (2021). *Maps and bases of transport modes [Mapas e Bases dos Modos de Transportes]*. https://www.gov.br/infraestrutura/pt-br/assuntos/dados-de-transportes/bit/bitmodosmapas#mapport

Noryani, M., Sapuan, S. M., & Mastura, M. T. (2018). Multi-criteria decision-making tools for material selection of natural fibre composites: A review. *Journal of Mechanical Engineering and Sciences, 12*(1), 3330–3353. https://doi.org/10.15282/JMES.12.1.2018.5.0299

Panagakos, G. (2015). Green corridors basics. In *Green transportation logistics: The quest for win-win solutions* (pp. 81–121). Springer International. https://doi.org/10.1007/978-3-319-—3_3

Planning and Logistics Company [Empresa de Planejamento e Logística]. (2021). *Downloads*. https://www.ontl.epl.gov.br/downloads

Rentizelas, A., Melo, I. C., Alves Junior, P. N., Campoli, J. S., Rebelatto, D. A., & do N. (2019). Multi-criteria efficiency assessment of international biomass supply chain pathways using Data Envelopment Analysis. *Journal of Cleaner Production, 237*, −0. https://doi.org/10.1016/J.JCLEPRO.2019.–0

Saen, R. F. (2005). Developing a nondiscretionary model of slacks-based measure in data envelopment analysis. *Applied Mathematics and Computation, 169*(2), 1440–1447. https://doi.org/10.1016/J.AMC.2004.10.053

Sarkis, J. (1997). Evaluating flexible manufacturing systems alternatives using data envelopment analysis. *The Engineering Economist, 43*(1), 25–47. https://doi.org/10.1080/——88

Seiford, L. M., & Zhu, J. (2002). Modeling undesirable factors in efficiency evaluation. *European Journal of Operational Research, 142*(1), 16–20. https://doi.org/10.1016/S0377-2217(01)–-4

Shen, H., Hu, L., & Lai, K. K. (2018). A mathematical programming model to determine objective weights for the interval extension of TOPSIS. *Mathematical Problems in Engineering, 2018*. https://doi.org/10.1155/2018/–01

Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research, 130*(3), 498–509. https://doi.org/10.1016/S0377-2217(99)–-5

World Bank. (2018). *Aggregated LPI 2012-2018*. https://lpi.worldbank.org/

Competitiveness of Nations and Inequality-Adjusted Human Development: Evaluating the Efficiency of Nations Using DEA and Random Forest Classification

Christopher Colin Campbell

Abstract Although popular indices like the Human Development Index (HDI) and Global Competitiveness Index (GCI) measure human development and competitiveness separately, no index directly considers their linkage, namely, the relative ability of countries to leverage their economic competitiveness to improve the human development of their citizens. This paper aims to combine data envelopment analysis and random forest classification to explore the relative performance of countries in terms of competitiveness and human development. In the first stage of the methodology, we evaluate 124 countries using data envelopment analysis (DEA), taking indicators from the GCI and IHDI (inequality-adjusted human development index) as input and output variables, respectively. In the methodology's second stage, we use random forest classification to identify the relative importance of input and output variables on the DEA results—specifically, whether countries were classified as efficient or inefficient. Our findings indicate that only 20 of 124 countries are efficient at using their competitiveness to generate human development, and that variables related to a country's innovation ecosystem are most important. The results suggest most countries fail to take full advantage of their economic resources amidst a period of rapid technological and social change; it also highlights huge disparities between different groups of countries (e.g. regions).

Keywords Competitiveness of nations · Human development · Random forests · Data envelopment analysis

The online version contains supplementary material available at [https://doi.org/10.1007/978-3-030-91851-4_5].

C. C. Campbell (✉)
HelloFresh, Birmingham, UK

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_5

1 Introduction

In this paper, we examine the relationship between human development and the competitiveness of nations using well-known indices—the IHDI and GCI—as their proxies.

'Human Development' is the idea that development should better peoples' lives by expanding the number of freedoms and opportunities at their disposal (Sen, 2001; Fukuda-Parr, 2003). The term is strongly associated with the work of Amartya Sen and, in particular, his capability approach. Still, topical literature suggests there are many sources of inspiration, including Aristotle (Seth & Santos, 2018), the Catholic Church (Seth & Santos, 2018), Jeremy Bentham, John Stuart Mill and John Rawls (Stanton, 2007). Even the United Nations Development Programme (UNDP), the institution which publishes the Human Development Index each year, describes the concept as having '[grown] out of global discussions . . . during the second half of the 20th Century' (HDRO, n.d.). These sweeping accounts imply that, while Sen's definition is central, human development has a fluid meaning that is subject to change and individual interpretation.

'Human development' first gained traction within the international development community during the 1990s among United Nations (UN) agencies and big international Non-Governmental Organizations (NGOs). At the time, there was mass recognition that development policy was too preoccupied with resource ownership and neoliberal economic theory; the resultant paradigm shift led to the creation of the Human Development Index, a measure which offers an alternative to development proxies, such as the Gross Domestic Product (GDP) or Gross National Income (GNI), which focus exclusively on economic productivity.

Today, the HDI is a leading tool for tracking and comparing countries' progress and, thus, we use it as our chosen proxy for human development in this paper. At the same time, however, readers should note that the HDI only represents a single interpretation of human development rather than a perfect or complete account. Indeed, the UNDP has never claimed their HDI is infallible. Quite the opposite, they deliberately designed the HDI to be 'flexible in both coverage and methodology [so that] gradual refinements [could be made]' (Kovacevic, 2010b, p. 1). Such refinements were made in 2010 when the UNDP updated the HDI's methodology and introduced its inequality-adjusted variant—the IHDI.

The HDI's current formulation looks at three dimensions—health, knowledge and standard of living (Conceição, 2019). And each of those dimensions is made up of either one or two indicators. 'Life expectancy at birth' is the indicator for health, 'Mean Years of Schooling' and 'Expected Years of Schooling' are the indicators for knowledge (50:50 weighting) and 'Gross National Income per capita' ($ PPP) is the indicator for 'standard of living'. After collecting the raw data, the UNDP normalizes them to produce indicator- and then dimension-scores that typically range from 0 to 1. The UNDP then obtains each country's final HDI score by taking the geometric mean of the separate dimension scores:

$$\mathrm{HDI} = \sqrt[3]{I_{\mathrm{Health}} \times I_{\mathrm{Education}} \times I_{\mathrm{Income}}}$$

where 'I_x' is one of the three dimensions.

Unlike the IHDI, HDI scores are not affected by the distribution of achievement across the given population (Conceição, 2019). HDI scores can appear misleading when a small proportion of the population is responsible for a high proportion of the achievement within a dimension, which can mask critically low levels of achievement elsewhere (Kovacevic, 2010a). The IHDI overcomes this by including a variable which penalizes inequality called 'A'. This is calculated as $A = 1 - g/\mu$, where g is the geometric mean and μ is the arithmetic mean of data for the given indicator (Kovacevic, 2010a). Subsequently, the final IHDI is obtained as shown below:

$$\mathrm{IHDI} = \left(\sqrt[3]{(1 - A_{\mathrm{Health}}) \times (1 - A_{\mathrm{Education}}) \times (1 - A_{\mathrm{Income}})}\right) \times \mathrm{HDI}$$

Accounting for inequality can drastically change a country's human development profile. For example, in 2019, IHDI scores were 19.5% lower on average than HDI scores, and some countries' rankings differed dramatically too. One such country was Brazil, whose IHDI rank was 23 places lower than its HDI rank. The extent of these discrepancies demonstrates the importance of taking inequality into account when measuring human development and our study.

The second key concept in our paper is the 'Competitiveness of Nations'. The basic notion of nations competing can be traced back through several chapters of classical and neoclassical economic theory; from mercantilism, through Adam Smith's theory of absolute advantage, David Ricardo's comparative advantage, the Heckscher-Ohlin model of factor endowment and onwards (García Ochoa et al., 2017; Mashabela & Raputsoane, 2018). Notwithstanding some continuity between these paradigms, the concept has evolved significantly to where it is today. In general, the competitiveness of nations is no longer treated as a zero-sum game (Schwab, 2019), and while international trade remains an important factor, judgements of success are no longer based on the ratio of imports to exports alone (García Ochoa et al., 2017). Contemporary conceptualizations are generally more complex, taking into account more micro- and macro-level factors (Ketels, 2016). Although, it is worth noting that they often fail to consider the military power of countries as mercantilism once did.

In our paper, 'Competitiveness of Nations' specifically refers to countries' economic capabilities relative to one another, with a twin focus on the well-being of citizens (Ülengin et al., 2011). Önsel et al. (2008) state that a nation is competitive if 'it can, under free and fair market conditions, produce goods and services that meet the standards of international markets while simultaneously expanding the real income of its citizens, thus improving their quality of life' (p. 222). Note that this definition treats economic productivity as a 'means' rather than an 'ends'—an important parallel with human development. The difference is that while patrons of human development assert the equal importance of health, education and the economy (which is visible in the HDI/IHDI), patrons of 'Competitiveness of Nations' are primarily concerned with economic means of development (Im & Choi, 2018).

The Global Competitiveness Index (GCI), published annually by the World Economic Forum (WEF), is one of two major indices that rank nations according to competitiveness. The latest version— the GCI 4.0—considers 'factors and attributes [driving competitiveness] in the [context] of the Fourth Industrial Revolution' (Schwab, 2019, p. vii). It comprises 103 indicators covering 12 'pillars' of competitiveness:

1.	Institutions	7.	Product market
2.	Infrastructure	8.	Labour market
3.	ICT adoption	9.	Financial system
4.	Macroeconomic stability	10.	Market size
5.	Health	11.	Business dynamism
6.	Skills	12.	Innovation capability

Once again, readers should note that the GCI only represents a single interpretation of the factors driving economic competitiveness between nations. Like human development, proclaiming a universal definition of the concept is problematic because it is so subjective and pluralistic. The contentious nature of the key concepts in our paper is a notable limitation of our work.

Following the collection of raw data, the WEF computes the GCI over two stages. Firstly, the WEF normalizes raw data to produce scores that typically range from 0 to 100. A score approaching 100 means the country is near an ideal situation 'where the factor no longer represents a constraint on productivity'. Conversely, a score approaching 0 indicates 'a completely unsatisfactory situation' (Schwab, 2019, p. 13). The equation below summarizes the normalization process:

$$\text{Normalized indicator score} = \left(\frac{\text{Actual value} - \text{Minimum value}}{\text{Maximum value} - \text{Minimum value}}\right) \times 100$$

Readers should note that the minimum and maximum values in the equation above vary between indicators and may reflect policy targets, naturally occurring minimums and maximums, or percentile figures derived from statistical analysis (Schwab, 2019, p. 13). The normalized data are aggregated over successive phases, proceeding from indicator-level up until each country's fully composite GCI score is obtained. Their aggregation method is to find the arithmetic mean of relevant subcomponents (Schwab, 2019, p. 13). At the end of the process, each country receives a score ranging from 0 to100.

In this paper, we assume that the primary objective of the nation's economy is the human development of its citizens. Therefore, we also assume that it is essential to examine the performance of countries in terms of how well they use their economic competitiveness to generate human development for their citizens. Building on previous work by Ülengin et al. (2011) and drawing on data from the 2019 editions of the GCI and IHDI, we propose a hybrid methodology—made up of data envelopment analysis (DEA) and random forest classification—to achieve this function.

Our specific objectives are listed below:

1. Apply DEA to determine the relative efficiency of countries in terms of their ability to leverage economic competitiveness to generate human development, taking GCI and IHDI indicators/dimensions as input and output variables, respectively.
2. Apply random forest classification to determine the relative importance of DEA variables on country performance.
3. Identify possible explanations for the first and second stage results, the implications for practice and future research, and the limitations of the study.

The rest of the chapter is organized as follows. The second section provides the theoretical context for the reader, introducing the human development and competitiveness concepts. The third section outlines the methodology, from the selection of countries for the dataset, to our decision to use DEA and random forest classification instead of alternatives. The fourth section summarizes the findings of our model. The fifth section discusses the study's practical and theoretical implications and its limitations. Finally, the concluding chapter summarizes the key takeaways, reiterates the aim and value of our study, before recommending areas for future research.

2 Literature Review

Few studies have focused on the relationship between competitiveness of nations and human development, and fewer still have examined the relationship in a manner similar to what we propose here. Cetinguc et al. (2018) explored the relationship between the GCI and HDI but their theorization focused on how human development could be applied to produce competitiveness, rather than vice versa like our paper. Nonetheless, the paper's findings support the existence of a relationship between the GCI and HDI, with the authors concluding that countries should invest more in their human capital to foster the 'innovativeness' and competitiveness of their economies. Bucher (2018) confirmed a strong correlation between the HDI and GCI but did not speculate any further on the nature of the relationship. There was also no discussion of the relative performance of countries, or the relative importance of the underlying factors.

Very much the forerunner and inspiration of this study, Ülengin et al. (2011) used data from the 2005 editions of the GCI and HDI to build DEA models that assessed how well nations generated human development from economic competitiveness. Subsequently, they applied ANN analysis to investigate the relative importance of sub-factors (i.e. the sub-indices of the GCI and HDI). Like us, they prescribed a hierarchy and direction of influence between the variables, asserting—on philosophical grounds—that economic competitiveness is only meaningful if it improves human development. The study also employed a super-efficiency DEA model to issue distinct scores and ranks to efficient countries.

The results indicated that wealthier countries with more developed economies tended to have higher efficiency scores. Africa was the worst-performing region. Surprisingly, South American countries—including Venezuela and Argentina—performed exceptionally well, with an average score higher than Europe and North America. The authors noted that the average scores of Europe and North America, though high and approaching one, were lower than anticipated. The ANN analysis indicated that GDP per Capita was by far the most influential sub-factor, followed by Life Expectancy and Efficiency Enhancers (a GCI sub-index covering market size, adoption of technology, plus the quality of financial and labour markets).

Kılıç and Kabak (2019) proposed two DEA models, having assumed the possibility of a bi-directional relationship between economic competitiveness and human development. One model examined countries' efficiency at producing economic competitiveness from human development, while the other examined the opposite direction of influence. One of the distinguishing elements of the study was the assumption of a three-year time lag between cause and effect. Hence, the authors paired input data from 2007 to 2014 with output data from 2010 to 2017. In addition, they adopted a time window approach to calculate yearly scores, using averages that took neighbouring years into account. Finally, the study used cluster analysis to investigate the stability of countries' DEA performances over the period (2010–2017) and to provide additional insight about the relationship between the variables.

The results indicated that the GCI-to-HDI DEA model was more reliable than its HDI-to-GCI counterpart, leading Kılıç and Kabak (2019) to conclude that the predominant direction of influence is from competitiveness to human development. This assertion is significant when previous studies have elected to depict human development as a determinant of economic competitiveness (see: Cetinguc et al., 2018; Bucher, 2018).

Cluster analysis indicated high levels of stability, meaning countries tended to remain in the same band of achievement between 2010 and 2017. Kılıç and Kabak (2019) interpreted this as proof of the model's veracity, but it could also imply that nations' capabilities had become crystallized. This fits with the pattern of the DEA results, which echoed the findings of Ülengin et al. (2011) albeit with fewer surprises: the best-performing nations were almost exclusively highly developed and affluent.

In their other study, Kılıç and Kabak (2020) used DEA alongside Fuzzy Analytical Network Process (FANP). Again, their goal was to investigate a bi-directional relationship between economic competitiveness and human development. However, this time, the authors compared the results of two DEA models—one for each direction of influence—with those from a separate composite index weighted according to FANP. Once again, the authors assumed a time lag of 3 years between cause and effect, using GCI and HDI data from 2012 and 2015.

Kılıç and Kabak (2020) reasoned that FANP offered a means of incorporating the complexity inherent in multi-criteria decision-making, where inter-dependencies and hierarchies exist between and within variables, alongside uncertain human decision-makers. We agree that FANP is an exciting alternative to the prescriptive and somewhat rudimentary weighting of subcomponents in the GCI and HDI. That said, we also recognize its limitations. For example, in Kılıç and Kabak (2020)'s application, FANP involved respondents making highly complex determinations in reductive numerical terms. Moreover, the questions, which covered multiple social science disciplines, were directed at just two experts. We feel a more targeted consultation of a larger field of experts would make the approach more credible.

Kılıç and Kabak (2020)'s main conclusion was that competitiveness has a greater effect on human development than vice versa, reinforcing the findings of their 2019 paper. This was due to the authors observing a stronger correlation between the results of the GCI-to-HDI DEA model and the FANP index. It is also notable that the GCI-to-HDI model showed surprising high performers, including Algeria and Venezuela, alongside the likes of Australia and Norway. The FANP results were less surprising by comparison.

This study distinguishes itself from the existing literature in the following ways. Firstly, it uses data from the 2019 editions of the HDI and GCI, which is doubly significant because the indices' methodologies were updated in 2010 and 2019, respectively. Thus, it offers a fresh snapshot and analysis of country performance. This study also separates itself by using the inequality-adjusted HDI to populate the output side of the DEA model. As we have already noted, accounting for inequality can dramatically change one's perception of a country's progress. With that in mind, the adoption of the IHDI is a clear methodological advancement. Finally, this study is the first to combine random forest classification with DEA in this field. In our methodology below, we elaborate further on why we selected random forest classification over alternatives.

In addition to making an original contribution to topical literature, this paper also has practical relevance for policymaking, governance and commerce (alongside similar studies and indices in general). As world-leading indices, the HDI and GCI are already magnificent tools but, by using DEA to unify them, we extend their utility. Likewise, when combined with DEA, machine learning offers another convenient way of extracting further insight from their rich data. Ultimately, through the proffer of this tool, we hope to promote government accountability and efficacy via the benchmarking of nations' achievements or the identification of areas for further improvement. Equally, we wish to support businesses as they weigh the conditions of countries vying for their investment.

3 Methods

Our methodological framework combined DEA and random forest classification, adapting and improving on Ülengin et al. (2011). Our first objective was to use DEA to assess the relative efficiency of 124 countries, measuring their ability to convert economic competitiveness into human development. Our second objective was to identify the relative importance of variables affecting countries' DEA outcomes. To this end, we trained and deployed a random forest classification model. Figure 1 provides an overview of the methodology.

3.1 Selection of Countries

Our analysis included all 124 countries featuring in *both* the 2019 GCI and IHDI. The complete list of these countries is in the electronic companion which can be downloaded from the book's website, alongside respective scores and ranks. Unlike Ülengin et al. (2011), we decided not to filter or cluster the eligible countries further because choosing criteria for doing so would have been highly subjective and perhaps controversial. Furthermore, we did not feel it was imperative since the WEF and UNDP present all countries together on their indices.

A larger sample of countries offered methodological advantages too. Firstly, concerning DEA, it increased our chances of 'capturing high-performance units that would determine the efficient frontier and improve the discriminatory power' of our model (Sarkis, 2007, p. 1–2). It also facilitated having more input and output variables (Dyson et al., 2001). Concerning the second stage of our methodological framework—random forest classification—a larger base meant the training and test subsets could also be significant, which helped us train the model (Beleites et al., 2013). Finally, it also reduced the risk of 'overfitting' and, in general, gave us extra space to experiment with different parameters and techniques (Riley et al., 2020).

3.2 Stage 1: Data Envelopment Analysis

DEA is a linear programming technique for determining the relative efficiency of a set of entities, referred to as decision-making units (DMUs). It is a non-stochastic, non-parametric alternative to econometric models such as regression analysis (Ülengin et al., 2011). Convenience is one of its major advantages, since it does not require 'assumptions regarding the statistical properties of variables' and offers the researcher a significant amount of discretion (Ülengin et al., 2011, p. 19). In addition, it gives the researcher the control over extraneous constraints, plus the selection and weighting of variables, though this places an extra burden on the researcher to make the model meaningful (Ruggiero, 1998; Ülengin et al.,

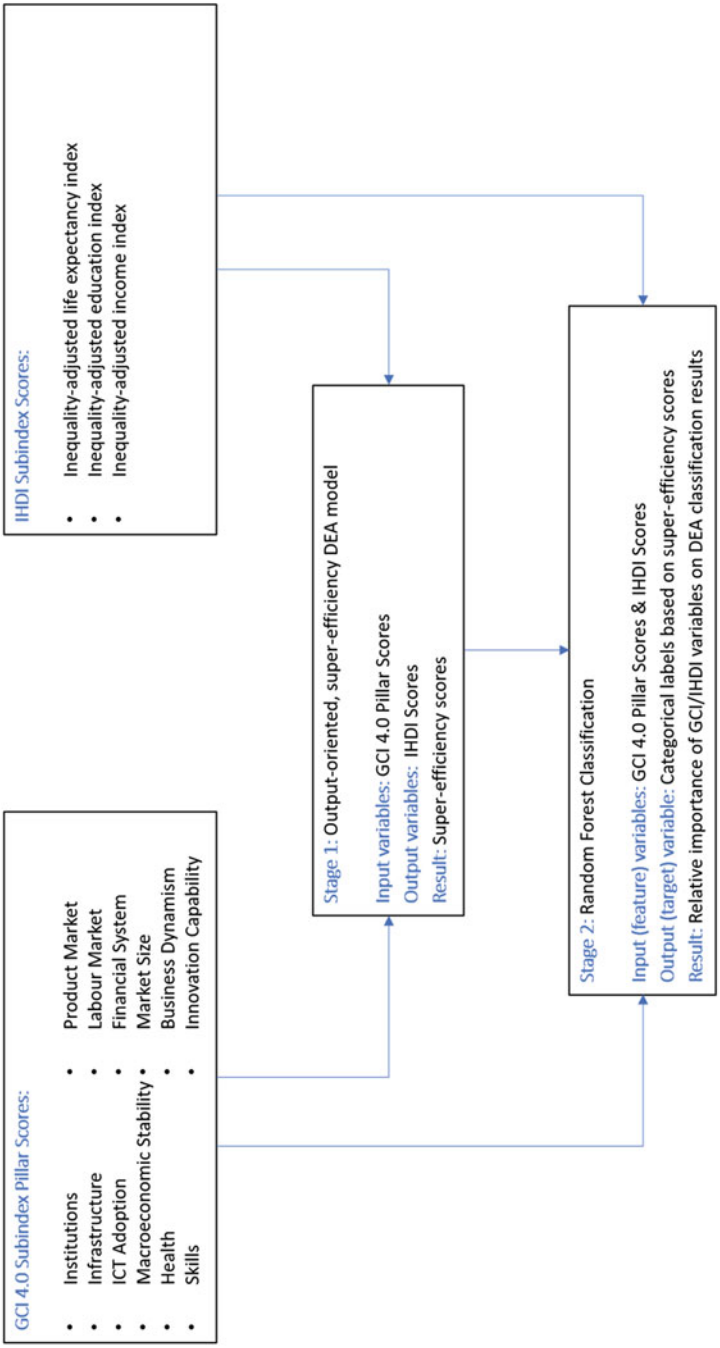


Fig. 1 Summary of the methodology

Table 1 GCI 4.0 pillars and categories

Enabling environment	**Markets**
1. Institutions	7. Product Market
2. Infrastructure	8. Labour Market
3. ICT Adoption	9. Financial System
4. Macroeconomic Stability	10. Market Size
Human capital	**Innovation ecosystem**
5. Health	11. Business Dynamism
6. Skills	12. Innovation Capability

Table 2 Revised DEA output variables

Dimension of human development	Ülengin et al. (2011)	This paper
Health	Life expectancy	Inequality-adjusted life expectancy index (life expectancy at birth)
Knowledge	Combined gross enrolment ratio (primary, secondary and tertiary)	Inequality-adjusted education index (expected years of schooling; mean years of schooling)
Standard of Living	GDP per capita	Inequality-adjusted income index (GNI per Capita)

2011). Nonetheless, DEA remains a popular option and continues to appear in many contexts, including this field of inquiry (see: Ülengin et al., 2011; Liu et al., 2013a; Mariano et al., 2015; Kılıç & Kabak, 2019, 2020).

Our DEA model consisted of 124 DMUs (countries), plus 12 input and 3 output variables from the GCI and IHDI, respectively. We view the output of human development as the goal of all nations, so we employed an output-oriented model. Regarding the input side, the WEF formerly grouped the 12 pillars of competitiveness (GCI) under three categories:

- Basic requirements
- Efficiency enhancers
- Innovation and sophistication factors.

Ülengin et al. (2011), Kılıç and Kabak (2019, 2020) adopted these categories as inputs for their DEA model, but they are obsolete since the emergence of the GCI 4.0. Today, the 12 pillars are arranged into the four categories as follows:

However, instead of adopting these categories as inputs, we used the 12 pillars for a more granular picture of relative importance. And in a significant departure from previous studies, we populated the output side of our model with variables from the IHDI instead of the HDI. Table 2 illustrates the difference versus Ülengin et al. (2011), whose output variables came from the pre-2010 HDI.

We used the IHDI to account for the distribution of achievement across the population. The IHDI's formulation penalizes countries in proportion to the unevenness of their outcomes: countries with higher levels of inequality are penalized more. The input and output data were from 2019 editions of their respective indices.

Fig. 2 Inversion of raw input data for DEA

The raw input and output data were already in index form, so there was no need for normalization or standardization to address issues related to scale and magnitude (Sarkis, 2007). However, since DEA is a measure of efficiency, models customarily reward lower consumption of inputs and higher production of outputs (Lewis & Sexton, 2004). In other words, input and output variables are usually defined as costs and benefits, respectively. While our output variables fulfilled this stipulation in their raw form, input variables from the GCI were inverted to reflect cost instead of benefit (see Fig. 2):

Any score below 1 after this transformation was assigned a score of 1 instead, to avoid input variables near or equal to zero distorting DEA results (see Appendix D).

Like Ülengin et al. (2011), we based our model on the classic Charnes–Cooper–Rhodes (CCR) blueprint (Charnes et al., 1978) and thus assumed constant returns to scale. The alternative Banker–Charnes–Cooper (BCC) framework (Banker et al., 1984) assumes variable returns to scale. 'Returns to scale' determines the shape of the efficiency frontier and affects the evaluation of DMUs. We took the view that having constant returns to scale was more appropriate because it implied that marginal gains in competitiveness should lead to equally proportional benefits for the citizens of the given country.

Our model also employed super-efficiency. DEA efficiency scores typically range from zero to one, with efficient units scoring one and inefficient units scoring less than one. So, while this facilitates the sortation of efficient DMUs (scores = 1) from inefficient ones (scores <1), efficient units are not differentiated (Andersen & Petersen, 1993). Super-efficiency models address this limitation by enabling individual scores of 1 *or* above for efficient units too.

Overall, we employed the same output-oriented, CCR-based, super-efficiency model used by Ülengin et al. (2011):

$$\text{Maximize } \eta_0, \text{ subject to :}$$

$$x_{i0} - \sum_{j=1;\mathrm{j}\neq 0}^{n} \mu_j x_{ij} \geq 0 \; i = 1, \ldots, m,$$

$$\eta_0 y_{r0} - \sum_{j=1;\mathrm{j}\neq 0}^{n} \mu_j y_{rj} \geq 0 \; r = 1, \ldots, s,$$

$$\mu_j \geq 0 \quad j = 1, \ldots, n$$

The model assumed there were n comparable DMUs, which all use m inputs x_{ij} $(i = 1, \ldots, m)$ to produce r outputs y_{rj} $(r = 1, \ldots, s)$. The super-efficiency value for DMU_0 was subsequently obtained as the value of $\frac{1}{\eta_o}$. Efficient DMUs obtained scores of 1 or above, while inefficient DMUs obtained score of less than 1. The model performed two functions. Firstly, it sorted DMUs into two groups, efficient and inefficient, and, secondly, it ranked all DMUs from most efficient to least.

The model was built in AMPL, using the CPLEX solver (see Appendix E). AMPL uses a syntax which '[closely matches] that of the algebraic, symbolic representation of a linear programming model' and thus readily accommodated our needs (Green, 1996, p. 1). The code for the model drew on an example from Cooper et al. (2007).

3.3 *Stage 2: Random Forest Classification*

In the second stage of the methodology, we trained a random forest to estimate the relative importance of the DEA variables on a country's classification as either efficient or non-efficient. A *random forest* is an ensemble learning algorithm that combines the predictive power of multiple, independently formed decision trees to perform classification or regression tasks (Breiman, 2001). Their 'randomness' has two sources. First, the decision trees that constitute the forest are trained with random subsets of data; second, each layer of the tree's node-splitting process uses a random subset of feature variables (Breiman, 2001). Hence, the trees are encouraged to take their own random approach to classification. Crucially, when classification outcomes differ between trees, the final prediction is based on the trees' aggregate decision, as displayed in Fig. 3.

As a classification tool, random forests boast several advantages. They perform better on classification tasks than alternatives like neural networks and support vector machines (Cutler et al., 2007; Fukuda et al., 2016), yet are remarkably user-friendly (Liu et al., 2013a). Like DEA, they do not require assumptions about the statistical properties of data (Liu et al., 2013b). Data pre-processing requirements are low (Liu et al., 2013b), and, aside from the configuration of parameters, they function with minimal human input (Lebedev et al., 2014). Other advantages include:

- resilience to overfitting
- their utility for determining the relative importance of feature variables (Cutler et al., 2007; Fukuda et al., 2016)
- 'their ability to model complex interactions between predictor variables' (Cutler et al., 2007, p. 2783)
- their modest computational processing requirements (Lebedev et al., 2014)

So, for these reasons and despite an artificial neural network featuring in Ülengin et al. (2011) to good effect, we preferred to use a random forest.

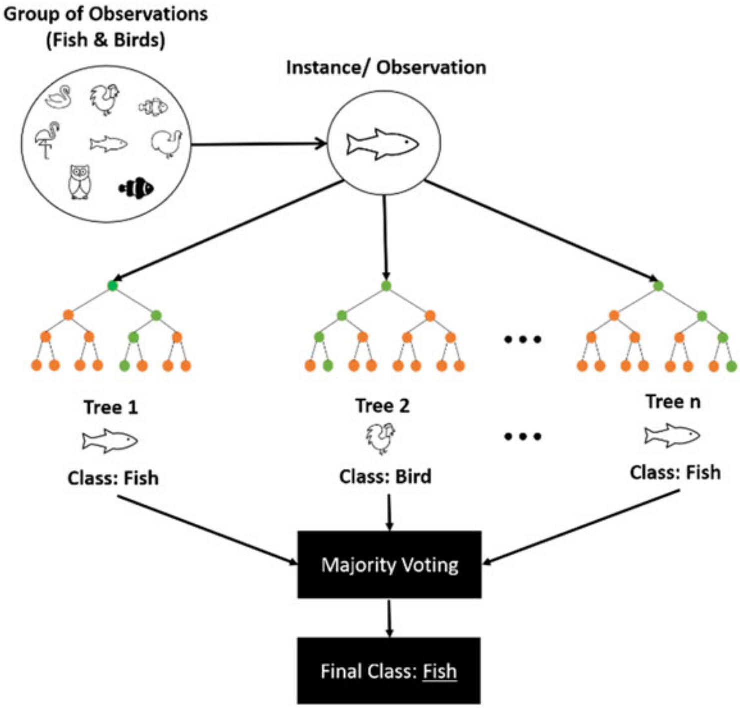


Fig. 3 Random forest classification

We built the random forest using the Python library 'Scikit-learn' (Pedregosa et al., 2011) (see Appendix F). Then, we trained it to predict whether a country was efficient or non-efficient using DEA variables and results. The model included 15 feature variables and one target variable. The feature variables were the input (GCI) and output (IHDI) variables used during DEA, while the target variable was an encoded version of countries' super-efficiency scores. Efficient countries were assigned scores of 1, while non-efficient countries were assigned scores of 0.

We used k-fold cross-validation to train the model and measure its accuracy, meaning the data was randomly split into k parts or overlaps, where the value of k is user-defined (Yadav & Shukla, 2016). We used stratification so that each fold preserved the same ratio of efficient to inefficient countries as the overall dataset. We then trained using '$k - 1$' parts and used the remaining one part for testing. This process was repeated k times until each part had been used for training $k - 1$ times and for testing once. Finally, we used the average testing performance as an indication of the model's overall accuracy.

When tuning the model for higher accuracy, we focused on two parameters—the number of folds (i.e. k-values) and the number of trees. After exploring the effects of different combinations, we decided that the optimum combination was 13 folds and 250 trees (see Fig. 4), which produced a cross-validation accuracy of 96.84%.

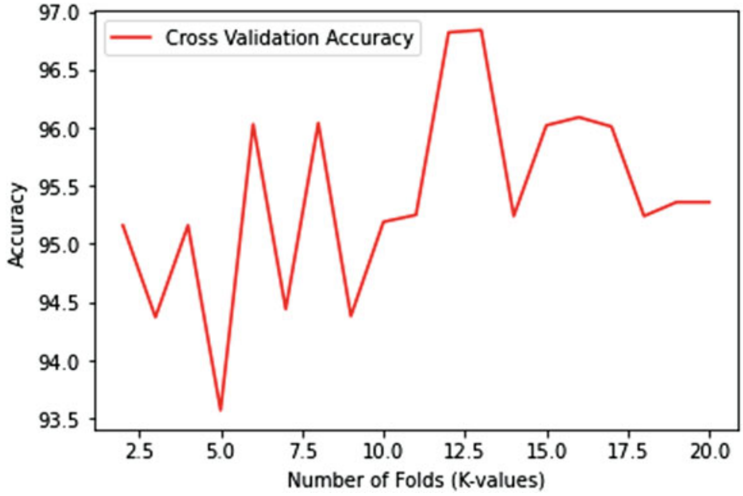


Fig. 4 Cross-validation accuracy with different k-values (number of trees = 250)

Following the advice of Probst et al. (2019) and given the modest size of our dataset plus the objectives of our analysis, we used Scikit's default settings for all other parameters (Pedregosa et al., 2011).

Last, we determined the relative importance of feature variables to the classification task using Scikit's designated 'feature importance' functionality (Pedregosa et al., 2011). The functionality offered two alternatives: computation could either be impurity- or permutation-based. The impurity-based option suited our study because our feature variables were continuous with similar, high levels of cardinality (Altmann et al., 2010).

4 Findings

4.1 Stage 1: Data Envelopment Analysis

The results revealed that most countries are inefficient at using their economic competitiveness to generate human development for their citizens. The median DEA score was 0.491, which means at least 50% of countries are less than halfway to reaching efficiency. Furthermore, there were only 20 efficient countries that obtained DEA scores greater than or equal to one. The remaining 104 countries all obtained scores of less than one, meaning they were inefficient. Appendix G provides a complete list of countries' super-efficiency scores and ranks.

Table 3 displays the best- and worst-performing countries according to the DEA results. The United States (USA) is the most efficient nation with an exceedingly

Table 3 Best- and worst-performing nations

			Bottom 20 countries		
Top 20 countries	Super-efficiency score (SE)	Super-efficiency rank		Super-efficiency score (SE)	Super-efficiency rank
United States	3.654	1	Guinea	0.261	105
Switzerland	1.650	2	Uganda	0.259	106
Japan	1.640	3	Eswatini	0.255	107
Korean Rep.	1.508	4	Lesotho	0.252	108
Germany	1.389	5	Malawi	0.251	109
Singapore	1.341	6	Zimbabwe	0.246	110
Netherlands	1.296	7	Ethiopia	0.234	111
China	1.276	8	Gambia	0.228	112
Sweden	1.230	9	Yemen	0.227	113
France	1.227	10	Angola	0.225	114
Hong Kong	1.215	11	Burundi	0.225	115
Finland	1.205	12	Mali	0.223	116
Denmark	1.153	13	Cameroon	0.215	117
Israel	1.112	14	Benin	0.214	118
Iceland	1.063	15	Burkina Faso	0.207	119
Norway	1.042	16	Mauritania	0.198	120
Australia	1.033	17	Mozambique	0.193	121
New Zealand	1.020	18	Haiti	0.183	122
United Kingdom	1.017	19	Congo (DR)	0.179	123
Canada	1.010	20	Chad	0.142	124

high score of 3.654. Switzerland (SE = 1.650) and Japan (SE = 1.64) are next. Chad is the least efficient nation with a score of 0.142, followed by Congo (SE = 0.179) and Haiti (SE = 0.183). The difference between a country's score and the efficiency frontier (SE = 1) signals the magnitude of a particular country's over- or under-achievement.

Table 4 provides further information on the best- and worst-performing countries. It shows that the best performers had better than average scores for every input and output variable and that the opposite was true for the worst performers.

Figure 5 shows the distribution of countries' efficiency scores over bands. With 22 countries, the most frequent band was $0.2 \leq \text{SE} < 0.3$. Next were the $1 \leq \text{SE}$ and $0.5 \leq \text{SE} < 0.6$ bands, with 20 countries falling into each. On the other hand, the least frequent bands were $0.8 \leq \text{SE} < 0.9$ (1 country) and $0.7 \leq \text{SE} < 0.8$ (2 countries). Overall, the distribution had a bimodal shape. Two distinct peaks were separated by a low-frequency trough covering the $0.6 \leq \text{SE} < 0.9$ range, containing only nine countries. It is unclear whether this shape is evidence of a deeper trend. If we assume it is, there are multiple possible explanations. One explanation is that an external influence, outside the scope of the model and beyond national control, keeps countries in two separate groups and makes it difficult for them to traverse the $0.6 \leq \text{SE} < 0.9$ range. It could also reflect a lack of internal resources and capabilities. Finally, there may be some other x-factor lacking among countries with scores below 0.6 but abundant in countries with scores above 0.9.

4.2 Correlation Analysis of Country Ranks

We used scatter plots and correlation analyses to investigate the degree of consistency between a country's rank on either the IHDI or GCI and its super-efficiency rank. We specifically used Kendall's test for the latter because ranking data is discrete and ordinal (Cliff, 1996).

The results indicate a strong relationship between a country's rank on either the GCI (Fig. 6; $\tau = 0.86$) or the IHDI (Fig. 7; $\tau = 0.75$) and its super-efficiency rank. The scatters shown in Figs. 6 and 7 are homoscedastic and portray positive linear relationships. Countries with better GCI or IHDI ranks generally obtained better super-efficiency ranks. As a result, these countries fell in the bottom-left corners of the graphs. On the other hand, countries with high (bad) GCI or IHDI ranks generally obtained high super-efficiency ranks and thus fell in the top-right corners of the graphs.

Table 5 shows the results of Kendall's test: the tau coefficient and p-values. Tau coefficients vary between −1 and 1 (Cliff, 1996). A negative value signifies an inverse relationship, while a positive value signifies a positive relationship. The tau coefficient reading was 0.86 for the relationship between GCI and super-efficiency ranks and 0.75 for IHDI and super-efficiency ranks. With p-values virtually equal to zero, these results indicate strong, positive associations. However, there is a stronger

Table 4 DEA variables for best- and worst-performing nations

	United States	Switzerland	Japan	Mean	Haiti	Congo (DR)	Chad
Input (GCI) variables							
Institutions	71.17	77.51	71.67	54.66	30.87	32.78	35.42
Infrastructure	87.90	93.16	93.16	64.83	26.88	29.23	30.53
ICT adoption	74.35	78.58	86.20	54.33	28.14	19.11	10.77
Macroeconomic stability	99.77	100.00	94.89	79.12	60.16	31.39	75.00
Health	83.02	99.94	100.00	74.83	50.82	41.60	35.87
Skills (Workforce)	82.47	86.72	73.28	60.88	41.48	42.30	29.04
Product market	68.55	63.80	70.36	54.90	37.81	44.73	35.43
Labour market	77.98	79.48	71.54	59.77	49.07	48.30	42.22
Financial system	90.99	89.72	85.94	61.77	44.02	42.53	37.31
Market size	99.53	66.23	86.86	54.93	33.92	43.27	37.09
Business dynamism	84.21	71.55	75.03	59.91	14.07	40.47	29.67
Innovation capability	84.15	81.20	78.31	43.09	18.90	17.97	22.66
Output (IHDI) variables							
Inequality-adjusted life expectancy index	0.85	0.94	0.96	0.72	0.46	0.40	0.31
Inequality-adjusted education index	0.85	0.88	0.84	0.58	0.28	0.35	0.16
Inequality-adjusted income index	0.70	0.82	0.85	0.56	0.21	0.23	0.31

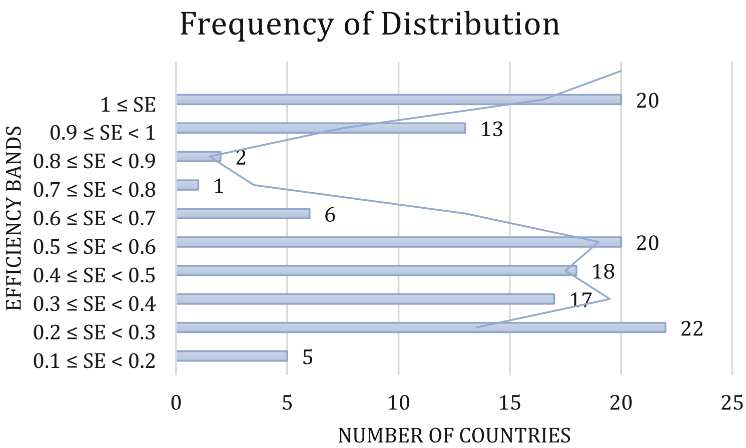


Fig. 5 Frequency of distribution across efficiency bands

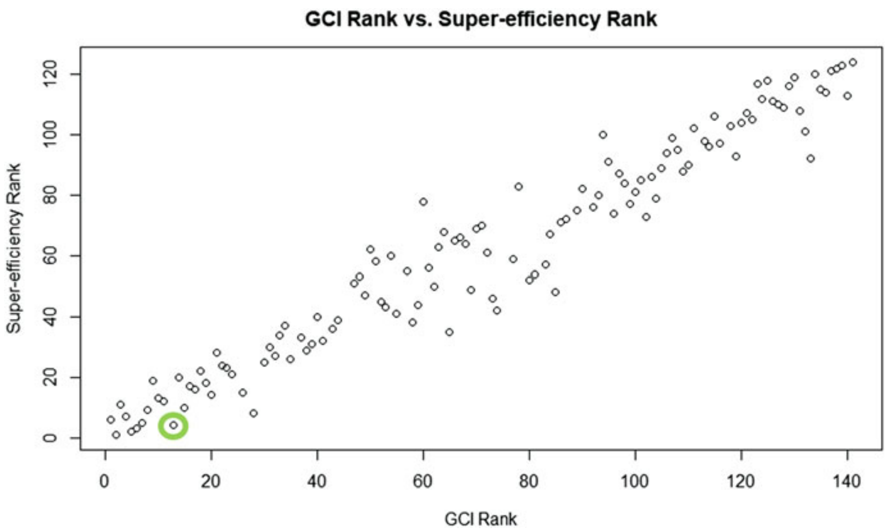


Fig. 6 Scatter plot of GCI rank and super-efficiency rank

Table 5 Results of Kendall's test

Relationship	Kendall's Tau coefficient (τ)	p-value (p)
GCI rank vs. super-efficiency rank	0.86	$p < 2.2e{-}16$
IHDI rank vs. super-efficiency rank	0.75	$p < 2.2e{-}16$

relationship between GCI rank and super-efficiency rank than between IHDI rank and super-efficiency rank, which explains why Fig. 6 is less scattered than Fig. 7.

Despite high consistency overall, super-efficiency ranks did not always correspond with GCI and IHDI ranks. China is the most extreme outlier. Its datum point

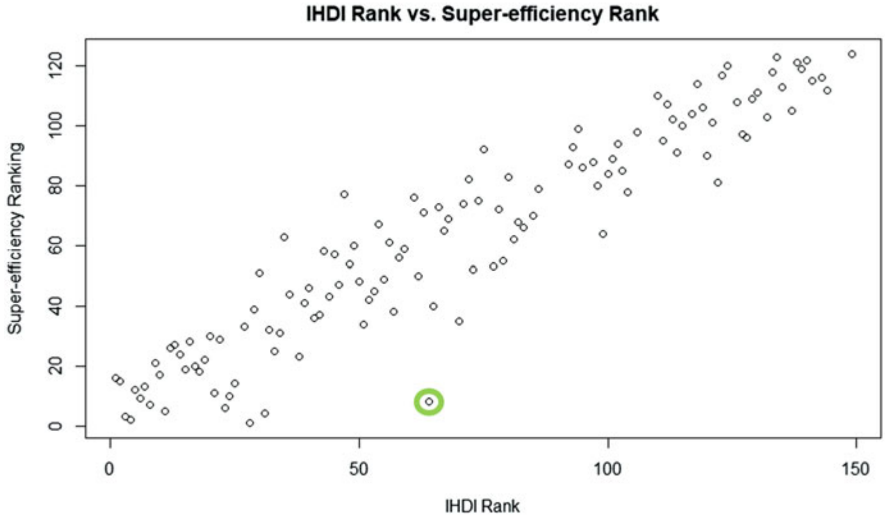


Fig. 7 Scatter plot of IHDI rank and super-efficiency rank

is circled in green in Figs. 6 and 7. Being on the bottom side of both scatters means China's DEA performance was better than expected, given its GCI and IHDI ranks. China obtained an eighth place super-efficiency rank but was only ranked 28th on the GCI and 64th on the IHDI. The implication is that China is highly efficient at leveraging its global competitiveness to achieve human development.

4.3 Analysing Group Performance

Although we avoided filtering or clustering countries in the pre-analysis, we considered group performance when interpreting and discussing results. Firstly, we considered geographic region, classifying countries according to their continent:

- Africa
- Asia
- Australia/Oceania
- Europe
- North America[1]
- South America

Secondly, we considered countries' affiliations with major intergovernmental organizations:

[1] Central America and the Caribbean were considered parts of North America.

- Emerging 7 (E7)
- Group of 7 (G7)
- Organization for Economic Co-operation and Development (OECD)

Thirdly, we considered countries 'Level of Human Development'. This classification was based on the 2019 HDR, with fixed cut-offs applied to HDI scores:

- 0.800 or above for very high human development
- 0.700–0.799 for high human development
- 0.550–0.699 for medium human development
- Less than 0.550 for low human development

Lastly, we considered the political regime of the countries. This classification was based on data from the V-Dem Project (Roser, 2013; Coppedge et al., 2019, 2021). Countries were grouped into 4 possible regime categories ranging from most to least democratic[2]:

- Liberal Democracy—Complete Democracy
- Electoral Democracy—Predominantly democratic with autocratic features
- Electoral Autocracy—Predominantly autocratic with democratic features
- Closed Autocracy—Complete Autocracy

Please refer to Appendix H for further detail on how specific countries were classified.

Table 6 summarizes group performance. The second and third columns show the ratio of efficient to inefficient countries within each group. The fourth column shows the geometric means of the groups' super-efficiency scores (GMSE). The use of geometric means (instead of arithmetic means) guaranteed that one country's good performance could not compensate for the bad performance of another, thus giving a better account of the group's central tendency.

Geographic Regions

Our analysis indicates that Africa is the least efficient geographic region. In addition to having no efficient countries, it also had the lowest average score (GMSE = 0.269). Its highest performing nation—Mauritius—only managed an efficiency score of 0.567 (Appendix G). Europe was the best-performing region. While having the highest number of countries at 42, the region had a high GMSE (0.805). Europe also had the most favourable ratio of efficient to inefficient countries outside of Australia/Oceania. Although Australia/Oceania had the highest GMSE (1.027), the group comprised two countries—Australia and New Zealand.

Intergovernmental Organizations

The performance of the E7 group was mediocre. The GMSE was 0.591, and China was the only efficient member of the group. We had anticipated slightly better results because economic efficiency is one of the connotations of 'E7' status.

[2] For further information on the features of democracy evaluated by the V-Dem Project, please refer to: Coppedge et al. (2021, pp. 254).

Table 6 Summary of DEA performance by group

	No. of countries	No. of efficient countries	Geometric mean of super-efficiency scores (GMSE)
Geographic region			
Africa	33	0	0.269
Asia	27	5	0.529
Australia/Oceania	2	2	1.027
Europe	42	11	0.805
North America	12	2	0.519
South America	8	0	0.463
Intergovernmental organization			
E7	7	1	0.591
G7	7	6	1.394
OECD	35	17	1.005
Level of human development			
Very High	48	19	0.914
High	34	1	0.479
Medium	21	0	0.323
Low	21	0	0.233
Political regime			
Closed Autocracy	9	2	0.500
Electoral Autocracy	36	1	0.343
Electoral Democracy	44	1	0.445
Liberal Democracy	35	16	0.912

With high scores and good ratios of efficient to inefficient countries (G7—6:7; OECD—17:16), the performances of the G7 and OECD groups were strong. These results felt unsurprising because the countries in these organizations tend to have wealthy societies, stable economies and high levels of human development.

Level of Human Development

As expected, the *Very High* group had the best mean performance (GMSE = 0.914) while the *Low* group had the worst (GMSE = 0.233). The *High* group also performed better than the *Medium* group.

Figure 8 shows how the average disparity in efficiency changed across consecutive levels of development. The most significant disparity was between countries with very high and high levels of human development. By contrast, the smallest disparity was between countries with medium and low levels of human development. The implication was that, as a country moves from low levels of human development to very high levels of human development, it becomes increasingly difficult to make the associated leaps in efficiency. This explanation also fits the bimodal distribution mentioned above in Fig. 5. The upper and lower bounds of this range (0.9 and 0.6) fit with the mean performances observed for the Very High (0.9) and High (0.5) groups. These findings suggest that the 0.6 to 0.9 range is critical.

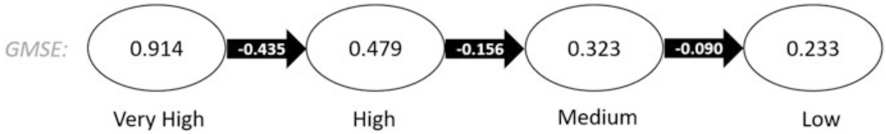


Fig. 8 Efficiency disparity across consecutive 'levels of development'

Political Regime
According to our results, the two most efficient political regime types are Liberal Democracy and Closed Autocracy, while the least efficient are Electoral Autocracy and Electoral Democracy. Thus, the suggestion here is that efficiency benefits from either maximizing or minimizing democracy. As a caveat, it is worth highlighting that the sample size for closed autocracies was relatively small (9 countries), which meant China's SE score could inflate the group's average despite using the geometric mean.

4.4 *Stage 2: Random Forest Classification*

Using the same selection of 124 countries, we trained a random forest comprising 250 decision trees for binary classification. We trained the forest to separate efficient countries from non-efficient countries using the DEA model's inputs and outputs. The forest achieved a cross-validation accuracy of 96.84% ($k = 13$).

After training the forest, we evaluated the relative importance of feature variables on the classification task. Figure 4 shows the calculation results, and Table 7 shows the full names of the feature variables. We found that GCI variables were far more influential in deciding a country's classification as efficient or inefficient, accounting for 91.8% of 'relative importance'. By contrast, IHDI variables only accounted for 8.2%. The single most influential variable was *Innovation Capability,* with a relative importance of 16.7%. *Business Dynamism* (15.7%) and *Institutions* (12.3%) followed. The least influential variables were *Health* (2.2%) and *Market Size* (2.1%).

We also considered the relative importance of the four GCI categories superseding the pillars. Table 8 shows *Innovation Ecosystem* had the highest relative importance with 32.4%, while the *Enabling Environment* was second with 26.8%.

5 Discussion

5.1 *Implications for Policy and Business*

The DEA model tended to reward countries with good GCI *and* IHDI performances; hence, super-efficiency rankings were strongly correlated with IHDI and GCI

Table 7 Key for horizontal labels in Fig. 9

GCI variables		IHDI variables	
INST	Institutions	INQLIFE	Inequality-adjusted life expectancy index
INFR	Infrastructure	INQEDU	Inequality-adjusted education index
ICT	ICT adoption	INQINC	Inequality-adjusted income index
MSTAB	Macroeconomic stability		
HEA	Health		
SKI	Skills (Workforce)		
PROD	Product market		
LAB	Labour market		
FIN	Financial system		
MAR	Market size		
DYN	Business dynamism		
INNOV	Innovation capability		

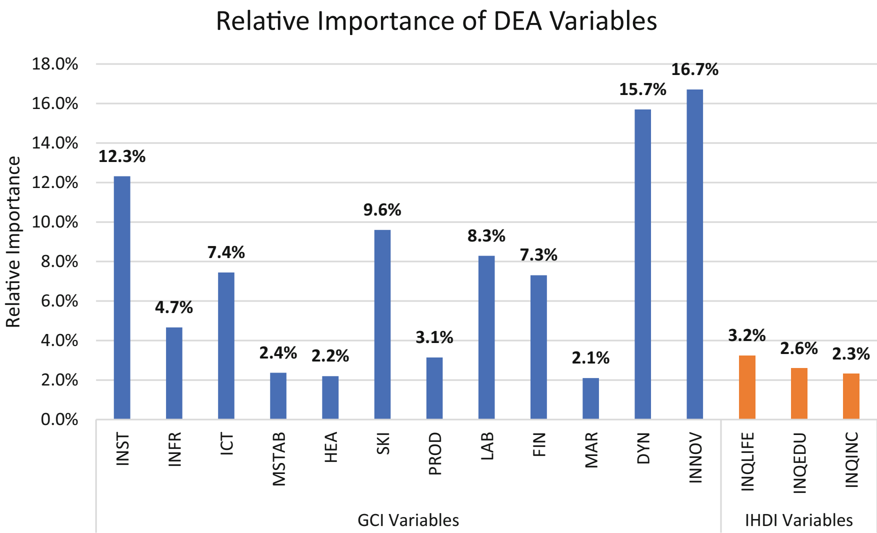


Fig. 9 Relative importance of criteria to DEA classification

Table 8 Relative importance of GCI categories

GCI super-category	Relative importance
Enabling environment	26.8%
Human capital	11.8%
Markets	20.8%
Innovation ecosystem	32.4%

rankings. However, there were exceptions like China, whose DEA performance outshone their IHDI and GCI performance. Overall, the model indicated that some countries are indeed better at leveraging their competitiveness to generate human development than others, reinforcing the findings of previous studies (see: Ülengin et al., 2011; Kılıç & Kabak, 2019, 2020).

Governments—and other parties invested in the betterment of a given country—should learn from the policies and practices of countries with better DEA performances than theirs, particularly those with similar characteristics. DEA models allow for the identification of 'peer units'—units that occupy a similar area of the efficiency frontier as a given unit under consideration (Bogetoft & Otto, 2010, p. 93; Ülengin et al., 2011, Kılıç & Kabak, 2019, 2020). Individuals and institutions may use 'peer units' or their discretion to determine the best candidates for comparison. For example, one might choose to focus on countries with similar GCI or IHDI scores; countries with similar resource endowments (see: Auty, 1998; Sachs, 1999); countries from the same region; or those with similar socio-cultural values, political systems and histories.

On the other hand, businesses could use the DEA model to direct corporate social responsibility (CSR) initiatives and other investments, using a country's efficiency score as a proxy for its potential.

The second part of our methodology identified the relative importance of variables underpinning DEA results. Notably, our results depicted trends *across* countries. Consequently, we cannot comment on how the relative importance of particular variables fluctuates between and within individual countries. Nonetheless, this information could also support decision-makers in government, policymaking and commerce.

The variables with the most significant influence on the DEA results were Innovation Capability and Business Dynamism. The GCI indicators that make up these variables (pillars) are shown in Table 9.

The high relative importance of a country's *Innovation Capability* and *Business Dynamism*—variables affecting the *Innovation Ecosystem*—seems timely given the advent of the fourth industrial revolution. Recent literature gives the impression that the relative importance of a country's innovation ecosystem will increase in the near

Table 9 Business Dynamism and Innovation Capability indicators (Schwab, 2019)

Business dynamism	Innovation capability
Diversity of workforce	Cost of starting a business
State of cluster development	Time to start a business
International co-inventions per million of the population	Insolvency recovery rate
Multi-stakeholder collaboration	Insolvency regulatory framework
Research and Development	Attitudes towards entrepreneurial risk
Scientific Publications Score	Willingness to delegate authority
Patent Applications per million of the population	Growth of innovative companies
Research and Development Expenditure	Companies embracing disruptive ideas
Research Institutions prominence	
Commercialization	
Buyer Sophistication	
Trademark Applications per million population	

future (Schäfer, 2018; Yang et al., 2019). There is an ongoing debate on whether the fourth industrial revolution is relevant to one of the most inefficient regions—Sub-Saharan Africa—given the region's socio-economic characteristics (Ayentimi & Burgess, 2019). However, the prevailing opinion appears to be that the region can benefit if governments direct their activities and resources towards their innovation ecosystems (Amankwah-Amoah et al., 2018; Asongu & Nwachukwu, 2018; Ogwo, 2018; Ayentimi & Burgess, 2019). All-in-all, our results reinforce the notion that countries should pay close attention to the fourth industrial revolution, showing that there are implications for human development.

Our results also emphasized the value of a countries' *Enabling Environment*, and in particular its *Institutions*. Again, this feels timely in the context of the fourth industrial revolution. Institutions are a clear theme in the literature cited above (Schäfer, 2018; Yang et al., 2019); they all press the idea that institutions (especially those from the public sector) will be especially crucial during the period of transformation ahead. Our results reinforce this view. In general, it is important to note that different areas of competitiveness are interdependent, so focusing exclusively those with the highest relative importance may not yield dividends.

Many in the West view democracy as a necessary ingredient for becoming an advanced society but our results challenge this. Although liberal democracies had the highest average DEA performance by far, electoral democracies—the most common political system in our dataset—performed very poorly on average. Although electoral democracies were superior to electoral autocracies, they did not outperform closed autocracies. Our results indicated that each type of political system is capable of producing efficiency.

5.2 Stratified world system

Our findings revealed two distinct classes of country, in terms of DEA performance. In Fig. 5, there was a bimodal distribution of efficiency scores, with two peaks separated by a trough spanning the 0.6–0.9 efficiency range. Figure 3 corroborated this. It showed that, as a country moves from low levels of human development to very high levels of human development, it becomes increasingly difficult to make the associated leaps in efficiency. The largest gap, by far, was between countries with very high human development, with an average DEA score of 0.914, and countries with high human development, which had an average of 0.479. Countries with high human development would have to increase their efficiency by approximately 91% to reach the standard of those with very high human development. Although further observation would help to corroborate this further, IHDI and GCI scores do not fluctuate greatly from year to year so it is likely that this is evidence of a trend rather than an anomaly (Conceição, 2019; Schwab, 2019).

The bimodality of our results throws into question the assumption of homogeneity that underpinned our selection of countries. It is possible that we should have followed the blueprint set by Ülengin et al. (2011), as well as Kılıç and Kabak (2019), by either clustering or filtering countries' pre-selection.

5.3 Limitations

The methodology entailed multiple assumptions and limitations. Firstly, it assumed that the countries under assessment were comparable in terms of internal resources and activities, as well as environmental factors. It also inherited the limitations and assumptions of the GCI and IHDI, having relied on these indices heavily as proxies for competitiveness and human development, respectively. Given the logistical challenge of collecting and verifying data across multiple countries simultaneously, both indices are likely to suffer from observational error. In previous sections, we also discussed how aspects of both indices are subjective, and therefore contestable. Generally, we viewed these flaws as inescapable by-products of decisions involved in the design of multidimensional indices. However, we were particularly concerned about the manner in which the GCI derives data from the Executive Opinion Survey because it produces highly subjective yet outwardly quantitative data. We were also briefly concerned that some of the DEA variables proxied the same or highly correlated phenomena. For example, 'health' is covered on both the input and output side of the model. Fortunately, it is extremely rare for inter-variable correlation to significantly impact DEA scores (Dyson et al., 2001).

6 Conclusions

This paper was predicated on the idea that the objective of a nation's economy competitiveness is to enhance the welfare of its citizens. It explored the relationship between human development, as measured by the IHDI, and competitiveness of nations, as measured by the GCI. In particular, it set out to determine (a) the relative performance of countries in leveraging their competitiveness to produce human development and (b) the relative importance of the factors, which facilitate that process.

The results were bleak. Most countries do not maximize the human-development-producing potential of their economies, and at least half are extremely inefficient, needing to more than double their current performance to become efficient. We found that the most important explanatory factors were related to the innovation ecosystem of the country, a discovery which felt timely given the ongoing buzz around the fourth industrial revolution.

The implications of our work are manifold. However, the main implication, particularly for policy, is that there is no substitute for tracking the connection between competitiveness and human development directly. Measuring human development

(IHDI) and competitiveness (GCI) in isolation is not the same as measuring 'the ability of a country to produce human development from its competitiveness'. In addition, we have also provided a practical tool, in the form of our hybrid DEA-random forest model, for such a task. While our results are deserving of attention and warrant further analysis, it is important to acknowledge the limitations of our methodology, in particular, those stemming from the HDI and GCI. They, like many other composite indices, are doomed to fallibility. This means our methodology is fallible too.

Further research is required to verify the underlying causes of some of the patterns we have observed. For instance, to explain the bimodality discovered in the frequency distribution of DEA/super-efficiency scores. Another area of intrigue concerns the surprising inefficiency of electoral democracies; more detailed analysis regarding the characteristics of political systems that nurture efficiency would be welcomed. Lastly, we recommend further studies of the relationship between countries' innovation ecosystems, their competitiveness and human development outcomes. With the fourth industrial revolution looming large, it is doubly important to understand this process so that as many countries as possible can take advantage of the opportunities to come.

References

Altmann, A., et al. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics*. Oxford University Press, *26*(10), 1340–1347.

Amankwah-Amoah, J., Osabutey, E. L. C., & Egbetokun, A. (2018). Contemporary challenges and opportunities of doing business in Africa: The emerging roles and effects of technologies. *Technological Forecasting and Social Change*. Elsevier, *131*, 171–174.

Andersen, P., & Petersen, N. C. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science*. INFORMS, *39*(10), 1261–1264.

Asongu, S. A., & Nwachukwu, J. C. (2018). Openness, ICT and entrepreneurship in sub-Saharan Africa. In *Information Technology & People*. Emerald Publishing Limited.

Auty, R. M. (1998). *Resource abundance and economic development improving the performance of resource-rich countries*. UNU World Institute for Development Economics Research (UNU/WIDER).

Ayentimi, D. T., & Burgess, J. (2019). Is the fourth industrial revolution relevant to sub-Sahara Africa? *Technology Analysis & Strategic Management*. Taylor & Francis, *31*(6), 641–652.

Banker, R., Charnes, A., & Cooper, W. (1984). Some models for estimating technical and scale effficiencies in data envelopment analysis. *Managment Science, 30*, 1078–1092.

Beleites, C., et al. (2013). Sample size planning for classification models. *Analytica Chimica Acta*. Elsevier, *760*, 25–33.

Bogetoft, P., & Otto, L. (2010). *Benchmarking with dea, sfa, and r*. Springer Science & Business Media.

Breiman, L. (2001). Random forests. *Machine learning*. Springer, *45*(1), 5–32.

Bucher, S. (2018). The global competitiveness index as an indicator of sustainable development. *Herald of the Russian Academy of Sciences, 88*(1), 44–57. https://doi.org/10.1134/S1019331618010082

Cetinguc, B., et al. (2018). The relationships among the prominent indices: HDI-GII-GCI. In *Industrial engineering in the industry 4.0 era* (pp. 187–194). Springer.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Cliff, N. (1996). Answering ordinal questions with ordinal data using ordinal statistics. *Multivariate Behavioral Research*. Taylor & Francis*, 31*(3), 331–350.

Conceição, P. (2019). *Human Development Report 2019: beyond income, beyond averages, beyond today*. United Nations Development Programme.

Cooper, W. W., Seiford, L. M., & Tone, K. (2007). Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software (2nd ed.). Springer.

Coppedge, M., et al. (2019). V-dem dataset V9. *SSRN Electronic Journal*. Elsevier BV. https://doi.org/10.2139/ssrn.3441052

Coppedge, M., Gerring, J., Knutsen, C. H., et al. (2019). V-dem country coding units v9 (V-dem): Varieties of democracy project. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3441067

Coppedge, M., Gerring J., Knutsen, C. H., et al. (2021). V-dem methodology v9: Varieties of democracy (V-dem) project. Available at: https://www.v-dem.net/static/website/img/refs/methodologyv111.pdf

Cutler, D. R., et al. (2007). Random forests for classification in ecology. *Ecology*. Wiley Online Library*, 88*(11), 2783–2792.

Dyson, R. G., et al. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*. Elsevier*, 132*(2), 245–259.

Fukuda, S., et al. (2016). Random forests hydrodynamic flow classification in a vertical slot fishway using a bioinspired artificial lateral line probe. In *International conference on intelligent robotics and applications* (pp. 297–307). Springer.

Fukuda-Parr, S. (2003). The human development paradigm: Operationalizing Sen's ideas on capabilities. *Feminist Economics, 9*(2–3), 301–317. https://doi.org/10.1080/1354570022000077980

García Ochoa, J. J., León Lara, J. d. D., & Nuño de la Parra, J. P. (2017). Propuesta de un modelo de medición de la competitividad mediante análisis factorial. *Contaduria y Administracion, 62*(3), 792–809. https://doi.org/10.1016/j.cya.2017.05.002

Green, R. H. (1996). DIY DEA: Implementing data envelopment analysis in the mathematical programming language AMPL. *Omega*. Elsevier*, 24*(4), 489–494.

Im, T., & Choi, Y. (2018). Rethinking national competitiveness: A critical assessment of governmental capacity measures. *Social Indicators Research*. Springer Netherlands*, 135*(2), 515–532. https://doi.org/10.1007/s11205-016-1494-z

Ketels, D. C. (2016). *Review of competitiveness frameworks* (pp. 1–58). Irish National Competitiveness Council. Available at: http://www.competitiveness.ie/Publications/2016/ReviewofCompetitivenessFrameworks.pdf

Kılıç, H., & Kabak, Ö. (2019). Analysis of relation between human development and competitiveness using fuzzy ANP and DEA. In *International conference on intelligent and fuzzy systems* (pp. 859–866). Springer.

Kılıç, H., & Kabak, Ö. (2020). Analysis of relation between human development and competitiveness using fuzzy ANP and DEA. In C. Kahraman, S. Cebi, & Cevik Onar, Sezi & Oztaysi, Basar & Tolga, A. Cagri & Sari, Irem Ucal (Eds.), *Intelligent and fuzzy techniques in big data analytics and decision making*, (pp. 859–866), Springer.

Kovacevic, M. (2010a). Measurement of inequality in Human Development–A review. *Measurement*, 35.

Kovacevic, M. (2010b). *Review of HDI critiques and potential improvements*. UNDP.

Lebedev, A. V., et al. (2014). Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *NeuroImage: Clinical*. Elsevier*, 6*, 115–125.

Lewis, H. F., & Sexton, T. R. (2004). Data envelopment analysis with reverse inputs and outputs. *Journal of Productivity Analysis*. Springer*, 21*(2), 113–132.

Liu, J. S., et al. (2013a). A survey of DEA applications. *Omega (United Kingdom)*. Elsevier*, 41*(5), 893–902. https://doi.org/10.1016/j.omega.2012.11.004

Liu, M., et al. (2013b). Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. *Sensors and Actuators B: Chemical*. Elsevier*, 177*, 970–980.

Mariano, E. B., Sobreiro, V. A., & Rebelatto, D. A. d. N. (2015). Human development and data envelopment analysis: A structured literature review. *Omega (United Kingdom)*. Elsevier*, 54*, 33–49. https://doi.org/10.1016/j.omega.2015.01.002

Mashabela, J., & Raputsoane, L. (2018). Munich Personal RePEc Archive Important factors in a nations international competitiveness ranking Important factors in a nations international competitiveness ranking, (86477).

Ogwo, B. A. (2018). Re-visioning technical vocational education and training (TVET) for the youth in sub-Saharan Africa (SSA) and the Sustainable Development Goals (SDGs): Prospects and promises within the framework of the ubuntu paradigm. In *Re-Visioning Education in Africa* (pp. 155–173). Springer.

Önsel, Ş., et al. (2008). A new perspective on the competitiveness of nations. *Socio-Economic Planning Sciences, 42*(4), 221–246. https://doi.org/10.1016/j.seps.2007.11.001

Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*. JMLR.org*, 12*, 2825–2830.

Probst, P., Wright, M. N., & Boulesteix, A. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. Wiley Online Library*, 9*(3), e1301.

Riley, R. D., Ensor, J., Snell, K. I. E., Harrell, F. E., Martin, G. P., Reitsma, J. B. et al. (2020). Calculating the sample size required for developing a clinical prediction model, *BMJ 368*, m441. https://www.bmj.com/content/368/bmj.m441

Roser, M. (2013). *Democracy*, *our world in data*. Available at: https://ourworldindata.org/democracy. Accessed 21 September 2020.

Ruggiero, J. (1998). Non-discretionary inputs in data envelopment analysis. *European Journal of Operational Research*. Elsevier*, 111*(3), 461–469.

Sachs, J. D. (1999). Resource endowments and the real exchange rate: a comparison of Latin America and East Asia. In *Changes in exchange rates in rapidly developing countries: Theory, practice, and policy issues* (pp. 133–154). University of Chicago Press.

Sarkis, J. (2007). Preparing your data for DEA. In *Modeling data irregularities and structural complexities in data envelopment analysis* (pp. 305–320). Springer.

Schäfer, M. (2018). The fourth industrial revolution: How the EU can lead it. *European View*. SAGE Publications*, 17*(1), 5–12. https://doi.org/10.1177/1781685818762890

Schwab, K. (2019). *The Global Competiitiveness Report 2019*. World Economic Forum. Available at: http://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2019.pdf

Sen, A. (2001). *Development as freedom*. Oxford Paperbacks.

Seth, S., & Santos, M. E. (2018). *Multidimensional inequality and human development, OPHI working paper 114*. University of Oxford. https://doi.org/10.2139/ssrn.3113294

Stanton, E. A. (2007). Human development index: A history. *Political Economy Research Institute Working Paper Series, 127*(February), 1–36.

Ülengin, F., et al. (2011). The competitiveness of nations and implications for human development. *Socio-Economic Planning Sciences*. Elsevier Ltd*, 45*(1), 16–27. https://doi.org/10.1016/j.seps.2010.10.001

United Nations Human Development Report Office. (n.d.) *What is human development? | Human development reports*. Available at: http://hdr.undp.org/en/content/what-human-development. Accessed 6 July 2020.

Yadav, S., & Shukla, S. (2016). Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In *2016 IEEE 6th international conference on advanced computing (IACC)* (pp. 78–83). IEEE.

Yang, H., Kim, S. Y., & Yim, S. (2019). A case study of the Korean government's preparation for the fourth industrial revolution: public program to support business model innovation. *Journal of Open Innovation: Technology, Market, and Complexity*. Multidisciplinary Digital Publishing Institute*, 5*(2), 35.

Communicating Value in Healthcare Marketing from a Social Media Perspective

Sandy Çağlıyor, Petek Tosun, and Nimet Uray

Abstract Sustainable healthcare policies and a developed healthcare industry are vital to countries' competitiveness and productivity. The ongoing transformations in healthcare services and advances in health technologies and analytics make it clear that there is a pressing need for more collaborative and interdisciplinary efforts in the industry. This study aims to explore the effectiveness of online marketing communication for healthcare services in Turkey with regard to the value-driven marketing approach utilized by leading chain hospitals through an examination of two research questions: (1) *Which messages are emphasized in the social media marketing communications of hospitals?* (2) *Which factors increase engagement with healthcare consumers on social media?* To that end, we compiled the Facebook and Twitter posts of three of the largest hospital chains in Turkey for the last 5 years along with the interaction metrics of the posts, ultimately generating a dataset consisting of 9212 posts in total. Using Latent Dirichlet Allocation, we identified four main topics: Posts on holidays and special days/weeks promoting healthy lifestyles, informative posts about the symptoms and treatments of illnesses, posts containing statistics about diseases, and posts including news about the hospital in question. In the following stage, we carried out predictive analysis using three tree-based machine learning algorithms (decision trees, random forests, and gradient boosting trees) to predict total interaction and relative variable importance. Our model performed at an accuracy rate of 70%. The findings of this study indicate that contextual factors such as the number of followers may have more predictive power than content or interactivity factors. Hospitals use social media to improve their brand reputation and increase public awareness about health and critical diseases. The posts about holidays and special days and using links in the posts resulted in the most interaction. Message source was identified as an important factor, so different social media platforms should be treated as separate mediums in the design of marketing communication strategies and the different dynamics

S. Çağlıyor (✉) · P. Tosun · N. Uray
Kadir Has University, Istanbul, Turkey
e-mail: sendi.cagliyor@khas.edu.tr; petek.tosun@khas.edu.tr; nimet.uray@khas.edu.tr

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_6

of those platforms should be considered instead of posting the same content on various platforms. As such, this research has valuable implications for marketing managers and administrators working in healthcare in terms of the design of their online marketing communication strategies.

Keywords Healthcare marketing · Marketing communication · Social media · Latent Dirichlet allocation · Random forest · Gradient boosting tree · Decision tree

1 Introduction

Social media marketing is becoming increasingly prevalent in business management, particularly in the healthcare sector. As more and more consumers seek out health-related information and advice on social media (Koumpouros et al., 2015), in the last decade hospitals have increasingly turned to social media as a means of communicating with their patients and consumers (Sharma & Gupta, 2019). Through their social media accounts, healthcare institutions promote their brands, share information about well-being, and highlight the advantages of the treatments they provide (Radu et al., 2017). Compared to other mediums of communication, social media marketing provides a number of additional advantages, including the ability to engage with certain communities (Shawky et al., 2019). Moreover, social media makes it easy for consumers to obtain information about hospitals, doctors, and the experiences of other patients (Hackworth & Kunz, 2011), and when consumers have doubts about the trustworthiness of public news sources for specific topics, they tend to rely more on information that is shared on social media (Jang & Baek, 2019).

Advances in digital communications have created a need for observational research that deepens our understanding of the impacts and place of social media in the healthcare sector (Schillinger et al., 2020). The healthcare sector functions as a value network consisting of hospitals, medical laboratories, pharmaceutical companies, medical equipment vendors, and health insurance companies (Sharma & Gupta, 2019). In such an environment, social media facilitates the building of relationships with patients, the transmission of accurate information, and the acquisition of immediate consumer feedback, especially for hospitals (Hackworth & Kunz, 2011). Increased digitalization has driven people to spend more time on social media, which in turn makes it possible for healthcare brands to create online consumer engagement (Haenlein et al., 2020). Social media communication is an effective way for hospitals to act in a socially responsible manner and guide people towards healthier patterns of behavior.

When it comes to information concerning health issues, people tend to think that messages which are conveyed by institutions and include factual information are more believable than those communicated by their peers (van der Meer & Jin, 2020). One of the core responsibilities of hospitals is to provide high-quality healthcare to patients, and that can be taken up in a broader sense to include the dissemination of health-related information that is both correct and proper for the situation at hand. And in light of current developments, social media communications can also provide psychological support to frontline healthcare workers and the community at large (Cheng et al., 2020).

Within that context, this study focuses on the social media communications of hospitals from the perspective of consumer engagement and examines the value that is provided by marketing communication content. Agarwal et al. (2020) have defined value-centered marketing in the healthcare sector as a set of processes that are employed in the delivery and communication of value associated with well-being and treatment. Hospitals interact with their patients and consumers on their social media accounts by sharing content about a wide variety of issues, including disease avoidance and maintaining health. Marketing communication content generates value by associating it with specific outcomes and messages. In line with that framework, the research questions of this study are as follows: (1) *Which messages are emphasized in the social media marketing communications of healthcare providers and/or hospitals*? (2) *Which factors increase engagement with healthcare consumers*? In order to explore those questions, we examined the social media accounts of the three largest hospital chains in Turkey by way of machine learning algorithms. Through the use of web scrapers, we compiled the Facebook and Twitter posts of those hospitals for the last 5 years, and then we employed latent Dirichlet allocation (LDA) for the purposes of topic detection and tree-based machine learning algorithms to carry out our predictive analyses, which will be further discussed in the methodology section. First, however, we provide a summary of the literature on the issue.

The study will continue with the literature review section that includes a synthesis of social media marketing in the healthcare sector and consumer engagement on social media and a brief overview of the Turkish healthcare sector. After the explanation of the research methodology, the findings will be presented. The chapter will conclude with the discussion of findings and conclusion sections.

2 Literature Review

2.1 Social Media Marketing in the Healthcare Sector

In the last decade, hospitals have been increasingly engaging with consumers via their social media accounts, and as a consequence, social media marketing is on the rise in the healthcare sector (Sharma & Gupta, 2019). Making use of an effective marketing strategy that utilizes social media platforms is essential for hospitals to acquire a competitive edge in the healthcare market (Hackworth & Kunz, 2011). As noted by Popović and Smith (Popović 2010), the one-way pattern of communication that predominated in health-related television shows in the 1960s and 1970s, in which inaccessible doctors discussed illnesses, gave way to the medical talk shows of the 2000s that focused on doctor–patient dialogues. Undoubtedly, the underlying reason for the popularity of such two-way interactions on television was steeped in a need for timely, accurate information about health issues. As early as 2011, the US-based *Mayo Clinic* included contact information, videos of doctors explaining issues to patients, discussions raised through the posing of questions, and invitations to consumers to share their comments on Facebook, and all the while *Scripps Health* was communicating daily via Twitter to inform people about health issues (Hackworth & Kunz, 2011).

Shawky et al. (2019) have pointed out that social media has proven to be effective in reaching audiences and developing communications with consumers, and in that way, it has contributed to the communication objectives of organizations. Healthcare service providers use social media to communicate with patients, promote their brands, and attract human resources (Bejtkovsky, 2020). Furthermore, healthcare institutions use social media for advertising, sharing informative content about well-being, and posting messages about the advantages of certain treatments (Radu et al., 2017). Videos, photos, widgets, blogs, and personalized messages are just some of the elements of the rich marketing communication environment that social media platforms provide to marketers and consumers (Popović & Smith, 2010).

Social media marketing makes it possible for physicians and healthcare institutions to satisfy their patients' needs, as indicated by the fact that more than 75% of consumers seek out health-related information on the web (Koumpouros et al., 2015). It has also been found that 62% and 53% of consumers search the internet for health advice and specific products, respectively (Koumpouros et al., 2015). In the past, marketers merely had to increase message frequency to drive up brand awareness, but in today's connected digital world where social media is increasingly used by patients, communities, and other user groups, it is not enough to simply push messages through targeted channels (Popović & Smith, 2010). Hackworth and Kunz (2011) have suggested that healthcare marketers need to share information about their brands, interact with consumers, monitor patient feedback, and regularly keep up with online conversations about their brands. Moreover, it has been found that satisfaction with health-related information obtained online is positively correlated with patient loyalty, and patient loyalty in turn is positively correlated with purchase intentions (Sharma & Gupta, 2019). Given that situation, social media has become a crucial medium for both retaining existing patients and acquiring new ones as well.

On the other hand, studies have indicated that some doctors and healthcare managers are not comfortable with using social media marketing and still place their trust in conventional marketing channels (Koumpouros et al., 2015). In their examination of the social media marketing activities of healthcare institutions in Poland, Gregor and Gotwald (2013) found that only 18% of healthcare institutions made use of social media, and they traced that low rate of usage led to managerial mistrust of social media marketing (71%), a lack of skilled staff (22%), and managers' assumption that patients would simply not be interested (19%). However, hospitals can provide a major service in terms of improving public health by sharing health-related content and information on their social media accounts, and as such, they not only act as providers of health services but also socially responsible institutions. The corporate social responsibility of hospitals vis-à-vis patients and society has the potential to improve brand advocacy, brand trust, patient-hospital identification, and the positive word-of-mouth intentions of patients (Limbu et al., 2020). Moreover, as Agarwal et al. (2020) have pointed out, patients are increasingly becoming empowered consumers of healthcare services in a value-delivery framework that consists of three dimensions: precision in treatment, the preferences of consumers, and customer-centric processes. In today's highly digitalized world, social media needs to become an embedded part of patient acquisition as well as processes of informing, retaining, and engaging consumers of healthcare.

Consumer and patient engagement on social media make it possible for patients to not only be well-equipped with information but also empowered (Hewitt, 2011). The social media campaigns of hospitals, including those that encourage a healthy diet and exercise as a way to avoid obesity and diabetes, can be taken up as social marketing actions that attempt to influence people's behavior for the benefit of society (Kotler & Zaltman, 1971; Hewitt, 2011). Similarly, Gregor and Gotwald (2013) have noted that healthcare institutions use social media, in particular Facebook, to increase public awareness and prevent diseases. The engagement of hospitals with existing and potential clients on social media has become one element of social marketing that aims to improve public health primarily through the dissemination of information. However, while it is true that the use of social media does offer such benefits, healthcare organizations must also consider legal, medical, and regulatory issues when they utilize social media as a part of integrated marketing communications (Popović & Smith, 2020).

2.2 Consumer Engagement with Social Media in Health Services

Healthcare institutions are increasingly making use of social media to communicate and interact with consumers who have become more conscious about issues such as well-being, critical diseases, and medical diagnostics (Sharma & Gupta, 2019). Since consumers have the option of sharing their opinions in the form of online comments or reviews regarding products and services, social media offers a rich and interactive communication environment. Although websites are the main source of information for health-related topics, consumers also search social media platforms such as YouTube and Facebook to access information about their health-related concerns (Koumpouros et al., 2015; Cangelosi et al., 2019). Studies have shown that Facebook is the most commonly used social media platform for sharing content-based information and raising awareness about social marketing programs that aim to improve public health (Gregor & Gotwald, 2013; Shawky et al., 2019). It has been found that single young adults and people whose employers provide health insurance benefits tend to think that social media is a crucial source of information about healthcare (Cangelosi et al., 2019), and the number of followers that a given page has on social media is often considered to be a primary indicator of how brands interact with consumers (Wang & Jin, 2010).

Consumer comments on social media can be a valuable source of feedback for managers. For example, Chatterjee et al. (2020) have used text-mining to analyze consumer reviews regarding health products and healthcare e-commerce, and they demonstrated that an e-commerce company's service quality and the perceived quality of the physical and online facilities that are integrated into an omnichannel model play a critical role in terms of customer satisfaction in fitness and nutrition services. Moreover, consumers can act as service providers and beneficiaries when they share information with online health communities (Stadtelmann et al., 2019) because their opinions, advice, and comments influence others. A survey conducted about the new patients of a dental clinic revealed that 79% of them were influenced by Internet content in their selection of the clinic, as indicated by the fact that Facebook and Google ads were major sources of information for them (Radu et al., 2022).

Engaging with existing and potential patients on social media platforms contributes to brand equity for healthcare institutions. Open and sincere communication with patients on social media can contribute to the development of brand trust, as perceptions of transparency suggest that the hospital is responsible for its delivery of services (Limbu et al., 2020). Moreover, many consumers find it difficult to ascertain the medical accuracy of the content on websites (Popović & Smith, 2010). When prestigious healthcare institutions share information, answer questions, and engage with people on social media, they are in a better position to build bonds with consumers. In turn, if consumers are satisfied with the richness of information provided by healthcare institutions in online environments, they are more likely to spread positive word-of-mouth (Sharma & Gupta, 2019).

Social media can have both positive and negative impacts on health-related attitudes, behaviors, and norms in society (Schillinger et al., 2020). El-Awaisi et al. (2020) have examined social media posts to explore public perceptions of healthcare sector employees and they pointed out that social media content can generate a wide variety of emotions such as frustration, relief, and gratitude. Patients' levels of involvement in online information about health are significantly correlated with the perceived quality of that content, and people tend to be more selective regarding online content that is posted by healthcare institutions (Sharma & Gupta, 2019). In that context, social media has the potential to influence public health within the scope of a continuum of factors, including disease control, treatment, and the dissemination of disinformation (Schillinger et al., 2020).

2.3 The Healthcare Sector in Turkey

Turkey is a developing country with a population of 81.4 million people and a GDP of $28,423 per capita (OECD, 2021a). For the year 2019, the proportion of health expenditures in relation to GDP was 4.40% and 8.8% in Turkey and the average of OECD countries, respectively (OECD, 2021a). The rates of population growth and health expenditures in Turkey for the 2007–2019 period are shown in Fig. 1.

In parallel with the country's population growth, the number of doctors, nurses, and other healthcare employees has also been increasing in Turkey, the figures for which are shown in Fig. 2. The number of physicians per 100,000 patients has increased from 138 in 2002 to 193 in 2019 (Saglik, 2020). The number of doctors per 1000 people in Turkey was 1.9 in 2017; in comparison, those figures were 0.8 in India, 2.0 in China, 2.4 in Mexico, 3.4 in France, 4.0 in Italy, and 4.3 in Germany (OECD, 2021c).

In 2019, there were 1538 inpatient medical care institutions in Turkey. The distribution of public and private institutions is shown in Fig. 3. The share of private hospitals in the health sector increased from 29% in 2007 to 37% in 2019, and chain hospitals contributed much to that transformation in the healthcare sector.

According to statistics provided by the Turkish Ministry of Health, more than 70% of Turkish citizens have said that they are satisfied with the health services they have received, and the average satisfaction rate was 72.59% for the 2010–2019 period (Saglik, 2020). Although the majority of patients have expressed satisfaction with the health services they have received, the Turkish healthcare sector nonetheless still has great potential to grow and improve.

Within that context, the current study aims to improve our understanding of how hospitals use social media as a marketing communication tool. It utilizes a unique research framework and methodology to reveal the main reasons for social media marketing communication in the Turkish healthcare sector. Moreover, this study will explore online consumer engagement with hospitals and provide

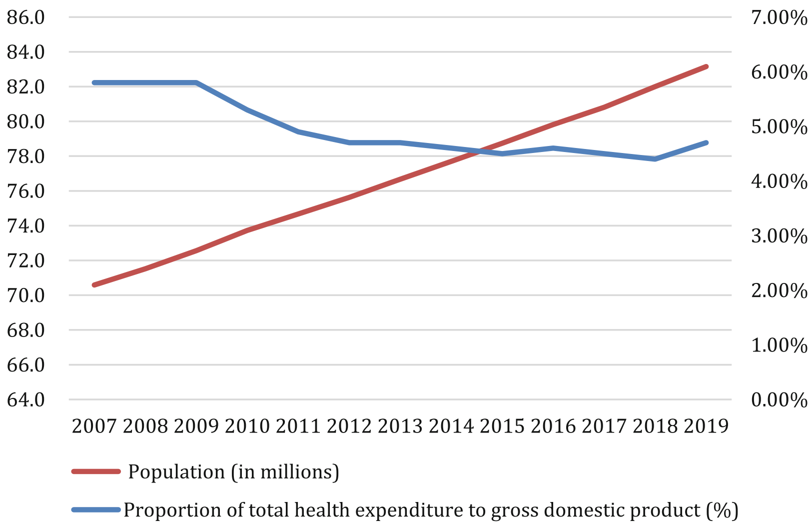


Fig. 1 Population growth and health expenditures in Turkey. (Source: www.saglik.gov.tr)

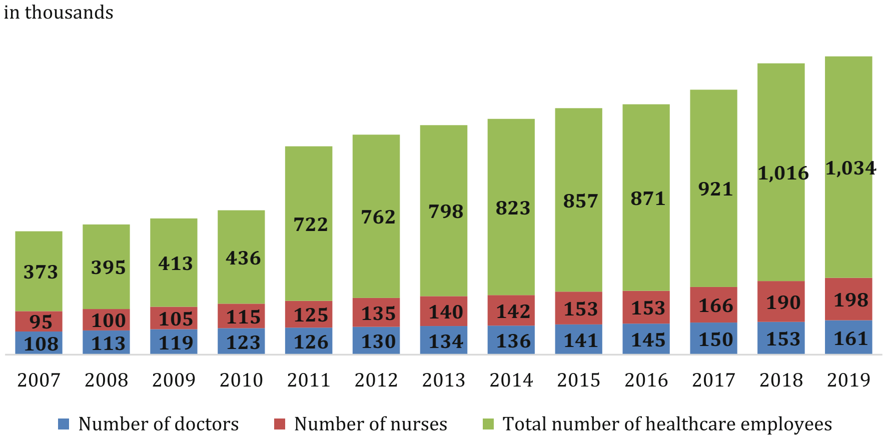


Fig. 2 The number of healthcare sector employees for the years 2007–2019. (Source: www.tuik.gov.tr)

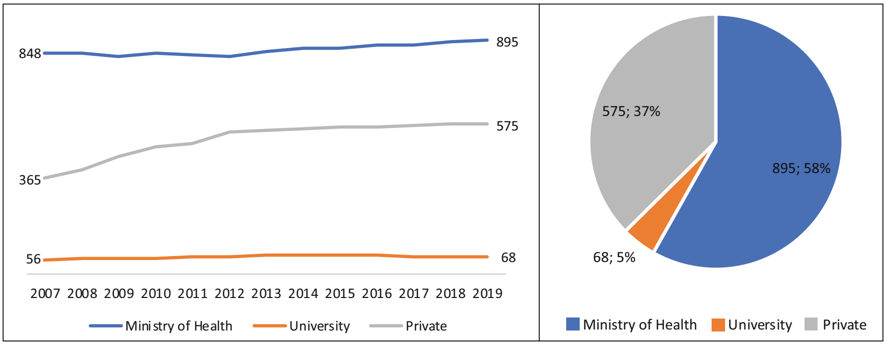


Fig. 3 The total number of inpatient medical institutions in Turkey. (Source: www.tuik.gov.tr)

valuable findings for marketing practitioners. This chapter aims to contribute to the literature by focusing on a relatively less studied topic, social media marketing in the healthcare sector. Furthermore, it bridges a gap between data science and marketing by applying machine learning methodology and interpreting the findings from a social science perspective. The next section will summarize the research methodology.

3 Methodology

This study aimed to explore the effectiveness of online marketing communication for healthcare services in Turkey. In accordance with our aims, this study is based on a descriptive research design. We examined the social media accounts of the three largest hospital chains by way of machine learning algorithms. As illustrated in Fig. 4, the methodology is based on a sequence of successive steps. Although they will be explained in more detail in the following sections, to mention briefly these steps can be summarized as data collection, topic detection, data transformation, and prediction analysis. The first step, in which posts made by Turkey's three largest hospital chains (Acıbadem Healthcare Group, Medical Park Hospital Group, and Medicana Healthcare Group) in the last 5 years were extracted from Facebook and Twitter with the help of two web scrapers along with the interaction metrics of the posts, viz. number of likes, number of shares or retweets, and number of comments, as well as metadata such as video duration, number of video views, and dates.

After cleaning and removing duplicates, our final dataset consisted of 9212 posts, and following preprocessing, topic modeling was applied to the textual data. In the data transformation process, we computed the topics we obtained as new categorical variables, and other metrics such as total interaction, word count, and the presence/lack of a hashtag, link, or question were determined, whereupon the total interaction score was discretized into four classes. By employing three machine learning algorithms—a random forest, gradient boosting tree, and decision tree—we carried out prediction analyses. After the predictive performance of the models' accuracy was evaluated, we examined the importance of the random forest variable with statistical attributes. Topic detection modeling and all predictive analyses were carried out on Knime 4.1.0.

The available contextual factors, content, and interactivity factors were used in the analysis in a manner that is consistent with the literature (Schultz, 2017). In order to represent the contextual factors, we employed a categorical variable indicating the hospitals and a continuous variable indicating the number of followers, since the former provides information about the social media platforms on which posts were made and the latter is considered to be a primary factor that impacts total interaction on brand accounts (Wang & Jin, 2010). In this study, we used the topics generated by LDA as content-related factors and interactivity factors; interactivity factors were coded as three binary variables indicating whether the post contained a hashtag, a link, or a question.

3.1 Data Collection and Formation of the Dataset

Social media and other platforms based on user-generated content have proven to be a great resource for researchers in many disciplines, as there is a plethora of publicly available data about human behavior (Fiesler et al., 2020). Data from social media platforms can be extracted manually (Aydin, 2020) or automatically with a help of a web scraper (Dewi et al., 2019). The process of automatically collecting data from social media and other websites and transforming unstructured data into a structured format is known as web scraping, as well as web harvesting and web data extraction, and the tools used for those purposes are referred to as web crawlers,

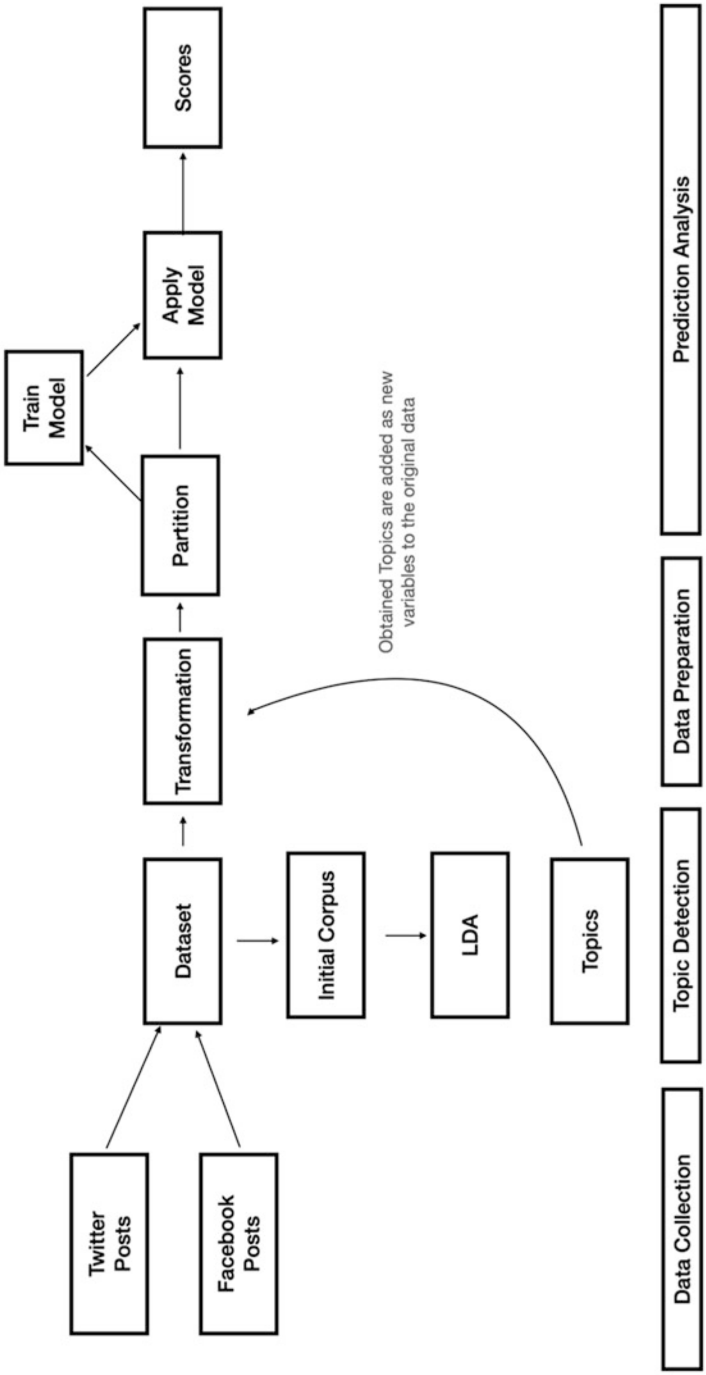


Fig. 4 The workflow

scrapers, robots, bots, and spiders. While those terms are not exactly synonyms, however, they are often used interchangeably since they denote tasks that are quite similar (Algiriyage et al., 2013). While automated scrapers drastically reduce the amount of time needed for data retrieval, they can have certain limitations, as they may require constant adjusting so they can handle the differing formats of websites and they may necessitate a certain amount of technical know-how, especially when websites have complex structures; furthermore, social media platforms may impose restrictions to prevent the collection of data or limit the amount of data that can be gathered by automated means (Batrinca & Treleaven, 2014). Nonetheless, while such limitations prevented us from accessing all the messages posted by hospital chains in Turkey for the time period 2006–2020, we were able to obtain enough data to construct a sufficiently large dataset for our analyses.

After carrying out a preliminary examination of the social media usage of the hospital chains in question, we decided to focus our efforts on postings made via Facebook and Twitter. While some of the hospitals also had Instagram and LinkedIn accounts, there were either fewer messages posted on those accounts or they paralleled postings that were made on Facebook. We used two web scraping tools, Octoparse and Outwit Hub, for the extraction of data from the official Twitter and Facebook pages of the Acıbadem Healthcare Group (@acibademsaglik), Medical Park Hospital Groups (@MedicalParkHG), and Medicana Healthcare Group (@medicanasaglik). Octoparse is a machine learning-based visual tool with a point-click interface that can run extractions on a cloud with multiple servers or a local device (Octoparse, 2021), and Outwit Hub is also a web scraping tool but it works by breaking down the elements on a webpage into their different constituents (Outwithub, 2021). In order to extract the Twitter posts of the accounts, we created yearly filters through the use of advanced search functions for each account and scraped them individually with Octoparse. Because Twitter sets limitations for the collection of past data, we utilized a yearly filter to create a more homogeneous dataset. For each account, we extracted text, timestamps, numbers of likes, numbers of comments, numbers of retweets, and the image links of tweets, and timestamps were converted into dates. For Facebook, once we had obtained the URLs of the posts, their contents were scraped with the use of Outwit Hub. As an extension to the "like" button, Facebook also offers a reaction feature (in the form of a heart, sad face, angry face, etc.), which we also took into account. For those posts we collected text, numbers of reactions, numbers of likes, numbers of shares, dates, and comments. When posts included videos, we were also able to obtain the number of views and duration of the videos, but instead of an exact date or time stamp, the scraper only harvested the year of the post. After the results were compiled and duplicates were removed, the overall data set consisted of 9212 posts, of which only 742 contained visual but not textual content. In total, we were able to obtain 6147 Facebook posts and 3065 Twitter posts for the three hospitals. Tables indicating the distribution and descriptive statistics of the scraped posts can be found below. Table 1 presents a summary of the descriptive statistics for the number of likes, retweets/shares, and comments for our whole dataset, and Table 2 shows the composition of the posts in terms of the social media platform used and the hospital that posted them, along with the number of followers for each one. Lastly, Table 3 illustrates the distribution of the posts with regard to the number of posts by year, which hospital made them, and the social media platform that was used.

3.2 Topic Detection

In order to determine what kinds of messages were emphasized in the social media marketing communications of the hospitals in question, we employed LDA Topic Detection modeling for the text corpora we obtained—in other words, the textual content of the posts. LDA is a tri-level hierarchical Bayesian model that is frequently

Table 1 Summary statistics for the social media metrics in this study

	Min.	Max.	Mean	Median	Std. Deviation
Likes	0	18, 823	174.32	0	16.51
Shares/retweets	0	1400	16.27	2	58.49
Comments	0	746	2.26	21	574.50

Table 2 Number of followers and number of posts

	# of Followers: Twitter	# of Followers: Facebook	# of Posts: Twitter	# of Posts: Facebook
Acıbadem	28, 263	908,489	1544	2370
Medical Park	14, 483	219,548	854	1965
Medicana	8884	220,635	676	1805

Table 3 Distribution of posts (by hospital per year)

	Hospital	Facebook	Twitter	Total	
2016	Acıbadem	504	407	911	2212
	Medical Park	665	292	957	
	Medicana	220	124	344	
2017	Acıbadem	567	435	1002	2051
	Medical Park	559	187	746	
	Medicana	189	114	303	
2018	Acıbadem	536	344	880	2002
	Medical Park	174	134	308	
	Medicana	646	168	814	
2019	Acıbadem	305	122	427	1353
	Medical Park	236	126	362	
	Medicana	424	140	564	
2020	Acıbadem	473	236	708	1594
	Medical Park	329	106	435	
	Medicana	321	130	451	
Total		6148	3065		9212

used in topic modeling because of its ability to identify and describe latent thematic structures in compiled text documents (Blei et al., 2003). The input for LDA is a textual corpus in which each document is considered as a collection of topics and each word in the document corresponds to one of those topics (Lash & Zhao, 2016). LDA clusters each topic with a set of words that best describes the group in which each item is modeled as a finite mixture over an underlying set of topics.

The initial step in topic modeling is data cleaning and preprocessing. In our study, we started that process by removing links from the corpus and then eliminating terms that did not contain content. We followed the usual preprocessing steps such as stop word filtering, punctuation erasing, case conversion, number filtering, RegEX filtering, stemming, and lemmatization through the use of Knime's related nodes. The ZemberekNLP (Natural Language Processing) Library was used for stemming and POS tagging. The data cleaning process can be difficult for user-generated comments because of typos, spelling mistakes, and the usage of "social media jargon," but in our case the cleaning and preprocessing was fairly straightforward since the corpus was obtained from the posts of official accounts, which tend to be more standard in terms of grammar and spelling. All the same, it wasn't enough to just preprocess the nodes, as some manual interventions were also needed, which we carried out with a dictionary filter node. We also found 147 reply posts in our dataset, which we excluded.

In the LDA model, the number of topics (K) and the values for parameters such as document-topic density (α) and topic word density (β) are not defined beforehand. Therefore, as suggested in the literature, we tested for different K values and different combinations of α and β so that we could choose the best model in terms of the LDA evaluation, and the outcomes were evaluated on the basis of their significant differences and interpretability (Maier et al., 2018). Among all the combinations, the most representative model was obtained when $K = 4$, $\alpha = 0.03$, and $\beta = 0.01$.

3.3 Prediction Analysis

Due to their ability to process high-dimensional and complex data and reveal hidden patterns (Wuest et al., 2016), machine learning algorithms are increasingly being used in prediction models. In studies that utilize machine learning algorithms, several different algorithms are usually used instead of a single algorithm and the best-performing model is chosen. Since the performance of models may vary as a result of the structures of the problem, variables, and the dataset, there is no across the board formula for determining which model should be selected (Delen & Sharda, 2009; Hur et al., 2016). Tree-based models have proven to be a potent solution for machine learning problems in various fields (Asadi et al., 2014). In this study, three tree-based machine learning algorithms—a decision tree (DT), random forest (RF), and gradient boosting tree (GBT)—were employed for the predictive analysis. The partitioning ratio of the training set to test ratio was fixed at 70:30 for all the algorithms and stratified sampling was adopted. In that way, a stratification variable target value was selected.

The first algorithm used in the model was a DT, which is a supervised machine learning algorithm that consists of a root node, branches, and leaf nodes, and while each node represents an attribute, each branch represents a decision (Naik & Samant, 2016). The ability of such algorithms to handle both continuous and categorical data, as well as their insensitivity to outliers and missing values (Breiman Breiman et al., 1984) and their utility in providing insights about variable importance, make DTs ideal for such cases (Tso & Yau, 2007). Despite some disadvantages such as overfitting problem, repetition problem, or fragmentation problem, decision tree is still a highly preferred algorithm due to its speed and simplicity (Patel & Rana, 2014). There are different quality measures for decision trees. In this study, we chose the Gini Index, also known as Gini Impurity, as the quality measure (Kingsford & Salzberg, 2008). The minimum number of records per node was set at two and a reduced error pruning option was selected.

The second algorithm we used was a random forest, which is an ensemble model consisting of simple decision trees that can generate a response to predictor values. Random forest algorithms primarily use bagging, by means of which each model runs individually, and in our study the result was calculated by the mean of each individual output. Unlike decision trees, random forests do not learn from a single tree but from a number of different simple trees, and that reduces the risk of learning from noise (Breiman, 2001). In this way, it allows to obtain more accurate results especially for noisy data, and avoids the problem of overfitting, but it may be slower and more difficult to interpret because of higher number of trees (Reis et al., 2018). In this study, information gain ratio was chosen as a split criterion for the tree option, and we did not interfere with the tree depth and minimum node size. As for the forest option, 100 models were selected.

The third algorithm, a gradient boosting tree, is also an ensemble model, but unlike bagging it uses boosting to achieve a strong learner (Zhang & Haghani, 2015). Like bagging algorithms, boosting trees also consist of simple trees, but unlike bagging, in which each model runs individually, boosting trees compute a sequence of trees, and each tree learns from the preceding one. Like random forest, GBDT can better prevent the possibility of overfitting, and also has superior robustness compared to decision tree because its performance is less likely to be affected by outliers and irrelevant features (Cui et al., 2018). In terms of tree options, the tree depth was set to 4, and in terms of boosting options, the number of the model was set to 100 and the learning rate was set to 0.1.

Several metrics can be used to measure the predictive performance of a machine learning algorithm. In this study, the average percentage hit rate (APHR), which is the ratio of true positives to the total number of samples, was used to evaluate performance.

$$\mathrm{APHR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{1}$$

Although machine learning algorithms often have higher predictive power than linear models, they may fall short with regard to explaining the effect of independent variables on dependent variables. Nevertheless, even if they do not fully describe the impact of independent variables on dependent variables where the direction is unclear, the contribution of an individual variable to the predictive power of the model—in other words, the relative variable importance—can be calculated by way of additional scruties, such as sensitivity analysis (Sharda & Delen, 2006). On the other hand, unlike other popular machine learning algorithms like support vector machines and neural networks, tree-based methods can establish a relationship between dependent and independent variables without requiring additional analysis (Zhang & Haghani, 2015). In Knime, a Random Forest Learner node includes an output showing how often a single feature is used at different levels of the tree. Frequent selection as the best split indicates that the variable has a strong informative effect, and hence the importance of the variables in the model can be compared relatively in terms of the contributions they make to the predictive power of the model by dividing the split number by the number of candidates at its level and adding the values.

$$\text{Variable importance score} = \Sigma \frac{\text{Splits}_{\text{Level}_\text{i}}}{\text{Candidates}_{\text{Level}_\text{i}}} \tag{2}$$

3.4 Variables

Online consumer engagement and interaction are considered important metrics as they not only indicate the response of a given customer to a specific issue, but also provide clues about how much the consumer wants to interact with the brand as well as the customer's feelings about and attitude towards the brand (Demmers et al., 2020). There is an extensive body of literature dedicated to improving our understandings of the antecedents and consequences of customer engagement on social media networks. In order to capture customer engagement or interaction on social media platforms, most of these studies use behavioral metrics such as the number of likes, comments, and shares generated in response to posts made by brands. Those metrics can be evaluated separately (Fiok et al., 2020) or combined into a single metric (Liu et al., 2019). Alternatively, they can also be transformed into an engagement rate by dividing the total number of interactions by the number of followers and generating a percentage.

As regards target value, we tested both customer engagement and total interactions. As is frequently done in machine learning studies, we approached the subject from a classification point of view in which the target value is discretized into classes. Since a wide range and the distribution of values complicates discretization, we first needed to compute the normalization. Even though we carried out several trials of different normalization and binning techniques, we were unable to effectively discretize customer engagement. We tried using classes with different ranges, but in every attempt, all the algorithms yielded low-quality results, so we decided to go with total interaction as the target value. By computing the log of values, a wide range can be compressed to a narrow range. After several trials with different binning techniques and class numbers, we found that 3 and 4 were the most favorable number of classes. Posts that received a total number of likes, shares, or comments in the range of 0–10 were designated as class C, those with 11–100 were designated as class B, and those that received total interaction higher than 101 were designated as class A for the analysis with 3 classes. For the analysis with 4 classes, the range was set as follows: 0–10, 11–100, 101–1000, >1000. Since machine learning algorithms can handle outliers, we did not intervene in them. Furthermore, at this point, we should clarify that the correlation between the number of likes and the total interaction score does not affect the performance of the model, in which there are no assumptions associated with them, unlike in the case of linear regression models. However, it is to be expected that the number of followers will be the best split in most of the sub-sampling.

As stated earlier, the available contextual, content, and interactivity factors constituted the dependent variables in the prediction analysis. The number of followers, the hospital's name, and the social media platforms utilized were taken up as contextual factors, and the topics generated by the LDA topic modeling were used as content-related variables. Details about the results yielded by the LDA will be presented in full in the findings section, but it will be useful to discuss them briefly here to give an idea about the classes of the topics. We identified four topics in total: greetings on special days, healthy ways of living, information about serious diseases, and messages promoting videos or events in which doctors provide information about particular medical conditions. For posts that did not contain textual data, we appended a dummy variable. The majority of the non-textual posts consisted of photographs of symposiums or events organized by hospitals, and the remainder were health-related posts in which text was superimposed over the photo, and captions were not used.

The impacts of interactivity factors such as links, questions, polls, hashtags about customer engagement, and total interaction have been investigated at length. Although there are variations in terms of the metrics used, it has been found that platforms, the industry in which the brand operates, and the use of interactive elements enhance customer interaction and engagement (Schultz, 2017). For example, Kujur and Singh (2016) have demonstrated that while including links in posts has a negative impact on likes, it has a positive effect on increasing share numbers. In this study, we coded interactivity factors as three binary variables indicating whether posts contain a hashtag, a link, or a question.

In terms of variables, this study has two drawbacks. First of all, we could only identify the year that videos were posted but not the date or time, so we were unable to use variables related to time or post frequency. For dates and times, Facebook uses the elements: “<abbr title=”” data-time=”” data-shorten=”” class=”” >. We modified the Xpath element of the scraper, but it still did not work. Another type of data that could not be extracted was related to video content on Twitter. Although the scraper successfully detected photo content, it could not detect videos on Twitter,

Table 4 Key variables of the study

Variable	Type	Coding
Total interaction	Categorical	4 Classes
Social media platform	Categorical	2 Classes
Content	Categorical	4 Classes and a dummy
Number of followers	Continuous	Number
Hashtag	1/0 Binary	1/0 Binary
Question	1/0 Binary	1/0 Binary
Link	1/0 Binary	1/0 Binary
Video duration	Categorical	3 Classes (<4 min.; ≤4–8 min.; ≥8 min.
Number of views	Continuous	Number

Table 5 Topics and top words

Topic_0	Topic_1	Topic_2	Topic_3
Holiday	Water	Health	Doctor
Day	Health	Cancer	Professor
Week	Drink	Treatment	Specialist
Happy	Eat	Control	Question
World	Caution	Symptom	Information
Health	Protect	Heart	Broadcast

and as a consequence vividness factors could not be included in the analysis. We did not, however, experience such difficulties with Facebook video data, and since we had enough Facebook posts (6147), we were able to do a second analysis by filtering the Twitter data and adding the vividness variables (Table 4).

4 Findings

4.1 LDA Topic Detection

After we evaluated various models with a different number of topics (K), document-topic density (α) and topic-word density (β) values four topics emerged as a result of the LDA topic modeling. This section presents our findings as well as the average and maximum total engagement ratios per topic and some notable examples, and the next section offers an interpretation of those findings. All the posts that are presented as examples in this section (with the exception of one) were originally in Turkish and have hence been translated. Table 5 shows the top representative words with the highest term weights for each topic.

All of the terms in this table were translated from Turkish by the authors.

The first topic, Topic_0, mainly includes posts that were observances of national or religious holidays, commemorations of special global days and days/weeks dedicated to the awareness of certain causes, or hospital-related issues such as obituaries and messages posted following natural disasters. 19% of the total posts (1745) fell within this category. The engagement average for this group was found to be 0.13%, and the average interaction rate was 280. The post with the highest engagement rate had a score of 8.91% and the highest total interaction scored 19,555 (that was the highest score not only in this group but also in the data set as a whole). It was dated October 27, 2018 and posted on Facebook by Medical Park. In terms of visual content, this post included a photo of Atatürk in front of the Turkish National Flag.

> Happy 95th anniversary of our Republic! #29October #Republic Day

Medical Park also posted a message that had one of the lowest scores in this group. The message was posted on Twitter on December 30, 2019:

> Do you remember phonebooks? On special days you would go through the pages of those phonebooks and call people one by one. Call your loved ones, share your good wishes and increase your happiness in the new year. □ #newyear #health #healthforeveryone #medicalpark

The other posts that had high engagement scores included commemorations of the 10th of November (the anniversary of the death of Atatürk, who was the founder of the Republic of Turkey), the 18th of March (marking the victory of Turkish forces in the Battle of Gallipoli, also commemorated as Martyrs' Day), and the 25th anniversary of the founding of Acıbadem Hospital.

The second topic, Topic_1, included posts aimed at inspiring people to adopt healthier lifestyles, dealing with issues like quitting smoking, eating healthier, sleeping better, and exercising, as well as messages informing people about mild medical conditions like insulin resistance, gastroesophageal reflux disease (GERD), cramps, fungal infections, and allergies. The percentage ratio of the messages in this group was 27% with a total of 2533 messages, and the average engagement rate was 0.10% with an average total interaction of 245. A message posted on Facebook by Acıbadem on July 2, 2020 received the highest interaction rate in this group with a score of 9115.

> Cinnamon, cloves, basil, cumin, cardamom, mint, coriander, rosemary, garlic, ginger, and turmeric are all health-friendly spices we often use in our kitchens. They help prevent inflammation and also have detox properties.

The message that scored the highest engagement rate was a tweet posted by Medical Park on the 1st of February in 2020. The engagement rate for the post was 3.91%.

> If you want to stay healthy during the winter months and prevent weight gain:
>
> Eat more fruits and vegetables!
> Make sure you get enough protein.
> Drink plenty of fluids.
> Increase your physical activity!
> Eat less often.
> #healthforeveryone #medicalpark

Examinations of other messages in this group that had high engagement rates revealed that they were about healthy lifestyles and included recommendations about improving nutrition, recipes for healthy meals, and nutrition facts. Below are some examples of such posts that did not receive responses:

> Remember to take precautions against the harmful effects of the sun when the temperature is higher than the seasonal norm. #healthyweekend (Twitter, Medical Park, June 25, 2016)

> What should you do to get relief from cramps? (Twitter, Medicana, June 28, 2016)

The posts in the third group, Topic_2, were mostly messages aimed at raising awareness, and they provided information about the symptoms and treatment of serious diseases as well as statistical facts. The illnesses they covered include cancer, Alzheimer's, COPD, osteolysis, bipolar disorder, autism, obesity, and cardiovascular diseases. Containing a total of 3031 posts, this group accounted for 33% of the overall data with an average engagement rate of 0.08% and an average total interaction score of 178. The post with the highest interaction score (8695) was made by Acıbadem on Facebook on June 1, 2018:

> Scoliosis can be treated through early detection and in the course of treatment patients will retain their mobility. For detailed information, visit the website omurgasagligimerkezi.com.

In this group, the highest engagement score was 2.33% for a post made on Twitter by Medical Park on January 20, 2018:

> Make sure you keep up with your regular gynecological check-ups so your doctor can make an early diagnosis of cervical cancer. Don't be late for life—catch cancer early!
> #cervicalcancer #cancer #cervix #gynecology #women'shealth #health #smeartest #healthcare #medicalparkTwitter

Thirty-eight messages in this group did not receive any responses. Two such postings are as follows:

> Growth to abnormal heights, infertility, reduced visual acuity, weakness... These are all symptoms that may indicate the presence of pituitary gland tumors. (Facebook, Acıbadem, February 20, 2017)

> Narrowing of the coronary artery can occur suddenly, and if the flow of blood is not supplemented from other arteries, it can lead to a heart attack. When this occurs together with a disturbance in the rhythm of the heartbeat, the heart may not be able to supply enough blood or it may even stop completely, leading to death if an intervention is not carried out. (Facebook, Medicana, September 22, 2020)

The last topic includes hospital-related posts. The majority of the posts in this group consisted of messages promoting videos, events, or live broadcasts in which doctors discuss specific subjects. Additionally, there were posts providing information about the opening of new hospitals, health centers, and clinics, as well as other news about the hospitals. A total of 159 of the messages in this group received responses. When we calculated the average for total interactions, we excluded posts that did not get responses and we did not carry out predictive analyses of those posts.

In this group, the post with the highest total interaction score (9600) and highest customer engagement ratio (4.37%) was a video posted by Medical Park on Facebook in 2017. It was about someone who survived cancer as a child and

Table 6 Summary of the descriptive statistics of the topic groups

Topics	# of Posts	%	Mean Engagement (%)	Mean Interaction	Max. Engagement (%)	Max. Interaction	# without responses
0	1745	21	0.13	280	8.91	19, 555	17
1	2533	30	0.10	245	3.91	9155	14
2	3031	36	0.08	178	2.33	8695	35
3	1014	12	0.06	134	4.37	9600	58

Table 7 Accuracy of prediction analysis: 3 classes

Algorithm	Accuracy
DT	70.136%
RF	70.84%
GBT	70.732%

later in life volunteered to organize musical events for the children being treated at the pediatric oncology clinic of a Medical Park hospital. The caption of the video read as follows:

> Together it is possible. We believe in the importance of psychological support in the struggle against cancer. In the fight against despair and giving up, we say #TogetherItIsPossible.

Two other messages that also had high interaction scores are presented below. The second post was made in English.

> Medical Park and Liv Hospital are going public! (Facebook, Medical Park, 2018. Total Interaction Score: 7,900)

> At 14:00 on Natural TV, Dr. Ulas Sozener will talk about Kidney Transplantation. (Facebook, Medicana, 2018. Total Interaction Score: 8,900)

Below are two posts from this group that did not receive responses:

> The 19th episode of the TV series "Eşkıya Dünyaya Hükümdar Olmaz" was shot at our hospital. #eskiyadunyayahukumdarolmaz (Twitter, Medical Park, February 4, 2016)

> Prof. Dr. Vildan Çerçi explained the McKenzie technique, which provides treatment for waist and neck hernias without surgery. (Twitter, Medicana, September 8, 2016)

A summary of the descriptive statistics of the topic groups is provided in Table 6.

4.2 *Prediction Analysis*

In the first experiment, after excluding the reply posts we tested both the Twitter and Facebook posts along with independent variables such as the social media platform used, the number of followers, the hospital, questions, hashtags, links, and topics. As can be seen in Tables 7 and 8, for both experiments where the target value was discretized into three and four classes, all the algorithms yielded very close results. For three classes we obtained an approximate performance accuracy of 70% and for four classes it was 65%.

Table 8 Accuracy of prediction analysis: 4 classes

Algorithm	Accuracy
DT	65.378%
RF	65.016%
GBT	65.197%

Table 9 Accuracy of prediction analysis: Facebook

Algorithm	Accuracy
DT	69.075%
RF	68.588%
GBT	69.643%

Table 10 Average attribute statistics for 4 classes

Row ID	#splits (level 0)	#splits (level 1)	#splits (level 2)	#candidates (level 0)	#candidates (level 1)	#candidates (level 2)
Number of followers	19	39	65	22	57	108
Platform	28	21	17	28	52	111
No text	17	32	25	23	59	118
Hospital	7	26	59	29	43	110
Links	16	21	17	34	64	110
Topic 3	7	11	31	25	51	103
Topic 1	4	17	32	39	57	106

To make it possible to observe the impacts of vividness, we conducted another analysis of the Facebook posts for which we included variables such as the number of views and whether the post included a video or photo, and if it did include a video, the number of people who viewed it and the duration of the video (Table 9).

In order to observe the relative importance of the variables, namely the relative contributions to the models' predictive power, we used statistic attributes, dividing the number of splits to their candidates for each level and adding them together. As a means of avoiding the randomness factor, we repeated three individual analyses and took the average values. When the Facebook and Twitter posts were analyzed together, we found that the number of followers, the social media platform used, non-textual posts, and the hospital in question made the largest contributions to the predictive power of the models, while links, Topic 0, and Topic 1 did so to a lesser extent. The variables Topic 2, Topic 3, questions, and hashtags were found to be the variables that contributed the least. For Facebook, while the number of followers, views and lack of text made the most contributions, the hospital, links, video duration, and Topic 3 contributed moderately to the model. We found that questions, hashtags, Topic 0, Topic 1, and Topic 2 were the variables that contributed the least (Tables 10, 11 and 12).

Table 11 Average attribute statistics for 3 classes

Row ID	#splits (level 0)	#splits (level 1)	#splits (level 2)	#candidates (level 0)	#candidates (level 1)	#candidates (level 2)
Number of followers	19	41	66	22	57	107
Platform	28	19	18	28	52	109
No text	17	32	28	23	59	117
Hospital	9	28	60	29	43	108
Links	16	15	12	34	64	108
Topic 3	6	12	27	25	51	103
Topic 1	4	16	33	39	57	106

Table 12 Average attribute statistics for facebook

Row ID	#splits (level 0)	#splits (level 1)	#splits (level 2)	#candidates (level 0)	#candidates (level 1)	#candidates (level 2)
Number of followers	20	38	47	20	50	76
Views	10	20	35	17	38	85
No text	19	25	10	22	51	67
Hospital	5	20	36	17	40	73
Links	16	20	22	27	48	80
Video duration	8	13	30	26	35	84
Topic 3	3	18	26	14	48	108

5 Discussion

In order to improve our understanding of how hospitals use social media as a marketing communication tool, this study first sought to shed light on what kinds of posts they make on those platforms. For that reason, we conducted LDA topic modeling analysis, which is frequently used in text classification and is known for its ability to generate satisfactory outcomes. Our topic modeling analysis revealed that hospitals use social media for four main reasons: to send greetings for holidays and other special days, to promote healthy lifestyles, to disseminate information related to diseases, and to make news announcements about the hospitals themselves. In terms of post frequency, it was found that hospitals mostly use social media to circulate information about illnesses and promote healthy habits, followed by postings that include observations of holidays and special days/weeks and hospital-related news such as announcements about upcoming talks, conferences, and seminars. Although our study is unique in terms of its research framework and methodology, these findings are nonetheless consistent with other studies that have examined the content of hospitals' social media messages. For example, research by İlgün and Uğurluoğlu (2019) indicated that hospitals primarily use social media to create public awareness about health issues, commemorate special days/weeks, and make announcements about upcoming training sessions, congresses, conferences, and seminars, and also to promote hospital services and their departments (İlgün & Uğurluoğlu, 2019). In another study, Huang and Dunbar (2013) categorized the social media messages posted by hospitals as “hospital news,” “event announcements,” “patient stories,” “holiday salutations,” “public service announcements,” and “polls” (Huang & Dunbar, 2013).

In terms of total interaction and customer engagement, we found that greetings for holidays and special days/weeks resulted in the most interaction, followed by postings that target the promotion of healthy lifestyles, the dissemination of information about diseases, and hospital-related announcements. However, with regard to the contributions they made to the predictive power of the model, they were ranked as hospital-related announcements, the promotion of healthy habits, and observations of holidays and special days/weeks. The fact that their relative importance is low does not mean that those variables are insignificant; rather, when predicting interaction levels, contextual factors such as the number of followers were found to matter more than interactivity or content-related factors. In short, a larger number of followers means more visibility. Since sources were identified as an important factor, Twitter and Facebook should be treated as separate mediums in the design of marketing communication strategies, and the different dynamics of those platforms should be carefully considered instead of simply allowing the same material to be posted on various platforms.

The impacts of informative content have been studied for a wide variety of industries, and the outcomes have been found to vary (De Vries et al., 2012). In this study, contextual factors were found to be more critical than informative factors in terms of interaction, but it should be noted here that informative content in the healthcare industry differs from that of other industries. In the field of healthcare, informative content can refer to information about illnesses or medical conditions. Even if a follower of a page or someone else who reads the post is interested in the content, he or she may not want to “like” or respond to it because they might be concerned that their friends will see their reaction. The fact that a lack of text in posts was found to be significant may be indicative of two things. Non-textual postings are usually photos of events that are held at or organized by hospitals, but they also include some health-related posts in which text is superimposed over the photo without a caption,

As regards interactivity factors such as hashtags, questions, and links, we observed that hashtags and questions make less of a contribution than links or other factors. Scholars who have investigated the impact of links have arrived at differing conclusions. For example, Sabate et al. (2014) found no significant relationship between the number of likes and links, whereas Kujur and Singh (2016) found a negative relationship and speculated that links may have a negative impact on total interaction because they direct users to another website. Keeping both views in mind, we checked some of the posts that contained YouTube links to see if one or the other also held true for the healthcare industry. For instance, the Tweets mentioned in the results section that had an interaction score of 0 were viewed by 6721 people and received 42 likes on YouTube. All the same, further research is needed to arrive at concrete conclusions regarding this issue. Lastly, among all the variables, questions were found to have the least impact. However, it would be misleading to compare that finding with the results of studies examining other industries, as the vast majority of the posts that contained a question included answers as well, and as such the questions seemed to be rhetorical in nature. It can thus be concluded that in the healthcare sector, adding questions to posts may not increase interactivity, but it does increase noticeability.

As a final point, we obtained a performance accuracy of around 70% for all the predictive models in the study. Those results could be improved, however, by means of hyperparameter optimization or the application of other models such as neural networks, support vector machines, or fusion models, in addition to the inclusion of other variables.

6 Conclusion

In this study, we attempted to fill in several gaps in the literature. First of all, few studies have sought to determine how social media is used in the health sector, especially by hospitals, and what kinds of outcomes are produced. Through the development of a comprehensive model, we sought to build upon the existing knowledge on this subject. Secondly, there are two poles in social media studies in the literature. While the first group approaches the subject from a data science perspective and focuses on developing larger, more advanced datasets but ultimately fails to interpret the results from a marketing or social sciences point of view, studies taking up the subject from a social sciences standpoint try to explain human behavior but work with a limited amount of data and models that do not represent the complex nature of social media. Given that situation, we set ourselves the goal of creating a bridge between those two approaches, and through the use of NLP techniques and machine learning algorithms, we also interpreted the messages of posts from a marketing communication perspective.

Nonetheless, this study has some limitations that should be mentioned. First of all, the fact that we were unable to take into account the days and times of posts may have affected our results; as the literature suggests, they can have a significant impact on total interaction or customer engagement. Moreover, in this study, textual data consist of captions where the texts on the photographs were not included in the analysis. An evaluation of the pictures and videos that were included in posts not just in terms of their mere presence but also their visuality and content could also improve the performance of the models. Another noteworthy limitation is the fact that we only looked at two social media platforms, Facebook and Twitter, whereas YouTube and Instagram are also extensively used by hospitals as marketing communication tools. Further research that includes metrics from YouTube and/or Instagram could create a more inclusive understanding of how hospitals use social media and which factors affect interaction levels.

Moreover, future studies could include more hospitals to enlarge the sample size and examine interaction metrics separately with regard to comments and shares. Also, the hospitals in this study have international social media accounts, too. Comparing the hospitals' local and international accounts from a global marketing perspective could potentially contribute to the literature on medical tourism, which is a major sector in Turkey and as such could also tell an important story.

Last but not the least, the dynamics of the health sector in Turkey differ from those of many other service sectors because of the nature of the services they provide. In that regard, the meanings and impacts of interactions might differ in other industries. For example, a recent study has demonstrated that having a high number of likes or followers is not necessarily a good thing and that low customer engagement can negatively impact perceived account credibility (De Vries, 2019). As a consequence of those findings, we are left with a pressing question: Does that also hold true for the health sector, where trust is a key building block? If so, how would that affect the decision-making processes of consumers and their sense of loyalty?

Social media has become indispensable for almost every industry as a means of interacting with consumers. However, there is much to learn about how it can be used effectively. Companies need to think carefully when determining their social media strategies because they need to be compatible with their target audiences as well as their marketing and communication tactics and objectives.

References

Agarwal, R., Dugas, M., Gao, G., & Kannan, P. K. (2020). Emerging technologies and analytics for a new era of value-centered marketing in healthcare. *Journal of the Academy of Marketing Science, 48*, 9–23. https://doi.org/10.1007/s11747-019-00692-4

Algiriyage, N., Jayasena, S., Dias, G., Perera, A., & Dayananda, K. (2013). Identification and characterization of crawlers through analysis of web logs. In *2013 IEEE 8th international conference on industrial and information systems, ICIIS 2013 - conference proceedings*, 150–155. https://doi.org/10.1109/ICIInfS.2013.6731972

Asadi, N., Lin, J., & De Vries, A. P. (2014). Runtime optimizations for tree-based machine learning models. *IEEE Transactions on Knowledge and Data Engineering, 26*(9), 2281–2292. https://doi.org/10.1109/TKDE.2013.73

Aydin, G. (2020). Social media engagement and organic post effectiveness: A roadmap for increasing the effectiveness of social media use in hospitality industry. *Journal of Hospitality Marketing and Management, 29*(1), 1–21. https://doi.org/10.1080/19368623.2019.1588824

Batrinca, B., & Treleaven, P. C. (2014). Social media analytics: A survey of techniques, tools and platforms. *AI Soc, 30*(1), 89–116. https://doi.org/10.1007/s00146-014-0549-4

Bejtkovsky, J. (2020). Social media platforms as HR marketing tool in selected healthcare service providers. *Marketing and Management of Innovations, 1*, 294–302. https://doi.org/10.21272/mmi.2020.1-25

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation David. *Journal of Machine Learning Research, 3*(3), 993–1022. https://doi.org/10.1016/B978-0-12-411519-4.00006-9

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classificaiton and regression trees*. Chapman & Hall/CRC.

Cangelosi, J., Kim, D., Griffin, K., & Ranelli, E. (2019). The role of social media and social networking as marketing delivery systems for preventive health care information. *Atlantic Marketing Journal, 8*(1), Article 1. https://digitalcommons.kennesaw.edu/amj/vol8/iss1/1

Chatterjee, S., Goyal, D., Prakash, A., & Sharma, J. (2020). Exploring healthcare/health-product eCommerce satisfaction: A text mining and machine learning application. *Journal of Business Research, 131*, 815–825. https://doi.org/10.1016/j.jbusres.2020.10.043

Cheng, P., Xia, G., Pang, P., Wu, B., Jiang, W., Li, Y. T., Wang, M., Ling, Q., Chang, X., Wang, J., Dai, X., Lin, X., & Bi, X. (2020). COVID-19 epidemic peer support and crisis intervention via social media. *Community Mental Health Journal, 56*, 786–792. https://doi.org/10.1007/s10597-020-00624-5

Cui, H., Huang, D., Fang, Y., Liu, L., & Huang, C. (2018). Webshell detection based on random forest-gradient boosting decision tree algorithm. In *Proceedings – 2018 IEEE 3rd international conference on data science in cyberspace, DSC 2018*, 153–160. https://doi.org/10.1109/DSC.2018.00030

De Vries, E. L. (2019). When more likes is not better: the consequences of high and low likes-to-followers ratios for perceived account credibility and social media marketing effectiveness. *Marketing Letters, 30*(3), 275–291.

De Vries, L., Gensler, S., & Leeflang, P. S. H. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of Interactive Marketing, 26*(2), 83–91. https://doi.org/10.1016/j.intmar.2012.01.003

Delen, D., & Sharda, R. (2009). Predicting the financial success of hollywood. *Endüstri Mühendisliði Dergisi, 21*(1), 30–37.

Demmers, J., Weltevreden, J. W. J., & van Dolen, W. M. (2020). Consumer engagement with brand posts on social media in consecutive stages of the customer journey. *International Journal of Electronic Commerce, 24*(1), 53–77. https://doi.org/10.1080/10864415.2019.1683701

Dewi, L. C., Meiliana, & Chandra, A. (2019). Social media web scraping using social media developers API and regex. *Proc Comput Sci, 157*, 444–449. https://doi.org/10.1016/j.procs.2019.08.237

El-Awaisi, A., O'Carroll, V., Koraysh, S., Koummich, S., & Huber, M. (2020). Perceptions of who is in the healthcare team? A content analysis of social media posts during COVID-19 pandemic. *Journal of Interprofessional Care, 34*(5), 622–632. https://doi.org/10.1080/13561820.2020.1819779

Fiesler, C., Beard, N., & Keegan, B. C. (2020). No robots, spiders, or scrapers: legal and ethical regulation of data collection methods in social media terms of service. In *Proceedings of the international AAAI conference on web and social media*, *2020* (ICWSM), 187–196. Retrieved from www.aaai.org

Fiok, K., Karwowski, W., Gutierrez, E., & Ahram, T. (2020). Predicting the volume of response to tweets posted by a single twitter account. *Symmetry, 12*(6), 1–15. https://doi.org/10.3390/sym12061054

Gregor, B., & Gotwald, B. (2013). Potential of social media and mobile marketing in marketing communication management in health care institutions. *International Journal of Contemporary Management, 12*(3), 38–49.

Hackworth, B. A., & Kunz, M. B. (2011). Health care and social media: Building relationships via social networks. *Academy of Health Care Management Journal, 7*(2), 1–14.

Haenlein, M., Anadol, E., Farnsworth, T., Hugo, H., Hunichen, J., & Welte, D. (2020). Navigating the new era of influencer marketing: How to be successful on Instagram, TikTok, & co. *California Management Review, 63*(1), 5–25.

Hewitt, A. M. (2011). Aligning social media, social networking and social marketing: engagement as the key to healthcare marketing. *Proceedings of Northeast Business and Economics Association, 4*(2), 211–214.

Huang, E., & Dunbar, C. L. (2013). Connecting to patients via social media: A hype or a reality? *Journal of Medical Marketing, 13*(1), 14–23. https://doi.org/10.1177/1745790413477647

Hur, M., Kang, P., & Cho, S. (2016). Box-office forecasting based on sentiments of movie reviews and independent subspace method. *Information Sciences, 372*, 608–624. https://doi.org/10.1016/j.ins.2016.08.027

İlgün, G., & Uğurluoğlu, Ö. (2019). How Turkish private hospitals use social media: A qualitative study. *Journal of Social Service Research, 45*(1), 34–43. https://doi.org/10.1080/01488376.2018.1479339

Jang, K., & Baek, Y. M. (2019). When information from public health officials is untrustworthy: the use of online news, interpersonal networks, and social media during the MERS outbreak in South Korea. *Health Communication, 34*(9), 991–998. https://doi.org/10.1080/10410236.2018.1449552

Kingsford, C., & Salzberg, S. L. (2008). What are decision trees? *Nature Biotechnology, 26*(9), 1011–1012. https://doi.org/10.1038/nbt0908-1011

Kotler, P., & Zaltman, G. (1971). Social marketing: An approach to planned social change. *Journal of Marketing, 35*, 3–12.

Koumpouros, Y., Touliasa, T. L., & Koumpouros, N. (2015). The importance of patient engagement and the use of social media marketing in healthcare. *Technology and Health Care, 23*, 495–507. https://doi.org/10.3233/THC-150918

Kujur, F., & Singh, S. (2016). Social networking sites as a multimedia tool for brand popularity - an exploratory study. *Indian Journal of Science and Technology, 9*, 45. https://doi.org/10.17485/ijst/2016/v9i45/98151

Lash, M. T., & Zhao, K. (2016). Early predictions of movie success: The who, what, and when of profitability. *Journal of Management Information Systems, 33*(3), 874–903. https://doi.org/10.1080/07421222.2016.1243969

Limbu, Y. B., Pham, L., & Mann, M. (2020). Corporate social responsibility and hospital brand advocacy: Mediating role of trust and patient-hospital identification and moderating role of hospital type. *International Journal of Pharmaceutical and Healthcare Marketing, 14*(1), 159–174. https://doi.org/10.1108/IJPHM-04-2019-0029

Liu, X., Shin, H., & Burns, A. C. (2019). Examining the impact of luxury brand's social media marketing on customer engagement: Using big data analytics and natural language processing. *Journal of Business Research, 125*, 815–826. https://doi.org/10.1016/j.jbusres.2019.04.042

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., ... Adam, S. (2018). Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Communication Methods and Measures, 12*(2–3), 93–118. https://doi.org/10.1080/19312458.2018.1430754

Meer, T. G. L. A., & Jin, Y. (2020). Seeking formula for misinformation treatment in public health crises: The effects of corrective information type and source. *Health Communication, 35*(5), 560–575. https://doi.org/10.1080/10410236.2019.1573295

Naik, A., & Samant, L. (2016). Correlation review of classification algorithm using data mining tool: WEKA, Rapidminer, Tanagra, Orange and Knime. *Proc Comput Sci, 85*(Cms), 662–668. https://doi.org/10.1016/j.procs.2016.05.251

Octoparse. (2021). Web Scraping Tool & Free Web Crawlers | Octoparse. Retrieved January 20, 2021, from https://www.octoparse.com/

OECD. (2021a). Gross domestic product (GDP) (indicator). Retrieved January 3, 2021, from https://doi.org/10.1787/dc2f7aec-en

OECD. (2021b). Health expenditure. Retrieved January 24, 2021, from https://www.oecd.org/health/health-expenditure.htm#:~:text=Health%20spending%20as%20a%20share,second%20and%20third%20highest%20spenders

OECD. (2021c). Health workforce. Retrieved January 24, 2021, from http://www.oecd.org/els/health-systems/workforce.htm

Patel, B. R., & Rana, K. K. (2014). A survey on decision tree algorithm for classification. *IJEDR, 2*(1), 1–5.

Popović, K., & Smith, D. C. (2010). Tweeting@doctorwelby: Practical examples of social media in healthcare. *Journal of Communication in Healthcare, 3*(2), 138–151. https://doi.org/10.1179/175380710X12813488575925

Radu, G., Solomon, M., Gheorghe, C. M., Hostiuc, M., Bulescu, I. A., & Purcarea, V. L. (2017). The adaptation of health care marketing to the digital era. *Journal of Medicine and Life, 10*(1), 44–46.

Reis, I., Baron, D., & Shahaf, S. (2018). Probabilistic random forest: A machine learning algorithm for noisy data sets. *The Astronomical Journal, 157*(1), 16. https://doi.org/10.3847/1538-3881/aaf101

Sabate, F., Berbegal-Mirabent, J., Cañabate, A., & Lebherz, P. R. (2014). Factors influencing popularity of branded content in Facebook fan pages. *European management journal, 32*(6), 1001–1011.

Saglik (2020). Health statistics yearbook. Retrieved January 2, 2021, from https://www.saglik.gov.tr/TR,62400/saglik-istatistikleri-yilligi-2018-yayinlanmistir.html

Schillinger, D., Chittamuru, D., & Ramírez, A. S. (2020). From "infodemics" to health promotion: a novel framework for the role of social media in public health. *American Journal of Public Health, 110*(9), 1393–1396. https://doi.org/10.2105/AJPH.2020.305746

Schultz, C. D. (2017). Proposing to your fans: Which brand post characteristics drive consumer engagement activities on social media brand pages? *Electronic Commerce Research and Applications, 26*, 23–34. https://doi.org/10.1016/j.elerap.2017.09.005

Sharda, R., & Delen, D. (2006). Predicting box-office success of motion pictures with neural networks. *Expert Systems with Applications, 30*(2), 243–254. https://doi.org/10.1016/j.eswa.2005.07.018

Sharma, V., & Gupta, M. (2019). Correlation between attributes concerning social media marketing and private healthcare sector. *MANTHAN: Journal of Commerce and Management, 6*(1), 1–15. https://doi.org/10.17492/manthan.v6i1.182678

Shawky, S., Kubacki, K., Dietrich, T., & Weaven, S. (2019). Using social media to create engagement: a social marketing review. *Journal of Social Marketing, 9*(2), 204–224. https://doi.org/10.1108/JSOCM-05-2018-0046

Stadtelmann, M., Woratschek, H., & Diederich, C. (2019). Actor engagement in online health communities: A netnographic analysis of engagement practices and roles. *International Journal of Pharmaceutical and Healthcare Marketing, 13*(4), 500–515. https://doi.org/10.1108/IJPHM-06-2018-0033

Tso, G. K. F., & Yau, K. K. W. (2007). Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. *Energy, 32*(9), 1761–1768. https://doi.org/10.1016/j.energy.2006.11.010

Wang, R., & Jin, Y. (2010). An empirical study on the relationship between the followers' number and influence of microblogging. *Proc Int Conf E-Bus E-Gov ICEE, 2010*(2007), 2014–2017. https://doi.org/10.1109/ICEE.2010.509

Wuest, T., Weimer, D., Irgens, C., & Thoben, K. D. (2016). Machine learning in manufacturing: Advantages, challenges, and applications. *Prod Manuf Res, 4*(1), 23–45. https://doi.org/10.1080/21693277.2016.1192517

Zhang, Y., & Haghani, A. (2015). A gradient boosting method to improve travel time prediction. *Transport Res Part C: Emerg Technol, 58*, 308–324. https://doi.org/10.1016/j.trc.2015.02.019

Exploring the Third Wave of Feminism Through Hierarchical Clustering and Sentiment Analysis

Hüseyin Umutcan Ay, Semra Nazlı Günesen, and Tolga Kaya

Abstract Although much had changed throughout the feminist movement during the first wave (the 1800s–1960s) and second wave (1960s–1980s), with the inclusion of distinct minority groups into the idea, most of the contemporary approaches had been developed within the third wave. The purpose of this study is twofold: First, to investigate the works published during the third wave of feminism through unsupervised clustering methods. Second, to determine the driving emotional structures for each of these consequent clusters using sentiment analysis. To conduct the analyses, sizeable data is gathered using the literature published after the mid-1990s. The data is then cleaned and prepared according to the bag of words methodology before the usage of the hierarchical clustering technique. As a final step, sentiment analysis based on Plutchik's Wheel of Emotion has been used to illustrate the magnitudes of eight distinct emotions. Results demonstrate that there are clear points of distinction between the 3 waves' sentiment analysis and main ideas. One of the significant findings was the 3rd wave's feminism understanding being not only about women but all types of disadvantaged minority groups. Another important finding is that the aggression level of the ideas that do not change during the 3 waves increases significantly over time. This research contributes to the literature by providing an objective framework to analyze how the feminism ideology is evolved.

Keywords Text Mining · Natural Language Processing · Third Wave Feminism · Sentiment Analysis · Bag of Words Method

H. U. Ay (✉) · S. N. Günesen · T. Kaya
Istanbul Technical University, Sarıyer/Istanbul, Turkey
e-mail: ayh16@itu.edu.tr

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_7

1 Introduction

Feminism emerging from a movement advocating women's basic rights in the male-dominated society of the early 1800s has now evolved to a broader ideology that upholds equal rights and legal protection for all gender identities in all areas of life and even endorses civil rights and environmentalism. Although it has already been achieved to reshape the social fabrics of societies considerably almost all over the world, there is still room for improvement. Nowadays, women's suffrage and equal rights are no longer the driving motives of the movement. However, feminism has still been in demand as females and gender minorities are not treated fairly on any platform everywhere in the world, mainstream media continues to commercialize the over-sexualization of women's bodies, and enlightening societies on a global scale is still a significant need. For this reason, feminism will presumably remain on the agenda for all human beings in the world for at least a few more decades. The evolvement of feminism in the world over the years is usually categorized into three waves: the first wave (the 1800s to early 1960s), the second wave (1960s–1980s), and the third wave (1990s to today).

Within this research, to examine three consequent waves of feminism, machine learning algorithms will be used alongside the sentiment-based text mining methods to create unsupervised clusters of different topics. Since 3rd wave contains a wide variety of emotions and ideas, more detailed analyses are conducted to accurately indicate emotions and ideas behind the documents, which would create a basis for future researchers to pursue their works on this wave.

This research aimed to have an unbiased determination of different feminism concepts and to identify the distinctive ideas of each of the 3 waves by the usage of unsupervised clustering methods and also by visualizing the frequent terms within the waves. Another objective was to demonstrate the emotional contrasts among eras and analyze the emotional fluctuations towards the same concept over different ideology phases with the help of sentiment analysis methods. Since it would require a lot of time and effort to examine all works of feminism objectively, methods providing accurate brief information shorten the analysis process significantly. In this research, power of the data wanted to be highlighted by analyzing a complex concept like feminism.

2 Literature Review

2.1 History of the Feminism

2.1.1 The First Wave (1800s to Early 1960s)

Men and women in the society of the nineteenth century were assumed biologically created for predetermined separate roles and duties. The conception of women was evoking the role of wife and mother, physical weakness, emotionality, lower intellectual capacity, intuitiveness, irrationality, being able to work in manual jobs only, no right to vote, no opportunity to be educated at school/universities. Right to divorce was far more difficult for women compared to that of men (Riley, 1995). Besides, husbands had the right to own their wives' properties and salaries (Offen, 1988). Further, marital rape and physical abuse were legal for husbands (Teasys, 2019). Additionally, abortion was strictly forbidden by law and divorced women had no rights concerning their children (Burdett, 2001).

The first wave of feminist activism arose in the 1800s as a movement upholding abolition of slavery and women's right to have a voice in society. Towards the late nineteenth century, the movement focused on women's right to vote in particular and ended with the legislation of the 19th amendment granting women voting rights to the US Constitution in 1919 (Rampton, 2008).

Women's suffrage (right to vote), contract and property rights, right to divorce, anti-racism (many abolitionists and black rights champions spearheaded the movement) were prominent arguments raised by the first wavers (Cobble et al., 2014).

2.1.2 The Second Wave (Late 1960s to 1980s)

The society of the 1960s was still patriarchal and women were still treated as second class although they have some legal rights. They had the opportunity to be educated at school/university. They started to participate in the labor force but did not receive equal pay for the same work. They had the right to divorce although it was socially disapproved (Walters, 2005). Marital rape and physical abuse were legally forbidden but guilty husbands were rarely sentenced (Teays, 2019). Besides, women's body was used as a commercial objective in advertising (Dyhouse, 2010).

The second wave that unfolded in the 1960s was emphasizing the women's liberation movement for equal legal and social rights. After the Second World War, it rose in the context of civil rights and anti-war movements coinciding with the awakening of the leftist anti-system minority groups (68 generations) everywhere in the world. For this reason, the mode of the second wave was more radical (Jacob & Linoa, 2005). The movement driven by middle-class white women focused on passing the Equal Rights Amendment (guaranteeing legal gender equality) to the Constitution (Rompton, 2008). Additionally, Supreme Court in the USA declared that abortion legal by 1973 (Burdett, 2001).

Equal employment opportunity, equal pay for equal work, contraception (birth control), racial/sexual discrimination in the workplace, body positivity, civil rights (anti-war, anti-racism), right of abortion were the leading themes underlined by the second wavers (Rompton, 2008).

2.1.3 The Third Wave (1990s to Today)

With the third wave, women became more equal to men, and women's rights were better protected by law. Marriage and having children were no longer obligations for women. Rights concerning marriage were more equal (Savingny & Warner, 2015). Conversely, the LGBTQ movement was getting widespread beyond western societies (Annapurany, 2016).

The third wave of feminism appearing towards the mid of 1990s destabilized many constructs, including gender, body, sexuality, heteronormativity, and even identity politics (Rosenil, 2013). It was a time of empowered ladies, radical punk rock Riot Grrrls unfolded as a form of feminism rejecting the notion of being sex objects, and the LGBTQ movement was also embraced (Loomba & Sanchez, 2016). The movement has spent notable effort to address issues concerning homophobia, social class inequality, sexism, and racism (Walters, 2005). Besides, it became a lifestyle rather than a thought (Freedman, 2002). Additionally, the young feminist movement was emphasizing collective action including gender minorities, and assuming a universal female identity was avoided by the third wave feminists (Cobble et al., 2014).

Issues that have been emphasized by the third wavers are LGBTQ rights, body positivity, fat positivity, sex positivity, brain + beauty (women can have both brain and beauty), homophobia, social class inequalities, and sexual orientation (Rosenil, 2013).

2.2 Methodologies Applied in the Literature

To create a bridge between the feminist studies and the text mining literature, in this section widely used text mining techniques will be described while also emphasizing how those techniques would contribute to the contemporary feminist studies.

2.2.1 Corpus

A corpus can be described as, "A collection of linguistic data, either compiled as written texts or as a transcription of recorded speech" and the computer corpus as "A large body of machine-readable texts" (David, 1992). Since the gathered dataset in this study includes a large variety of books each having around 10,000 words at minimum, using such structure is crucial to address the sheer volume of texts. By converting the data into a corpus structure, it becomes possible to clear and analyze their contents to derive results faster than regular reading.

2.2.2 N-Gram Tokenization

N-gram tokenization is a method independent from language types and it identifies the problems caused by morphological issues that lower information retrieval performance. N-grams have been widely used for analyzing Asian languages such as Chinese and Japanese. However, it has also been shown to be an effective method to analyze European languages (McNamee & Mayfield, 2004). *Since the texts and movements created and promoted in the European languages had shaped the general direction of feminism in each of the eras, when analyzing the creation of new ideas, basing the study on the European texts will highlight the general characteristics of that era. Therefore, n-gram tokenization will be an effective tool to derive the development of the ideas within the feminist movement.*

2.2.3 Word Association

Word association is a method to investigate the relations between text data. These relations could be in the form of paradigmatic or syntagmatic. The purpose of paradigmatic relations is to find similarities between words after obtaining them from a bag of clouds. Syntagmatic relations are more interested in how frequently the two words are combined (Correia et al., 2018). By looking at the syntagmatic relations between the words it is possible to determine the general idea flow within the texts. The words that are combined frequently will communicate the stance of the author on a certain topic while also highlighting the general theme of the works.

2.2.4 Hierarchical Cluster Analysis (HCA)

HCA is one of the multifactorial exploratory approaches and it cluster individuals based on the distance between them. Agglomerative and divisive are two distinct types of HCA and they have to find the optimum number of clusters as common property. Agglomerative clustering takes a table as an input and converts it into a distance matrix. It disarranges the distance matrix to make individuals in distance objects accumulated into clusters. The processes compound individuals in gradually larger clusters until they all become united at the end. This process is often illustrated with a graphic form called "dendrogram" (Desagulier, 2014). Here, for the feminist study, HCA could be utilized to highlight how each of the authors is grouped in the way of thinking. Especially for the trend determination within the waves, the HCA method could be beneficial since it groups similar concepts according to their frequency of occurrences in each text.

2.2.5 Ward's Clustering Method

Ward's Clustering Method is the only type of agglomerative clustering approach working with classical sum-of-squares principle creating groups that minimize within-group dispersion and it can find clusters in multivariate Euclidean space (Murtagh & Legendree, 2014). Since focusing on determining the general characteristics of the ideas rather than ones of the outliers is more important to understand how the idea developed between the waves. Therefore, Ward's clustering method provides the most suitable agglomerative clustering option for capturing the general behavior.

2.2.6 Canberra Distance

The Canberra distance was introduced by Lance and Williams and it is often used for data cumulated around an origin. It is based on a generalized equation in the form of Eq. 1.

$$d^{\mathrm{CAD}}(i, j) = \sum_{k=0}^{n-1} \frac{\left|y_{i,k} - y_{j,k}\right|}{\left|y_{i,k}\right| + \left|y_{j,k}\right|} \tag{1}$$

Large number of zero entries are well handled by Canberra distance method (Desagulier, 2014) which occurs frequently when working with different types of books and texts. Since each author has their own choice of words and styles, the corpus becomes a type of scarce matrix with a lot of zero entries.

2.2.7 Sentiment Analysis and Plutchik's Wheel of Emotions

The sentiment analysis is mainly used to indicate the writer's emotions or the polarity of their statements depending on the scope of the analysis. The main idea behind the sentiment analysis is to compare the subject with the pre-existing dictionary to assign certain emotion ratings onto the text (Silge & Robinson, 2017). As a sentiment dictionary, the NRC emotion lexicon was used for the analysis (Mohammad & Turney, 2010). It assigns eight different emotions, indicated in a form of a wheel of emotions by Plutchik, to its pre-designated dictionary of 13,901 distinct words. Usage of Plutchik's "wheel of emotions" to simplify emotion classification as a binary problem for four opposing emotion pairs (Suttles, J., & Ide, N., 2013). By dividing the emotions of the authors under eight distinct categories, change in the general attitude towards a certain subject could be observed.

3 Processes Applied in the Research

In Fig. 1 an overview for the process is illustrated. The whole process constitutes five phases: data preparation, data cleaning, bag of words analysis, clustering, and sentiment analysis.

3.1 Data Preparation

Using the keywords like “feminism,” “women rights,” “patriarchy,” and “gender inequality,” 662 different books and articles about feminism had been retrieved for analysis, also including the well-known works by the renowned women rights activists such as Mary Wollstonecraft, Virginia Woolf, and Betty Friedan. Categorizing each of them by their publication dates, the works are then divided into three consequent waves of which boundaries are set by the previous researches on the waves of feminism. In Fig. 2, this process could also be seen step by step.

As could be observed from Table 1, the dataset lacks in terms the first and second wave documents due to their limited digitalization rates, constraining the analysis. However, as both the first and second wave documents are consisting

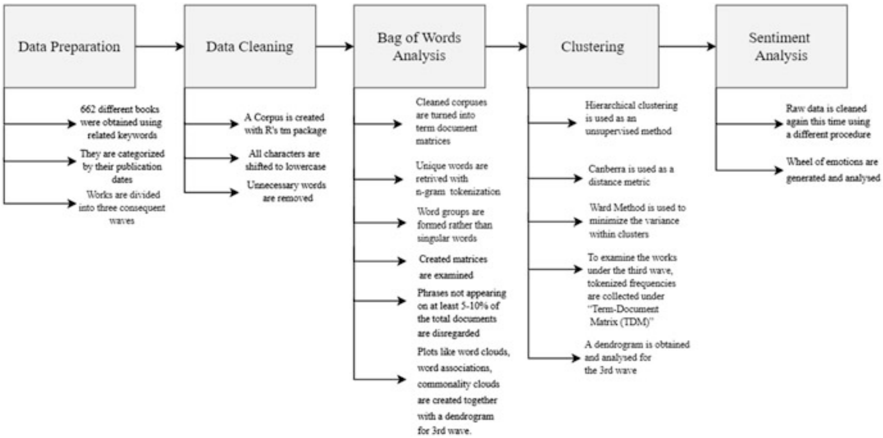


Fig. 1 Summary process flowchart

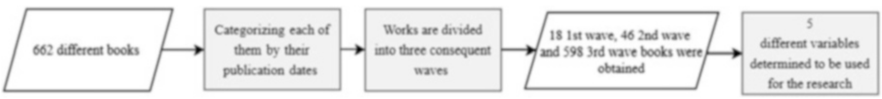


Fig. 2 Summary data preparation process flowchart

Table 1 Distribution of documents and word counts according to the waves of feminism

feminism_raw$wave			
Wave	First	Second	Third
Document	18	46	598
Word count	1,364,634	1,812,327	18,033,550

Table 2 Variable types and explanations

Variable	Type	Explanation
doc_id	Character	Unique documentation id ranging from 1 to 662
text	Character	Text of work subject to investigation
year	Numeric	Publication year
wave	Factor	Factor of 3
title	Character	Main title of the related work

highly important works that shaped the modern-day discussions on the feminism, they will be included to find the common patterns and subjects of their related era. Five different variables coded for the research are indicated under Table 2 where "doc_id" and "text" variables are used for constructing the analysis while the rest are regarded as a metadata component used for slicing the raw data.

3.2 Data Preparation

Before going into the research, text of the raw data requires cleaning to eliminate the insignificant subjects that are currently masking the importance of relevant and unique words. As the texts within the dataset are too large to handle with an ease, for the procedure R's tm package (Feinerer et al., 2008) is utilized to create a corpus which reduces the space requirements of data. After shifting to a corpus, the procedures provided in Fig. 3 are implemented to eliminate insignificant subjects from the text.

3.3 Bag of Words Analysis

For the bag of words analysis, cleaned corpuses are then turned into document term matrix and term document matrices only indicating the document id's and all of the unique words within the texts. When retrieving those unique words, n-gram

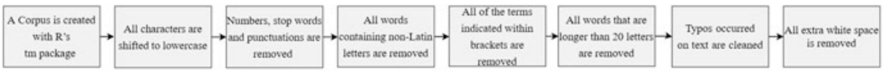


Fig. 3 Data preparation process flowchart

tokenization where $n = 2$ (bi-gram tokenization) is also utilized alongside the usual one-word mapping as McNamee and Mayfield states that the n-gram tokenization outperforms the word-based indexing for complex languages (2004). By taking the word groups rather than singular words, it is aimed to reach initial conclusions about the word associations which will indicate the general way of thinking lies behind the subject era.

Another issue with the created matrices is the sparsity. When the created matrices are examined, terms of non-sparse and sparse entries will appear indicating the common and different terms between the documents. Although their main aspects are expected to be the same, as each document is done including different topics, effect of high frequency-low occurrence terms is required to be reduced. For this, phrases not appearing on at least 5–10% (on average, depends on the sample size) of the total documents are disregarded from the matrix. Using the created matrices, plots like word clouds, word associations, commonality clouds and for third wave dendrogram are created for visualization.

3.4 Cluster Dendrogram

Different from the first two waves, the third wave is where most of the different ideas flourish within the feminist community and also where most of the data is from. Therefore, to paint a clear picture of the different topics for this wave; hierarchical clustering will be used as an unsupervised method to reduce the researcher's bias on the topic. As explained during the literature review, Canberra is used as a distance metric as the document term matrix has a lot of zero entries within, which Canberra is stated to be good at while clustering them. Also to reduce the bias derived from the complete linkage method, the ward method is emphasized instead, minimizing the variance within clusters. That way, it is expected to have more flexible clusters where the characteristics of the cluster indicate the general approach over the stated topic rather than focusing on an individual idea.

Alongside Ward's Method and Canberra distance, for examining the works under the third wave, tokenized frequencies are collected under "Term-Document Matrix (TDM)" where the documents are represented as columns, while the bi-gram tokenized words are represented as rows with the entries as frequencies of words in each document. However, due to the general keywords within the wave being the dominating ones in terms of frequencies, some form of penalty is required to highlight the occurrences of not-so-common words within the documents. For this, a penalty score suggested by Salton and Buckley (1988) will be used

$$\mathrm{idf}_i = \log_2 \frac{|D|}{|\{d \mid t_i \in d\}|} \tag{2}$$

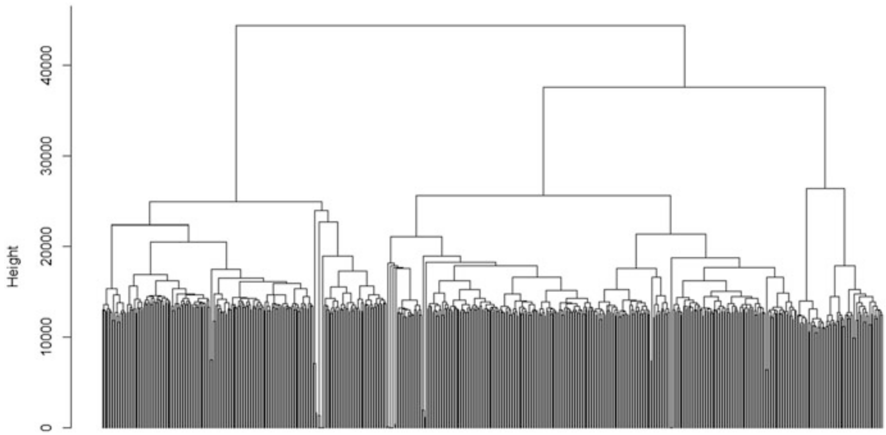


Fig. 4 Cluster dendrogram for the third wave

where t_i is the term in TDM, $|D|$ is the total number of documents, $|\{d \mid t_i \in d\}|$ is the number of documents including the term t_i, and idf_i is the "inverse document frequency."

By utilizing all of the aspects, the dendrogram on Fig. 4 is obtained which is rather hard to interpret with naked eye. However, it provides insights on at what distance the different documents are segmented into clusters. It is seen that most of the documents are grouped where only a few of the outlier documents could not be clustered with the rest of them. Those outliers could be the ones with an importing error or the ones promoting significantly different subjects. However, each of them belonging to a group of 1–3 documents does not allow for an effective interpretation. Therefore, they will not be included within the discussions although for the following analyses the total cluster number is selected as 10, they will not be included within the discussions.

3.5 Sentiment Analysis

As explained before, sentiment analysis is mainly used on texts to indicate the writer's emotions or the polarity of their statements depending on the scope of the analysis. The main idea behind the sentiment analysis is to compare the subject with the pre-existing dictionary to assign certain emotion ratings onto the text (Silge and Robinson, 2017).

For this research, sentiment analysis will be conducted to find relations with different waves of feminism and certain emotions. However, there is an issue with the pre-cleaned data as the initial approach was to remove the stop words from the text. And it was beneficial within the bag of words approach as their masking effect over the data was removed. But, as those stop words included also the amplifier words such as “very,” “much,” “really,” etc. the actual tone of the words is lost during the conversion process. Because of this, the raw data once again required to be cleaned, this time using the different procedure as follows:

1. All characters are shifted to lowercase.
2. Numbers are eliminated from text.
3. All end sentence punctuations are removed.
4. All words containing non-Latin letters are removed.
5. Within bracket texts are removed.
6. All words that are longer than 20 letters are removed to address the pdf conversion issues where two or more words are combined.
7. Typos occurred on text are cleaned by decreasing the 3 or more times repeated letters within words to 2.
8. All extra whitespace is removed to reduce the total space of the document.

The NRC package (Mohammad and Turney, 2010) is then used on the clean corpus data in tidy format and the radar chart is drawn indicating emotion weights.

4 Findings

For the explanations of the findings, although most of the contemporary improvements had happened within the third wave of feminism, still the basis for those ideas stems from the developments within the first two waves. Therefore, before explaining each of the clusters for the third wave feminism and their difference in terms of terminologies, a brief look over the books published during the first and second wave will also be provided.

4.1 First Wave of Feminism

When both words clouds in Fig. 5 are examined, it is clear that the first wave of feminism has a more intrinsic approach to the subject. The most frequent terms “mother,” “women,” “little girl,” “husband,” “men women,” “child” indicate that most of the arguments were about the place of the women within the society. For these texts, no frequent terms from more recent discussions like gender, race, and ethnicity appear within the word clouds pointing out these concepts had not reached their maturity yet. To investigate further at what phrases these most frequent terms are related with, the following association plots will be investigated.

As seen in Fig. 6 on the left, the term “women men” is mostly correlated with terms like women’s rights and public opinion, indicating that the main concern is solidifying the place of women within society. Moreover, the term “can hope” is parallel with the era’s more passive approach to defending rights. Also, an interesting term appears at the bottom of the list that could be completed as

Fig. 5 Commonality cloud (left) and bi-gram word cloud (right) for the first wave

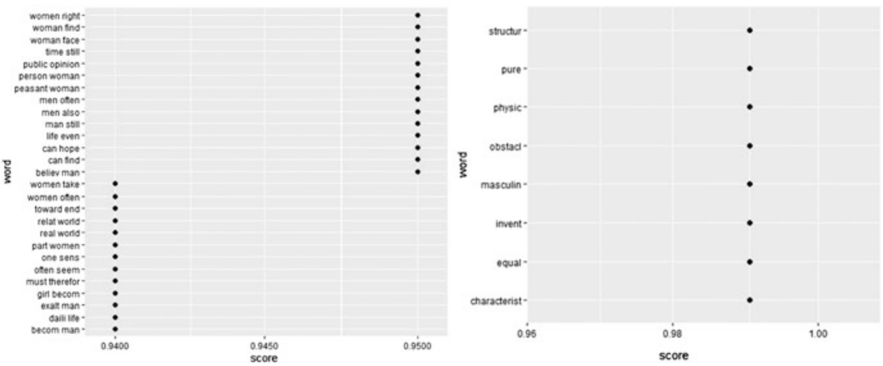


Fig. 6 First wave association plot for "women men" (left) and "women" (right)

"becoming a man" which indicates the instinctive self-defense mechanism women develop against the era's strong ignoring attitude. The terms "daily life" and "real world" might show women's being stuck in daily life with their chores and struggle to understand and find a place for their identity in the real world. Also, the opposition to patriarchy could be observed from the second association plot in Fig. 6 on the right as the term "women" is correlated with the terms "obstacle," "masculinity," and "equal." Moreover, the term "pure" highlights society's expectation from women to be the "virgin bride" and "perfect mother" figures. Therefore, it can be concluded that the first wave of feminism mostly revolved around achieving equal rights for women within the societal structure which also compliments the previous researches.

4.2 Second Wave of Feminism

When the texts from the second wave are investigated, it is clear that the ethnicity issues had come into the picture as “black” and “white” are appearing within most of the words inside the cloud. However, it also indicates that the discussion of equal rights for everyone is not yet global since no frequent words are indicating other ethnicities. Also, the term “woman suffrage” shows that the voting right of the woman is a widespread idea within the society as it also has a strong basis within the word cloud. Additionally, working conditions for women is another issue for this wave as terms like “capitalist society,” “women work,” and “work class” have taken place inside the most frequent phrases. It should also be noted that sexuality issues are becoming more frequent as the waves progress (Fig. 7).

When investigated, the association plot for “black women” shows the society’s attitude towards black people for the given period as the correlated words carry strong adjectives like “oppress,” “kill,” “rape,” and “blame victim.” Also, the education gap between black and white people is another illustrated subject (Fig. 8).

Fig. 7 Bi-Gram word cloud for second wave

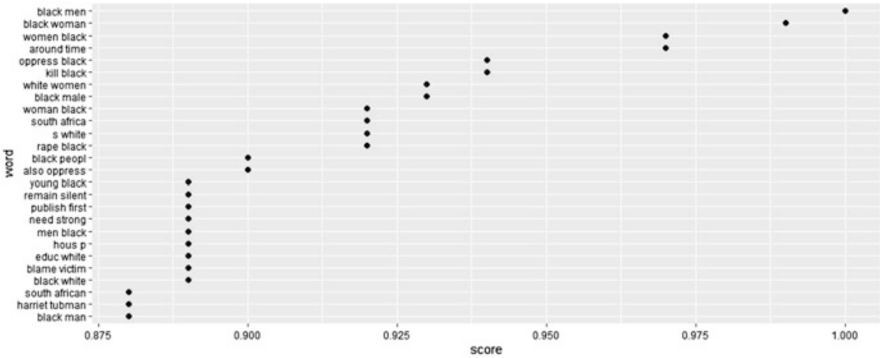


Fig. 8 Second wave association plot for “black women”

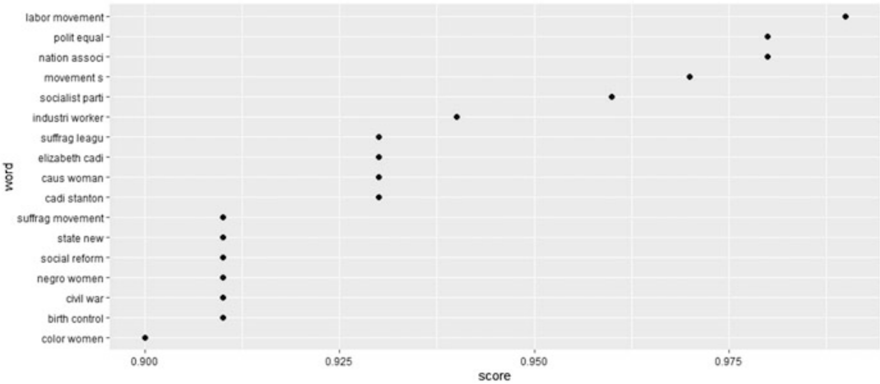


Fig. 9 Second wave association plot for “woman suffrage”

In the second wave period society was still patriarchal and although women started to join the labor force receiving equal pay was one of the main motives of the feminist movement. When the associated plot for woman suffrage is investigated there are clear patterns with emphasis on the concepts like “labor movement” and “industry worker.” It was also a period that coincides with just after the second war and leftist tendencies were getting stronger with the emphasis of anti-war movements. The highlighted concepts like “socialist party,” “social reform,” and “civil war” are visible pieces of evidence. “Political equality” continues to remain as a main motive for the feminist movement (Fig. 9).

During the second wave, women started to join the labor force. However, receiving equal pay and labor rights become some of the main arguments raised by feminists. Emphasis on the concepts such as “women hard,” “employ domestic,” “female labor,” and “woman worker” can be interpreted as visible pieces of evidence. The right to abortion was also one of the leading arguments raised by the second wavers. The “right birth” word association appearing is aligned with this consideration (Fig. 10).

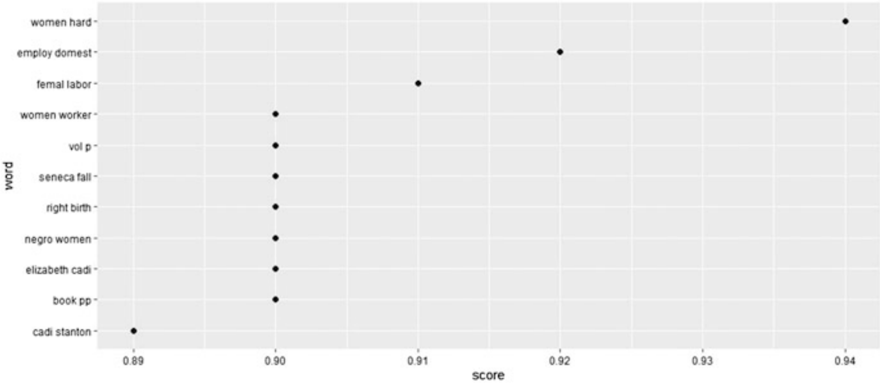


Fig. 10 Second wave association plot for "work women"

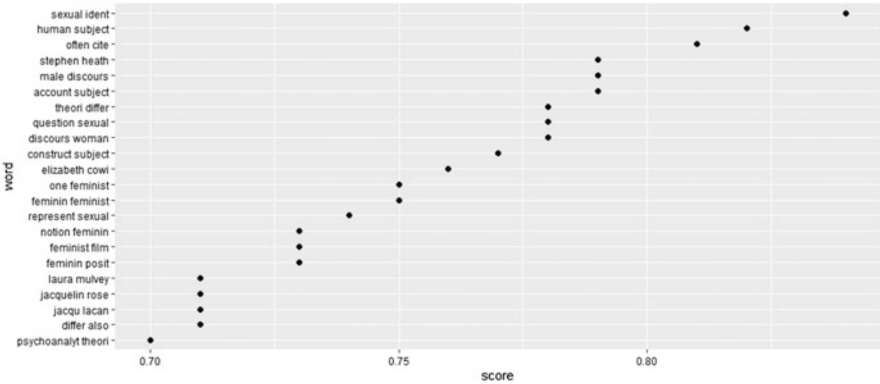


Fig. 11 Second wave association plot for "sexual differ"

The word associations like "sexual identity," "human subject," and "question sexual" are aligned with the issue of the female body being objectified in advertisements. Moreover, although marital rape and physical abuse were legally forbidden in this period guilty husbands were often not sentenced. The "male discourse" word association can be the reflection of this issue underlined by 2nd wavers (Fig. 11).

4.3 Emotional Framework of the First Two Waves

To establish the emotional framework of two consequent waves and show the changes among the waves, three of the most influential works from the related eras were selected. As a book of choice, "Vindication of Rights of Women" and "A Room of One's Own" are selected to represent the first wave as they are one

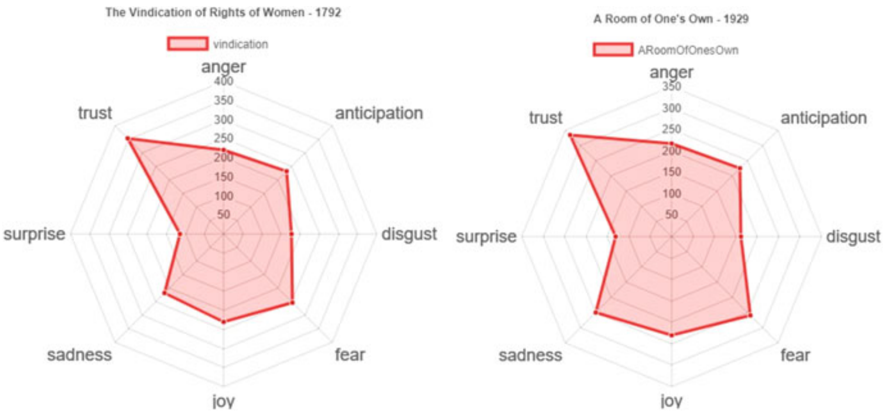


Fig. 12 Sentiment analysis on first wave books

of the most memorable books of their era, one starting the actual movement and other for bringing it to a new level. For the second wave, "The Feminine Mystique" by Betty Friedan is selected as it is considered to be the one that had started the second wave of feminism by introducing brand new ideas within the existing feminism framework. Although there may have been several different works having different emotional structures, it is assumed that the selected books were the most representative of their eras solely because of the impact they had on the overall movement.

Upon investigation, it is clear that both of the first wave books carry similar patterns in terms of emotions with "trust" being the densest emotion for both of the books. In terms of anger and anticipation, both works remain at lower levels. However, it could be also stated that "A Room of One's Own" carries a more melancholic ambiance within with higher sadness and fear components. Also, when examining the radar chart according to principles of Plutchik's wheel, the high levels of trust and fear within the second text combines into the emotion of submission, indicating that the second text has a more melancholic atmosphere in comparison. However, since both have low levels of surprise elements, it could be stated that the writers do not have any horror towards their environment highlighting their determination towards their path (Fig. 12).

When compared to the first wave, it is clear for the "Feminine Mystique" that the anger and anticipation level within the text had been increased indicating more aggressiveness of the word choices from the writer. Since with the second wave the ideas became stronger while also the injustices were communicated more freely, this type of aggressiveness is expected from the text. Once again as in all other texts, this one also has low surprise levels while having more sadness component compared to the others. By looking at the changes among the two waves, it could be referred that the tensions are rising within the feminist community, moving from the first wave to the second (Fig. 13).

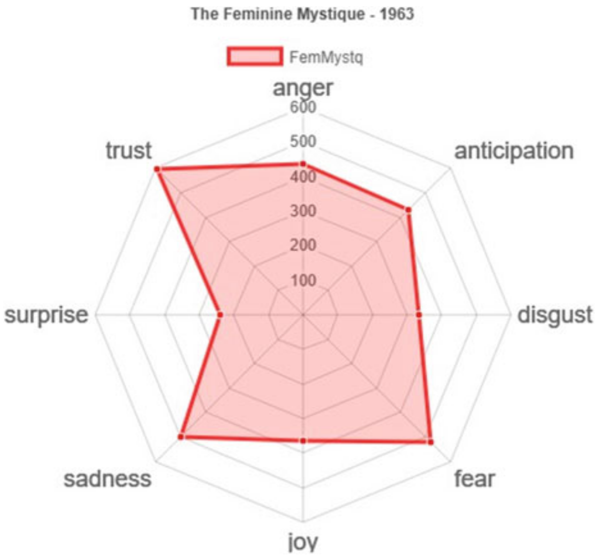


Fig. 13 Sentiment analysis on Feminine Mystique

4.4 Third Wave of Feminism

4.4.1 Association Plots for the Third Wave

Before dividing the third wave documents under several clusters, to obtain the overall picture, association plots were drawn using several distinct keywords like in the previous sections to show which topics are associated with the selected keywords.

Before dividing the third wave documents under several clusters, to obtain the overall picture, association plots were drawn using several distinct keywords like in the previous sections to show which topics are associated with the selected keywords.

The most striking term among “Middle East”s correlated words was “honor crime” which means violence or murder committed by male family members to defend their reputation. This finding puts forward that Middle Eastern Women have concerns about “living rights” which is the essence of human rights. Another interesting word was “Egyptian center” which might be associated with the Egyptian Center for Women’s Rights (ECWR). ECWR is a non-governmental organization working to support women’s efforts for defending their rights. In the books and articles used for the research preventive actions taken with the help of this organization might be mentioned (Fig. 14).

“Queer theory” was highly correlated with “epistemology closet” which is the name of the book “Epistemology of the Closet” written by Eve Kosofsky Sedgwick who is considered as one of the founders of queer theory. “Publish Butler” might be

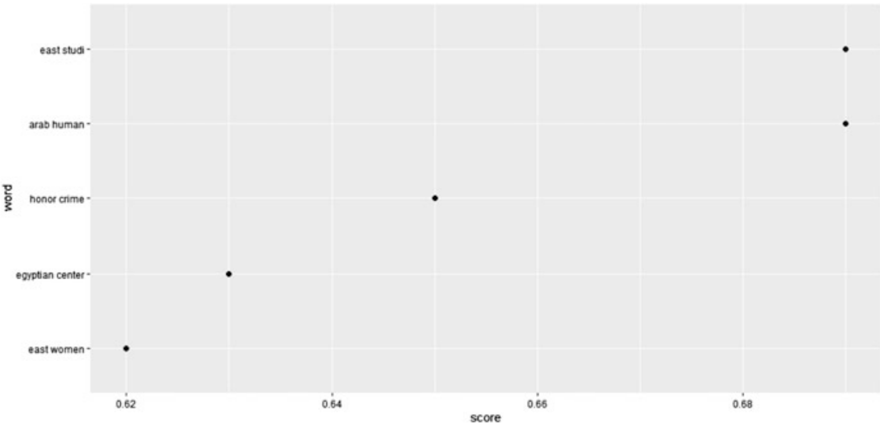


Fig. 14 Third wave association plot for "middle east"

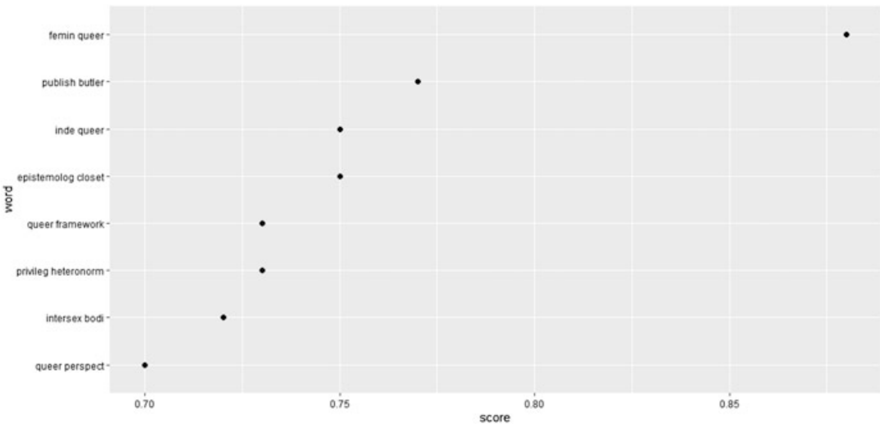


Fig. 15 Third wave association plot for "queer theory"

associated with Judith Butler who is also one of the early queer theorists. "Queer perspective" and "Queer framework" reflect 3rd wavers effort for redefining gender equality and make it about all groups that have been suffering from society's double standards. Moreover, "Priveleg heteronorm" is associated with the queer theory's criticism towards heteronormativity. Heteronormativity is the acceptance of heterosexuality as the default and only way of sexual expressions and it was questioned by queer theorists. "Intersex body" was another found association that is aligned with the queer theory being about challenging what is sex-based binaries and defending free expression of human sexuality for everyone (Fig. 15).

Although suffrage was a resolved issue in the 2nd wave of feminism, these ideas were still being discussed by the 3rd wavers. "Women suffrage" was strongly associated with Harriot Stanton Blanch who was a women's rights activist. The terms

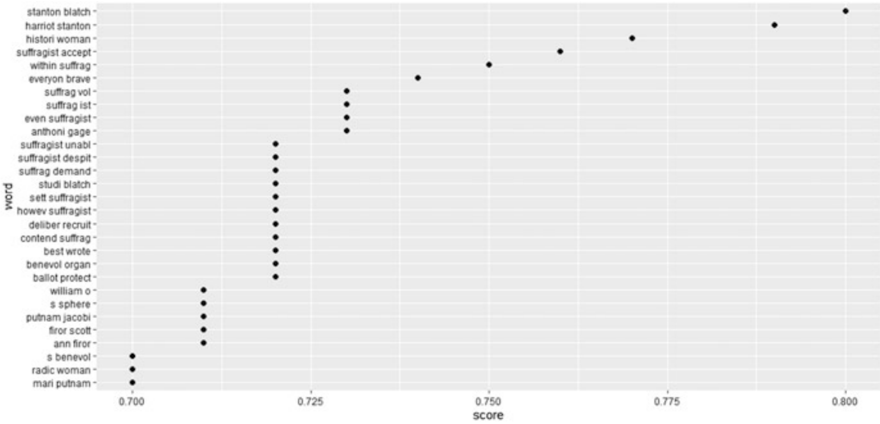


Fig. 16 Third wave association plot for "women suffrage"

"Suffragist accept," "everyon brave," "sufrag demand" express the determination of suffragettes when defending women's right to vote. Studies about this topic might be used by 3rd wavers for further researches about women's right to equal education and employment (Fig. 16).

Being one of the most heavily motives of 3rd wavers "birth control" was found correlated with "public contracept" which is not surprising because it is another term used for "birth control." "Women diaphragm" was an earlier birth control method that works by forming a physical barrier to prevent pregnancy. Moreover, terms like "mother confid," "support birth," and "decis latter" are linked with the era's belief of women should be able to freely choose to be a mother or not. 3rd wavers strongly supported women's freedom for sexuality and pregnancy prevention, these findings are aligned with these discussed topics. Another interesting association found was "two physician" and it might indicate the activists involved in the invention of the birth control pills' process. Margaret Sanger and Katharine McCormick are still being remembered as "mothers of birth control." Further, the terms "open clinic," "area clinic," and "clinic threaten" might have a linkage with the era's concerns about the sufficiency of health services provided for women (Fig. 17).

Discrimination and double standards of society towards African American people continued in the 3rd wave era. "African American" term is correlated with the name Cynthia Ann McKinney who is the elected first black woman to represent Georgia in the House. She criticized the Bush Administration heavily on various matters. Another associated word is "Combahe collect" which actually represents a Black feminist lesbian socialist organization called "Combahee River Collective." The organization was active from 1974 to 1980 and operated in Boston. Moreover, the expressions like "encourage African" and "among African" might be associated with researches taking successful African American people as a reference to encourage and motivate African American people. Another interesting correlated

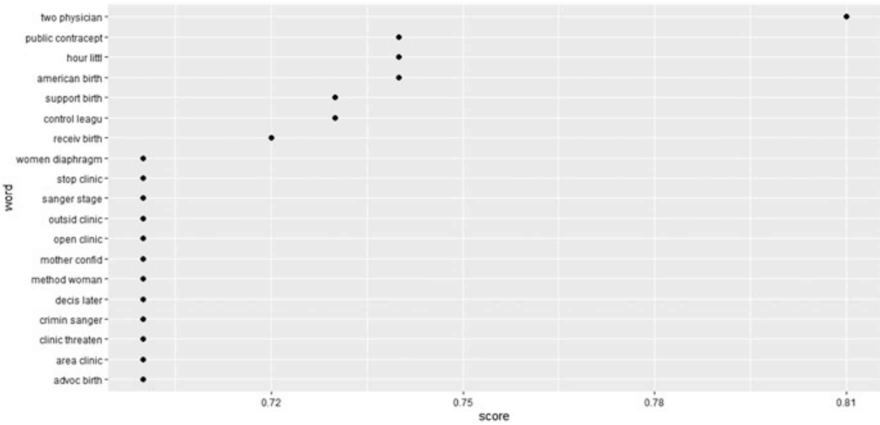


Fig. 17 Third wave association plot for "birth control"

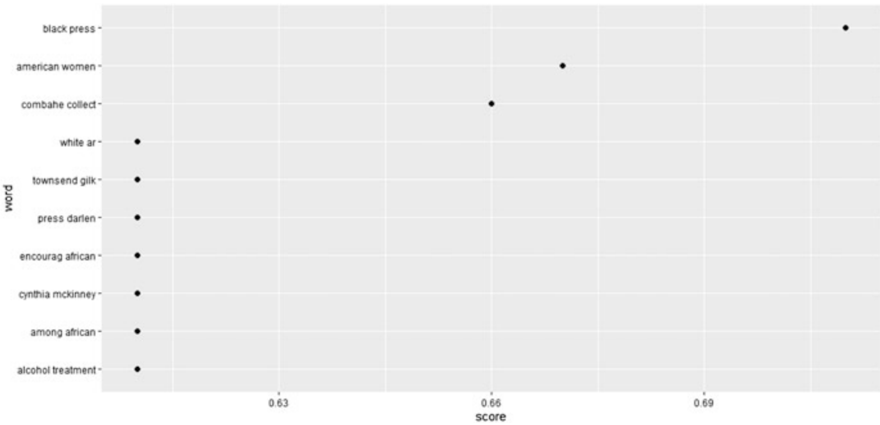


Fig. 18 Third wave association plot for "African American"

word was "alcohol treatment," this might reflect findings of reaching problems of African American people to alcohol and drug treatments (Fig. 18).

"Sexual assault" is another covered issue by the 3rd wavers and it is found highly correlated with the terms like "rape boy," "import consent," "rape culture," and "end rape." Those findings put forward that women being more open to share their experiences and raise their voices to search for justice. Consent was also a covered topic and its discussion is highly critical to define the border between normal sexual activity and sexual assault. Elliot Rodger who was the misogynist murderer in the 2014 Isla Vista massacre incident was a mentioned name, this case might have been used as a reference in researches. Another interesting association was "buddi system," which is a method used by women when going out to prevent sexual

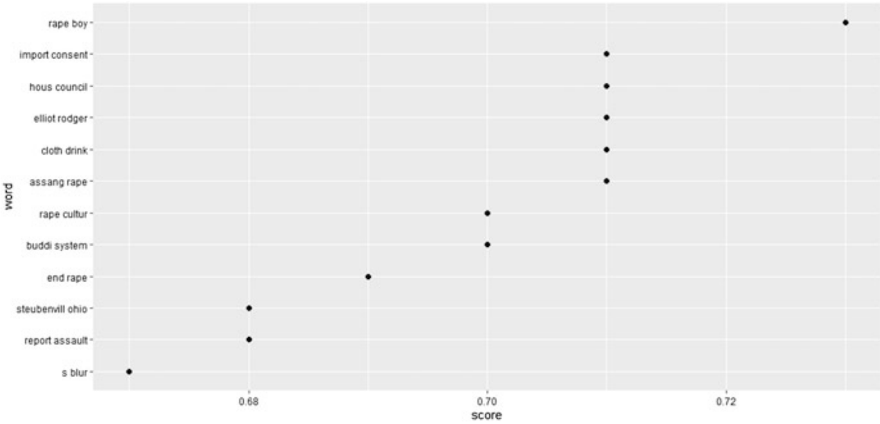


Fig. 19 Third wave association plot for "sexual assault"

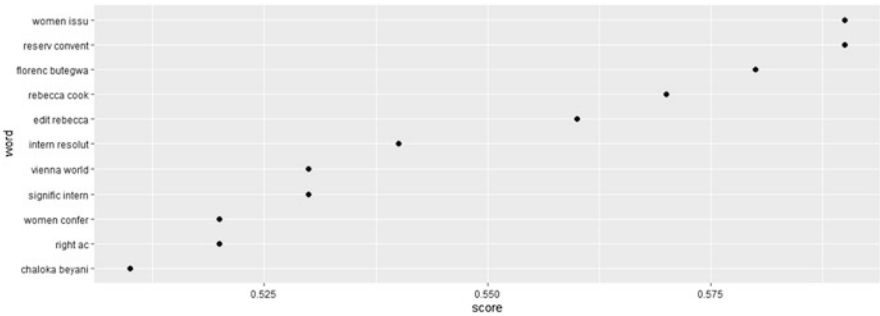


Fig. 20 Third wave association plot for "women right"

assaults. Violation and sexual assault is an ongoing issue discussed since earlier waves of feminism (Fig. 19).

"Women right" was found highly correlated with the name Rebecca Cook who is a researcher who contributed to the legal literature on women's rights. Moreover, the terms "Reserv convent," "Vienna world," and "women conference" are associated with the organization of women to get the idea of "women rights as human rights" accepted in "The United Nations World Conference on Human Rights in Vienna" in 1993. After this conference the linkage between feminist and human rights movements became unbreakable. Chaloka Beyani as also another mentioned name who is the UN Special Rapporteur on human rights. With all being said, it is clear that in the 3rd wave's era, "women's rights" is a discussed topic in various serious mediums and a legal framework is prepared on-demand to consider it as "human rights" (Fig. 20).

In conclusion, during the third wave, feminism had reached the masses like it did not before which also reflects the commonly stated topic that is provided above. As the media become mass spread, the feminist thoughts had expanded to cover up the problems of all of the minority groups while also including their problems into the feminist agenda which also corresponds to the sole description of the third wave.

4.4.2 Clustering and the Sentiment Analysis

Since the third-wave feminism ideology is no longer just about women but many other minority groups, there are various ideas to analyze. Moreover, most of the subject data is obtained from third-wave books and articles. Thus, to put forward all dimensions of this wave accurately, hierarchical clustering was applied to prevent the researcher bias on the topic. As a result, the cluster number is selected as 10 to have sizable clusters. However due to them having not enough sample size cluster 1 and cluster 3 had not been included within the discussions.

When investigated the word cloud in Fig. 21, it is observed that the main idea around the general topics had not been changed as much going from the second wave into the third wave. The concepts are still revolving around the women's effort to seek justice and equality around the concepts such as "human rights," "gender equality," and "social work." Moreover, violence was also a considered issue which puts forward it is an ongoing problem since earlier waves of feminism. Other interesting included terms were "economic development" and "per capita" which may be associated with issues of women from underdeveloped or developing countries. As it will be observed within the other clusters as well, it could be stated that the feminist movement had reached the global masses, which could be observed by the inclusion of the terms belonging to a geographical location like South Africa or the Middle East. Also, considering the previous radar chats and the sentiments included within the texts, the anger levels are observed to be increasing meaning that the aggressiveness of the texts is also increasing. It is not that surprising considering that the main ideas and areas of the problem remain the same throughout the eras. Not finding a solution to the problems which are spoken throughout decades is reflecting onto the writers in a form of growing levels of frustration.

For the cluster in Fig. 22, starting from the most frequent words which are "Woolf s" and "Virginia Woolf" which clearly illustrate that the works published by her protect their importance to this day as it is cited by the writers of the third wave. However, for this word cloud less frequent terms yield more interesting results. The inclusion of "sexual assault," "rape murder," "rape cases," and "get lost" words puts forward women's violence is a serious ongoing problem for the third wave era. However, it is important to note that the cluster includes much less anger and disgust as a general emotional framework compared to cluster 2 meaning that rather than being activating, the works falling under this cluster play more of an informative role, speaking about the general threats the women in society faces. The informative nature of the cluster also shows itself with the inclusion of other interesting terms like "economic policy," "monetary fund," and "wall street" which may be associated with the 2008 financial crisis and its reflections on women's life.

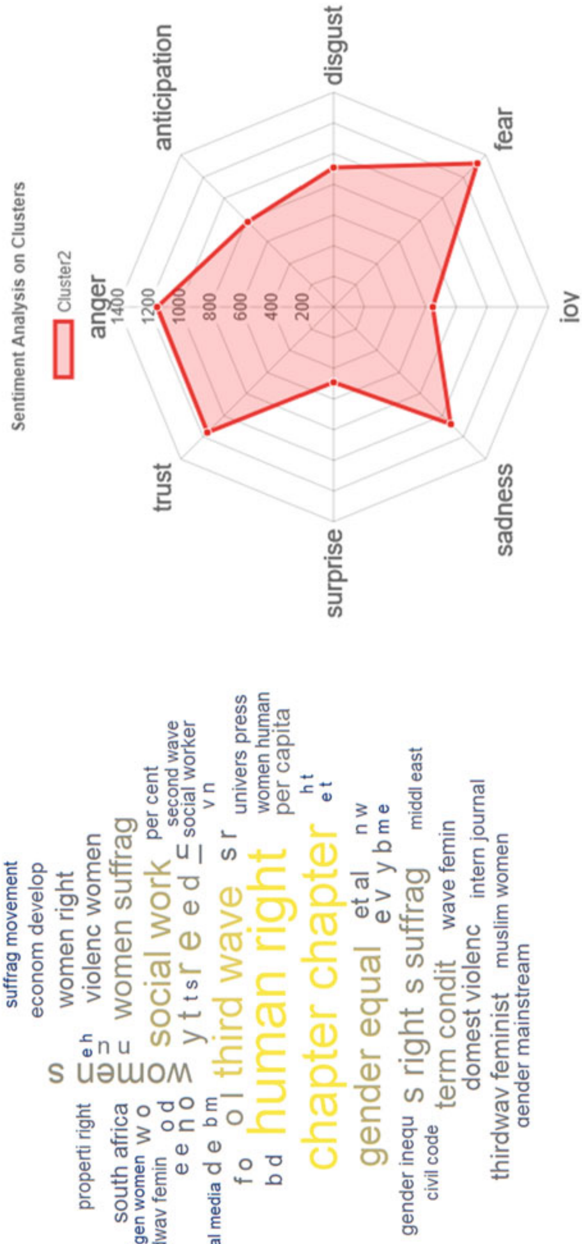


Fig. 21 Word cloud (left) and radar chart (right) for cluster 2

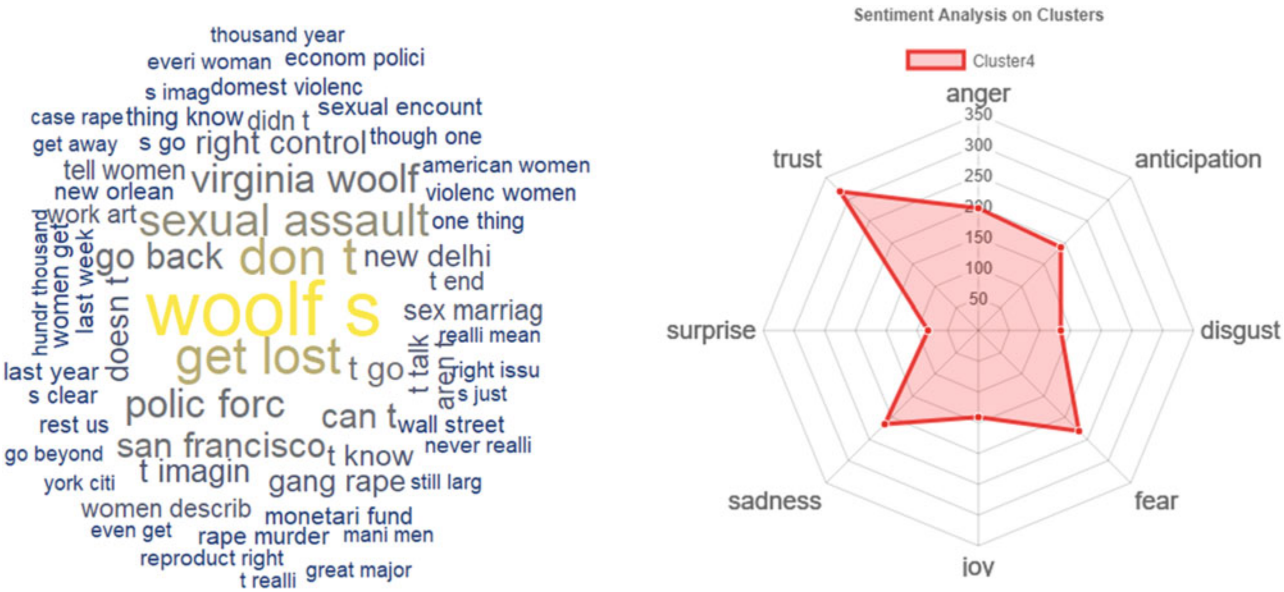


Fig. 22 Word cloud (left) and radar chart (right) for cluster 4

The fifth cluster includes “human rights,” “African American,” and “equal protection” which clearly shows the continuing discriminative attitude towards African American people. Still, as the discrimination towards the African American people was also being talked about within the third wave, an increase within the overall anger and disgust levels is observed for the works under cluster 5. According to Plutchik’s Wheel of Emotions, both of them together increase the feeling of contempt within the text itself. With the growing frustration at the feminist side, the insults may have been occurring within the text itself both coming from the writer itself or from the discriminative society that had been acting as such for centuries. Moreover, the name “Margaret Sanger” indicates that the right of birth control was also a discussed issue by the 3rd wavers differing from the common topics of the previous waves. To generalize, the texts under the cluster 5 address issues like women’s rights, the discrimination within contemporary society as it also includes social media as a frequent term. The works under the cluster are more aggressive and assertive towards their wants rather than being informative and neutral towards the subjects (Fig. 23).

The sixth cluster includes “new woman” which is the feminist ideal of the era. The works under this cluster issue most of the contemporary topics as content as “Human rights” was again the most frequently discussed issue caused by the search of justice and equality for not only women but all disadvantaged minority groups. The works include terms such as “lesbian” and “gay” which is aligned with 3rd wave feminism being a broader term covering minority groups including LGBTQ+. It shows that starting for arguing equal rights for women, feminism had started to shift in favor of many more disadvantaged groups within the society and starts to establish the idea of equality for all regardless of their race, gender, or sexual orientation. Moreover, the queer theory was another included term which is a critical discourse developed in the era. Once again, in terms of emotions, the anger and disgust levels are at high this time indicating the general behavior towards the LGBTQ+ community and the daily insults they are facing. However, the overall cluster has slightly less sadness level indicating the writers are expecting the change to come from the society with the less included feeling of “remorse” according to the wheel of emotion (Fig. 24).

In the seventh cluster, “Human rights” and “Gender equality” were again discussed issues. Inclusion of “percent” may be associated with 3rd wavers chance of conducting researches with numeric data. Alongside the commonly discussed topics of women’s rights, this time the texts within cluster 7 expand the coverage of the movement by talking about several distinct regions and the laws applied within those regions. The Islamic terms also found a place within the works under cluster 7 as the general subjects of feminism are investigated within the Islamic society. An important thing to notice is that the LGBTQ+ members are not included within these works although being the most disadvantaged group within those regions. This might be an indicator that some of the ideas under those regions had not been matured enough to be understanding towards half of the population let alone talking about the rights of the LGBTQ+ minorities. In terms of sentiments, just like the

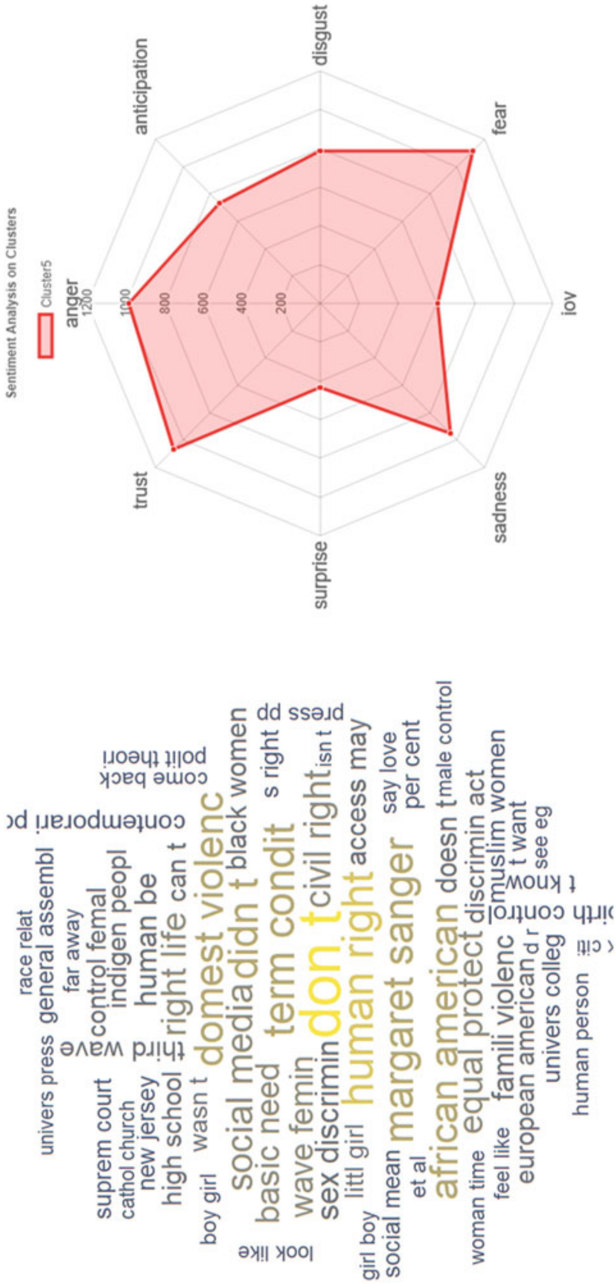


Fig. 23 Word cloud (left) and radar chart (right) for cluster 5

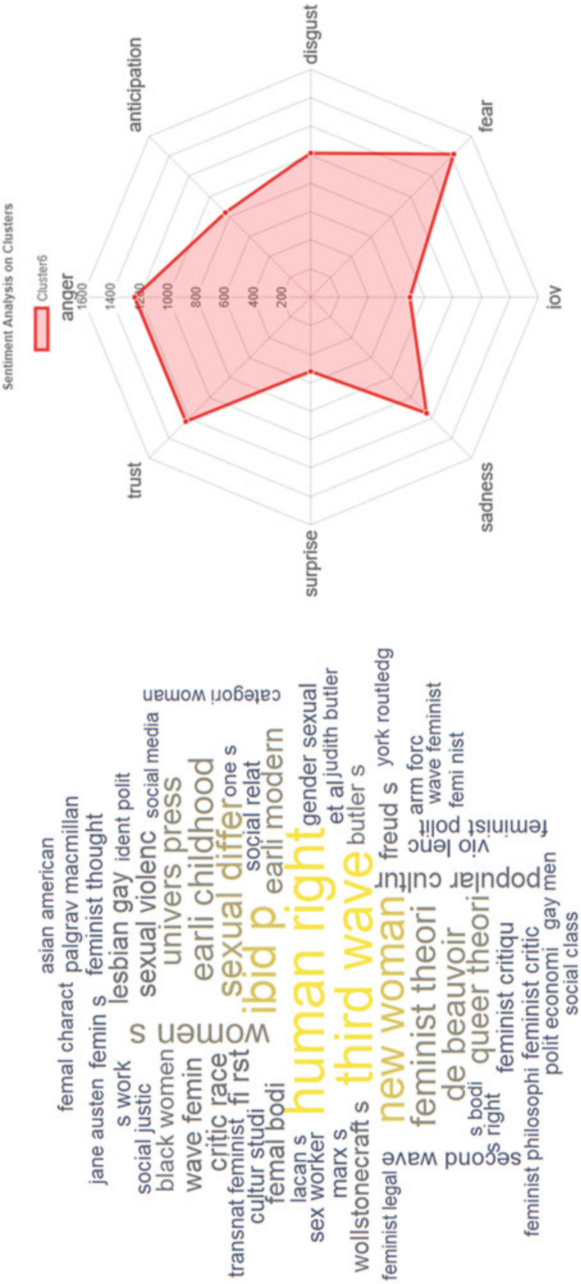


Fig. 24 Word cloud (left) and radar chart (right) for cluster 6

other clusters talking about the common rights the texts are aggressive and assertive about their subjects (Fig. 25).

Eight clusters in Fig. 26 include words such as, “black woman,” “white woman,” “black men,” “American women” which are aligned with the conducted researches that put forward the contrast of attitude towards different ethnicities by 3rd wavers. Although having similar anger levels as previous ones, it has slightly lower disgust and fear levels meaning that the texts have less submissive natured words within increasing the assertiveness of the texts within the cluster. “Woman suffrage” was also a covered matter and the name “Cady Stanton” might be associated with a group of 3rd wavers who embraced her ideas when discussing women’s right to vote. The media has found a place within the topics of cluster 8 as well, showing the effect of the media and the place it has within the conversations (Fig. 26).

The ninth cluster mostly indicates the daily issues of women’s lives as it includes the subjects as beauty myths and social standards. The terms “the mother says” and “mother said” were included twice which may indicate that 3rd wavers are not adopting their mothers’ beliefs and thoughts. Another important point is, different from the other clusters the documents include negative terms such as “didn’t,” “don’t,” “wasn’t,” “couldn’t,” etc. more often which carry close resemblances with the first wave word cloud. This means that cluster 9 includes more deeply rooted issues within the feminist thought which had not been able to find any solutions with time. It seems that things women can and cannot do stay to be one of the primary conversation topics of feminism even after a century. Therefore, it is not surprising to see such aggressive emotional structure within the texts (Fig. 27).

In the tenth cluster, the inclusion of “black women,” “African American women,” “Muslim women” is again aligned with 3rd wave feminism being not only about women but all religions and ethnicities that are wronged by society. “Population control,” “Birth control,” and “Reproduction health” words illustrated the era’s other important topics like the right of abortion and the right of birth control. Further, “Population growth” and “Climate change” were also covered matters which may reflect the researches associating women’s issues with global concerns. In general, cluster 10 talks about feminism as a reflection of contemporary societal problems. It is also seen that the addition of new terminologies within the feminist literature like “women of color” also happened within the third wave. The cluster has, in general, lower trust values decreasing the level of submission within the text itself (Fig. 28).

5 Discussion

When the word clouds and association plots were analyzed it is seen that they are highly aligned with the findings in literature review. However, although anti-racism was also a topic raised by the first wave’s activists there were not any significant findings on this topic on our research. Although they may have been in some of the topic that were discussed, still discussion of these were not widespread enough to reflect on the word cloud. Moreover, contraception (birth control) was a covered

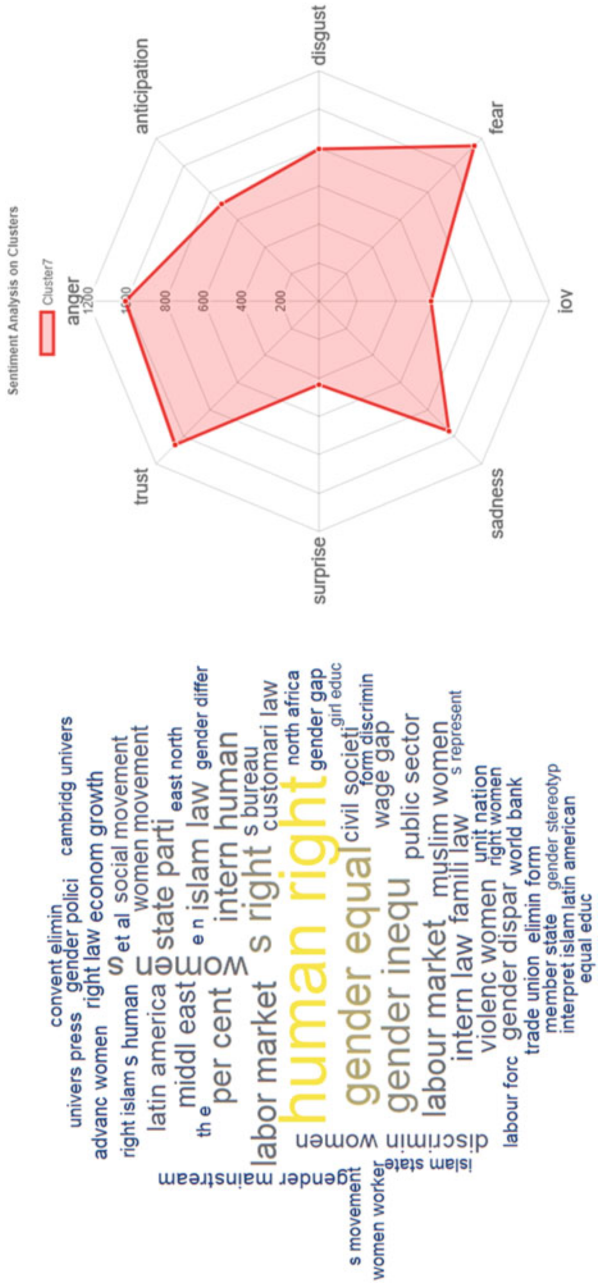


Fig. 25 Word cloud (left) and radar chart (right) for cluster 7

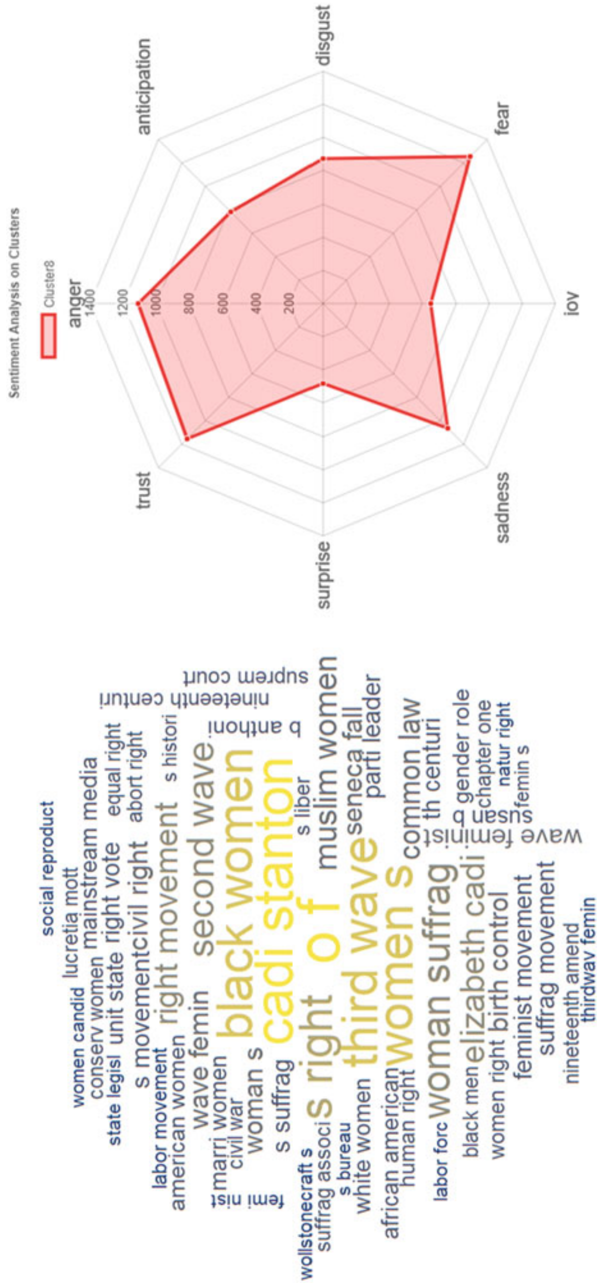


Fig. 26 Word cloud (left) and radar chart (right) for cluster 8

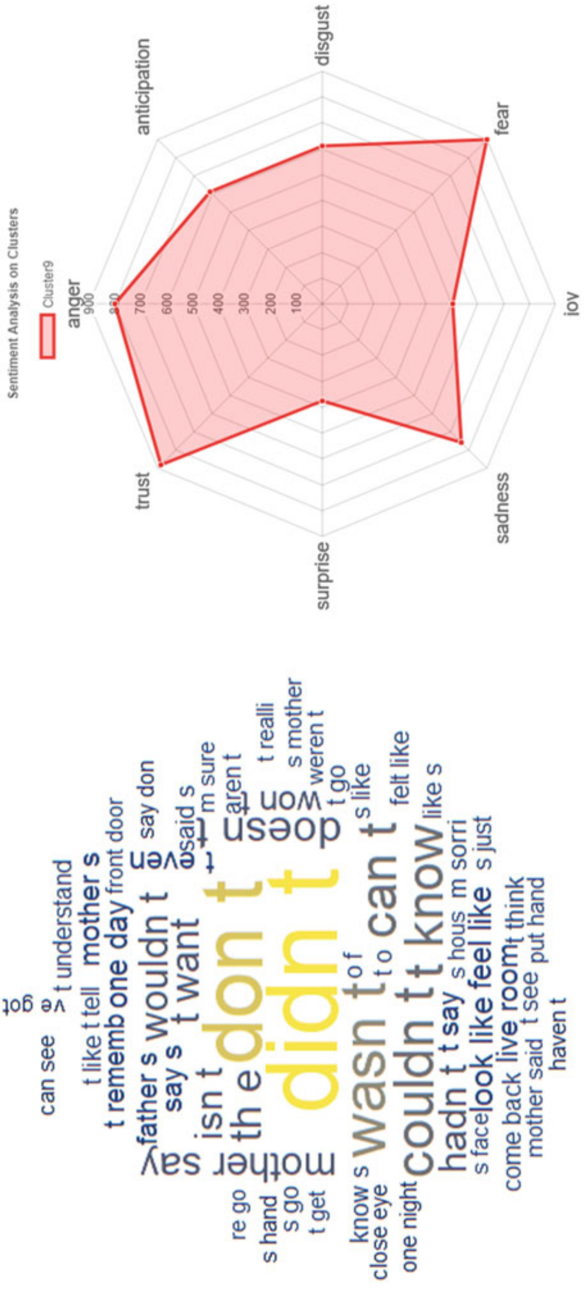


Fig. 27 Word cloud (left) and radar chart (right) for cluster 9

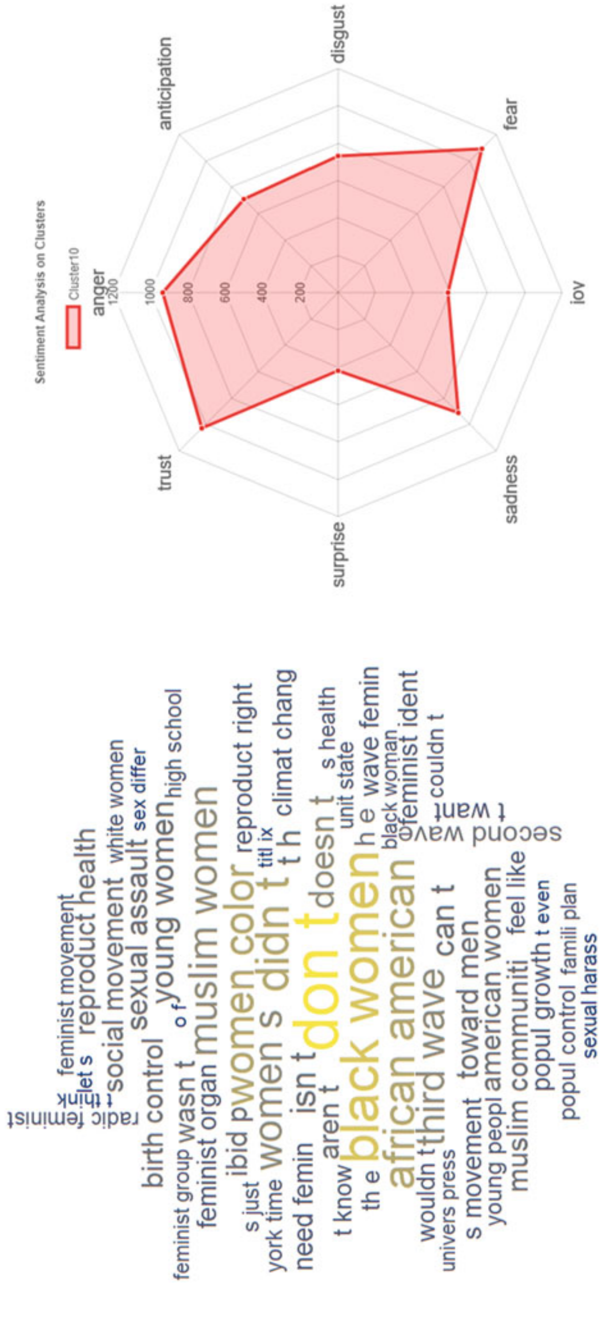


Fig. 28 Word cloud (left) and radar chart (right) for cluster 10

matter by 2nd wavers but there were again no related findings when documents were analyzed. This might show that discussions of some sensitive topics are postponed to be put forward in written documents. In contrast, there were no uncovered topics in word clouds and association plots of 3rd wave. Even the most controversial themes were discussed openly in the era's books and articles. This alone shows the progression within the society in comparison to the early stages where the topics were only consisted of obtaining basic human rights. However, still the human rights is one of the widely discussed subjects meaning that the society has more ways to go in terms of equality for all.

If the sentiment analysis for all of the waves has been regarded as a whole, the primary thing to notice is that no matter the topic or the era is, all of the texts have high trust values which in collaboration with the feeling of joy creates love. This shows that feminism had imposed the feeling of self-love on the disadvantaged groups within the society. Although some emotions vary according to the topic of interest, the feeling of self-love is always intact with the ideology which is also evident from the inclusion of other minority groups within the feminist thought with the third wave. The idea of self-love had spread to cover up all of the groups where the main emphasis is being human.

The aggressiveness within the texts, however, is an ongoing phenomenon that had increased significantly as time passes for a specific group of subjects. This is mainly due to the frustration of the feminists coming from the never-changing acceptance problems within the society. It is evident from the word clouds created during the study that although the subject group changes the main idea behind feminism stays the same: Equal rights for all. Since the topic had stayed the same for all of these eras, and that topic being the primary human right, it is not surprising to be creating such an aggressive wording in return. Therefore, it is important to understand the underlying cause for the behaviors some might call provocative during the feminist protests. Since the feminists are fighting for so long for such basic rights, it is not that hard to establish empathy with those groups. The findings indicate that the aggressive tone that is widely propagandized by the anti-feminist is almost all related to the basic human rights that is an ongoing structure since the early feminist movements had started. Also, another fact is that not all of the feminist texts share the common aggressiveness. For the topics that are relatively fresh, the tone of the texts is as calm as the ones compared with the first wave texts. They are informative rather than assertive since the claims need to be communicated to the masses.

6 Conclusions

During this research, the development of topics and emotional characteristics of feminism had been analyzed under the three consequent waves using the natural language processing principles where most of the emphasis is put on the third wave of feminism. Although the topics and the subject groups had been changed over time, it is observed that the main idea behind feminism had remained the same.

With time, the idea once started for obtaining equal rights for women had evolved into a more inclusive subject issuing equal rights for all of the disadvantaged groups on a more global scale. In terms of emotions, however, most of the works within the feminist framework had become more assertive and somewhat more aggressive towards the general topics while the idea of self-love remained the same throughout the waves.

As for the limitations of the study, since most of the earlier documents were scanned for digital storage, some of their contents had been lost during the conversion to a text format. For improving the study on earlier waves, this issue should be addressed to obtain more clear results and better word clouds. As some of the feminist literature also talks about the potential fourth wave that is currently happening, time-based topic changes might be investigated as further research to determine how much the topics had been changed within the third wave itself and if the contemporary works require new classification under the so-called fourth wave. Also, since the study had focused on the result of two centuries-old fights, it is as important to know how and under what instances such developments had occurred within the society. By combining the text mining principles that had been used in this research to match the contents with the political and other societal movements on a time and region basis, the catalysts of the social change could also be determined.

References

Annapurany, K. (2016). A prospective study of feminism waves, phases, issues and critical analysis. *International Journal of Applied Research, 2*(5), 424–426. ISSN Print: 2394-7500, ISSN Online: 2394-5869.

Burdett, C. (2001). *Olive Schreiner and the progress of feminism evolution, gender, empire*. Palgrave. isbn:978-1-349-39260-5.. https://doi.org/10.1057/9780230598973

Cobble, D. S., Gordon, L., & Henry, A. (2014). *Feminism unfinished: A short, surprising history of American Women's movements*. WW Norton & Co. isbn:0871406764. https://doi.org/10.1111/j.2050-5876.2014.00812.x.

Correia, A., Filomena, T., & Lobo, V. (2018). Statistical Methods for Word Association in Text Mining, Recent Studies in Risk Analysis and Statistical Modeling. Contributions to Statistics, Springer, Cham, 375–384

David, C. (1992). *An Encyclopedic Dictionary of Language and Languages*. Oxford:Blackwell.

Desagulier, G. (2014). Visualizing distances in a set of near synonyms: rather, quite, fairly, and pretty. Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy, 145–178

Dyhouse, C. (2010). *Glamour - women, history, feminism*. Palgrave Macmillan. isbn:9781848134072.

Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure in R. *Journal of Statistical Software, 25*(5), 1–54. https://www.jstatsoft.org/v25/i05/.

Freedman, E. B. (2002). *No turning back -the history of feminism and the future of women*. The Random House Publishing Group. eISBN 9780307416247.

Jacob, K., & Licona, A. C. (2005, Spring). Writing the waves: A dialogue on the tools, tactics, and tensions of feminisms and feminist practices over time and place. *NWSA Journal, 17*(1), 197–205, The Johns Hopkins University Press.https://www.jstor.org/stable/4317111.

Loomba, A., & Sanchez, M. E. (2016). *Rethinking feminism in early modern studies gender, race, and sexuality*. Routledge. isbn:9781472421753.

McNamee, P., & Mayfield, J. (2004). Character N-gram tokenization for European language text retrieval. *Information Retrieval, 7*(1/2), 73–97. https://doi.org/10.1023/b:inrt.0000009441.78971.be

Mohammad, S. M., & Turney, P. D. (2010). Emotions evoked by common words and phrases: Using Mechanical Turk to create an emotion lexicon. In *Proceeding of workshop on computational approaches to analysis and generation of emotion in text*, 26–34.

Murtagh, F., Legendre, & P. (2014). Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion?. Journal of Classification 31, 274–295

Offen, K. (1988, Autumn). Defining feminism: A comparative historical approach. *The University of Chicago Press Journals, 14*(1), 119–157. https://www.jstor.org/stable/3174664

Rampton, M. (2008). *Four waves of feminism*. Pacific Magazine, Pacific University, Oregon. https://www.pacificu.edu/magazine/four-waves-feminism

Riley, D. (1995). *"Am I that name?" Feminism and the category of "women" in history*. University of Minnesota Press. isbn:0816617309.

Roseneil, S. (2013). *Beyond citizenship? Feminism and the transformation of belonging - citizenship, gender and diversity.* Palgrave Macmillan. isbn:9781349340255. https://doi.org/10.1057/9781137311351.

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management, 24*(5), 513–552.

Savigny, H., & Warner, H. (2015). *The politics of being a woman feminism, media and 21st century popular culture*. Palgrave Macmillan. isbn:9781137384652.

Silge, J., & Robinson, D. (2017). *Text mining with R: A tidy approach.* O'Reilly Media.

Suttles, J., & Ide, N. (2013). Distant supervision for emotion classification with discrete binary values. Computational Linguistics and Intelligent Text Processing, 121–136. https://doi.org/10.1007/978-3-642-37256-8_11

Teays, W. (2019). *Analyzing violence against women*. Springer Nature Switzerland A.G. isbn:9783030059880.. https://doi.org/10.1007/978-3-030-05989-7

Walters, M. (2005). *FEMINISM a very short introduction*. Oxford University Press. isbn:9780192805102.

A Comparative Analysis of Forecasting Models on COVID-19

Müjde Erol Genevois and Michele Cedolin

Abstract The COVID-19 spread all around the world, causing more than a million deaths and reaching over 50 million confirmed cases. A forecast of these numbers is vital for the adequate preparations of health care capacities and for the governments to take the necessary decisions. In this study, it is aimed to predict the evolution of COVID-19 figures, employing alternative statistical models such as the Holt-Winters, ARIMA, and ARIMAX while using the time series corresponding to different parameters of this disease such as daily cases, daily deaths, and the stringency index. Considered are the John Hopkins University epidemiological world data and the top ten countries with the highest cases, along with China. The fitting of the time series and the upcoming 10 day~~s~~ projections resulted in a high level of accuracy, presented with alternative error metrics and comparisons between the situations of countries. Holt-Winters is the best performing model, while ARIMAX gives the worst accuracy results. Moreover, it was found that the use of coefficient determination and Bayesian information criterion alone are not suitable, and scale independent metrics should be employed when the data ranges differ. The results of this study would be useful to set up benchmark results for other studies and the projections may be used for medical, economic, and social precaution and preparation.

Keywords COVID-19 · Forecasting · Epidemiological forecasting · Holt-Winters · Econometric models · ARIMA · ARIMAX

M. Erol Genevois · M. Cedolin (✉)
Galatasaray University, Ortaköy/İstanbul, Turkey
e-mail: merol@gsu.edu.tr; mcedolin@gsu.edu.tr

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_8

1 Introduction

Since December 2019, the world is in combat with COVID-19, which started in Wuhan, China and spread to more than a 100 countries. According to the data collected by the World Health Organization (WHO)[1] on February 14th, 2021, there have been 108,153,741 confirmed cases of COVID-19, with 2,381,295 cases resulting in death~~s~~. While some countries are going through the second wave, some states started the vaccination process, and governments responded to the global pandemic with different measures.

Apart from the clinical researchers, academics approached the COVID-19 problem in different ways. While the pandemic spread and 1 year of living with COVID-19 passed, the social impact and economic aspects of the virus have become critical (Bruns et al., 2020). The diagnosis of the virus with the artificial intelligence image processing techniques became important (Bhattacharya et al., 2021). Similarities between the SARS and other epidemics were investigated (Peeri et al., 2020). Some part of the studies concentrated on estimating the cases and deaths per country, and a significant forecasting literature was formed.

Because there was no data available at the start of this epidemic, predicting and projections were difficult. However, the spread of COVID-19 is highly dangerous and necessitates strict plans and government policies. Therefore, forecasting confirmed cases and deaths in the future days is crucial in order to manage health resource capacities and put in place the necessary protection procedures. Consequently, this study tries to apply alternative forecasting models for the daily reported COVID-19 confirmed cases and deaths of the most affected 10 countries and China. It employs, namely Holt-Winters, ARIMA, and ARIMAX models, providing accuracy results in alternative error metrics.

The rest of the study is organized as follows: The second section consists of literature review. The third section presents the employed methods with the application, and the fourth section gives the concluding remarks and discussions.

2 COVID-19 Forecasting Literature Review

Forecasting the outcomes of a pandemic is important for governments in order to take the necessary restriction measures while preparing the appropriate health infrastructure (e.g. intensive care units for COVID-19 cases). The countries shared their data despite their discordance, with the public and WHO. Many researchers employed this data (worldwide or in specific countries) to forecast cases, deaths, and recovery numbers. In the last year, a solid forecasting literature was built where researchers alternated approaches such as machine learning approaches, statistical

[1] https://COVID19.who.int/

and epidemiological models. These articles focused on a selected country or a group of countries' case and death data (daily or cumulative), while some articles aimed to forecast the worldwide data generally in alternative forecasting horizons and training scales were written.

Al-Qaness et al. (2020) proposed an Adaptive Neuro-Fuzzy Inference System (ANFIS), which uses an enhanced flower pollination algorithm (FPA) with the help of the Salp Swarm Algorithm (SSA) to forecast the confirmed cases in China for the upcoming 10 days. Ankarali et al. (2020) employed ARIMA, Simple Exponential Smoothing, Holt's Two Parameter Model, and Brown's Double Exponential Smoothing Model to forecast 10 days of cumulative cases, cumulative deaths, daily cases, daily deaths, cumulative recovered and active cases of 25 countries, which exceed 1000 as cumulative cases in March 15. Ayinde et al. (2020) focused on the Nigeria data set and tried to forecast 2 months of confirmed cases, discharged cases, and death cases using classic, quadratic, cubic, and quartic versions of linear regression, logarithmic regression, logistic regression, and exponential linear regression. Ayyoubzadeh et al. (2020) predicted the COVID-19 cases in Iran using the Google Trends data. They employed the linear regression model and long short-term memory (LSTM) models and obtained a strong correlation for keywords like "hand sanitizer," "handwashing," and "antiseptic."

Benvenuto et al. (2020) employed ARIMA to forecast the next 2 days of COVID-19 confirmed cases and indicated that "if the virus does not develop new mutations, the number of cases should reach a plateau." Crokidakis (2020) employed a susceptible–infectious–quarantined–recovered (SIQR) model to estimate confirmed cases, ratio of infectious individuals, the reproduction number, and the epidemic doubling time of Brazil. Dandekar and Barbastathis (2020) built a neural network aided quarantine control model to test the impact of strict quarantine measures on the reproduction number in Wuhan. Their simulation results showed that rigid quarantine measures helped China on the new case numbers. Hernandez-Matamoros et al. (2020) built ARIMA models to forecast total case numbers per million, grouping countries according to their continents. Hu et al. (2020) used modified autoencoders for modeling the number of the cumulative and newly confirmed cases. They outlined the immense difference between the immediate and late interventions on total active cases and suggested a case ending time of January 10, 2021 under immediate aggressive interaction. Ibrahim et al. (2021) employed a variational Long Short-Term Memory (LSTM) autoencoder to forecast the spread of the coronavirus across the globe for the next day and 10 days ahead that employs historical data with the urban characteristics and stringency index measures. Ivorra et al. (2020) developed a new θ-SEIHRD model containing the characteristics of COVID-19, to identify the unknown parameters of the pandemic~~s~~ that fit the total cases of China, to estimate the reproduction rate, and to find the maximum number of hospitalized people.

Jia et al. (2020) employed a Logistic model, the Bertalanffy and the Gompertz model to estimate the new cases and death toll of China. Among these mathematical models, the logistic model is the best fitting-one. Kafieh et al. (2020) trained alternative machine learning models as random forest, multi-layer perceptron, LSTM with regular and extended features, and multivariate LSTM to estimate daily number of confirmed, death, and recovered COVID-19 cases. Kolozsvari et al. (2020) used recurrent neural networks with LSTM units to create prediction models of 17 countries' daily infection numbers per 100,000 habitants, outlining the effect of the repeated peaks of the epidemic. Kumar et al. (2020) employed ARIMA and Richard's model to estimate new cases, new deaths, and recovery rates of India. Liu et al. (2020) used related internet search activity in their combined mechanistic and machine learning model to estimate the real-time COVID-19 cases of the Chinese provinces. Liu et al. (2021) modeled the coronavirus in China, South Korea, Italy, Germany, and the UK, and under different scenarios, simulated their new cases. Milhinhos and Costa (2020) employed nonlinear regression to estimate the active cases and total deaths of Portugal and built a comparative model with South Korea, outlining the similarities. Pandey et al. (2020) employed SEIR and a regression model to predict the expected cases in India within 2 weeks.

Petropoulos and Makridakis (2020) employed exponential smoothing to forecast global confirmed cases, deaths, and recoveries with a forecasting horizon of 10 days. Roosa et al. (2020) used the generalized logistic growth model (GLM) and the Richards model to estimate 5-, 10-, and 15- days of cumulative cases of China. Sameni (2020) employed SEIR and the compartmental model to estimate the propagation. They simulated seven different scenarios and tried to find the reproduction and fatality rates of COVID-19. Xu et al. (2020) used the SEIQRPD model which divided the population into susceptible, exposed, infectious, quarantined, recovered, insusceptible, and deceased individuals to estimate the USA COVID-19 cases. Yang et al. (2020) used the SEIR model helped by a trained LSTM in SARS-2003, to predict COVID-19 peaks and sizes in China. Yonar et al. (2020) employed exponential smoothing and ARIMA to forecast the number of COVID-19 cases of the G8 countries. Table 1 summarizes the existing literature per country (forecasting target), the employed method, and the forecasting horizon.

As can be observed from Table 1, most of the studies focused on a single country data, with a forecasting horizon ranging from 2 to 15 days, while there are articles that aim only to fit the training data set. Along with the epidemiological models, regression models are widely used in the literature. Literature outlines that statistical models are simple but effective tools to forecast COVID-19 numbers with the highest proportion. Machine learning models such as LSTM or ANFIS, epidemiological models such as SEIR and combinations like SIR and SIQR are the other common approaches.

In this study, double exponential smoothing (Holt-Winters), ARIMA, and exogenous version ARIMAX models are employed to fit and forecast the daily case and daily death numbers of the selected countries and global data.

Table 1 COVID-19 forecasting literature

References	Country	Method	Forecasting horizon
Al-Qaness et al. (2020)	USA, China	ANFIS, FPA,SSA	10 days
Ankarali et al. (2020)	25 countries	ARIMA, exponential smoothing	10 days
Ayinde et al. (2020)	Nigeria	Linear regression models and versions	2 months (fitting)
Ayyoubzadeh et al. (2020)	Iran	LSTM, logistic regression	35 days (fitting)
Benvenuto et al. (2020)	World	ARIMA	2 days
Crokidakis (2020)	Brazil	SIQR	1 month (fitting)
Dandekar and Barbastathis (2020)	China	NN aided SIR	40 days (fitting)
Hernandez-Matamaros et al.	145 countries	ARIMA	15 days
Hu et al. (2020)	World	Modified auto-encoder	5 days
Ibrahim et al. (2021)	World	Variational-LSTM autoencoder	1–10 days
Ivorra et al. (2020)	China	θ-SEIHRD	1.5 months (fitting)
Jia et al. (2020)	China	Logistic, Bertalanffy, Gompertz models	2 months (fitting)
Kafieh et al. (2020)	Iran	RF, MLP, LSTM	10 days
Kolozsvari et al. (2020)	17 countries	RNN with LSTM	11–12 days
Kumar et al. (2020)	India	ARIMA, Richard's model	1 month
Liu et al. (2020)	7 countries	SIRU	2 days
Milhinhos and Costa (2020)	Portugal	Nonlinear regression	140 days (fitting)
Pandey et al. (2020)	India	SEIR, regression	2 weeks
Petropoulos and Makridakis (2020)	World	Exponential smoothing	10 days
Roosa et al. (2020)	China	Logistic growth, Richard's sub-epidemic wave models	5, 10, and 15 days
Sameni (2020)	USA	SIR	–
Xu et al. (2020)	USA	SEIQRP	2 weeks
Yang et al. (2020)	China	SEIR, LSTM	3 months (fitting)
Yonar et al. (2020)	G8 countries	ARIMA, exponential smoothing	10 days

3 Methodology and Application

3.1 Data Characteristics

The data employed in this study is available at the Coronavirus Research Center at the Johns Hopkins University website.[2] A total of 192 countries deal~~s~~ with the virus; however, in this study, only the first ten countries with the highest case numbers on February 9th, 2021 and China are considered. The first case dates differ among the countries, and the reporting process of these cases is somehow problematic. Therefore, the first day of the training set is selected as the day when cumulative cases reached "100" for each country, which is considered to be a more robust option. The training set length differs for each country and ends on January 30th, 2021. The remaining days are separated for the forecasting part. Table 2 shows the countries, with their cumulative case and death numbers, the initial date with data length and the fatality rate.

As is observed in Table 2, the countries have combatted the virus since March 2020. The USA is the most effected country by case and death numbers. The average fatality rate is 2.2%, while the maximum fatality rate is observed in China and Italy, 4.8% and 3.5%, respectively. The minimum fatality rates are in India and Turkey with 1.4% and 1.6%, which may be linked to the average population age of these countries. China is the virus-source (the virus' source) country. The virus spread after approximately 1 month to Europe, starting with Italy, France, the UK, Germany, and Spain. At last, it affected Russia and Turkey. The last affected countries had more time to prepare, while countries like Italy, which was the first effected, experienced more difficulties in the initial days of the spread of the virus. The countries show some similarities; however, they all have different COVID-19 wave~~s~~ lengths and population properties. In the appendix, charts belonging to daily case vs daily death numbers of the countries can be referenced to investigate these differences. To mathematically evaluate the resemblances between the case and death time series, a correlation test between the countries' data and world data is done. However, due to the initial day differences, the test is applied for the first 316 days of the pandemic. The outcomes for the new cases and new death~~s~~ correlations are as follows.

Table 3 shows that for most of the countries, a correlation between the country based new cases and worldwide new cases can be obtained; however, this hypothesis is not true for the daily death numbers of the countries. China is acting as an outlier in every aspect. The Spain data is corrupted since it contains negative values for the new deaths and new cases along with zeroes. In terms of the case numbers USA, UK, Germany, and Russia are highly correlated with the World. In terms of the death numbers, there is no country that is linearly correlated with the worldwide death numbers. To sum up, worldwide data is not an explanatory variable to yield better

[2] https://coronavirus.jhu.edu/map.html

Table 2 COVID-19 details of the countries

	Initial day	Total case	Total death	Data length	Fatality rate
World	28.01.2020	103, 554, 872	2, 268, 415	369	0.022
USA	7.03.2020	26, 558, 791	451, 416	330	0.017
India	18.03.2020	10, 790, 555	154, 701	319	0.014
Brazil	17.03.2020	9, 335, 247	227, 467	320	0.024
Russia	20.03.2020	3, 858, 234	73, 499	317	0.019
UK	4.03.2020	3, 844, 233	109, 308	333	0.028
France	3.03.2020	3, 307, 342	77, 628	334	0.023
Spain	5.03.2020	2, 871, 140	60, 161	332	0.021
Italy	25.02.2020	2, 586, 016	89, 768	341	0.035
Germany	4.03.2020	2, 253, 819	59, 642	333	0.026
Turkey	21.03.2020	1, 677, 723	26, 345	316	0.016
China	28.01.2020	98, 930	4783	369	0.048

Table 3 Correlation test results

	USA	**India**	**Brazil**	**Russia**	**UK**	**France**
New Case	0.917	0.203	0.693	0.811	0.836	0.517
New Death	0.494	0.232	0.636	0.742	0.179	−0.086
	Spain	**Italy**	**Germany**	**Turkey**	**China**	
New Case	0.696	0.663	0.832	0.547	−0.327	
New Death	−0.394	0.075	0.589	0.388	−0.395	

results when it comes to individual country data. The time series of the countries show obvious differences; therefore, they should be examined separately, and they should have their own model configurations. The figures in this study show the charts belonging to the worldwide data, and in the appendix the figures belonging to the other countries may be found. Figure 1 shows the New Case vs Daily Death numbers of the World and USA.

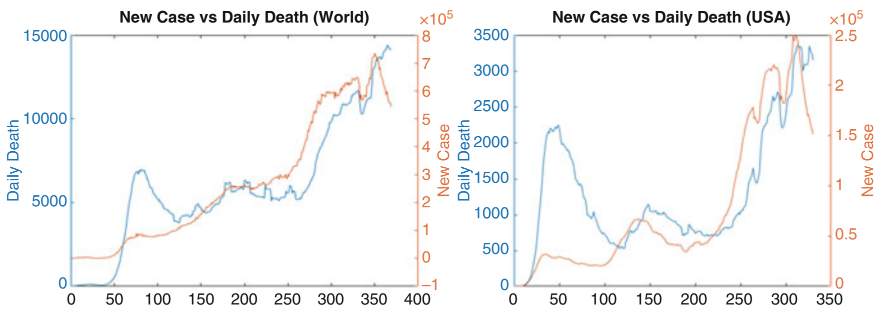


Fig. 1 Training data: left—World, right—USA

3.2 Error Metrics

Alternative error metrics are employed in the COVID-19 forecasting literature. All the statistical models based their study on R^2 which is the coefficient of determination that represents the proportion of the variance for dependent variable by the regression variable. RMSE and bic are the other error metrics that are used by Ankarali et al. (2020), Kumar et al. (2020), and Yonar et al. (2020). In this study, the results are provided according to these metrics. However, these metrics are scale dependent and are not suitable to compare the forecasting accuracies by countries. Therefore, the results are also shown in SMAPE and MAPE. The formulae of the employed metrics are provided next.

- *Bayesian Information Criterion (bic)*

The bic or Schwarz information criterion (SIC) is a criterion for model selection based on the likelihood function like AIC (Schwarz, 1978). The general notation is ~~as~~

$$\mathrm{BIC} = k \ln(n) - 2 \ln\left(\hat{L}\right) \tag{1}$$

- *Root Mean Squared Error (RMSE)*

RMSE or root mean-squared deviation (RMSD) is the square root of the averaged squared errors. It is scale dependent and highly sensitive to the outliers. When there is a set of time series, it is a difficult metric to interpret.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2} \tag{2}$$

- *Symmetric Mean Absolute Percentage Error (SMAPE)*

SMAPE is an accuracy measure based on percentage errors where the absolute difference between the A_t and F_t is divided by the half sum of absolute values of the A_t and F_t. This value is summed for every t and divided by the number of fitted points n.

$$\mathrm{SMAPE} = \frac{100\%}{n}\sum\nolimits_{t=1}^{n}\left|\frac{F_t - A_t}{\left(|A_t| + |F_t|\right)/2}\right| \tag{3}$$

The main advantage of the SMAPE is the interpretability (values range between 0 and 1) and the scale independency, which are necessities when dealing with multiple time series. The drawbacks are that when the actual value is zero, this metric is undefined because of the denominator.

- *Mean Absolute Percentage Error (MAPE)*

MAPE or mean absolute percentage deviation (*MAPD*) is a prediction measure where the difference between the actual value (A_t) and forecast value (F_t) is divided by the actual value. The absolute value of division is summed for every t and divided by the number of fitted points n. This value may be multiplied by 100% for a percentage error.

$$M = \frac{1}{n}\sum\nolimits_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{4}$$

- *Coefficient of Determination (R^2)*

The coefficient of determination denoted as R^2 is a widely used error metric in regression statistics, based on the proportion of variance in the dependent variable that may be justified by the independent variable. It is known also as the goodness of fit and it is the square of the correlation coefficient.

3.3 Holt-Winters Model

Holt-Winters is a statistical model that employs exponential smoothing to encode past values, used to predict the training data and forecasting. When the data is not stationary, in other words when there is a trend factor in data, simple exponential smoothing remains insufficient and the use of double exponential smoothing, or the Holt-Winters model becomes necessary (Holt, 1957). The COVID-19 data does not yet show any seasonality. Therefore, the seasonal parameter of the model is not included. With this adjustment, the method comprises the forecast equation with two smoothing equations for the level l_t and for the trend b_t, with corresponding parameters α and β between 0 and 1. The component form of the Holt-Winters model is

$$\hat{y}_{t+\mathrm{hIt}} = l_t + hb_t \tag{5}$$

$$l_t = \alpha\,(y_t) + (1-\alpha)\,(l_{t-1} + b_{t-1}) \tag{6}$$

$$b_t = \beta\,(l_t - l_{t-1}) + (1-\beta)\,b_{t-1} \tag{7}$$

The equations are done in MS Excel with generalized reduced gradient nonlinear solver method that looks at the slope of the objective function (decreasing selected error metrics) as the input values change and determine the optimality when the partial derivatives are zero (Abadie, 1978). Table 4 gives the results accuracy in R^2, RMSE, SMAPE, and MAPE.

As is observed from Table 4, for each time series, three alternative double exponential smoothing models are solved to decrease the RMSE, SMAPE, and

Table 4 Holt-Winters accuracy results and parameters

	World case	**World death**	**USA case**	**USA death**	**India case**	**India death**	**Brazil case**	**Brazil death**
α	1.000	1.000	0.801	1.000	0.885	1.000	1.000	0.913
β	0.525	0.224	0.589	0.512	0.383	0.050	0.022	0.043
RMSE	4621.620	149.704	2613.511	46.550	633.241	22.953	1528.380	31.745
α	1.000	0.893	0.817	0.936	0.926	0.942	1.000	0.970
β	0.523	0.470	0.881	0.691	0.325	0.083	0.117	0.162
SMAPE	0.026	0.027	0.021	0.030	0.023	0.039	0.042	0.040
alpha	0.962	0.893	0.849	0.936	0.922	0.942	1.000	0.969
beta	0.597	0.470	0.821	0.691	0.310	0.083	0.102	0.152
MAPE	0.025	0.027	0.020	0.030	0.023	0.039	0.042	0.039
R^2	0.9996	0.9983	0.9985	0.9970	0.9995	0.9958	0.9902	0.9897
	Russia case	**Russia death**	**UK case**	**UK death**	**France case**	**France death**	**Spain case**	**Spain death**
α	0.951	1.000	1.000	0.796	1.000	1.000	0.624	1.000
β	0.719	0.304	0.594	0.411	0.000	0.000	0.364	0.155
RMSE	98.039	4.565	452.017	15.948	1882.908	28.010	693.694	29.679
α	0.816	0.742	1.000	0.995	1.000	1.000	0.922	1.000
beta	1.000	0.531	0.648	0.321	0.015	0.001	0.166	0.155
SMAPE	0.009	0.032	0.028	0.058	0.108	0.097	0.081	0.123
α	0.816	0.845	1.000	0.996	1.000	1.000	0.881	0.908
β	1.000	0.357	0.597	0.299	0.024	0.001	0.233	0.420
MAPE	0.009	0.030	0.028	0.055	0.113	0.097	0.084	0.744
R^2	0.9999	0.9993	0.9991	0.9978	0.9739	0.9862	0.9928	0.9770
	Italy case	**Italy death**	**Germany case**	**Germany death**	**Turkey case**	**Turkey death**	**China case**	**China death**
α	1.000	1.000	0.887	0.867	1.000	0.967	1.000	0.999
β	0.568	0.441	0.050	0.033	0.549	1.000	0.132	0.000
RMSE	230.277	9.850	716.699	20.551	598.717	0.914	164.609	14.012
α	0.981	1.000	0.955	0.932	1.000	1.000	0.945	0.895
beta	0.444	0.127	0.308	0.229	0.949	1.000	0.486	0.672
SMAPE	0.029	0.066	0.050	0.086	0.019	0.016	0.113	0.094
α	0.974	1.000	0.949	0.957	1.000	1.000	0.882	–
β	0.439	0.129	0.305	0.179	0.949	1.000	0.662	–
MAPE	0.029	0.070	0.049	0.084	0.018	0.016	0.102	–
R^2	0.9995	0.9983	0.9908	0.9932	0.9943	0.9998	0.9526	0.8455

MAPE, respectively. The objective error metric highly effects the parameters α and β and the accuracy of the model. RMSE is a scale dependent measure, thus it is not suitable for comparison. When case predictions are observed, SMAPE ranges between 0.9% (Russia) and 11.3% (France). For the death predictions, the maximum SMAPE is 12.3% (Spain, corrupted data with negative values) and the minimum SMAPE is 1.6% (Turkey). For the world data, SMAPE and MAPE values are around 2.5%. For the all-time series, the R^2 shows the power of the correlation with 99.99%, with a poor discriminating power. SMAPE and MAPE values show

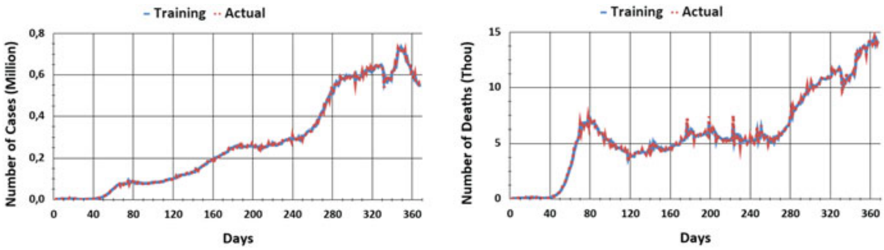


Fig. 2 Fitting Curves for the World: left—New Case, right—New Death

the suitability of the model to the COVID-19 data set. The fitting charts for the World are in Fig. 2.

The forecasting values by the parameters, optimized for SMAPE are in Tables 5 and 6 for daily cases and daily deaths, respectively.

3.4 ARIMA

The ARIMA model describes a univariate time series as a combination of autoregressive (AR) and moving average (MA) lags which capture the autocorrelation within the time series. The order of integration denotes how many times the series has been differenced to get a stationary series. An ARIMA(p,d,q) model where p is the autoregressive lag, d is the degree of differencing, and q is the number of moving average lags can be denoted as:

$$\Delta^D y_t = \sum\nolimits_{i=1}^{p} \varphi_i \Delta^D y_{t-i} + \sum\nolimits_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t, \quad \epsilon_t \sim N\left(0, \sigma^2\right) \tag{8}$$

The (p,d,q) parameters of the model are found by an iterative algorithm that tries to minimize the Bayesian information criterion (bic) values, considering the autocorrelation values. The sample and partial autocorrelation functions belonging to the World are given in Fig. 3.

ARIMA configurations and results for the new cases and new deaths are given in Tables 7 and 8, respectively. These and the following tables show the result by five different error metrics. Bic values are the goodness of fit measure that evaluate the performance of the selected model compared to other models. R^2 represents the proportion of the variance for a dependent variable that is explained by the independent variable. RMSE is the square root of the mean of the squared errors. The existing literature share their results with these three metrics; however, these metrics are scale dependent, and they are not suitable to compare accuracy results for different countries. Therefore, in this study, the scale independent error metric MAPE is calculated. The "Inf" values on MAPE are based on the instability at near

Table 5 Daily case forecasting (Holt-Winters)

Daily cases	World	USA	India	Brazil	Russia	UK	France	Spain	Italy	Germany	Turkey	China
31.01.2021	536,686	147,402	13,007	51,560	18,562	24,104	20,491	34,675	12,348	12,391	6798	125
1.02.2021	528,767	143,587	12,904	51,589	18,306	22,602	20,545	34,555	12,322	12,370	6943	118
2.02.2021	520,847	139,771	12,801	51,617	18,049	21,099	20,598	34,436	12,297	12,348	7088	111
3.02.2021	512,927	135,955	12,699	51,646	17,793	19,597	20,651	34,317	12,271	12,326	7232	104
4.02.2021	505,008	132,140	12,596	51,674	17,537	18,095	20,705	34,197	12,246	12,304	7377	97
5.02.2021	497,088	128,324	12,493	51,703	17,281	16,592	20,758	34,078	12,220	12,283	7522	90
6.02.2021	489,168	124,508	12,390	51,731	17,025	15,090	20,812	33,959	12,195	12,261	7667	84
7.02.2021	481,249	120,693	12,287	51,760	16,769	13,587	20,865	33,839	12,169	12,239	7811	77
8.02.2021	473,329	116,877	12,185	51,788	16,513	12,085	20,918	33,720	12,144	12,218	7956	70
9.02.2021	465,409	113,061	12,082	51,817	16,257	10,583	20,972	33,601	12,118	12,196	8101	63

Table 6 Daily death forecasting (Holt-Winters)

Daily Deaths	World	USA	India	Brazil	Russia	UK	France	Spain	Italy	Germany	Turkey	China
31.01.2021	14,113	3090	129	1079	515	1167	427	420	442	744	131	2
1.02.2021	14,090	3025	125	1087	509	1154	427	429	439	748	129	2
2.02.2021	14,067	2960	121	1094	504	1140	427	437	435	752	127	2
3.02.2021	14,043	2895	117	1102	498	1127	428	446	432	757	125	2
4.02.2021	14,020	2830	113	1109	493	1114	428	455	429	761	122	2
5.02.2021	13,997	2765	109	1117	487	1101	428	463	425	765	120	2
6.02.2021	13,973	2700	105	1125	482	1088	428	472	422	769	118	2
7.02.2021	13,950	2636	101	1132	476	1075	428	481	419	773	116	2
8.02.2021	13,927	2571	97	1140	471	1062	428	489	415	777	114	2
9.02.2021	13,903	2506	93	1147	465	1048	429	498	412	782	112	2

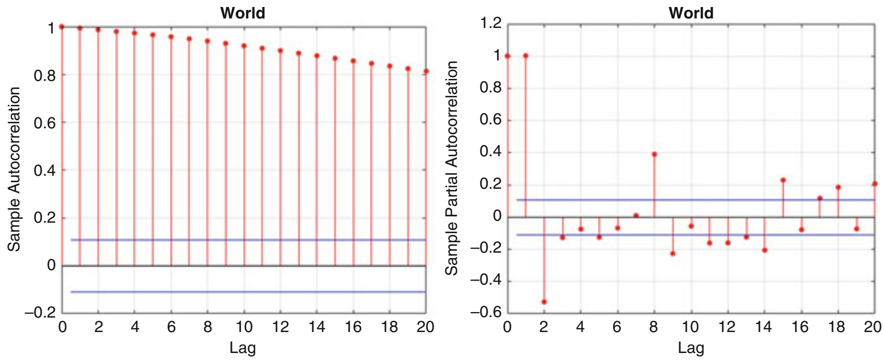


Fig. 3 ACF and PACF charts (World)

Table 7 ARIMA results for the new cases

New case	Configuration	bic	R^2	RMSE	SMAPE	MAPE
World	(7,1,6)	7161.60	0.9997	4141.62	0.0398	0.0367
USA	(7,1,6)	6000.20	0.9991	2390.19	0.0367	0.0587
India	(7,0,7)	4887.30	0.9997	511.21	0.0619	0.0479
Brazil	(7,1,7)	5471.40	0.9939	1297.05	0.0670	0.0836
Russia	(7,2,7)	3660.60	0.9999	76.52	0.0079	0.0080
UK	(7,2,7)	4873.20	0.9994	394.27	0.0401	0.0404
France	(6,1,7)	5785.40	0.9873	2140.64	0.1909	0.1483
Spain	(7,1,4)	5235.50	0.9944	649.56	0.1521	0.1370
Italy	(7,2,7)	4573.00	0.9996	204.75	0.0322	0.0317
Turkey	(6,0,7)	4741.40	0.9972	490.62	0.0259	0.0260
Germany	(6,1,7)	5185.10	0.9950	627.06	0.0805	0.0797
China	(7,1,7)	4574.80	0.9800	117.33	0.4119	0.6388

zero of the time series. To overcome this problem, the symmetric version SMAPE is considered.

When the configurations of the models are investigated, most of the models show a non-stationarity characteristic, which is supported also with the augmented Dickey–Fuller test. For the new cases, the algorithm does not integrate the Turkey and India data and for the daily deaths does not differ between the India and Spain data. A second degree of differentiation is only required for the new cases for Russia, the UK, and Italy, and new deaths for Spain. In general, a seven-lag order is selected by the model for the autoregressive and moving average degrees. However, when the data is decomposed, the seasonality is found to be approximately 0; therefore, a SARIMA model is not necessary. China gives the maximum error values, and the reliability of their values is often discussed in public, therefore in the comments, China will be excluded due to data instability. Most of the countries fit the ARIMA model quite well. The focus of the study is not decreasing the errors as much as

Table 8 ARIMA results for the new deaths

New death	Configuration	bic	R^2	RMSE	SMAPE	MAPE
World	(7,1,7)	4602.40	0.9989	122.43	0.0360	0.0348
USA	(7,1,7)	3335.80	0.9980	37.42	0.0378	0.0321
India	(7,0,7)	4887.30	0.9997	352.87	0.3741	4.2665
Brazil	(7,1,7)	2987.40	0.9937	25.32	0.0371	0.0403
Russia	(6,1,7)	1704.10	0.9996	3.49	0.0383	0.0336
UK	(7,1,7)	2678.30	0.9985	13.28	0.1186	Inf
France	(7,0,7)	2973.30	0.9932	20.40	0.1837	0.1675
Spain	(6,2,7)	3148.90	0.9824	27.31	0.3677	Inf
Italy	(7,1,7)	2428.30	0.9990	8.38	0.0638	0.0643
Turkey	(7,1,7)	664.80	0.9999	0.71	0.0147	0.0150
Germany	(7,1,7)	2708.80	0.9969	13.90	0.2560	Inf
China	(7,1,7)	2769.90	0.9611	10.17	1.1200	Inf

possible but providing an easy and fast fitting and forecasting solution and offering a comparative platform to the researchers and readers to discuss.

The R^2 values greater than 99% show the robustness of the model to explain the variance. The RMSE values may be used for each country to interpret fitting and forecasting intervals. The model performances over countries are done by SMAPE values. For the daily case numbers, the lowest SMAPE is for Russia with 0.79% and Turkey with 2.59%. France and Spain are the worst fitting countries, with 19.09% and 15.21%, respectively. Remaining countries and the world are within acceptable limits, their SMAPE ranging between 1% and 7%. The fitting of the death numbers is not as successful as the new cases fitting. In Table 6, the worst fitting countries are India and Spain with 37.41% and 36.77%, respectively. Turkey (1.47%) and Brazil (3.71%) are the best fitting countries using ARIMA. These countries may be grouped in alternative ways. One way of it is considering the fitting error closeness of the country with the world error term. The countries which have numerically close SMAPEs can be considered as coherent countries. When SMAPEs are too low, the countries may be grouped as negative coherent countries and when SMAPEs are too high, they may be grouped as positive coherent countries, where the necessity of building more sophisticated models arises. Table 9 gives this classification. Spain's data set is corrupted and contains negative values along zeroes, which reflects directly the model results.

Table 9 Classification of the countries by coherence to the world

	Coherent	Positive coherent	Negative coherent
New case	USA, India, Brazil, UK, Italy, Turkey, Germany	France, Spain, China	Russia
New death	USA, Brazil, Russia, UK, Italy	India, France, Spain, Germany, China	Turkey

The ARIMA 10-days forecasting outcomes are in Tables 10 and 11 for daily cases and daily deaths, respectively.

Tables 10 and 11 show that world daily case and death numbers of the virus reached a steady plateau for the first days of the February 2021. USA and UK case and death numbers are decreasing, while the situation is worsening for India and Brazil. Turkey and Russia have a slightly negative slope, where the numbers seem to decrease.

3.5 ARIMAX

ARIMAX is an extension of the ARIMA model where there are suitable explanatory variables that can be incorporated into fitting and forecasting problems. In practice, these additional exogenous variables X create a multivariate time series model instead of a univariate model and improve the prediction performance. An ARIMAX(p,d,q) model for a time series y_t with an exogenous series X can be written as

$$\Delta^D y_t = \sum_{i=1}^{p} \varphi_i \Delta^D y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \sum_{m=1}^{M} \beta_m X_{m,t} + \epsilon_t, \quad \epsilon_t \sim N\left(0, \sigma^2\right) \tag{9}$$

New cases and new deaths are correlated time series and may be meaningful for each other as an exogenous variable fit and forecast better. Another significant data is the stringency index of the countries. The stringency index reflects the government attitudes of the countries and is calculated as a function of school and workspace closures, cancellation of public events, restrictions on public gatherings, closures of public transport, stay at-home requirements, public information campaigns, restriction on internal movements and international travel controls.[3] To test the effectiveness of using these exogenous variables, the Granger-causality test is applied among the time series.

The Granger-causality test is a statistical hypothesis test to determine the usefulness of a time series for forecasting another series (Granger, 1969). A time series X is said to Granger-cause Y, when it provides statistically significant information about the future of the Y The notation is.

$$p\,[Y\,(t+1) \in A|I\,(t)] \neq p\,[Y\,(t+1) \in A|I_{-X}(t)] \tag{10}$$

where p is probability, A is an arbitrary non-empty set, and $I(t)$ and $I_{-X}(t)$ denote the information as of time t in the universe, and in the modified universe where X is excluded. In this study this test is employed to detect in which series ARIMAX can be employed. In total, six hypotheses are built. These hypotheses are, respectively,

[3] https://ourworldindata.org/COVID-government-stringency-index

Table 10 Daily case forecasting (ARIMA)

Daily case	World	USA	India	Brazil	Russia	UK	France	Spain	Italy	Germany	Turkey	China
2021-01-31	541,044	150,053	13,061	51,798	18,467	24,193	21,249	34,287	12,392	12,341	6642	112
2021-02-01	533,534	145,656	12,866	52,644	18,278	23,473	21,315	33,026	12,443	12,071	6631	96
2021-02-02	530,433	143,884	12,957	52,558	18,093	23,208	21,277	32,830	12,598	12,243	6542	80
2021-02-03	530,139	143,415	13,214	52,561	17,886	22,669	21,309	31,410	12,608	12,426	6301	69
2021-02-04	532,153	143,319	12,285	53,833	17,679	21,934	24,121	30,880	12,727	12,542	5921	58
2021-02-05	528,806	140,978	12,363	53,392	17,493	21,094	22,883	29,853	12,934	12,456	5854	41
2021-02-06	528,633	141,601	12,729	53,400	17,288	20,695	21,658	28,554	13,029	12,053	5867	23
2021-02-07	528,731	141,320	12,554	53,736	17,067	20,437	21,864	28,028	13,190	12,267	5787	10
2021-02-08	527,357	140,361	12,529	53,781	16,803	20,381	23,076	26,415	13,401	12,249	5707	2
2021-02-09	527,713	140,703	12,463	53,985	16,541	20,233	22,942	25,720	13,591	12,183	5641	13

Table 11 Daily death forecasting (ARIMA)

Daily deaths	World	USA	India	Brazil	Russia	UK	France	Spain	Italy	Germany	Turkey	China
2021-01-31	14,235	3131	173	1090	519	1171	430	442	447	781	131	0
2021-02-01	14,326	3119	228	1108	520	1157	443	426	440	694	130	−1
2021-02-02	14,230	3062	303	1105	517	1126	462	425	439	784	130	0
2021-02-03	14,346	3070	399	1123	513	1112	444	420	439	792	129	−2
2021-02-04	14,287	3053	507	1134	512	1096	499	425	438	696	129	−3
2021-02-05	14,369	3039	625	1143	514	1076	454	416	437	804	128	−3
2021-02-06	14,368	3062	749	1132	516	1051	465	433	435	764	127	−3
2021-02-07	14,416	3070	854	1147	516	1034	467	437	435	715	127	−4
2021-02-08	14,434	3069	940	1153	516	1019	461	440	433	822	127	−4
2021-02-09	14,447	3058	1007	1159	517	999	455	439	433	763	126	−4

Table 12 The Granger-causality test results on case to death and vice versa

Case → Death	h	*p*-value	Stat	Death → Case	h	*p*-value	Stat
World	**1**	0.001	10.657	World	**1**	1.12E−19	8.24E+01
USA	**1**	0.002	9.405	USA	**0**	2.68E−01	1.23E+00
India	**1**	0.000	19.041	India	**0**	5.90E−01	2.90E−01
Brazil	**0**	0.147	2.101	Brazil	**0**	6.60E−01	1.93E−01
Russia	**1**	0.000	36.213	Russia	**1**	3.62E−05	1.71E+01
UK	**1**	0.000	14.606	UK	**0**	6.78E−01	1.73E-01
France	**1**	0.000	98.141	France	**1**	2.00E−04	1.38E+01
Spain	**0**	0.090	2.883	Spain	**1**	1.85E−05	1.83E+01
Italy	**1**	0.000	27.930	Italy	**1**	1.18E−06	2.36E+01
Turkey	**0**	0.161	1.968	Turkey	**0**	4.37E−01	6.05E−01
Germany	**0**	0.869	0.027	Germany	**1**	1.96E−02	5.45E+00
China	**0**	0.585	0.299	China	**1**	9.10E−03	6.80E+00

Table 13 The Granger-causality test results on case to stringency and vice versa

Case → Stringency	h	*p*-value	Stat	Stringency → Case	h	*p*-value	Stat
USA	**0**	0.759	0.094	USA	**0**	0.484	0.489
India	**1**	0.015	5.974	India	**1**	0.000	25.564
Brazil	**0**	0.184	1.762	Brazil	**0**	0.927	0.008
Russia	**0**	0.829	0.047	Russia	**1**	0.021	5.310
UK	**0**	0.217	1.522	UK	**0**	0.145	2.127
France	**0**	0.104	2.651	France	**0**	0.055	3.694
Spain	**0**	0.755	0.097	Spain	**0**	0.374	0.790
Italy	**0**	0.069	3.313	Italy	**1**	0.000	13.800
Turkey	**0**	0.890	0.019	Turkey	**0**	0.996	0.000
Germany	**1**	0.007	7.356	Germany	**0**	0.054	4.500
China	**0**	0.705	0.143	China	**0**	0.915	0.011

"case" Granger-causes "deaths" and vice versa, "case" Granger-causes "stringency index" and vice versa, and "deaths" Granger-cause "stringency index" and vice versa.

Tables 12, 13, and 14 show the results of these tests, where h value 1 indicates the acceptance of the hypothesis, which does not neglect the Granger-cause effect between the time series for a *p*-value lower than 0.05.

The first hypothesis is based on the strong correlation idea between the case and death numbers. However, as can be observed from Table 13, only for seven countries "case" has a Granger-cause on the "death" numbers. Similarly, only for seven countries the "death" numbers can be employed to estimate the "case" numbers. In addition, these countries are not the same, and this Granger-cause cannot be generalized for countries; therefore, it will not be included in the ARIMAX model.

The second hypothesis is based on the effect of the government restrictions on the case number and vice versa. Although this idea makes sense in theory, when the test is applied, it is found that it does not make sense statistically. Only in two countries

Table 14 The Granger-causality test results on death to stringency and vice versa

Death → Stringency	h	p-value	Stat	Stringency → Death	h	p-value	Stat
USA	**0**	0.783	0.076	USA	**1**	7E−06	2E+01
India	**0**	0.098	2.746	India	**1**	3E−02	5E+00
Brazil	**0**	0.081	3.053	Brazil	**1**	1E−03	1E+01
Russia	**0**	0.965	0.002	Russia	**1**	6E−03	8E+00
UK	**0**	0.870	0.027	UK	**1**	3E−06	2E+01
France	**0**	0.331	0.945	France	**1**	5E−05	2E+01
Spain	**0**	0.227	1.459	Spain	**1**	4E−05	2E+01
Italy	**0**	0.337	0.921	Italy	**1**	9E−23	1E+02
Turkey	**0**	0.550	0.357	Turkey	**1**	2E−06	2E+01
Germany	**0**	0.645	0.212	Germany	**1**	4E−05	2E+01
China	**0**	0.991	0.000	China	**0**	7E−01	1E−01

Table 15 ARIMAX scores on new cases (stringency index as exogenous variable)

New case	Configuration	bic	R^2	RMSE	SMAPE	MAPE
India	(7,0,7)	4509.09	0.9998	583.90	0.0243	0.0242
Russia	(7,2,7)	3652.34	0.9996	77.64	0.0082	0.0084
Italy	(7,2,7)	4552.71	0.9999	198.12	0.0348	0.035

"case" is the Granger-cause of the stringency index, and only in three states the stringency has a significant effect on the "case" numbers estimation. These three countries will be modeled with ARIMAX to measure the impact on the forecasting accuracy.

As is clear from Table 14, death has no impact on the stringency index in each country, however when the vice versa situation is considered, for all the countries (except China), the stringency index is a Granger-cause of the death numbers, therefore should be used in the ARIMAX as an exogenous variable to increase the forecasting accuracy. Based on the Granger-causality test, the results of the ARIMAX model are given in Table 15.

The SMAPE values of the ARIMA model belonging to India, Russia, and Italy were 6.19%, 7.9%, and 3.22%, respectively. ARIMAX results shows that, the only significant contribution of the stringency index on the estimation process, obtained in India, by an added value of 3.76%. This can be considered as warning not to employ complex models when the forecasting accuracy satisfactory.

The Granger-cause effect between the stringency index and new deaths is common for countries. Table 16 shows the results of the ARIMAX model where the stringency index is considered as an exogenous variable to predict the new deaths.

Spain gives the worst performance. When the data of Spain is investigated the negative values of new deaths are observed. This corruption of the data set is reflected directly on the solutions. Therefore, this country necessitates a data cleaning process rather than a sophisticated model. MAPE does not perform well

Table 16 ARIMAX scores on new deaths (stringency index as exogenous variable)

New death	Configuration	bic	R^2	RMSE	SMAPE	MAPE
USA	(7,1,7)	3348.17	0.9980	37.62	0.0757	0.2032
India	(7,0,7)	2766.18	0.9967	17.42	0.1312	0.6431
Brazil	(7,1,7)	3008.07	0.9931	25.67	0.0798	1.1843
Russia	(6,1,7)	1756.71	0.9996	3.55	0.0696	0.0633
UK	(7,1,7)	2703.46	0.9982	13.32	0.7225	Inf
France	(7,0,7)	2998.08	0.9938	20.38	0.2933	0.5039
Spain	(6,2,7)	3161.37	0.9670	27.58	0.9295	Inf
Italy	(7,1,7)	2464.41	0.9989	8.44	0.1269	0.1566
Germany	(7,1,7)	2734.91	0.9968	13.98	0.3149	Inf
Turkey	(7,1,7)	760.35	0.9999	0.71	0.0160	0.0161

Table 17 Daily death forecasting (ARIMAX)

	USA	India	Brazil	Russia	UK	France	Spain	Italy	Germany	Turkey
2021-01-31	3105	129	1089	520	1201	408	549	445	784	131
2021-02-01	3048	110	1106	521	1226	397	660	432	694	130
2021-02-02	2935	116	1087	520	1241	388	809	426	777	130
2021-02-03	2877	123	1093	517	1278	337	967	423	783	129
2021-02-04	2789	134	1104	519	1330	356	1165	414	682	129
2021-02-05	2701	127	1100	522	1381	272	1373	403	787	129
2021-02-06	2635	110	1071	526	1443	236	1636	394	746	129
2021-02-07	2562	106	1080	528	1509	200	1821	385	687	130
2021-02-08	2478	104	1074	530	1592	151	2012	372	788	131
2021-02-09	2381	103	1068	533	1676	100	2207	363	727	132

because of the near zero values. UK is not suitable to be fitted with ARIMAX with a SMAPE of 72.25%, which is far greater than the simple ARIMA process (Table 17).

Tables 18 and 19 show the Holt-Winters outperforming performance for the new case and new deaths except for Spain. For the new deaths ARIMAX is an overfitting method and should not be used in the prediction of the COVID-19 numbers.

The ARIMA and Holt-Winter models may be used for fitting and forecasting the cases and deaths, they can be employed as benchmark results for alternative forecasting methods. Figures 4 and 5 draw the 10-days forecasting outcomes of the employed models with the test data for World, USA, and UK.

4 Conclusion

The COVID-19 studies are an ongoing literature in alternative branches. This study is among the first efforts that compiles forecasting research. COVID-19, having completed its first year, employs a satisfactory large training data set and shows

Table 18 Comparative results (new death)

	SMAPE			MAPE		
New death	Holt-Winters	ARIMA	ARIMAX	Holt-Winters	ARIMA	ARIMAX
World	0.027	0.036	–	0.027		–
USA	0.03	0.0378	0.0757	0.03	0.0321	0.2032
India	0.039	0.3741	0.1312	0.039	4.2665	0.6431
Brazil	0.04	0.0371	0.0798	0.039	0.0403	1.1843
Russia	0.032	0.0383	0.0696	0.03	0.0336	0.0633
UK	0.058	0.1186	0.7225	0.055	Inf	Inf
France	0.097	0.1837	0.2933	0.097	0.1675	0.5039
Spain	0.744	0.3677	0.9295	0.744	Inf	Inf
Italy	0.066	0.0638	0.1269	0.07	0.0643	0.1566
Germany	0.086	0.256	0.3149	0.084	Inf	Inf
Turkey	0.016	0.0147	0.016	0.016	0.015	0.0161
China	0.094	1.2	–	–	Inf	–

Table 19 Comparative results (new case)

	SMAPE		MAPE	
New case	Holt-Winters	ARIMA	Holt-Winters	ARIMA
World	0.026	0.0398	0.025	0.0367
USA	0.021	0.0367	0.02	0.0587
India	0.023	0.0619	0.023	0.0479
Brazil	0.042	0.067	0.042	0.0836
Russia	0.009	0.0079	0.009	0.008
UK	0.028	0.0401	0.028	0.0404
France	0.108	0.1909	0.113	0.1483
Spain	0.081	0.1521	0.084	0.137
Italy	0.029	0.0322	0.029	0.0317
Germany	0.05	0.0259	0.049	0.026
Turkey	0.019	0.0805	0.018	0.0797
China	0.113	0.4119	0.102	0.6388

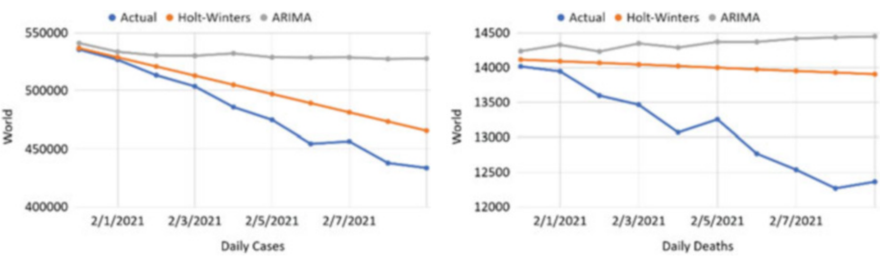


Fig. 4 Forecasting world COVID-19 data: left—New case, right—New death

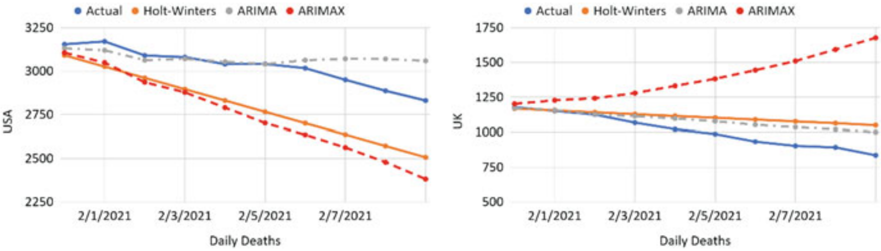


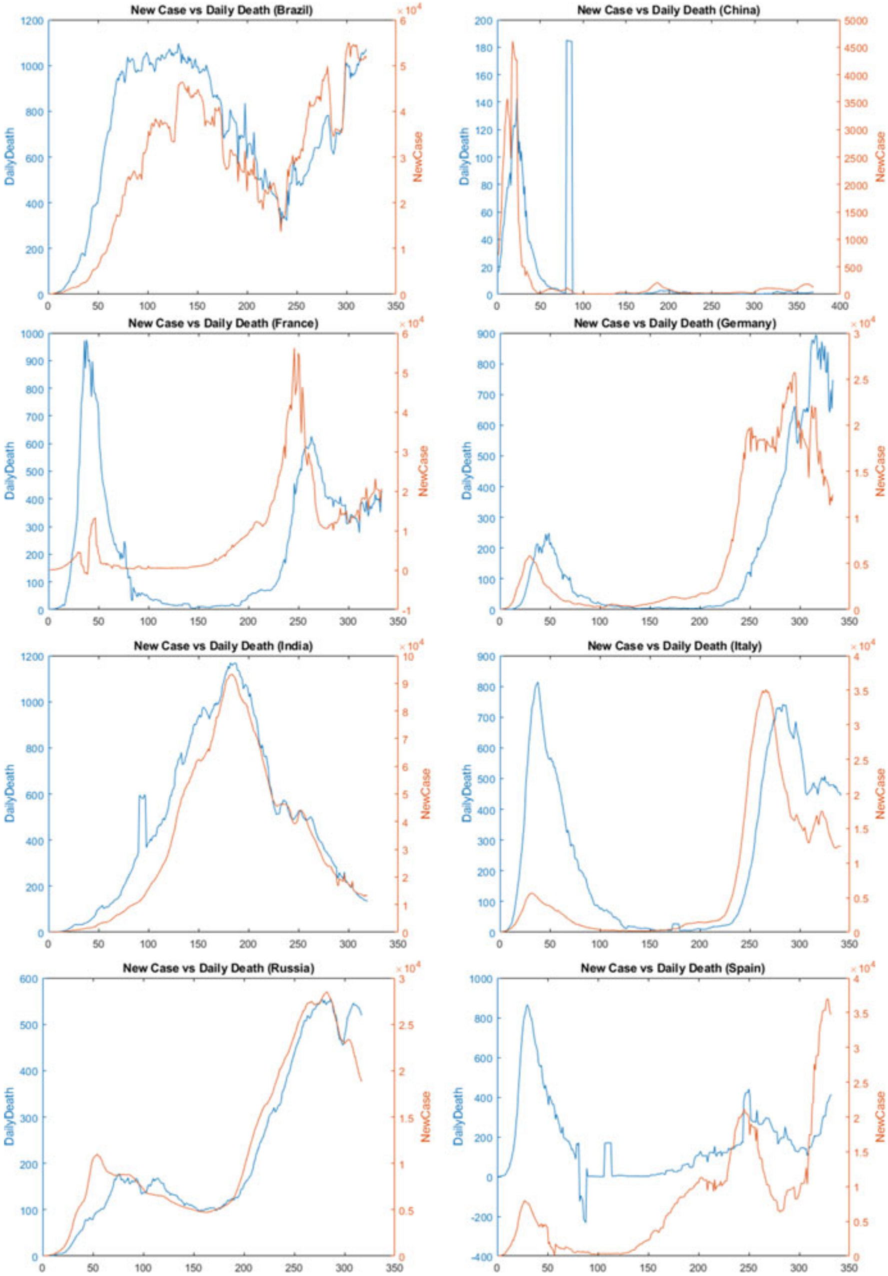
Fig. 5 Forecasting daily deaths: left—USA, right—UK

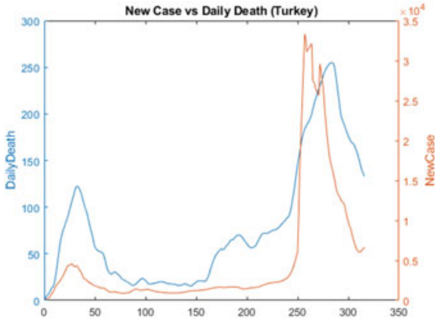
the accuracy results of simple but successful statistical forecasting models on a total of 24 time series (12 for new cases and 12 for new deaths). This paper employs three different models, those being the Holt-Winters, the ARIMA, and the ARIMAX models, with five different error metrics, bic, R^2, RMSE, SMAPE, and MAPE. All the models provide satisfactory results where percentage errors are generally lower than 10% and R^2 is approximately 99.9% showing the power of regression-based models. In general, the Holt-Winters (known as double exponential smoothing) outperforms the ARIMA, and although an introduction of an exogenous variable in the estimation process exists, ARIMAX is the lowest performing model, still with the acceptable results for most of the countries (see Figs. 4 and 5).

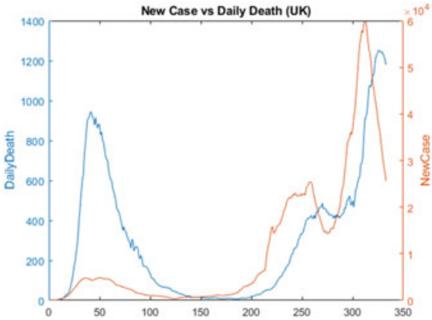
The correlation of the most effected countries' data is calculated with the world data. The Granger-causality tests show the importance of the correct exogenous variable selection. Statistically, the new cases and new deaths are dependent variables; however, in the estimation process they cannot be used for each other as auxiliary inputs. The stringency index consisting of government attitudes towards combatting the virus, statistically does not affect the case numbers; however, it has a Granger-cause effect in death numbers.

With the available data set and all the parameter details, this study provides reproducible results, where outcomes may be used by other researchers as benchmark results. Further researchers may classify the countries according to their response to statistical models, and with a more focused attention, such as data cleaning or machine learning approaches, they can improve the fitting and forecasting accuracy performances. The finding of a meaningful exogenous variable in the estimation would be beneficial to increase the ARIMAX performance.

Appendix


New Case vs Daily Death (Brazil)
New Case vs Daily Death (China)
New Case vs Daily Death (France)
New Case vs Daily Death (Germany)
New Case vs Daily Death (India)
New Case vs Daily Death (Italy)
New Case vs Daily Death (Russia)
New Case vs Daily Death (Spain)
DailyDeath
NewCase
$\times 10^4$





References

Abadie, J. (1978). The GRG method for nonlinear programming. *Design and Implementation of Optimization Software*, 335–362.

Al-Qaness, M. A., Ewees, A. A., Fan, H., & Abd El Aziz, M. (2020). Optimization method for forecasting confirmed cases of COVID-19 in China. *Journal of Clinical Medicine, 9*(3), 674.

Ankarali, H., Erarslan, N., & Pasin, O. (2020). Modeling and Short-Term Forecasts of Indicators for COVID-19 Outbreak in 25 Countries at the end of March. *medRxiv*. https://doi.org/10.1101/2020.04.26.20080754

Ayinde, K., Lukman, A. F., Rauf, R. I., Alabi, O. O., Okon, C. E., & Ayinde, O. E. (2020). Modeling Nigerian COVID-19 cases: A comparative analysis of models and estimators. *Chaos, Solitons & Fractals, 138*, 109911.

Ayyoubzadeh, S. M., Ayyoubzadeh, S. M., Zahedi, H., Ahmadi, M., & Kalhori, S. R. N. (2020). Predicting COVID-19 incidence through analysis of Google Trends data in Iran: data mining and deep learning pilot study. *JMIR Public Health and Surveillance, 6*(2), e18828.

Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., & Ciccozzi, M. (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset. *Data in Brief, 29*, 105340.

Bhattacharya, S., Maddikunta, P. K. R., Pham, Q. V., Gadekallu, T. R., Chowdhary, C. L., Alazab, M., & Piran, M. J. (2021). Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey. *Sustainable Cities and Society, 65*, 102589.

Bruns, D. P., Kraguljac, N. V., & Bruns, T. R. (2020). COVID-19: Facts, Cultural Considerations, and Risk of Stigmatization. *Journal of Transcultural Nursing, 31*(4), 326–332.

Crokidakis, N. (2020). Modeling the early evolution of the COVID-19 in Brazil: Results from a Susceptible–Infectious–Quarantined–Recovered (SIQR) model. *International Journal of Modern Physics C (IJMPC), 31*(10), 1–7.

Dandekar, R., & Barbastathis, G. (2020). Neural Network aided quarantine control model estimation of COVID spread in Wuhan, China. arXiv preprint arXiv:2003.09403.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424–438.

Hernandez-Matamoros, A., Fujita, H., Hayashi, T., & Perez-Meana, H. (2020). Forecasting of COVID19 per regions using ARIMA models and polynomial functions. *Applied Soft Computing, 96*, 106610.

Holt, C. E. (1957). Forecasting seasonals and trends by exponentially weighted averages (O.N.R. Memorandum No. 52). Carnegie Institute of Technology. https://doi.org/10.1016/j.ijforecast.2003.09.015

Hu, Z., Ge, Q., Li, S., Boerwincle, E., Jin, L., & Xiong, M. (2020). Forecasting and evaluating intervention of COVID-19 in the World. arXiv preprint arXiv:2003.09800.

Ibrahim, M. R., Haworth, J., Lipani, A., Aslam, N., Cheng, T., & Christie, N. (2021). Variational-LSTM Autoencoder to forecast the spread of coronavirus across the globe. *PLoS One, 16*(1), e0246120.

Ivorra, B., Ferrández, M. R., Vela-Pérez, M., & Ramos, A. M. (2020). Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) considering its particular characteristics. The case of China. *Communications in Nonlinear Science and Numerical Simulation, 88*, 105303.

Jia, L., Li, K., Jiang, Y., & Guo, X. (2020). Prediction and analysis of Coronavirus Disease 2019. arXiv preprint arXiv:2003.05447.

Kafieh, R., Arian, R., Saeedizadeh, N., Minaee, S., Yadav, S. K., Vaezi, A., & Javanmard, S. H. (2020). COVID-19 in Iran: a deeper look into the future. medRxiv. https://doi.org/10.1101/2020.04.24.20078477

Kolozsvari, L. R., Berczes, T., Hajdu, A., Gesztelyi, R., TIba, A., Varga, I., … & Zsuga, J. (2020). Predicting the epidemic curve of the coronavirus (SARS-CoV-2) disease (COVID-19) using artificial intelligence. medRxiv. https://doi.org/10.1101/2020.04.17.20069666

Kumar, P., Singh, R. K., Nanda, C., Kalita, H., Patairiya, S., Sharma, Y. D., … & Bhagavathula, A. S. (2020). Forecasting COVID-19 impact in India using pandemic waves Nonlinear Growth Models. MedRxiv. https://doi.org/10.1101/2020.03.30.20047803

Liu, D., Clemente, L., Poirier, C., Ding, X., Chinazzi, M., Davis, J. T., … & Santillana, M. (2020). A machine learning methodology for real-time forecasting of the 2019-2020 COVID-19 outbreak using Internet searches, news alerts, and estimates from mechanistic models. arXiv preprint arXiv:2004.04019.

Liu, Z., Magal, P., & Webb, G. (2021). Predicting the number of reported and unreported cases for the COVID-19 epidemics in China, South Korea, Italy, France, Germany, and United Kingdom. *Journal of Theoretical Biology, 509*, 110501.

Milhinhos, A., & Costa, P. M. (2020). On the progression of COVID-19 in Portugal: A comparative analysis of active cases using non-linear regression. *Frontiers in Public Health, 8*, 495.

Pandey, G., Chaudhary, P., Gupta, R., & Pal, S. (2020). SEIR and Regression Model based COVID-19 outbreak predictions in India. arXiv preprint arXiv:2004.00958.

Peeri, N. C., Shrestha, N., Rahman, M. S., Zaki, R., Tan, Z., Bibi, S., … Haque, U. (2020). The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned? *International Journal of Epidemiology, 49*(3), 717–726.

Petropoulos, F., & Makridakis, S. (2020). Forecasting the novel coronavirus COVID-19. *PLoS One, 15*(3), e0231236.

Roosa, K., Lee, Y., Luo, R., Kirpich, A., Rothenberg, R., Hyman, J. M., … Chowell, G. (2020). Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. *Infectious Disease Modelling, 5*, 256–263.

Sameni, R. (2020). Mathematical modeling of epidemic diseases; a case study of the COVID-19 coronavirus. arXiv preprint arXiv:2003.11371.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*(2), 461–464.

Xu, C., Yu, Y., Chen, Y., & Lu, Z. (2020). Forecast analysis of the epidemics trend of COVID-19 in the USA by a generalized fractional-order SEIR model. *Nonlinear Dynamics, 101*(3), 1621–1634.

Yang, Z., Zeng, Z., Wang, K., Wong, S. S., Liang, W., Zanin, M., … He, J. (2020). Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *Journal of Thoracic Disease, 12*(3), 165.

Yonar, H., Yonar, A., Tekindal, M. A., & Tekindal, M. (2020). Modeling and Forecasting for the number of cases of the COVID-19 pandemic with the Curve Estimation Models, the Box-Jenkins and Exponential Smoothing Methods. *EJMO, 4*(2), 160–165.

An Exploration of the Influence of Innovations on Organizational Performance: A Dynamic Capabilities Perspective

Aristides Matopoulos and Emel Aktas

Abstract This research explores the impact of dynamic innovation capabilities on firm performance, investigating how alliances and networks improve the focal firm's capabilities. We examine four innovation capabilities and their effect on profitability, performance, and growth of companies in the food industry. Drawing on the extant literature on innovation and dynamic capabilities theory, we develop six hypotheses to explain the relationship between different innovation capabilities and firm performance. We test the proposed research model and hypotheses using partial least squares structural equation modeling (PLS-SEM) with primary data from the food industry in Greece. Results indicate that dynamic innovation capabilities play a crucial role in improving profitability, growth, and overall firm performance through ordinary innovation capabilities. This study provides critical insights into innovation capabilities in the food industry. These insights are significant because prior studies have not investigated the relative effects of different innovation capabilities on profitability, growth, and overall firm performance, particularly the interrelationships between dynamic and ordinary innovation capabilities. The generalization of the study results may be limited due to the sample size.

Keywords Innovation · Dynamic Capabilities · PLS-SEM · Food industry · Greece

---

A. Matopoulos (✉)
Department of Engineering Systems & Supply Chain Management College of Engineering and Physical Sciences, Aston University, Birmingham, UK
e-mail: a.matopoulos@aston.ac.uk

E. Aktas
Logistics, Procurement and Supply Chain Management, Cranfield University, Cranfield School of Management, Cranfield, UK
e-mail: emel.aktas@cranfield.ac.uk

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_9

1 Introduction

From Schumpeter's (1934) innovation in product, process, customer, input, and organization to Daft's (1978) organizational innovation or some years later to Damanpour's (1991) administrative innovation, the multi-faceted view of innovation has been well recognized in the literature. More recently, researchers such as Roberts and Amit (2003), for example, have proposed channel innovation, while Birkinshaw et al. (2008, p.825) expanded the capability of a firm to innovate to include management innovation. All these different innovations have resulted from specific innovation capabilities or mixes of them that are needed to develop a particular innovation (Forsman, 2011; Boly et al., 2014). By innovation capability, we refer to the firm's ability to continuously transform knowledge and ideas into something beneficial to the firm and its stakeholders (Lawson & Samson, 2001).

Despite the increased interest in identifying and conceptualizing different innovation capabilities, comparatively little research has been conducted on the interrelations between innovation capabilities and the effect on organizational performance. Damanpour et al.'s (1989), Subramanian and Nilakanta's (1996), and Oke's (2007) research focused on service organizations (e.g., libraries, banks). Limited work has been done in the manufacturing/processing sector (Lee & Kang, 2007; Kirner et al., 2009; Gunday et al., 2011; Wu et al., 2016), where the results only partly support the idea that the innovation capability matters in organizational performance. Also, in all the above research efforts, the focus was on the relationship between a few innovations and a single performance aspect, but perhaps more importantly, a firm's capability to innovate was narrowly viewed and linked only to new things but not necessarily organizational performance.

In this chapter, we take an expanded, value-oriented view of innovation which links innovation to new value created for customers instead of new products or processes (Chen et al., 2011). We shed light on the mediating effects of certain innovation capabilities on others and performance aspects (i.e., growth, profitability, and overall performance). Hence, we essentially address the following research questions:

1. How can different innovations be explained under Dynamic Capabilities Theory?
2. What are the interrelations among different innovation capabilities?
3. How do different innovation capabilities affect organizational performance?

To address these questions, we view innovation as a capability and develop a model that builds on Dynamic Capabilities Theory which is a well-suited theory to the study of innovation (Lawson & Samson, 2001). Unlike the Resource-Based View (Barney, 1991), which assumes that differences across firms are due to differences arising from having access to firm-specific resources (valuable, rare, inimitable, and non-substitutable resources); the dynamic capabilities framework (Teece & Pisano, 1994; Teece, 2007) advocates that in addition to resources firms need processes (e.g., product development) to manipulate them into value-adding strategies (Lee & Kelley, 2008). Antecedents to innovation capability can therefore be found not only at the firm but also at the network level (Rothaermel & Hess, 2007), as alliances and networks can improve the focal firm's capabilities by providing potential information and resource advantage (Powell et al., 1996; Kogut, 2000), which also justifies the three research questions presented earlier.

We also use a specific domain, the food industry, as our means to understand the interrelations between different innovation capabilities and the impact on organizational performance. There are two main reasons behind this choice. Firstly, because the industry has strong links with various food (e.g., suppliers in all stages of the food chain) and non-food sectors (e.g., chemicals, food technology, packaging, machinery) all with different characteristics (Menrad, 2004); it is possible that innovation may come from outside the firm's boundaries. Secondly, the speed of product innovation in the industry has changed dramatically in the last decades (in response to shrinking product life cycles). Every year between 5000 and 10,000 "new" products are offered to the market in Australia, New Zealand, and the UK, and in the USA, this figure amounts to 18,000 a year (Bruin & Jongen, 2003; Winger & Wall, 2006; Gov.UK, 2014). These characteristics are what make the food industry a good candidate for exploring the scope of developing alternative capabilities to innovate.

This chapter's contribution to the literature is threefold: First, it offers an operationalization of different innovation capabilities using the concept of value-based innovation and distinguishes dynamic innovation capabilities and ordinary innovation capabilities. Second, it explores the link between dynamic and ordinary capabilities and performance in the field of innovation, adding to the not so vast body of literature with empirical evidence from the manufacturing/processing sector. Third, we demonstrate an application of PLS-SEM on empirical data. We model latent constructs representing dynamic and ordinary capabilities of the organization as higher-order constructs and test the relationships between them on a sample from the food industry. Hence, our chapter also shows the use of higher-order PLS-SEM applied to dynamic capabilities and firm performance.

We begin our chapter by reviewing the literature on innovation from a capabilities perspective and linking this to the different innovations. Next, we present our research model and proposed hypotheses followed by the research methodology and the results. We finally synthesize our findings and discuss research and managerial implications.

2 Theoretical Framework

2.1 Innovation as a Capability

Capabilities can be distinguished as ordinary (or operational) or dynamic (Collis, 1994; Winter, 2003; Helfat et al., 2007). Even though the line between the two is blurry, ordinary capabilities could be seen as lower order which *"enable a firm to* *perform an activity on an ongoing basis using the same techniques on the same scale to support existing products and services for the same customer population"* (Helfat & Winter, 2011: p. 1244). Conversely, dynamic capabilities have the capacity to sense (which means identifying and assessing opportunities outside the company), to seize (mobilizing resources to capture value from those opportunities), and to maintain competitiveness (through enhancement and reconfiguration of tangible and intangible assets) (Teece, 2007). According to Lawson and Samson (2001), dynamic capabilities emphasize management capabilities and inimitable combinations of resources that cut across many functions (e.g., R&D, product and process development, manufacturing) of the company and essentially expand the firm's boundaries.

Analogously the capability of a firm to innovate includes both ordinary and dynamic elements. A firm's capability to innovate across products, processes, or solutions could be classified as ordinary because it is driven by the company and is based on the same/similar techniques and practices (e.g., new products are primarily variations of existing ones). On the contrary, the capability of a firm to innovate "outside its comfort zone" by creating non-traditional channels or by finding new ways of working together with external partners or by building new alliances could be classified as dynamic as it requires a search for new knowledge and information outside the firm (i.e., sensing capabilities) which represents a high degree of change and uncertainty (Lee & Kelley, 2008; Wu et al., 2016).

In the following sections, we analyze the capability of a firm to innovate by considering four main innovations based on Sawhney et al. (2006) and Chen et al. (2011): network-oriented innovation, customer-oriented innovation, offering-oriented innovation, and operations-oriented innovation. An analysis of what is included in each innovation, also supported by literature, is provided below. We close this section with a summary of profitability, growth, and overall performance as the aspects of performance that are expected to improve as firms become more innovative in their network, offerings, operations, and customer relations.

2.2 Network-Oriented Innovation

Network-oriented innovation builds on the fact that value can be created outside the firm through relationships with external partners (Jacobides et al., 2006; Debruyne, 2014; Kim & Lui, 2015). These external partners are not only related to the distribution *channels* used by the company but also to the overall business ecosystem. For example, network-oriented innovation can be achieved by creating new ways to manage direct and indirect traditional channels of distribution as well as by creating non-traditional channels where traditional intermediaries are assigned new roles (Chen et al., 2011). But network-oriented innovation may also be the result of identifying new ways of working together with partners in the supply chain or by building new alliances or partnerships, or even by establishing new outsourcing arrangements with partners (Rothaermel & Hess, 2007).

2.3 Customer-Oriented Innovation

Customer-oriented innovation focuses on the *relationship* and the *interaction* built with the customer, which may result in a better understanding of customers' expectations and an improved overall customer experience. Customer-oriented innovation may come in different forms. Moore (2004) ties innovations to the product life cycle, Sawhney et al. (2006) consider pricing innovation, while other researchers refer to innovation in *communication*, which improves understanding customer segmentation. Debruyne (2014), for example, discusses how KLM, with its online frequent flier communities, tap into various customer segments on an ongoing basis and improve its customer experience. Often customer-oriented innovation may be the result of identifying and creating new ways to communicate with customers or new ways to promote a firm's products and services.

2.4 Offering-Oriented Innovation

Offering-oriented innovation builds around the unique functionality of the product or service, but at the same time does not treat products as stand-alone solutions. Often products are parts of a wider end-to-end solution to customers' problems, either because they are combined with services or because they are combined with unique functionalities (Teece, 2007). Offering-oriented innovation is often a result of the use of shared *platforms* or sets of technologies and tools that provide either a cost advantage in the development of new products or speed advantage. Indeed, through offering-oriented innovation, firms can increase product variety and meet diverse customer needs while maintaining cost and process efficiencies. In the innovation literature, this concept has been well recognized. For example, Wheelwright and Clark (1992) describe how "product platforms" can meet the needs of different customers simply by modifying, adding, or subtracting various features. Sood and Tellis (2005) also identify platform innovation, while Sawhney et al. (2006) consider *solution* innovation. More recently, Gawer & Cusumano (2014, p. 4) refer to external platforms, which "serve as foundations upon which a larger number of firms can build further complementary innovations, in the form of specific products, related services, or component technologies." In the food industry, there have also been examples of offering-oriented innovation. International Flavors and Fragrances (IFF), for example, is a global supplier of specialty flavors to the food industry and has built a toolkit that enables its customers to modify flavors for themselves, which IFF then manufactures (Von Hippel, 2005).

2.5 Operations-Oriented Innovation

Operations-oriented innovation takes a more inward view of innovation by focusing on how work is done and hence on the significant improvement (not just lower-level changes) of a firm's *processes*, resulting in lower-cost offerings and superior economic value. Operations-oriented innovation can be achieved by employing new processes and methods or new mechanisms (e.g., purchasing input materials and services) or using advanced information technologies to restructure and simplify the supply chain and the associated processes. Regarding process innovation, Davenport (1993) was one of the first strong advocates of their essential role, followed by Pisano (1997). Pisano's "Development Factory" work explored the effect of process development in the pharmaceutical sector, highlighting its importance on innovation performance. For some authors (Hamel, 2006; Birkinshaw et al., 2008), process innovation is very much related to *management* innovation, which may refer to implementing new management practices to improve organizational efficiency or change the organizational structure of the firm. Examples of management innovation include Toyota's lean production system or General Electric's six sigma (Mol & Birkinshaw, 2012).

We argue that looking at innovation from these four different perspectives is important because it helps the company to understand whether there are interdependencies among the four different innovations and how they are likely to be affecting each other.

2.6 Profitability, Overall Performance, and Growth

According to Damanpour et al. (2009), the adoption of innovation is a means towards organizational change and adjustment of external and internal functions to respond to environmental demands, operate efficiently and effectively, and maintain or improve performance. It is accepted widely, and there has been supporting evidence in the literature that innovation positively influences the performance of firms, both large companies, and SMEs. Walker's (2004) review of 30 empirical studies from 1984 to 2003 also showed that innovation influences performance positively. Organizational performance refers to how well a firm achieves its goals. Past research has measured organizational performance using primarily financial indicators such as return on investment (ROI), market share, and profit margin (Li et al., 2006; Sánchez & Pérez, 2005). This has been particularly the case in the innovation literature, where the sales (or turnover) generated from innovations has been one of the most common measures of business performance (Griffin, 1997). In this research, in addition to market-based indicators (e.g., market share growth, sales growth, new customers), we include the traditional, well-established, and widely used accounting-based measures (e.g., profitability, return on sales; see Rangus & Slavec, 2017). This, coupled with managers' perceptions of the organization's overall performance, provides a more rounded view.

3 Research Model and Hypotheses

We explain below the construction of these hypotheses.

3.1 Network-Oriented Innovation and Operations-Oriented Innovation

Networks and partnerships with suppliers play an important role in improving a firm's processes by improving or adapting existing processes or by developing new ones. From a dynamic capabilities point of view, such networks and partnerships with suppliers enable the firm to sense and seize opportunities outside its boundaries. The research by Nishiguchi and Ikeda (1996) on the Japanese automotive industry has shown that much of the process innovation achievements were at the OEM level, and the improvements of supply chain practices in the industry in the 1990s were the result of suppliers being equipped with an impressive range of self-developed technologies which in turn increased their ability to innovate and to be operationally flexible and adaptable. More recently, research by Roy et al. (2004), Lee et al. (2011), and Yam et al. (2011) has supported that supplier involvement and alliances with external firms are the routes to generating process innovation. This is particularly important for sectors that small and medium-sized enterprises dominate. Avermaete et al.'s (2004) research on the food industry revealed, for example, that suppliers of equipment were the primary sources of process innovation in small food manufacturing firms. Based on the above, our hypothesis is that:

- H1. Network-oriented innovation positively affects operations-oriented innovation.

3.2 Network-Oriented Innovation and Offering-Oriented Innovation

In recent years, alliances and networks have become an integral part of a firm's business environment, particularly regarding innovation development. From a dynamic capabilities point of view, network-oriented innovation can improve the sensing capability of the firm and enable the firm to acquire capabilities outside the firm (McEvily & Marcus, 2005). A company should constantly monitor its own industry and outside its industry for product technology to increase product innovativeness (Kotabe & Scott Swan, 1995). Cohen et al.'s (2000) investigation of Saturn's success in the 90s concluded that the firm's impressive performance in offering excellent after-sales service resulted from a differently designed service supply chain strategy executed by channel partners (dealers). On the same subject, Cousins et al.'s (2011) research showed that interaction and technical exchange with suppliers had a significant positive effect on product development performance. Hence, the following hypothesis follows:

- H2. Network-oriented innovation positively affects offering-oriented innovation.

3.3 Customer-Oriented Innovation and Operations-Oriented Innovation

The notion that firms can improve their processes and introduce new management practices by having better relationships and by interacting better with their customers is gaining attention in innovation studies. Indeed, information systems and technology do support process innovation (Khosrow-Pour, 2006; Tarafdar & Gordon, 2007). But also more recently, Lin et al.'s (2010) examination of 107 Taiwanese computer manufacturers on the effects of Customer Relationship Management (CRM) systems on innovation capabilities revealed that the adoption of technology-based CRM has positive effects on manufacturers' operations. We, therefore, hypothesize that:

- H3. Customer-oriented innovation positively affects operations-oriented innovation.

3.4 Customer-Oriented Innovation and Offering-Oriented Innovation

The role of customers and their contribution to innovation processes and new product offerings have been well-supported in the literature (Chesbrough 2003; O'Hern & Rindfleisch, 2009; Fitjar & Rodríguez-Pose, 2013). In dynamic capability terms, customers, through search activities, are among the first to sense the potential for new offerings (Teece, 2007). According to Foss et al. (2011), firms gain from being orientated toward, perhaps even working directly with, their customers and other users of their products. Food manufacturers, for example, draw heavily on market information from customers for developing innovations (Stewart-Knox & Mitchell, 2003). Hoyer et al. (2010, p. 283) argue that consumers take an active and central role as participants in the new product design process because "*they are now able to easily communicate these ideas to the company through Internet websites, e-mail, and social networks.*" In other words, the improved relationship and the interaction built with the customer will result in a better understanding of customers and lead to improvements in offerings. As a result, the following hypothesis is proposed:

- H4. Customer-oriented innovation positively affects offering-oriented innovation.

3.5 Operations-Oriented Innovation, Profitability, Growth, and Overall Performance

The literature has well documented the relationship between operations-oriented innovation and profitability, growth, and overall performance. The research by Geffen and Rothenberg (2000) on the performance of the US automobile assembly plants revealed that partnerships between original equipment manufacturers (OEMs) and their suppliers improve manufacturing operations and help meeting production quality and cost goals subsequently supporting OEM's growth. Davenport (2013) argues that through process innovation, firms can support low-cost producer strategies passing the savings on to customers, which can further increase growth. Similarly, Camisón and Villar-López (2014) showed that organizational innovation favors the development of technological innovation capabilities. More recently, Lin et al. (2016) study of Chinese firms showed that management innovation facilitates changes, including technical innovation. Hence, the following hypotheses follow:

- H5a. Operations-oriented innovation positively affects profitability.
- H5b. Operations-oriented innovation positively affects overall performance.
- H5c. Operations-oriented innovation positively affects growth.

3.6 Offering-Oriented Innovation, Profitability, Growth, and Overall Performance

The research by Langerak and Hultink (2005) investigated the impact of new product development on profitability in 233 manufacturing firms, showing that it indeed enhanced speed and profitability and, hence, firm's financial performance. Eggert et al.'s (2011) survey of German mechanical engineering companies suggested that when companies with high product innovation activity also offered services that support the product, this directly increased firm profitability. A similar positive relationship was found in Gemser and Leenders's (2001) research of two Dutch manufacturing industries, namely home furniture and precision instruments, but also in Wright et al.'s (2005) work, where it was found that product innovation has a positive effect on performance in hostile environments. In a more recent survey of

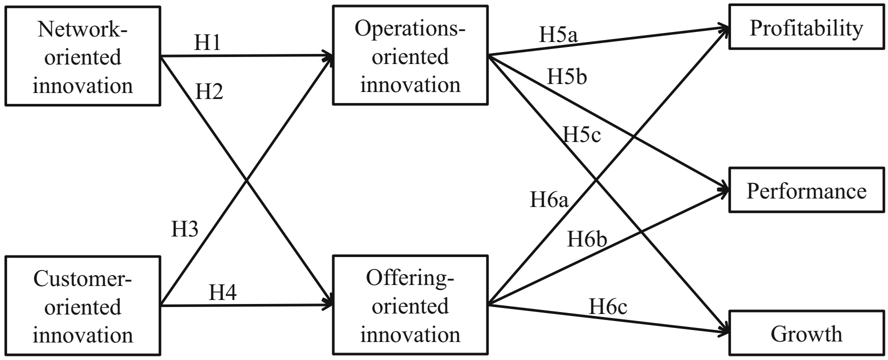


Fig. 1 The proposed research model and hypotheses

451 Spanish firms by Jiménez-Jiménez and Sanz-Valle (2011), the data also revealed a positive impact of innovation (including product innovation) on performance. As a result, we hypothesize the following:

- H6a. Offering-oriented innovation positively affects profitability.
- H6b. Offering-oriented innovation positively affects overall performance.
- H6c. Offering-oriented innovation positively affects growth.

Figure 1 shows the proposed research model describing the interactions between various innovation capabilities as well as growth, profitability, and overall performance.

The model presented in Fig. 1 suggests a cause-and-effect relationship between the constructs that are to the left of the other constructs, i.e., network-oriented and customer-oriented innovations lead to operations-oriented and offering-oriented innovation, then operations-oriented offering-oriented innovations result in higher profitability, overall performance, and growth. In dynamic capability theory terms, we argue that network-oriented and customer-oriented innovations are dynamic capabilities that affect ordinary (offering-oriented and operations-oriented) innovation capabilities.

4 Research Methodology

In this research, we draw data from the food and drink manufacturing sector in Greece, one of the country's largest and most important industries. *Food* processing is a critical sector in Greece, accounting for 10% of employment, holding the most significant share in terms of the value of production (19.8%) and gross value added (24.4%). In comparison, it ranks second in turnover (19.7%) (IOBE, 2017). The Greek food industry has traditionally played a central role in processing agricultural raw materials and providing food supply to the nation. Most of the food manufacturing companies (approximately 14,000) are family-based, with more than 97 percent of Greek enterprises categorized as "micro" (fewer than ten people employed and an annual turnover under €2 million) (IOBE, 2017). Many specialize in the production of food from local agriculture. The importance of the sector for the economy and the reliance of the sector on micro-companies make the food sector in Greece a good proxy for studying the development of innovation capabilities and the impact on performance.

4.1 Data Collection

We adopted the survey used for different innovations and performance by Chen et al. (2011). The constructs network-oriented innovation, customer-oriented innovation, operations-oriented innovation, offering-oriented innovation, and overall performance were assessed on a seven-point Likert scale with statements where respondents chose between strongly disagree, disagree, somewhat disagree, neither disagree nor agree, somewhat agree, agree, and strongly agree. Growth and profitability were also assessed on a seven-point Likert scale where respondents assessed their growth and profitability with respect to their competitors using comparators much worse, worse, moderately worse, neutral, moderately better, better, and much better. The survey tool is given in Appendix A.

At the beginning of the research, we established a database of 1200 food manufacturers and processors operating in the sector through industry associations. These companies operated in nine different subsectors: meat and meat products, processed fish and marine products, processed fruits and vegetables, fats and oils, dairy products, flour and grain mill products, fresh and processed salads and dressings, bakeries, snack and confectioneries, drinks and beverages.

The primary data collection started in October 2010 and finished in February 2011. Before commencing primary data collection, the data collection tool was piloted in September 2010 with two companies and one of the senior managers of the association of exporters, the majority of the members of which are food companies. The companies participated in this pilot voluntarily. Following the piloting of the data collection tool, some of the wordings and definitions were rephrased because the data collection tool was based on literature in English but was administered in Greek. Initially, the data collection tool was prepared in English and then translated to Greek. Then the Greek version was translated back to English to assure consistency of concepts and underlying ideas and to minimize differences in meaning. So multiple rounds of translations took place to ensure the data question tool was valid, i.e., the meaning of the questions in English and Greek was the same. In the pilot stage, the scales did not change. Some of the definitions for the types of innovation were further clarified for potential participants.

Questionnaires were sent through a link to an online survey. There were three rounds of data collection. The participants received the first reminder about 3 weeks

Table 1 Food subsectors represented in the sample

Sector	Frequency	Percentage
Bakeries, snacks, and confectioneries	26	27.1
Processed fruits and vegetables	17	17.7
Dairy products	11	11.5
Drinks and beverages	11	11.5
Meat and meat products	10	10.4
Fats and oils	6	6.2
Processed fish and marine products	4	4.2
Other	11	11.5

after sending the survey link. The second reminder was sent in about 3 weeks following the first reminder. There was a third and final reminder, which was sent approximately 3 weeks after the second reminder.

The survey included 37 questions, including questions about the company. The average time to complete the questionnaire was 25 minutes. The companies were given a phone call in January to facilitate the collection of responses. The participants were also given the opportunity to receive the results of the survey. Moreover, there was a prize, a gift card of $150.00 value, for the participants. At the end of the data collection, 154 questionnaires were returned with a response rate of 13%. After removing responses with missing values, data from 96 companies were used in the partial least squares structural equation model (PLS-SEM).

Table 1 presents the sectors represented in the sample, with bakeries, snacks, and confectioneries having the highest frequency of responding companies, followed by processed foods and vegetables. There were also 11 companies; participants responded they did not belong to any of the listed food subsectors.

Our survey participants had roles in sales, production, marketing, R&D, procurement, and logistics departments. Most of them held roles in sales, production, and marketing, which are the three departments highly involved in developing and introducing new products to the market together with the R&D department. These participants were directors or managers of their departments because they would be in a position to answer questions about innovation and organizational performance. It should be noted that not all companies that were represented in the research had a dedicated R&D department for developing new products.

Table 2 presents the years the company has been in operation at the time of data collection, and more than half of the companies represented in the research were relatively recently established with less than 20 years in operation. This also can be read as an indicator of the dynamism in the Greek food sector.

Table 2 Participant companies' years in operation

Sector	Frequency	Percentage
Up to 20 years	49	51.0
20–40 years	30	31.2
40–60 years	11	11.5
60–80 years	2	2.1
More than 80 years	4	4.2

4.2 Higher-Order Partial Least Squares Models

Hierarchical constructs can be defined as constructs that involve more than one dimension (Wetzels et al., 2009). Higher-order constructs allow for more theoretical parsimony and reduce model complexity (MacKenzie et al., 2005). This is regarded as theoretical utility as the theory requires general constructs consisting of specific dimensions or facets (Edwards, 2001). Partial least squares structural equation modeling allows for conceptualization of a hierarchical model through repeated use of indicator variables (Tenenhaus et al., 2005).

In this research, we conceptualize network-oriented innovation as a higher-order construct comprising chain and channel innovation. In a similar vein, customer-oriented innovation involves innovation in communication, relationship, and interaction with customers. Operations-oriented innovation involves innovation in management and processes, whereas offering-oriented innovation involves innovation in platforms and solutions. While network-oriented, customer-oriented, operations-oriented, and offering-oriented innovation constructs are second-order constructs, constructs related to profitability, performance, and growth of firms are first-order constructs since they do not involve other latent constructs.

5 Results

The analyses were performed with SmartPLS 3 (Ringle et al., 2014) on a 64-bit MacBook Pro with Intel Core i7 2.2 GHz processor and OS X version 10.9.4. We present in the following subsections the measurement model and the structural model and interpret the findings from these models.

5.1 The Measurement Model

In our PLS model, we assume all latent variables are measured by reflective indicators, and this assumption requires reflective indicators to have a strong mutual association, i.e., they will be highly correlated. This is because reflective measurement assumes that the latent variable is the cause of reflective indicators.

Another requirement is that reflective indicators should not load higher on another construct than their own construct.

Common-method variance We checked the common-method variance to ensure the data had no issues regarding the response bias since each organization was represented by a single respondent (Prajogo & McDermott, 2014). Following on from Prajogo and McDermott (2014), we checked the number of factors extracted from the 37 indicators in a principal component analysis and found seven factors were extracted in the unrotated solution. These seven factors explained 72% of the variation in the data. We also forced only one factor to be extracted from the data, which resulted in 36% of the variation being explained. Hence, the common-method variance did not pose a significant problem.

Indicator reliability In the case of reflectively measured constructs, loadings above 0.70 indicate that the construct explains over 50% of the indicator's variance (Sarstedt et al., 2014). Table 6 in Appendix A provides an overview of the constructs and their corresponding indicators with loadings. According to our assessment of the indicator reliability, all 66 indicators, which are all reflective, have loadings above 0.70 (Table 6). Hence, the indicators in the measurement model show satisfactory reliability levels.

Internal consistency reliability The composite reliability scores above 0.70 are considered "satisfactory to good" in line with Jöreskog's (1971) and Hair Jr et al. (2013). The measurement model's internal consistency reliability is achieved owing to composite reliability scores of 0.891 and higher (Table 4, composite reliability column).

Convergent validity A construct is expected to explain the variance in its items. The average variance extracted (AVE) scores show the convergence of each construct with its measurement items. The AVE score is calculated using the mean of squared loadings for all indicators associated with a construct. It should be 0.50 or higher, meaning that the construct explains at least 50% of the variance in its indicators. The AVE scores in Table 3 (column average variance extracted) are higher than the critical threshold value of 0.50, supporting the measures' convergent validity.

Discriminant validity Discriminant validity determines the extent to which a construct is empirically distinct from other constructs. The Fornell–Larcker (1981) criterion compares the square root of the AVE of each construct with the inter-construct correlations with all other constructs. Table 4 shows that the square root of the AVE of each construct exceeds the inter-correlations, supporting discriminant validity (Wetzels et al., 2009).

Table 3 Composite reliability and average variance extracted

Constructs	Composite reliability	Average variance extracted
Chain	0.907	0.765
Channel	0.904	0.760
Communication	0.910	0.772
Customer oriented	0.928	0.587
Growth	0.957	0.882
Interaction	0.940	0.839
Management	0.899	0.748
Network oriented	0.899	0.599
Offering oriented	0.920	0.592
Operations oriented	0.902	0.606
Overall performance	0.913	0.779
Platform	0.927	0.761
Process	0.905	0.760
Profitability	0.891	0.804
Relationship	0.920	0.793
Solution	0.921	0.745

Table 4 Inter-correlations of the latent variables for first-order constructs

Construct	1	2	3	4	5	6	7	8	9	10	11	12
1. Chain	0.874											
2. Channel	0.573	0.872										
3. Communication	0.293	0.295	0.879									
4. Growth	0.399	0.267	0.309	0.939								
5. Interaction	0.198	0.202	0.591	0.169	0.916							
6. Management	0.329	0.331	0.436	0.548	0.393	0.865						
7. Overall performance	0.369	0.298	0.416	0.685	0.318	0.499	0.883					
8. Platform	0.344	0.391	0.478	0.270	0.409	0.274	0.275	0.872				
9. Process	0.232	0.259	0.533	0.317	0.587	0.610	0.363	0.414	0.872			
10. Profitability	0.305	0.302	0.306	0.698	0.215	0.551	0.747	0.251	0.285	0.897		
11. Relationship	0.411	0.338	0.675	0.437	0.539	0.402	0.536	0.642	0.507	0.396	0.890	
12. Solution	0.400	0.293	0.545	0.339	0.369	0.302	0.461	0.573	0.434	0.374	0.708	0.863

Notes: Square root of AVE on diagonal

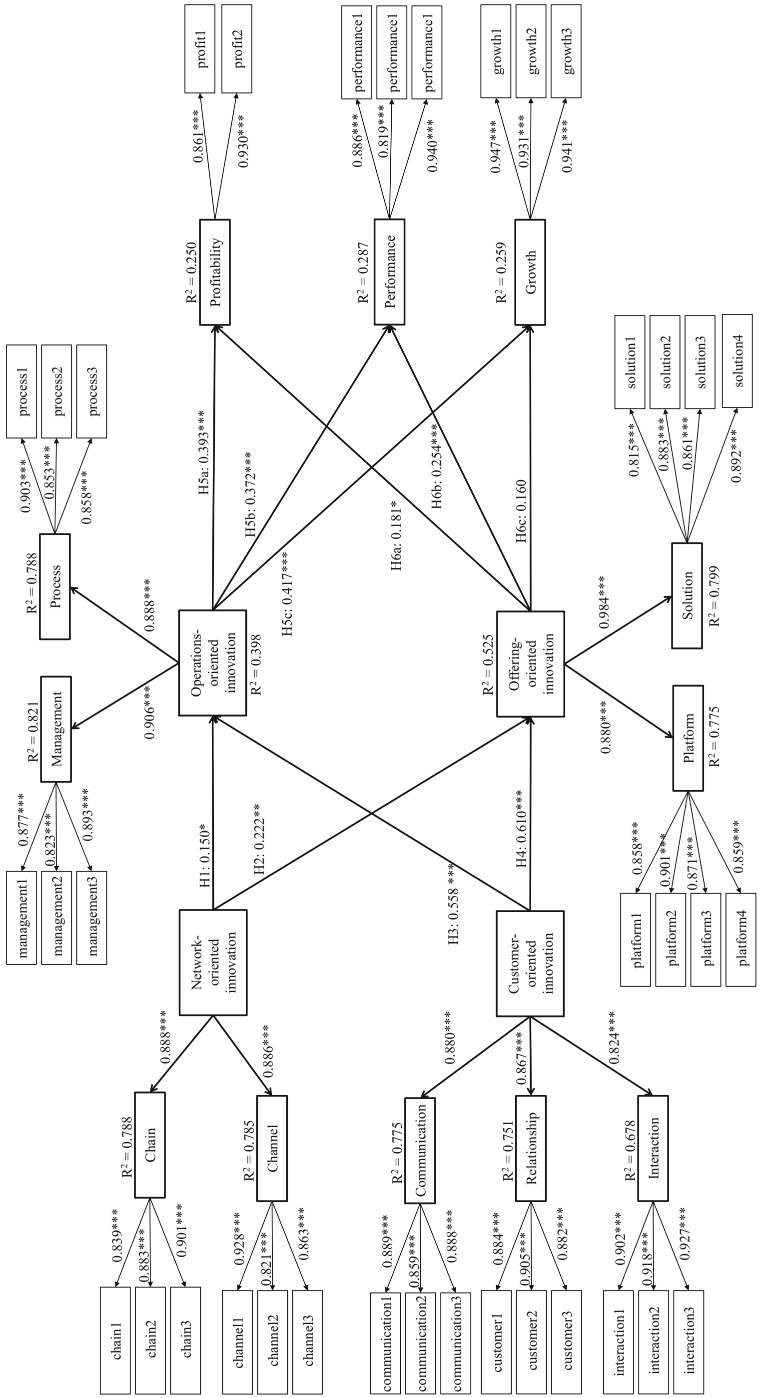
5.2 The Structural Model

Once the measurement model requirements are satisfied in terms of the common-method variance (the model does not suffer from this problem), indicator reliability (all measurement items have at least 0.7 or higher loadings), internal consistency reliability (all composite reliability scores are above 0.7), convergent validity (all AVE are above 0.5), and discriminant validity (all inter-construct correlations are smaller than the square root of the construct's AVE), we continue with the structural model. The structural model fit statistics are as follows: Root mean square error of approximation (RMSEA) = 0.099 and standardized root mean square error = 0.08.

Predictive Relevance (R^2 and Q^2) Predictive relevance is established by the coefficient of determination (R^2), cross-validated redundancy (Q^2), and the path coefficients. Figure 2 shows the predictive power of endogenous constructs with R^2 values reported inside blue circles. The R^2 values in Fig. 2 are moderate for operations-oriented ($R^2 = 0.398$) and offering-oriented ($R^2 = 0.525$) innovation and comparably weak for profitability ($R^2 = 0.250$), overall performance ($R^2 = 0.287$) and growth ($R^2 = 0.259$). However, all R^2 values are significant ($p < 0.01$) and considering the antecedents of these constructs, it is possible to conclude R^2 value is satisfactory. We used blindfolding to evaluate the model's predictive relevance for each endogenous construct. We ran the blindfolding procedure with an omission distance of seven. Cross-validated redundancy values for all five endogenous constructs were well above zero (operations-oriented innovation: 0.216; offering-oriented innovation: 0.302; profitability: 0.188; overall performance: 0.206; growth: 0.217), providing support for the model's predictive relevance.

To check the significance and relevance of the structural model, we ran the bootstrapping procedure (96 cases, 5000 samples, no sign changes option) and found that nine of ten structural relationships (Table 5) are significant ($p < 0.10$).

The results in Table 5 highlight the important role of customer-oriented innovation on operations-oriented and offering-oriented innovation with significant and relatively strong path coefficients (0.558 and 0.610, respectively). Table 5 also shows the role of operations-oriented innovation in achieving profitability, overall performance, and growth with significant and moderately strong path coefficients (0.393, 0.372, and 0.417, respectively). Surprisingly offering-oriented innovation has a significant effect on profitability and overall performance but not on growth.



Notes: *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$ (two-sided test)

Fig. 2 Theoretical model and analysis results

Table 5 Significance and relevance of path coefficients

Hypothesis	Path coefficient
H1: Network Oriented → Operations Oriented	0.150^{*}
H2: Network Oriented → Offering Oriented	0.222^{**}
H3: Customer Oriented → Operations Oriented	0.558^{***}
H4: Customer Oriented → Offering Oriented	0.610^{***}
H5a: Operations Oriented → Profitability	0.393^{***}
H5b: Operations Oriented → Overall Performance	0.372^{***}
H5c: Operations Oriented → Growth	0.417^{***}
H6a: Offering Oriented → Profitability	0.181^{*}
H6b: Offering Oriented → Overall Performance	0.254^{***}
H6c: Offering Oriented → Growth	0.160

Notes: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$ (two-sided test)

6 Discussion

The results support all the hypotheses but one, namely *H6c: Offering-oriented innovation positively affects growth*, and they provide new insights into the impact of different innovation capabilities on three aspects of organizational performance, i.e., profitability, growth, and overall performance. These insights are particularly significant because prior studies have not investigated the relative effects of different innovation capabilities on profitability, growth, and overall firm performance, in particular, the interrelationships between dynamic and ordinary innovation capabilities (H1–H4).

The results show that dynamic (network-oriented and customer-oriented) innovation capabilities have a positive impact on ordinary (operations-oriented and offering-oriented) innovation capabilities. Positive and significant path coefficients suggest that firms can improve their ordinary innovation capabilities through progress in dynamic innovation capabilities. This is in line with the recent findings of Macher and Mowery (2009) and Protogerou et al. (2012) on the impact of dynamic capabilities on operational capabilities. In a similar vein, operations-oriented innovation has a positive and significant impact on profitability, overall performance, and growth, whereas offering-oriented innovation has a positive and significant impact on profitability and overall performance. There was no evidence to support a positive relationship between offering-oriented innovation and growth, which may be investigated again in a similar context with different data.

Second, these results are important because they show the relationship between different innovation capabilities, which are higher-order constructs comprised of multiple latent constructs. The study shows successful implementation of higher-order PLS in innovation capabilities in terms of the data analysis tool. It is possible to hypothesize each different innovation capability as reflected on the lower order constructs, reported in Fig. 2. Although different innovation capabilities and their impact on firm performance have been addressed in previous research concerning dynamic capabilities, no research has investigated the interaction between dynamic and ordinary innovation capabilities as explanatory constructs for firm performance. Similar use of PLS-SEM can be found in Wilden et al. (2013), where the authors test how competitive intensity and organizational structure influence the effects of dynamic capabilities on firm performance.

7 Conclusion

The findings of the study suggest that there are cause-and-effect relationships between different innovation capabilities as hypothesized. In other words, network-oriented and customer-oriented innovations lead to operations-oriented and offering-oriented innovation, which positively affects firm performance. The results showed that customer-oriented, operations-oriented, and offering-oriented innovation were significant drivers of profitability, performance, and growth. The effect of network-oriented innovation on operations-oriented and offering-oriented innovation was positive. It is important to recognize the vital role of operations-oriented innovation as a driver of firm profitability, overall performance, and growth and the important role customer-oriented innovation plays in operations-oriented and offering-oriented innovation.

7.1 Theoretical Contribution

Research on innovation in the food industry has taken a rather "traditional," narrow view by placing enormous emphasis on a firm's capabilities to innovate across products and/or processes (Traill & Meulenberg, 2002; Baker, 2007; Karantininis et al., 2010; Matopoulos & Bourlakis, 2011; Baregheh et al., 2012). In this research, we moved away from the traditional firm-based product/process innovation debate using Chen et al.'s (2011) value-based view of innovation as the basis for describing and classifying innovation capability, linking innovation to new value created for customers.

Based on this approach, we sought to understand the link between dynamic and ordinary innovation capabilities and their influence on organizational performance. The findings of the study empirically support the hypotheses that dynamic innovation capabilities (network and customer oriented innovation) positively affect operations-oriented innovation capabilities, which in turn positively affect performance (profitability, growth and overall performance). The study contributes to the ongoing debates on the nature of dynamic capabilities (e.g., Eisenhardt & Martin, 2000; Easterby-Smith et al., 2009; Wilden et al., 2013) and on the link between dynamic capabilities and ordinary capabilities, and the role and performance consequences of dynamic capabilities (Cepeda & Vera, 2007; Protogerou et al., 2012). Concerning the first debate, our research helps to clarify the nature of dynamic capabilities in the context of innovation, bridging these two fields together.

The integration of these two fields helps to realize that dynamic capabilities do not have to be viewed only as firm-level capabilities (Pavlou & El Sawy, 2011). Dynamic capabilities could be the way to approach innovation, providing more depth to the explanation of non-firm-driven innovation and its impact on firm-based innovation.

As for the link between dynamic and ordinary capabilities, very few empirical analyses have attempted to explore their performance effects; so, our research provides support for the arguments of Macher and Mowery (2009) and Protogerou et al. (2012) that dynamic capabilities have a significant but indirect effect on organizational performance.

7.2 Managerial Implications

The goal of the chapter was to expand the view of a firm's capability to innovate and to explore the interrelations among different innovation capabilities and their influence on organizational performance. We argue that in the food sector, but also in other sectors with similar characteristics, where, for example, the level of research and development (R&D) expenditures is not as high as in traditional manufacturing (e.g., electronics or automotive) and where new products emanate mainly from variations of older ones (Wijnands et al., 2007), there is a bigger need to expand the pool of innovation capability so as to include non-firm-based innovation capabilities. What is more important is to understand the impact of these different innovation capabilities on profitability, growth, and overall performance. This could be of particular importance not only for SMEs but also for larger firms where resources are also scarce, and the outcome of traditional R&D-based innovation is not always translated into new value for the customer. Finally, this research may be of value for firms coming from smaller economies (like the Greek economy, which belongs in the economic periphery of the European Union). In such contexts, innovation capabilities may be an answer to the complex problem of economic downturn, and perhaps companies should be focusing on their dynamic innovation capabilities in the first place to entertain higher profitability, overall performance, and growth.

7.3 Limitations

Despite its contributions, the study presents some limitations. First, the data used was provided by a single respondent in each firm, in most cases the sales/production manager. Another limitation of the research is the origin of the data. Although we believe, for the reasons described in the introduction, that the food industry is ideal for exploring the interrelationships between different innovation capabilities, exploring other sectors will help generalize the results further. In a similar vein, the research needs to be extended to an international context (e.g., USA, Northern Europe, and the Far East) to check whether culture and local/national food supply chain configurations lead to different findings. Also, it is probably worth exploring whether the different innovation capabilities are independent or mutually exclusive.

Finally, an online questionnaire with an average completion time of 25 minutes could explain the 38% of responses being lost due to missing data. In fact, we were aware of the length of the survey, but to be able to answer the research question, a survey at such length was needed. Our sample size was 96; greater insights could be gained by expanding this to a larger sample so that subcategories of the sector can be analyzed and compared in terms of the innovation capabilities and the effect on performance.

Acknowledgments The authors wish to thank Jiyao Chen for his thorough and helpful feedback on an earlier version of this manuscript.

Appendix A Data Collection Tool and Indicator Loadings

Table 6 Constructs and corresponding indicators with their loadings

Constructs/indicators	Loading
Chain innovation	
We have employed new mechanisms for sourcing our input materials and services	0.839
We have found creative ways to improve the efficiency of our supply chain	0.883
We have introduced new procurement methods in our supply chain	0.901
Channel innovation	
We have created new ways to manage our direct and indirect channels of distribution	0.928
We have created new self-service distribution channels	0.821
We have changed the role of the intermediaries we use in our channels	0.863
Network-oriented innovation (higher-order component)	
We have employed new mechanisms for sourcing our input materials and services	0.778
We have found creative ways to improve the efficiency of our supply chain	0.733
We have introduced new procurement methods in our supply chain	0.813
We have created new ways to manage our direct and indirect channels of distribution	0.837
We have created new self-service distribution channels	0.722
We have changed the role of the intermediaries we use in our channels	0.753
Communication innovation	
We lead our competitors in new ways to promote our products and services	0.889
We are recognized for the creative ways we have developed our brands	0.859
Our products and brands are positioned in a way that clearly differentiates them from competitors	0.888

(continued)

Table 6 continued

Constructs/indicators	Loading
Interaction innovation	
Compared to our closest competitors, we have created a better customer experience at every stage in the buying cycle	0.902
We lead our competitors in managing customer interactions with our company	0.918
We have created a single face for interacting with customers through different channels	0.927
Relationship innovation	
We have successfully identified new customer segments	0.884
We have found creative ways of segmenting our customers	0.905
We successfully serve customer needs that our competitors have not identified	0.882
Customer-oriented innovation (higher-order component)	
We lead our competitors in new ways to promote our products and services	0.774
We are recognized for the creative ways we have developed our brands	0.768
Our products and brands are positioned in a way that clearly differentiates them from competitors	0.778
Compared to our closest competitors, we have created a better customer experience at every stage in the buying cycle	0.752
We lead our competitors in managing customer interactions with our company	0.791
We have created a single face for interacting with customers through different channels	0.771
We have successfully identified new customer segments	0.762
We have found creative ways of segmenting our customers	0.749
We successfully serve customer needs that our competitors have not identified	0.751
Management innovation	
We have implemented new management practices to improve organizational efficiency	0.877
Compared to our closest competitors, our organizational design allows us to respond more quickly to new opportunities and threats	0.823
We have introduced new methods of organizing work responsibilities	0.893
Process innovation	
We have restructured internal business processes to significantly reduce operating costs	0.903
We lead our competitors in improving the efficiency of our internal business processes	0.853
We have changed our internal operating processes in the past 3 years	0.858
Operations-oriented innovation (higher-order component)	
We have implemented new management practices to improve organizational efficiency	0.738
Compared to our closest competitors, our organizational design allows us to respond more quickly to new opportunities and threats	0.798
We have introduced new methods of organizing work responsibilities	0.809
We have restructured internal business processes to significantly reduce operating costs	0.812

(continued)

Table 6 continued

Constructs/indicators	Loading
We lead our competitors in improving the efficiency of our internal business processes	0.758
We have changed our internal operating processes in the past 3 years	0.750
Platform innovation	
We have found new ways to use common sets of technologies to develop our products and services	0.858
Our shared platform allows us to introduce new products and services faster than our competitors	0.901
Our shared platform gives us significant cost advantage in developing new products and services	0.871
We have created proprietary platforms to sustain our competitive advantage	0.859
Solution innovation	
We have combined products and services to create integrated solutions for our customers	0.815
We lead our closest competitors in creating integrated solutions for our customers	0.883
We provide an entire solution for the end-to-end problems of our customers	0.861
We lead our competitors in offering customized solutions for customers	0.892
Offering-oriented innovation (higher-order component)	
We have found new ways to use common sets of technologies to develop our products and services	0.728
Our shared platform allows us to introduce new products and services faster than our competitors	0.814
Our shared platform gives us significant cost advantage in developing new products and services	0.793
We have created proprietary platforms to sustain our competitive advantage	0.732
We have combined products and services to create integrated solutions for our customers	0.725
We lead our closest competitors in creating integrated solutions for our customers	0.807
We provide an entire solution for the end-to-end problems of our customers	0.743
We lead our competitors in offering customized solutions for customers	0.805
Profitability	
Profitability	0.861
Return on sales	0.930
Performance	
The overall performance of the business met expectations last year	0.886
The overall performance of the business last year exceeded that of our major competitors	0.819
Top management was very satisfied with the overall performance of the business last year	0.940
Growth	
Market share growth	0.947
Sales growth in existing markets	0.931
New customer acquisition	0.941

References

Avermaete, T., Viaene, J., Morgan, E. J., Pitts, E., Crawford, N., & Mahon, D. (2004). Determinants of product and process innovation in small food manufacturing firms. *Trends in Food Science & Technology, 15*(10), 474–483.

Baker, D. K. (2007). New product introductions in the food industry – results from a Danish survey. Technical report, Institute of Food and Resource Economics, Copenhagen. Report No: 189.

Baregheh, A., Rowley, J., Sambrook, S., & Davies, D. (2012). Food sector SMEs and innovation types. *British Food Journal, 114*(11), 1640–1653.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120.

Birkinshaw, J., Hamel, G., & Mol, M. J. (2008). Management innovation. *Academy of Management Review, 33*(4), 825–845.

Boly, V., Morel, L., Assielou, N., & Camargo, M. (2014). Evaluating innovative processes in French firms: Methodological proposition for firm innovation capacity evaluation. *Research Policy, 43*(3), 608–622.

Bruin, S., & Jongen, T. R. (2003). Food process engineering: The last 25 years and challenges ahead. *Comprehensive Reviews in Food Science and Food Safety, 2*(2), 42–81.

Camisón, C., & Villar-López, A. (2014). Organizational innovation as an enabler of technological innovation capabilities and firm performance. *Journal of Business Research, 67*(1), 2891–2902.

Cepeda, G., & Vera, D. (2007). Dynamic capabilities and operational capabilities: A knowledge management perspective. *Journal of Business Research, 60*(5), 426–437.

Chen, J., Sawhney, M., & Guth, W. (2011). The value-oriented view of innovation: Defining and measuring business innovation, *Academy of Management Meeting*, San Antonio.

Chesbrough, H. W. (2003). *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business School Press.

Cohen, M. A., Cull, C., Lee, H. L., & Willen, D. (2000). Saturn's supply-chain innovation: High value in after-sales service. *Sloan Management Review, 41*(4), 93–101.

Collis, D. J. (1994). Research note: How valuable are organisational capabilities? *Strategic Management Journal, 15*(S1), 143–152.

Cousins, P. D., Lawson, B., Petersen, K. J., & Handfield, R. B. (2011). Breakthrough scanning, supplier knowledge exchange, and new product development performance. *Journal of Product Innovation Management, 28*(6), 930–942.

Daft, R. L. (1978). A dual-core model of organisational innovation. *Academy of Management Journal, 21*(2), 193–210.

Damanpour, F. (1991). Organisational innovation: A meta-analysis of effects of determinants and moderators. *Academy of Management Journal, 34*(3), 555–590.

Damanpour, F., Szabat, K. A., & Evan, W. M. (1989). The relationship between types of innovation and organisational performance. *Journal of Management Studies, 26*(6), 587–602.

Damanpour, F., Walker, R. M., & Avellaneda, C. N. (2009). Combinative effects of innovation types and organisational performance: A longitudinal study of service organisations. *Journal of Management Studies, 46*(4), 650–675.

Davenport, T. H. (1993). *Process innovation: Reengineering work through information technology*. Harvard Business Press.

Davenport, T. H. (2013). *Process innovation: Reengineering work through information technology*. Harvard Business Press.

Debruyne, M. (2014). *Customer innovation: Customer-centric strategy for enduring growth*. Kogan Page Publishers.

Easterby-Smith, M., Lyles, M. A., & Peteraf, M. A. (2009). Dynamic capabilities: Current debates and future directions. *British Journal of Management, 20*(s1), S1–S8.

Edwards, J. R. (2001). Multidimensional constructs in organisational behavior research: An integrative analytical framework. *Organisational Research Methods, 4*(2), 144–192.

Eggert, A., Hogreve, J., Ulaga, W., & Muenkhoff, E. (2011). Industrial services, product innovations, and firm profitability: A multiple-group latent growth curve analysis. *Industrial Marketing Management, 40*(5), 661–670.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: What are they? *Strategic Management Journal, 21*(10–11), 1105–1121.

Fitjar, R. D., & Rodríguez-Pose, A. (2013). Firm collaboration and modes of innovation in Norway. *Research Policy, 42*(1), 128–138.

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research*, 382–388.

Forsman, H. (2011). Innovation capacity and innovation development in small enterprises. A comparison between the manufacturing and service sectors. *Research Policy, 40*(5), 739–750.

Foss, N. J., Laursen, K., & Pedersen, T. (2011). Linking customer interaction and innovation: The mediating role of new organisational practices. *Organisation Science, 22*(4), 980–999.

Gawer, A., & Cusumano, M. A. (2014). Industry platforms and ecosystem innovation. *Journal of Product Innovation Management, 31*(3), 417–433.

Geffen, C. A., & Rothenberg, S. (2000). Suppliers and environmental innovation: The auto-motive paint process. *International Journal of Operations and Production Management, 20*(2), 166–186.

Gemser, G., & Leenders, M. A. A. M. (2001). How integrating industrial design in the product development process impacts on company performance. *Journal of Product Innovation Management, 18*(1), 28–38.

GOV.UK. (2014). Food product innovation: A background paper. Available at https://www.gov.uk/government/publications/food-and-drink-in-the-uk-investment-opportunities/food-and-drink-in-the-uk-investment-opportunities. Accessed 19 Feb 2014.

Griffin, A. (1997). Pdma research on new product development practices: Updating trends and benchmarking best practices. *Journal of Product Innovation Management, 14*(6), 429–458.

Gunday, G., Ulusoy, G., Kilic, K., & Alpkan, L. (2011). Effects of innovation types on firm performance. *International Journal of Production Economics, 133*(2), 662–676.

Hair, J. F., Jr., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2013). *A primer on partial least squares structural equation modeling (PLS-SEM)*. SAGE Publications.

Hamel, G. (2006). The why, what, and how of management innovation. *Harvard Business Review, 84*(2), 72–83.

Helfat, C. E., Finkelstein, S., Mitchell, W., Peteraf, M. A., Singh, H., Teece, D. J., & Winter, S. G. (2007). *Dynamic capabilities: Understanding strategic change in organisations*. Blackwell Publishing.

Helfat, C. E., & Winter, S. G. (2011). Untangling dynamic and operational capabilities: Strategy for the (N) ever-changing world. *Strategic Management Journal, 32*(11), 1243–1250.

Hoyer, W., Chandy, R., Dorotic, M., Krafft, M., & Singh, S. S. (2010). Consumer co-creation in new product development. *Journal of Service Research, 13*(3), 283–296.

IOBE. (2017). Food and drink industry: Facts and figures 2016 (in Greek). Available at http://iobe.gr/research_dtl_en.asp?RID=140. Last accessed 18 May 2017.

Jacobides, M. G., Knudsen, T., & Augier, M. (2006). Benefiting from innovation: Value creation, value appropriation and the role of industry architectures. *Research Policy, 35*(8), 1200–1221.

Jiménez-Jiménez, D., & Sanz-Valle, R. (2011). Innovation, organisational learning, and performance. *Journal of Business Research, 64*(4), 408–417.

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika, 36*(2), 109–133.

Karantininis, K., Sauer, J., & Furtan, W. H. (2010). Innovation and integration in the agri-food industry. *Food Policy, 35*(2), 112–120.

Khosrow-Pour, M. (2006). *Cases on information technology and business process reengineering*. Idea Group.

Kim, Y., & Lui, S. S. (2015). The impacts of external network and business group on innovation: Do the types of innovation matter? *Journal of Business Research, 68*(9), 1964–1973.

Kirner, E., Kinkel, S., & Jaeger, A. (2009). Innovation paths and the innovation performance of low-technology firms—An empirical analysis of German industry. *Research Policy, 38*(3), 447–458.
Kogut, B. (2000). The network as knowledge: Generative rules and the emergence of structure. *Strategic Management Journal, 21*(3), 405–425.
Kotabe, M., & Scott Swan, K. (1995). The role of strategic alliances in high-technology new product development. *Strategic Management Journal, 16*(8), 621–636.
Langerak, F., & Hultink, E. J. (2005). The impact of new product development acceleration approaches on speed and profitability: Lessons for pioneers and fast followers. *IEEE Transactions on Engineering Management, 52*(1), 30–42.
Lawson, B., & Samson, D. (2001). Developing innovation capability in organisations: A dynamic capabilities approach. *International Journal of Innovation Management, 5*(3), 377–400.
Lee, K., & Kang, S.-M. (2007). Innovation types and productivity growth: Evidence from Korean manufacturing firms. *Global Economic Review, 36*(4), 343–359.
Lee, H., & Kelley, D. (2008). Building dynamic capabilities for innovation: An exploratory study of key management practices. *R&D Management, 38*(2), 155–168.
Lee, S. M., Lee, D., & Schniederjans, M. J. (2011). Supply chain innovation and organisational performance in the healthcare industry. *International Journal of Operations and Production Management, 31*(11), 1193–1214.
Li, S., Ragu-Nathan, B., Ragu-Nathan, T., & Subba Rao, S. (2006). The impact of supply chain management practices on competitive advantage and organisational performance. *Omega, 34*(2), 107–124.
Lin, R.-J., Chen, R.-H., Kuan, K., & Chiu, S. (2010). Customer relationship management and innovation capability: An empirical study. *Industrial Management & Data Systems, 110*(1), 111–133.
Lin, H. F., Su, J. Q., & Higgins, A. (2016). How dynamic capabilities affect adoption of management innovations. *Journal of Business Research, 69*(2), 862–876.
Macher, J. T., & Mowery, D. C. (2009). Measuring dynamic capabilities: Practices and performance in semiconductor manufacturing. *British Journal of Management, 20*(s1), S41–S62.
MacKenzie, S. B., Podsakoff, P. M., & Jarvis, C. B. (2005). The problem of measurement model misspecification in behavioral and organisational research and some recommended solutions. *Journal of Applied Psychology, 90*(4), 710–730.
Matopoulos, A., & Bourlakis, M. (2011). Identifying innovation strategies: Insights from the Greek food manufacturing sector. *International Journal of Innovation and Regional Development, 3*(2), 159–173.
McEvily, B., & Marcus, A. (2005). Embedded ties and the acquisition of competitive capabilities. *Strategic Management Journal, 26*(11), 1033–1055.
Menrad, K. (2004). Innovations in the food industry in Germany. *Research Policy, 33*(6), 845–878.
Mol, M., & Birkinshaw, J. (2012). Relating management innovation to product and process innovation: Private rents versus public gains. In T. Pitsis, A. Simpson, & E. Dehlin (Eds.), *Handbook of organisational and managerial innovation* (pp. 13–35). Edward Elgar Publishing Limited.
Moore, G. (2004). Darwin and the demon. *Harvard Business Review, 82*(7/8), 86–92.
Nishiguchi, T., & Ikeda, M. (1996). Suppliers' process innovation: Understated aspects of Japanese industrial sourcing. In T. Nishiguchi (Ed.), *Managing product development* (pp. 206–230). Oxford University Press.
O'Hern, M., & Rindfleisch, A. (2009). Customer co-creation: A typology and research agenda. In N. K. Malholtra (Ed.), *Review of marketing research* (Vol. 6, pp. 84–106). M.E. Sharpe.
Oke, A. (2007). Innovation types and innovation management practices in service companies. *International Journal of Operations and Production Management, 27*(6), 564–587.
Pavlou, P. A., & El Sawy, O. A. (2011). Understanding the elusive black box of dynamic capabilities. *Decision Sciences, 42*(1), 239–273.
Pisano, G. P. (1997). *The development factory: Unlocking the potential of process innovation*. Harvard Business Press.

Powell, W. W., Koput, K. W., & Smith-Doerr, L. (1996). Interorganisational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly, 41*(1), 116–145.

Prajogo, D., & McDermott, C. M. (2014). Antecedents of service innovation in SMEs: Comparing the effects of external and internal factors. *Journal of Small Business Management, 52*(3), 521–540.

Protogerou, A., Caloghirou, Y., & Lioukas, S. (2012). Dynamin capabilities and their indirect impact on firm performance. *Industrial and Corporate Change, 21*(3), 615–647.

Rangus, K., & Slavec, A. (2017). The interplay of decentralization, employee involvement and absorptive capacity on firms' innovation and business performance. *Technological Forecasting and Social Change*, 195–203.

Ringle, C. M., Wende, S., & Becker, J.-M. (2014). Smartpls 3.0. http://www.smartpls.com

Roberts, P. W., & Amit, R. (2003). The dynamics of innovative activity and competitive advantage: The case of Australian retail banking, 1981 to 1995. *Organisation Science, 14*(2), 107–122.

Rothaermel, F. T., & Hess, A. M. (2007). Building dynamic capabilities: Innovation driven by individual-, firm-, and network-level effects. *Organisation Science, 18*(6), 898–921.

Roy, S., Sivakumar, K., & Wilkinson, I. F. (2004). Innovation generation in supply chain relationships: A conceptual model and research propositions. *Journal of the Academy of Marketing Science, 32*(1), 61–79.

Sánchez, A. M., & Pérez, M. P. (2005). Supply chain flexibility and firm performance: A conceptual model and empirical study in the automotive industry. *International Journal of Operations and Production Management, 25*(7), 681–700.

Sarstedt, M., Ringle, C. M., Smith, D., Reams, R., & Hair, J. F., Jr. (2014). Partial least squares structural equation modeling (PLS-SEM): A useful tool for family business researchers. *Journal of Family Business Strategy, 5*(1), 105–115.

Sawhney, M., Wollcott, R., & Arroniz, I. (2006). The innovation radar: A 360-degree view of business innovation. *MIT Sloan Manage Review, 47*(3), 75–81.

Schumpeter, J. A. (1934). *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle* (Vol. 55). Transaction Publishers.

Sood, A., & Tellis, G. J. (2005). Technological evolution and radical innovation. *Journal of Marketing, 69*(3), 152–168.

Stewart-Knox, B., & Mitchell, P. (2003). What separates the winners from the losers in new food product development? *Trends in Food Science & Technology, 14*(1), 58–64.

Subramanian, A., & Nilakanta, S. (1996). Organisational innovativeness: Exploring the relationship between organisational determinants of innovation, types of innovations, and measures of organisational performance. *Omega, 24*(6), 631–647.

Tarafdar, M., & Gordon, S. R. (2007). Understanding the influence of information systems competencies on process innovation: A resource-based view. *Journal of Strategic Information Systems, 16*(4), 353–392.

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal, 28*(13), 1319–1350.

Teece, D., & Pisano, G. (1994). The dynamic capabilities of firms: An introduction. *Industrial and Corporate Change, 3*(3), 537–556.

Tenenhaus, M., Vinzi, V. E., Chatelin, Y.-M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics and Data Analysis, 48*(1), 159–205.

Traill, W. B., & Meulenberg, M. (2002). Innovation in the food industry. *Agribusiness, 18*(1), 1–21.

Von Hippel, E. (2005). *Democratizing Innovation*. MIT Press.

Walker, R. (2004). Innovation and organisational performance: Evidence and a research agenda. Advanced Institute of Management Research Working Paper, (002).

Wetzels, M., Odekerken-Schroder, G., & Van Oppen, C. (2009). Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. *Management Information Systems Quarterly, 33*(1), 177–195.

Wheelwright, S. C., & Clark, K. B. (1992). Creating project plans to focus product development. *Harvard Business Review, 70*(2), 70–82.

Wijnands, J., Banse, M., van Berkum, S., & Poppe, K. (2007). The external environment of the food industry. In J. Wijnands, B. van der Meulen, & K. Poppe (Eds.), *Competitiveness of the European Food Industry: An economic and legal assessment*. European Communities, Italy. Available at http://ec.europa.eu/enterprise/sectors/food/files/competitiveness_study_en.pdf. Accessed 26 Aug 2014.

Wilden, R., Gudergan, S. P., Nielsen, B. B., & Lings, I. (2013). Dynamic capabilities and performance: Strategy, structure and environment. *Long Range Planning, 46*(1), 72–96.

Winger, R., & Wall, G. (2006). Food product innovation: A background paper. Available at http://www.fao.org/docrep/016/j7193e/j7193e.pdf. Accessed 10 June 2014.

Winter, S. G. (2003). Understanding dynamic capabilities. *Strategic Management Journal, 24*(10), 991–995.

Wright, R. E., Palmer, J. C., & Perkings, D. (2005). Types of product innovations and small business performance in hostile and benign environments. *Journal of Small Business Strategy, 15*(2), 33–44.

Wu, H., Chen, J., & Jiao, H. (2016). Dynamic capabilities as a mediator linking international diversification and innovation performance of firms in an emerging economy. *Journal of Business Research, 69*(8), 2678–2686.

Yam, R., Lo, W., Tang, E. P., & Lau, A. K. (2011). Analysis of sources of innovation, technological innovation capabilities, and performance: An empirical study of Hong Kong manufacturing industries. *Research Policy, 40*(3), 391–402.

From E-Satisfaction to E-Repurchase Intention: How Is E-Repurchase Intention Mediated by E-Satisfaction and Moderated by Traditional Shopping Attitudes?

Esra Demirbaş, Gülberk Gültekin Salman, and Nimet Uray

Abstract Successive developments in information technologies have brought important developments in the business world, one of which is e-commerce. Undoubtedly, consumers' continuous adoption of online shopping, which has been specially accelerated as the result of the pandemic, is not likely to end or reduce after the Covid-19 passes, increasing volume and transaction in e-retailing make e-business more challenging.

Under these circumstances, the most important requirement of sustainable development and profitability in e-business management is to retain loyal customers rather than one-time buyers. Hence, in an e-commerce setup, understanding the repurchase intention of consumers is essential in sustaining growth. Most previous studies have focused on one or two factors, ignoring the whole picture, depicting the most effective factors both e-satisfaction and e-repurchase intention. The main purpose of this study is to investigate the relationships between e-service quality, information quality, e-satisfaction, and e-repurchase intention by involving customer decision-making styles in the context. An online retailer, belonging to a large brick-and-mortar Turkish company was chosen to conduct the survey. Consequently, the data collected from the conveniently selected sample

This study was produced from E. DEMIRBAS' doctoral thesis at BAU, GSSS under N. URAY's & G. SALMAN's supervision.

E. Demirbaş (✉)
İstanbul Yeni Yüzyıl Üniversitesi, Istanbul, Turkey
e-mail: esra.demirbas@yeniyuzyil.edu.tr

G. Gültekin Salman
Bahcesehir Üniversitesi, Istanbul, Turkey
e-mail: gulberk.salman@eas.bau.edu.tr

N. Uray
Kadir Has University, Istanbul, Turkey
e-mail: nimet.uray@khas.edu.tr

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_10

among the members of that e-retailer was used to test the research model using structural equation modeling. The results revealed efficiency, fulfillment, privacy, and information quality to influence both e-satisfaction and e-repurchase intention whereas after-sales e-services influencing e-satisfaction. Meanwhile, e-satisfaction mediates the relationship between the service quality of a website and e-repurchase intention. Furthermore, novelty and recreational shopping style attitude moderates the relationship between e-satisfaction and e-repurchase intention.

Keywords E-service quality · Information quality · E-customer satisfaction · E-repurchase intention · Shopping styles

1 Introduction

In Turkey, the number of e-retailers is on the rise and the country has a large young population, which means that there is a significant growth potential for e-commerce, and the lives of many people have shifted online during the Covid-19 pandemic period, further contributing to the prospects for expansion. According to the Turkish Ministry of Trade, the volume of e-commerce in Turkey grew 64% in the first half of 2020 relative to the same period of the last year. On a sector-by-sector basis, growth rates for e-commerce in 2020 were 434% for food and supermarkets, 116% for software, 95% for home and gardening products, 90% for home appliances, 58.5% for electronics, and 45% for clothing. The share of e-commerce in the first half of 2020 was 14.2% of total trade in Turkey, which is a notable increase from the same period of the previous year, when that figure stood at 8.4% (e-ticaret.gov.tr, accessed January 10, 2021), demonstrating both the importance and popularity of online shopping. As the figures above indicate, the Covid-19 pandemic has driven consumers to increasingly turn to online purchases, and as consequence companies have had to invest in e-commerce to ensure that their customers have the best shopping experiences possible. In such an intensively competitive environment, insights about how to develop and maintain customer loyalty are of paramount importance, as they may determine whether companies can survive.

As some earlier studies have shown, one of the key factors for surviving in an intensely competitive e-environment is the development of strategies that focus on services. Companies must deliver superior service experiences to their customers to ensure that they will engage in repurchase behavior and remain loyal (Gounaris et al., 2010). To achieve high levels of customer satisfaction, companies must provide high-quality services, as that often leads to favorable behavioral intentions (Brady & Robertson, 2001). Numerous researchers have examined the concept of e-service quality (e-SQ), the attributes of which are significantly associated with customer satisfaction and repurchase intention, but information quality has rarely been integrated into the extant studies in the literature. Furthermore, the interrelationship between customer satisfaction and repurchase intention may influence the shopping style attitudes of consumers, an issue that is of utmost importance in e-commerce settings. This work considers those aspects neglected in the previous studies, thus contributing to the literature by suggesting and testing a more integrative model related to e-service quality and repurchase intention relations. This study aims to provide insights by taking an integrative approach, investigate the effects of e-service quality and information quality on e-satisfaction and e-repurchase intention. Moreover, understanding how e-satisfaction mediates the relationships between service quality and information quality, as well as e-repurchase intention, has never been more important. As the last step, this study also examines the moderating effect of the shopping styles described by Kendall and Sproles (1986) vis-à-vis traditional shopping attitudes.

This chapter starts with describing a conceptual framework of the e-service quality, e-satisfaction, e-repurchase intention, information quality, and shopper types. It continues onto the research methodology where the data collection method, the sampling method, measures, the demographics, and shopping-related characteristics of the respondents were covered followed by the findings of this empirical study. The empirical part of this study was based on field research conducted with the customers of one of the largest and most aggressive brick-and-mortar and brick-and-click retail companies in Turkey. The tests of the hypothesis and the model were interrogated with structural equation modeling via Amos and SPSS v21.

The last part of the study discusses the confirmed direct effects of efficiency, fulfillment, privacy, and information quality on both e-satisfaction and e-repurchase intention. At the same time, e-recovery services were found to have a direct impact on e-satisfaction, whereas e-repurchase intention was indirectly affected. This study also concluded that e-satisfaction mediates the relationship between the dimensions of e-service quality and information quality as well as e-repurchase intention. Lastly, our research indicated that hedonic shopping attitudes, or hedonic consumer styles, moderated the relationship between e-satisfaction and e-repurchase intention for consumers making purchases from e-Company XYZ. This study contributes to the literature by emphasizing the importance of satisfaction to ensure repurchase behavior along with the direct effects of efficiency, fulfillment, privacy, and information quality to survive in an intensely competitive e-environment.

2 Conceptual Framework

2.1 E-Service Quality

Service can be defined as the efforts of an organization directed toward delivering high-quality experiences to consumers to satisfy their needs (Chang et al., 2015). Taherdoost et al. (2012) define e-services as "the provision of interactional, content-centered and electronic-based service over electronic networks" (p. 75). The importance of service quality in electronic commerce has been discussed by scholars for many years. Santos (2003) and Yang and Jun (2002) pointed to service quality as a crucial determinant of success or failure in electronic commerce. Reichheld and Schefter (2000) described e-loyalty as a "secret weapon" and noted that focusing on service quality is of critical importance in e-commerce. Grönroos et al. (2000) posited that e-SQ is comprised of two dimensions, functional and process quality. Zeithaml et al. (2002) defined e-SQ as "the extent to which a website facilitates efficient and effective shopping, purchasing, and delivery of products and services" (p. 362). According to Kim et al. (2006), poor service quality negatively affects online retailers because it leads to a lack of trust, because of which shoppers may leave websites without completing their transactions. When service quality increases, however, customer experiences are enhanced at every touchpoint (Rosenbaum & Losada, 2017; Sahin et al., 2017). Collier and Bienstock (2006) have noted that when service quality is poor, customers "are just one click away from switching to another e-retailer" (pp. 260–261) because of a concomitant decrease in repurchase intention and loyalty. Thus, it can safely be concluded that e-SQ is an important issue in today's business world.

Abelse et al. (1999) proposed six operational criteria for e-SQ, listing them as "use, content, structure, linkage, search and appearance" (p. 40). Yang and Fang (2004) proposed similar determinants for e-SQ about the notion of SERVQUAL developed by Parasuraman et al. (1985) in terms of "reliability, responsiveness, access, ease of use, attentiveness, credibility, and security" (p. 313). Accessibility is an indispensable feature of virtual stores, as they are open 24 h a day. Given that situation, such stores should have advanced technical infrastructure and employ user-friendly navigation systems. Since customers access and connect to web stores and carry out their transactions themselves, websites should operate smoothly and efficiently so that customers can easily manage their purchases. Santos (2003) explained e-SQ concerning "consumers' overall evaluation and judgment of the excellence and quality of e-service offerings in the virtual market space" and added that it can "potentially increase attractiveness, hit rate, customer repurchase intention, stickiness, positive WOM and maximize the competitive advantages of an online company" (pp. 233–236).

As indicated by the studies mentioned above, there is a large body of research on e-service quality. Wolfinbarger and Gilly (2003) pointed to the use of e-mails as a means of measuring customer perceptions of e-tailing quality. Wolfinbarger and Gilly's (2003) E-S-QUAL scale, was later modified by Blut et al. (2015) with the inclusion of four specific dimensions. Later, Blut (2016) used those same dimensions with additional items appended to the scale. "WebQual" was developed as a way to deal with the evaluation of transactions, while "SITEQUAL," developed by Yoo and Donthu (2001), concerns the matter of site quality. The work of Szymanski and Hise (2000) took up the various aspects of web pages to evaluate transaction-specific issues about sites.

All those studies examined the various features of e-SQ and offered different insights about how it can be evaluated. Building on the SERVQUAL model, Parasuraman et al. (2005) developed multi-item scales in two stages, E-S-QUAL and E-Recs-QUAL, as a means of capturing electronic service quality as a whole and measuring customer perceptions of service quality so that e-service businesses can provide superior service quality to enhance both customer satisfaction and repurchase intention. E-S-QUAL refers to the e-core service quality to be delivered to online customers (Parasuraman et al., 2005) who meet no routine cases in the websites. Parasuraman et al. (2005) broadly defined e-SQ as encompassing "all phases of customer interactions with a website: The extent to which a website facilitates efficient and effective shopping, purchasing and delivery" (p. 217). As a concept dealing with e-core service quality, E-S-QUAL has four dimensions (pp. 220–221), which are presented below.

Efficiency is defined as the ease and speed with which a website can be accessed and used (Parasuraman et al., 2005, p. 220) as well as the functionality of the service, which makes it possible for transactions to be completed in a convenient manner (Kesharwani, 2020).

Fulfillment is related to the extent to which a site's promises about order deliveries and item availability are realized (Parasuraman et al., 2005, p. 220). Also, according to Blut (2016), it refers to delivery times and the accuracy of customer orders. It should be noted, however, that fulfillment can only be assessed once payment has been made. That is especially crucial since post-payment dissonance occurs more frequently in online contexts because customers only experience the tangibility of products after delivery (Liao & Keng, 2013).

System availability refers to correct technical operations (Parasuraman et al., 2005, p. 221).

Privacy means the safety of websites and the protection of customer information (Parasuraman et al., 2005, p. 220) and credit card payments (Blut, 2016). Privacy is crucial for ensuring the credibility and quality of websites (Wang et al., 2015). Privacy and security features are indicators of the effectiveness of a website (Schmidt et al., 2008; Fortes & Rita, 2016; Fortes et al., 2017). When they purchase items, customers must provide private information such as their address, credit card number, and so forth (Holloway & Beatty, 2008), and that can lead to concerns about protection against fraud; for that reason, security and privacy are major aspects of online service quality (Rita et al., 2019).

Parasuraman et al. developed the concepts E-S-QUAL and E-Rec-QUAL as pre- and post-web service approaches to offering superior service quality to enhance both customer satisfaction and repurchase intention, those constructs were taken up as the primary start-up variables of the conceptual model of this study.

2.2 E-Satisfaction, E-Repurchase Intention, and E-Service Recovery

Several studies have highlighted the four dimensions of E-S-QUAL discussed above, including the work of Kim et al. (2006), who pointed out the importance of those aspects of online retailing and summarized them in terms of "simplicity of using the site, ease of finding information and fast check-out with minimum effort" (p. 55). They also emphasized how people may leave a webpage and seek out others if they encounter difficulties in doing searches, downloading material, and seeing items. The issue of privacy and security, particularly concerning the protection of personal and financial information, especially credit cards, represents yet another major sufficiency variable. Kim et al. (2006) noted the critical role that privacy plays in online shopping in terms of purchase intention, satisfaction, and overall site quality. Fulfillment and reliability bear "the strongest predictor of customer satisfaction and quality and the second strongest indicator of loyalty/intention to repurchase" (Zeithaml et al., 2002, p. 364). Yang and Fang (2004) described fulfillment within the scope of "accurate orders and keeping service promises" (p. 302), while Kim et al. (2006) defined it as one of the main service quality components for determining customer satisfaction or dissatisfaction. For system availability, Watcher (2002) recommended that e-retailers promptly resolve functionality problems such as missing links and broken buttons as they may lead to frustration during the process of browsing and purchasing, ultimately driving customers to leave the website without completing their transactions.

Satisfaction is critical for keeping consumers loyal and preventing them from defecting to other e-retailers. As Jiang and Rosenbloom (2005) have pointed out, "the only truly loyal customers are totally satisfied customers" (p. 152). Other studies have also investigated the relationship between e-satisfaction and e-service quality (Connolly et al., 2010; Gounaris et al., 2010; Schaupp, 2010; Xiao, 2016; Kaya et al., 2019; Zarei et al., 2019; Vo et al., 2020).

All the above studies clearly emphasize the impacts of the four dimensions of E-S-QUAL (efficiency, fulfillment, system availability, and privacy) on the satisfaction of e-consumers. When websites live up to those expectations, consumers tend to be more satisfied and are more likely to return to them for future purchases.

Therefore, the following hypotheses will be tested in this study:

H1: The quality level (including efficiency, fulfillment, and privacy dimensions) of websites' e-services positively affects consumer e-satisfaction.

- *H1a: Efficiency, as a dimension of the quality level of websites' e-services positively affects consumer e-satisfaction.*
- *H1b: Fulfillment, as a dimension of the quality level of websites' e-services positively affects consumer e-satisfaction.*
- *H1c: Privacy, as a dimension of the quality level of websites' e-services positively affects consumer e-satisfaction.*

King et al. (2016) suggest that consumers will continue to purchase from online retailers that maintain a high level of service quality. In that line of thinking, when consumers have positive experiences with online shopping, they will return to that website in the future (Collier & Bienstock, 2006). Ha et al. (2010) explain the relationship between satisfaction and repurchase intention by referring to attribution theory, and they point out that "consumer satisfaction judgments in a repurchase situation are updated spontaneously when previously formed satisfaction evaluations are available from memory and experience, with an exceeding expectation that means satisfaction facilitates customers' repurchase intention" (p. 1002).

As such, the factors that directly influence repurchase intention aside from information quality will be tested with the following hypotheses:

H2: The quality level (including efficiency, fulfillment, and privacy dimensions) of websites' e-services positively affects the e-repurchase intention of consumers.

H2a: Efficiency, as a dimension of the quality level of websites' e-services positively affects the e-repurchase intention of consumers.
H2b: Fulfillment, as a dimension of the quality level of websites' e-services positively affects the e-repurchase intention of consumers.
H2c: Privacy, as a dimension of the quality level of websites' e-services positively affects the e-repurchase intention of consumers.

Since independent variables also affect satisfaction and the results of several studies have demonstrated that there is a direct relationship between satisfaction and repurchase intention (Fullerton & Taylor, 2002; Cole & Steven, 2006; Srivastava & Sharma, 2013; Abdullah et al., 2018), the mediating role that satisfaction plays between the dimensions of e-service quality and repurchase intention (Tandon et al., 2017; Lestari & Ellyawati, 2019) will be tested with the hypothesis below:

H3: E-satisfaction mediates the relationship between the service quality of a website and e-repurchase intention.

Ease of access to customer services and affirmative, responsive attitudes are critical in e-services. Griffith and Krampf (1998) describe those issues as key indicators that have a direct relationship with the factors of trust, repurchase intention, commitment, and word-of-mouth, all of which pave the way to achieving success in e-retailing. Parasuraman et al. (2005) define e-recovery service quality as the non-routine activities of websites that also have a strong effect on e-consumer satisfaction. They can be grouped into three dimensions: *responsiveness*, meaning the effective handling of problems and hence returns to the site; *compensation*, which is the ability of a site to solve customers' problems; and *contact*, which involves providing assistance services over the telephone or via online representatives (p. 220). When problems arise, being able to contact a customer service agent by phone or online has a critical effect on online shopping (Kim et al., 2006). Collier and Bienstock (2006) emphasize that there are direct impacts when an online company undertakes service recovery efforts, as they may "create a 'terrorist' customer who disseminates negative information about that retailer or an 'apostle' customer who actively encourages others to use that retailer" (p. 265). In light of those issues, the following hypothesis will be tested:

H4: When the e-service recovery quality of a website is high, consumers' e-satisfaction while shopping online will also be high.

2.3 Information Quality

Information is a key component of all purchasing activities because it is assumed that people make rational decisions based on the information available to them. Korten (2009) defines it as “the one resource that is non-depletable and increases its real-wealth value when widely shared” (p. 135). Additionally, Thompson (2002) has described websites as “the most popular source of information” in online purchasing (p. 365). Since there are no physical salespersons to answer customers’ questions in online settings, verbal and/or visual information becomes critical (Kim et al., 2006). Information quality signifies the value of the products offered on online platforms (Yang et al., 2005), as well as the production of the website and the outputs (Al Debei, 2014). Gao (2005) pointed out that “consumers with a higher level of domain expertise will search for more information between sites because they can effectively locate the information and evaluate it in the search process” (p. 33). Accordingly, the offering of high-quality information on a website will lead customers to make good purchase decisions. In a similar vein, Rust and Lemon (2001) conceptualize e-services as an information service. Li and Suomi (2007) describe e-services as the provision of different experiences through an “interactive flow of information” (p. 176) and ensuring the reliability, relevance, accuracy, timeliness, and thoroughness (Ahn et al., 2007; Chen et al., 2011) of information, as well as its correctness, currency, and completeness (Lin, 2010) along with consistency and dependability (Yang et al., 2005).

Zeithaml et al. (2002) describe high-quality information as information that is “relevant, accurate, timely, customized and complete” (p. 364). Moreover, information quality relates to how customers perceive the information provided by e-retailers (Mun et al., 2013). Lynch and Ariely (2000) emphasize that high-quality information and the ability to search for prices and product features can raise satisfaction levels through the contributions of experience and product purchases, resulting in revisits and repurchase intention. Vo et al. (2020) have highlighted the importance of the timeliness and accuracy of the information on websites as crucial elements for building trust and satisfaction. Mai (2012) points out that “information quality becomes a product of the degree to which the exchange and production of meaning have been successful” (p. 687).

The model developed by Doll and Torkzadeh (1988) for measuring end-user satisfaction has five determinants: content, format, ease of use, accuracy, and timeliness (p. 268). Although the first three were examined in this study within the scope of the E-S-QUAL scale, information quality including the last two dimensions of accuracy and timeliness were taken up as separate variables that affect e-satisfaction and e-repurchase intention because of the key role they play in the assessment of websites. For that reason, the following two hypotheses will be tested:

H5: As the information quality of an e-store increase, there is a concomitant increase in the (a) e-satisfaction (H5a), and (b) e-repurchase intentions of consumers (H5b).

2.4 Shopper Types

Several studies on shopper types (Kau et al., 2003; Huang, 2003; Rohm & Swaminathan, 2002; Kendall & Sproles, 1986) were reviewed to identify the decision-making styles of Turkish e-consumers. The extensive summary of different types of shopping style attitudes developed by Kendall and Sproles (1986) within the rubric of a customer style index (CSI) was found to be the most suitable for the basis of this research, but we were also able to narrow down that scope via studies about consumers in other countries.

Examinations of consumers' decision-making styles have contributed much to our understanding of their moderating effects on repurchase intention. Kendall and Sproles (1986) defined the decision-making styles of consumers as "a mental orientation characterizing a consumer's approach to make choices" (p. 283), and their CSI, which was based on cognitive and affective characteristics (p. 268), is widely used to define the characteristics of consumers about different products and/or home countries. The eight basic characteristics of decision-making defined by Kendall and Sproles (1986, p. 269) are as follows:

1. Perfectionist or high-quality consciousness
2. Brand consciousness
3. Novelty fashion consciousness
4. Recreational, hedonistic shopping consciousness
5. Price and "value for money" shopping consciousness
6. Impulsiveness
7. Confusion arising from an overabundance of choices (due to a proliferation of brands, stores, and consumer information)
8. Habitual, brand-loyal orientation to consumption

Over the years, Kendall and Sproles' typology was used in different studies, and it has received particular attention in recent times as well (Chang et al., 2020; Ceylan & Alagoz, 2020; Raskovic et al., 2020; Ozturk & Sahin, 2020). This study adopts the CSI model developed by Kendall and Sproles for online purchasing to test the following hypothesis:

H6: In the course of shopping, the relationship between e-satisfaction and e-repurchase is affected in differing ways by different decision-making styles.

3 Research Method

3.1 Data Collection Method and Sampling

Our descriptive research was designed to bring to light the effects of e-service quality and information quality on e-satisfaction and e-repurchase intention. To assess those relationships, we created an online questionnaire to gather the preliminary data. The online questionnaire was based on the survey results of a pretest that was carried out with 62 people. Lastly, the online questionnaire was customized to fit the characteristics of the virtual environment of the company that was involved in the study. The company, which is an online retailer operating under the auspices and brand name of a large Turkish brick-and-mortar firm, sells products such as clothing, shoes, and accessories. For this study, the e-retailer will be referred to as Company XYZ per the terms of a mutual agreement of confidentiality. The company agreed to place the questionnaire on its website and take part in the survey on the condition that its name and data be kept confidential.

The data was obtained utilizing a judgmental sampling technique (Malhotra, 2009) from the customers of Company XYZ. When the study was carried out, the click and mortar division had 800,000 registered members, 50,000 of which (6.25%) company management described as being active because they kept up with Company XYZ's campaigns regularly, as evidenced by their interest in the company's messages and the frequency with which they visited its website. The company referred to them as "responsive customers." Company management sent those 50,000 active members a message stating that they would receive a gift card worth TRY ten for each completely answered questionnaire. In total, 1334 members took part in the survey, and 1075 completely and correctly answered questionnaires were analyzed, indicating a 2.7% response rate and a data validity rate of 81%. While Company XYZ surveyed over 1 week, most responses were received in the first 3 days. Consequently, enough responses were received, and the survey was terminated.

3.2 Measures

During the study, the e-SERVQUAL model developed by Parasuraman et al. (2005), was utilized including its e-core and e-recovery components, as a means of observing e-service quality. As noted earlier, E-S-QUAL has four dimensions—efficiency, fulfillment, privacy, and system availability—including 22 items that are measured with "Likert-type 5-point scales ranging from 1 (strongly disagree) to 5 (strongly agree). The Cronbach's alpha values were 0.94 for efficiency; 0.83 for system availability; 0.89 for fulfillment; and 0.83 for privacy, with the CFA ranged from 0.67 to 0.83" (Parasuraman et al., 2005, pp. 220–221). The e-service recovery component of the E-S-QUAL model includes three dimensions, "responsiveness, compensation, and contact with 11 items. The Cronbach's alpha values were 0.88 for responsiveness; 0.77 for compensation; 0.81 for contact, with the CFA ranged from 0.68 to 0.73" (Parasuraman et al., 2005, p. 220). In the study by Parasuraman et al. (2005), satisfaction was adopted from the perceived value on a ten-point semantic differential scale and intent to repurchase was based on a loyalty intention Likert-type five-point scale.

Table 1 The variables used in this study and their sources

Construct	Reference in the literature
E-S-QUAL	Parasuraman et al. (2005)
E-RecS-QUAL	Parasuraman et al. (2005)
Information Quality	Doll and Torkzadeh (1988)
E-satisfaction	Parasuraman et al. (2005)
E-repurchase Intention	Parasuraman et al. (2005)
Shopper Types	Kendall and Sproles (1986)
Demographic Variables	Parasuraman et al. (2005); Kim et al. (2006)

A study by Doll and Torkzadeh (1988) about end-user computing satisfaction utilized five determinants for information quality: content, timeliness, format, accuracy, and ease of use. Since some of the items in the e-core component of the e-SERVQUAL questions are part of the abovementioned three determinants of information quality, Doll and Torkzadeh's accuracy and timeliness determinants i.e., those not interrogated via the e-SERVQUAL scale—were observed within the scope of information quality to prevent repetition. The Cronbach's alpha values were 0.91 for accuracy and 0.82 for timeliness (Doll & Torkzadeh, 1988).

The eight dimensions of Sproles and Kendall's (1986) traditional study on consumer decision-making styles were adopted as moderating variables in this research. The Cronbach's alpha values of those dimensions were 0.76 for recreational (hedonic) shopping, 0.75 for brand-conscious shopping, 0.74 for both novelty fashion and perfectionist shopping, 0.55 for overly confused shopping, 0.53 for habitual and brand-loyal shopping, and 0.48 for price-value conscious and impulsive shopping. Although the last two were under the 0.5 threshold, they were included to clarify the attitudes of the sample.

As an interval scale, the five-point Likert scale is mainly used for the constructs of research models, while ten-point semantic differential scales are used to measure the satisfaction levels of respondents. Additionally, nominal and ratio scales are also used to measure visit and shopping frequencies, average spending, and the sociodemographic characteristics of samples.

The variables that were tested in the model and their literature sources are shown in Table 1.

3.3 Demographic and Shopping-Related Characteristics of the Sample

Women made up 70.5% of the total sample. Almost half of the respondents (44.5%) were between 32 and 38 years old, 26% were between 25 and 31 years old, 19% were between 39 and 45 years old, 5% were between 25 and 31 years old, and the rest were more than 45 years old. About education, 19% of the participants had completed high school, 67% of them had attended university, and 11% had done graduate studies. As for professions, 19% of the participants held administrative posts, 52% were salaried employees, 9% were business owners, 9% were housewives, and the remainder did not work.

In terms of technological background, most of the respondents (87%) had intermediate-advanced and advanced levels of computer literacy, and a major segment of the respondents (90%) were competent internet users, having used the internet for more than 7 years (92.2%). More than half of the participants (57.6%) engaged in e-shopping several times a month and 24% did so several times a week.

4 Findings

4.1 Exploratory Factor Analysis (Independent Variables and Moderating Variables)

The questions were translated and back translated to ensure validity issue. Thus, exploratory factor analysis was conducted. We used IBM-SPSS v21 to conduct the factor analysis for each of the independent variables (E-Rec-QUAL, E-S-QUAL, and information quality) and the moderator variable (shopping styles separately to reduce the number of variables included. As Table 2 indicates, both E-Rec-QUAL and information quality were found to be unidimensional variables, and the scale of E-S-QUAL was summarized in terms of three factors: efficiency, fulfillment, and privacy. Lastly, the variable of shopping style attitudes was divided into six dimensions, each of which described a different type of shopping: price-conscious, habitual, overwhelmed by choices, non-perfectionist, novelty, and recreational, which is a combination of the two separate components of novelty and brand-conscious shopping. Subsequently, EFA was applied to all the independent and moderating variables mentioned above. The summary findings of the EFA are shown in Table 2. A good KMO of 0.958 indicated the suitability of inter-dimension correlation in an adequate sample volume for conducting factor analysis with the significance of Bartlett's test of sphericity at 0.00 (Durmuş et al., 2010) and total variance explained at 74%. All the dimensions were found to be reliable with Cronbach Alpha values exceeding 0.70.

4.2 Exploratory Factor Analysis (E-Satisfaction and E-Repurchase Intention)

Factor analysis was applied to the two endogenous variables of the model: e-satisfaction and e-repurchase intention. Tables 3 and 4 provide an overview of the analysis for e-satisfaction and e-repurchase. A KMO value of 0.5 in Table 3

Table 2 EFA results

Dimension	Codes	Statement	Factor loading (%)	Factor variance explained (%)	Cronbach's alpha
E-Rec-QUAL					
E-service recovery	ERQ5	It takes care of problems promptly	0.814	18.68	0.959
	ERQ6	This site compensates me for the problems it creates	0.801		
	ERQ4	It tells me what to do if my transaction is not processed	0.793		
	ERQ8	This site provides a telephone number to reach the company	0.763		
	ERQ9	This site has customer service representatives available online	0.763		
	ERQ3	This site offers a meaningful guarantee	0.76		
	ERQ10	It offers the ability to speak to a live person for information if there is a problem.	0.756		
	ERQ7	It compensates me when what I ordered doesn't arrive on time	0.75		
	ERQ2	This site handles product returns well.	0.74		
	ERQ1	It provides me with convenient options for returning items	0.694		
E-S-QUAL					
Efficiency	ESQ6	This site is simple to use	0.78	13.897	0.941
	ESQ7	This site enables me to get on to it quickly	0.753		
	ESQ2	It makes it easy to get anywhere on the site	0.752		
	ESQ5	It loads its pages fast	0.743		
	ESQ4	Information at this site is well organized	0.736		
	ESQ3	It enables me to complete a transaction quickly	0.728		
	ESQ8	This site is well organized	0.718		
	ESQ1	This site makes it easy to find what I need	0.663		
	ESQ17	This site launches and runs right away	0.532		

(continued)

Table 2 (continued)

Dimension	Codes	Statement	Factor loading (%)	Factor variance explained (%)	Cronbach's alpha
Fulfillment	ESQ10	It quickly delivers what I order	0.769	6.457	0.936
	ESQ9	It delivers orders when promised	0.768		
	ESQ13	It makes accurate promises about the delivery of products	0.662		
	ESQ11	It has in stock the items the company claims to have	0.645		
	ESQ12	It is truthful about its offerings	0.611		
Privacy	ESQ15	It does not share my personal information with other sites	0.739	4.252	0.940
	ESQ16	This site protects information about my credit card	0.719		
	ESQ14	It protects information about my web shopping behavior	0.662		
Information quality					
	INFQ2	Do you think the output is presented in a useful format?	0.711	6.254	0.962
	INFQ3	Does the system provide up-to-date information?	0.703		
	INFQ4	Do you feel the output is dependable?	0.69		
	INFQ5	Does the system provide sufficient information?	0.677		
	INFQ1	Are you satisfied with the accuracy of the system (site)?	0.674		
Shopping style attitude					
Confused by over choice	SHPSTY23	Sometimes its hard to choose which stores to shop at	0.847	6.593	0.843
	SHPSTY24	The more I learn about products, the harder it seems to choose	0.81		
	SHPSTY22	There are so many brands to choose from that often I feel confused	0.79		
	SHPSTY19	Often, I make careless	0.674		
	SHPSTY25	I change brands I order regularly	0.663		

(continued)

Table 2 (continued)

Dimension	Codes	Statement	Factor loading (%)	Factor variance explained (%)	Cronbach's alpha
Novelty and recreational	SHPSTY8	I usually have one or more outfits of the very newest style	0.753	5.144	0.774
	SHPSTY10	I keep my wardrobe up to date with the changing fashions	0.678		
	SHPSTY9	To get variety, I shop at different stores and choose different brands	0.65		
	SHPSTY13	Going shopping is one of the enjoyable activities of my life	0.648		
	SHPSTY12	I enjoy shopping just the fun of it	0.611		
Price consciousness	SHPSTY18	I look carefully to find the best value for the money	0.763	4.612	0.758
	SHPSTY16	I buy as much as possible at sale prices	0.755		
	SHPSTY17	The lower price products are usually my choice	0.745		
Habitual	SHPSTY27	Once I find a product or brand I like, I stick with it	0.828	4.46	0.759
	SHPSTY26	I have favorite brands I buy over and over	0.782		
	SHPSTY28	I go to the same e-stores, each time I e-shop	0.682		
Non-perfectionist	SHPSTY2	I don't give my purchases much thought and care	0.848	3.655	0.832
	SHPSTY3	I shop quickly, buying the first product or brand I find that seems good enough	0.828		
	Total variance explained			74.004	
	KMO-Kaiser-Meyer-Olkin measure of sampling adequacy			0.958	
	Significance			0.00	

Table 3 Factor analysis results of e-satisfaction

Dimension	Code	Statement	Factor loading	Factor variance explained (%)	Cronbach's alpha
Satisfaction	STF1	The overall convenience of using this site.	0.96	92.234	0.914
	STF2	How much does this site fulfill your expectations?	0.96		
	KMO (Kaiser-Meyer-Olkin) measure of sampling adequacy			0.5	
	Significance			0	

Table 4 Factor analysis results of e-repurchase intention

Dimension	Code	Statement	Factor loading	Factor variance explained (%)	Cronbach's alpha
E-repurchase intention	RPCH2	Would you recommend this site to someone seeking your advice?	0.933	81.549	0.941
	RPCH1	Would you say positive things about this site to other people?	0.931		
	RPCH3	Would you encourage friends and others to do business on this site?	0.925		
	RPCH4	Would this site be your first choice for future transactions?	0.881		
	RPCH5	Would you do more shopping on this site in the upcoming months?	0.842		
	KMO (Kaiser-Meyer-Olkin) measure of sampling adequacy			0.852	
	Significance			0	

indicated the weak suitability with the significance of Bartlett's test of sphericity at 0.00 and the factors' variance was at 92%. Table 3 also shows that the reliability of the construct of Cronbach Alpha is 0.914.

Findings of exploratory factor analysis for e-repurchase intention are presented in Table 4. KMO of 0.852 with the significance of Bartlett's test of sphericity at 0.00 indicates a good fit with the factor's variance explained at 81.5%. All the dimensions were found to be reliable with Cronbach's alpha values exceeding 0.70.

Both factor analyses resulted in statistically good values for the KMO, Bartlett's test of sphericity, and reliability. Both variables appeared as one component and all the original items, which had been two for e-satisfaction and five for e-repurchase intention, were retained under the same factor.

4.3 Test of the Research Model

Three confirmatory factor analyses (CFAs) were conducted to test the relationships between the variables of the model as is presented in Table 5, including the fit indices values.

First, as a means of measuring the impact of e-service quality on e-customers' satisfaction, CFA was carried out with five latent variables including the four dimensions (efficiency, privacy, fulfillment, and e-recovery service) of the quality level of websites, and information quality, and e-satisfaction as an endogenous variable. The correlation coefficients of the latent variables and the AVE and CR values are presented in Table 6.

As a result of the path analysis, all the paths between efficiency, fulfillment, and e-service recovery, as the dimensions of e-service quality and information quality, were found to have positive relationships with e-satisfaction at a 5% significance level. Privacy with a critical ratio -1.460 being under the threshold t value of 1.96, did not affect e-satisfaction at a 5% significance level. Information quality, which had a regression coefficient of 0.213, was the most influential variable for e-satisfaction, followed by fulfillment, e-recovery, and efficiency. Additionally, the critical ratio values of efficiency, fulfillment, e-recovery services and information quality were above the threshold t value of 1.96 at a 5% significance level. H1 except privacy and H4 were thus supported, as indicated by the figures in Table 7, and in Table 14.

It can thus be said that an efficient and fulfilling website that operates with high e-service recovery levels and includes high-quality information increases consumers' e-satisfaction.

Another CFA was conducted with the five latent variables of efficiency, privacy, fulfillment, e-services recovery, and information quality; and e-repurchase intention as an endogenous variable. The analysis revealed that the effect of e-recovery on e-repurchase intention had a significance level of 0.25%, which is beyond the threshold level of 5%, but it was retained in the model for the mediating analysis (Baron & Kenny, 1986). The correlation coefficient of the latent variables and AVE and CR values are shown in Table 8.

Four exogenous variables, with efficiency, fulfillment, and privacy being dimensions of e-service quality and information quality, were found to have positive relationships with e-repurchase intention at a 5% significance level with critical ratio values higher than 1.96. E-service recovery had the least influence on the e-repurchase intention at a 30% significance level. Information quality was found to be the most influential variable in e-repurchase intention with a regression coefficient of 0.438, followed by fulfillment, efficiency, and privacy. Additionally, the critical ratio values of efficiency, fulfillment, privacy, and information quality were above the threshold t value 1.96 at a 5% significance level, as is displayed in Table 9.

H5 was supported based on the results presented in Tables 7 and 9, and H2 was supported as well. Therefore, it can be concluded that an efficient, fulfilling,

Table 5 CFA results of the model

Dimensions	Codes of questions	Standardized regression coefficients
E-recovery	ERQ6	0.862
	ERQ5	0.891
	ERQ4	0.866
	ERQ10	0.766
	ERQ9	0.771
	ERQ8	0.805
	ERQ7	0.823
	ERQ2	0.849
	ERQ3	0.825
	ERQ1	0.799
Efficiency	ESQ2	0.804
	ESQ7	0.870
	ESQ5	0.744
	ESQ4	0.782
	ESQ3	0.823
	ESQ8	0.838
	ESQ1	0.743
	ESQ17	0.706
	ESQ6	0.887
Information quality	INFQ1	0.881
	INFQ5	0.911
	INFQ4	0.947
	INFQ3	0.922
	INFQ2	0.910
Fulfillment	ESQ11	0.804
	ESQ12	0.922
	ESQ13	0.924
	ESQ10	0.796
	ESQ9	0.807
Privacy	ESQ16	0.921
	ESQ15	0.934
	ESQ14	0.897
E-satisfaction	STF1	0.915
	STF2	0.923
E-repurchase intention	RPCH1	0.960
	RPCH2	0.966
	RPCH3	0.899
	RPCH4	0.767
	RPCH5	0.710
Model fit indices	*Actual values*	
CMIN/DF	3.316	
GFI	0.902	
CFI	0.967	
NFI	0.953	
RMSEA	0.046	

Table 6 Correlation coefficients and AVE and CR values: e-satisfaction is endogenous

	E-recovery	Fulfillment	Efficiency	Information quality	Privacy	E-satisfaction
E-recovery	(0.823)∗					
Fulfillment	0.786	(0.853)∗				
Efficiency	0.750	0.720	(0.802)∗			
Information quality	0.755	0.674	0.762	(0.914)∗		
Privacy	0.739	0.661	0.638	0.631	(0.919)∗	
AVE**	0.68	0.73	0.64	0.84	0.84	0.84
CR***	0.96	0.93	0.94	0.96	0.94	0.92
Cronbach's Alpha	0.959	0.936	0.941	0.962	0.940	0.913

*The root square of AVE; a correlation coefficient below the root square of the AVE verifies discriminant validity
*The threshold is ≥0.5
**The threshold is ≥0.7, $p \leq 0.05$

Table 7 Regression coefficients: e-satisfaction is endogenous

Path to e-satisfaction	Regression coefficients	Critical ratio	Significance level
E-service recovery	0.142	2.436	0.015
Efficiency	0115	2.295	0.022
Information quality	0.213	4.112	***
Fulfillment	0.190	4.055	***
Privacy	−0.058	−1.460	0.144

Table 8 Correlation coefficients and AVE and CR values: e-repurchase intention is endogenous

	E-recovery	Fulfillment	Efficiency	Information quality	Privacy	E-repurchase intention
E-recovery	(0.826)∗					
Fulfillment	0.786	(0.853)∗				
Efficiency	0.750	0.720	(0.803)∗			
Information quality	0.755	0.674	0.762	(0.914)∗		
Privacy	0.739	0.661	0.638	0.631	(0.917)∗	
AVE**	0.68	0.73	0.64	0.84	0.84	0.75
CR***	0.96	0.93	0.94	0.96	0.94	0.94
Cronbach's alpha	0.959	0.936	0.941	0.962	0.940	0.941

*The root square of AVE; a correlation coefficient below the root square of the AVE verifies discriminant validity
*The threshold is ≥0.5
**The threshold is ≥0.7, $p \leq 0.05$

Table 9 Regression coefficients: e-repurchase intention is endogenous

Path to e-repurchase intention	Regression coefficients	Critical ratio	Significance level
Efficiency	0.123	3.016	0.003
Information quality	0.438	10.289	***
Fulfillment	0.196	5.131	***
Privacy	0.075	2.325	0.020
E-recovery	0.053	1.119	0.263

and secure website that offers high-quality information increases the e-repurchase intention of consumers, supporting H2, as is presented in Table 14.

The results of earlier CFA analyses made it possible (Baron & Kenny, 1986; Civelek, 2018) to observe how e-satisfaction influences the relationships between e-service quality and e-repurchase intention. Hence, a third CFA was carried out with e-satisfaction as a mediator.

The third CFA indicated that the impact of e-recovery on e-repurchase intention completely disappeared because of the regression coefficient between e-recovery and e-repurchase intention, the significance level of which was 46%. At the same time, the relationship between privacy and e-satisfaction was maintained, although the significance level (15) was beyond 5% because of the statistically significant relationship between privacy and e-repurchase intention. The correlation coefficients of the latent variables and AVE and CR values are presented in Table 10.

As Table 11 indicates, the regression coefficients of efficiency, information quality, and fulfillment diminished whereas the weight of privacy increased, and the coefficient of e-recovery became statistically insignificant in the case of the mediating effect of satisfaction. Figure 1 shows the relationships between the variables of the tested research model.

While the above results demonstrate that e-recovery does not directly affect e-repurchase intention, customers who are e-satisfied because of e-recovery services would tend to display e-repurchase intention; in other words, e-recovery indirectly affects e-repurchase intention. Nevertheless, there is a negative direct relationship between privacy and e-satisfaction at a significance level of 15%. Although the other latent variables have a direct impact on e-repurchase intention, they also mediate satisfaction in influencing e-repurchase intention.

As it is presented in Table 12; efficiency, fulfillment, privacy, and information quality were found to maintain a positive relationship with e-repurchase intention, whereas e-recovery relations deteriorated in cases of satisfaction mediation at a 5% significance level with critical ratio values higher than 1.96. E-satisfaction was also found to have a positive relationship with e-repurchase intention with a statistically valid critical ratio. Therefore, H3, which claims that satisfaction mediates efficiency, fulfillment, privacy, and information quality relations with e-repurchase intention, is supported at a 5% significance level, and it partially affects their relationships.

Table 10 Correlation coefficients and AVE and CR values: e-satisfaction and mediators

	E-recovery	Fulfillment	Efficiency	Information quality	Privacy	E-satisfaction	E-repurchase intention
E-recovery	(0.827)*						
Fulfillment	0.786	(0.853)*					
Efficiency	0.752	0.721	(0.802)*				
Information quality	0.756	0.674	0.763	(0.914)*			
Privacy	0.739	0.661	0.638	0.631	(0.92)*		
AVE**	0.68	0.73	0.64	0.84	0.84	0.84	0.75
CR***	0.96	0.93	0.94	0.96	0.94	0.92	0.94
Cronbach's alpha val.	0.959	0.936	0.941	0.962	0.940	0.913	0.941

*The root square of AVE; a correlation coefficient below the root square of the AVE verifies discriminant validity
*The threshold is ≥ 0.5
**The threshold is ≥ 0.7, $p < 0.95$

Table 11 Composite results of the research model

Path	Standardized regression coefficients		
E-recovery . . . E-satisfaction	0.154		0.152
Efficiency . . . E-satisfaction	0.211		0.127
Information Quality . . . E-satisfaction	0.205		0.202
Fulfillment . . . E-satisfaction	0.211		0.210
Privacy . . . E-satisfaction	−0.065*		−0.063*
E-recovery . . . E-repurchase Intention.			0.029**
Efficiency . . . E-repurchase Intention.		0.055	0.100
Information Quality . . . E-repurchase Intention.		0.121	0.366
Fulfillment . . . E-repurchase Intention.		0.399	0.171
Privacy . . . E-repurchase Intention.		0.205	0.091
E-satisfaction . . . E-repurchase Intention		0.080	0.164
Model fit indices			
CMIN/DF	3.824	3.543	3.316
GFI	0.904	0.901	0.902
CFI	0.964	0.966	0.967
NFI	0.952	0.953	0.953
RMSEA	0.051	0.049	0.046

$p < 005$. *Significance level 0.15%; **Insignificant

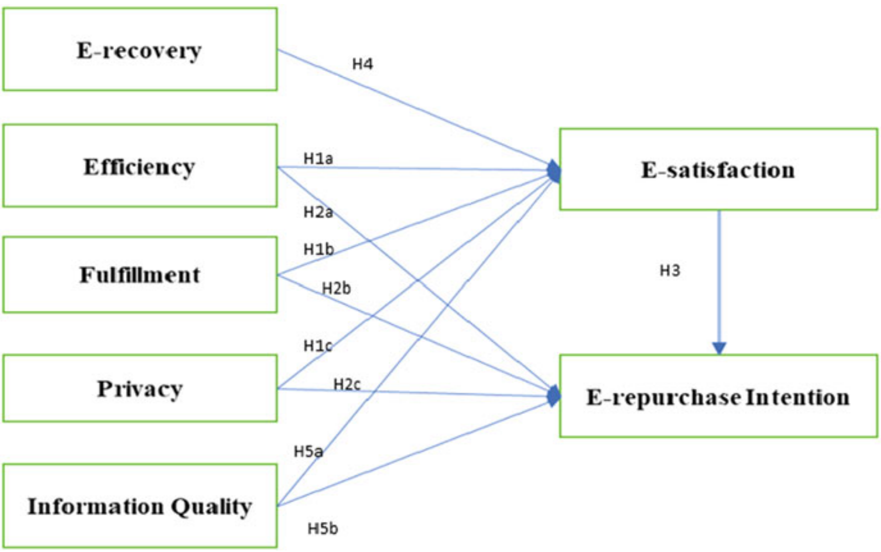


Fig. 1 The mediating effect of e-satisfaction between the variables of e-service quality, information quality, and e-repurchase intention

Table 12 Regression coefficients

	Regression coefficients	Critical ratio	Significance level
Path to e-satisfaction			
E-recovery	0.141	2.396	0.017
Efficiency	0123	2.402	0.016
Information quality	0.211	4.034	***
Fulfillment	0.191	4.025	***
Privacy	−0.056	−1.414	0.157
Path to e-repurchase intention			
Efficiency	0.101	2.501	0.012
Information quality	0.402	9.537	***
Fulfillment	0.164	4.325	***
Privacy	0.085	2.692	0.007
E-recovery	0.028	0.606	0.545
E-satisfaction	0.172	6.181	***

4.4 The Moderating Effects of Shopping Attitudes on E-Satisfaction and E-Repurchase Intention

As the last step of this empirical study, the moderating effects of shopping attitudes on e-satisfaction and e-repurchase intention were tested separately. Before testing the effects of the moderator, the standardized values of all the variables, which are referred to as Z variables in the models, and the interaction of the predictor(s) and moderator were calculated with the program SPSS v21, and the Z and interaction values were tested with AMOS v21.

Table 13 demonstrates that; the significance level of the interaction between e-satisfaction and shopping attitudes, which were defined as being price-conscious, habitual, non-perfectionist, and confused by an overabundance of choices, was beyond the significance level of 0.10 whereas the interaction of satisfaction and novelty and recreational shopping attitude was under the significance level of 0.10. Thus, H6, which asserts that the relationship between e-satisfaction and e-repurchase will be affected in different ways by different decision-making styles of shopping, is supported.

Figure 2 is the finalized model of this study which shows all relationships between the variables of the tested model at a 5% significance level. As it is seen in the path analysis; efficiency, fulfillment, and information quality directly influence both e-satisfaction and e-repurchase intention whereas e-service recovery does not directly impact e-repurchase intention, however, it has an indirect influence on e-repurchase intention via e-satisfaction which can be interpreted that customers who are e-satisfied from e-recovery services of an e-store, would tend to exhibit e-repurchase intention. Meanwhile, privacy does not directly influence e-satisfaction, but it does e-repurchase intention, in other words people keep repurchasing intention through secure websites. Moreover, satisfaction has also mediation role in the

Table 13 Regression coefficients

Path to Z repurchase intention	Regression coefficients	Significance level
Z satisfaction	0.479	***
Z price consciousness	0.305	***
Interaction of e-satisfaction and price consciousness	−0.006	0.807
Z satisfaction	0.480	***
Z habitual attitude	0.328	***
Interaction of e-satisfaction and habitual attitude	0.003	0.887
Z satisfaction	0.513	***
Z non-perfectionist attitude	0.135	***
Interaction of e-satisfaction and non-perfectionist attitude	−0.005	0.843
Z satisfaction	0.528	***
Z confused by overabundance of choices attitude	0.072	0.006
Interaction of e-satisfaction and confused by overabundance of choices attitude	−0.034	0.195
Z satisfaction	0.494	***
Z novelty and recreational	0.223	***
Interaction of e-satisfaction and novelty and recreational attitude	−0.039	0.087

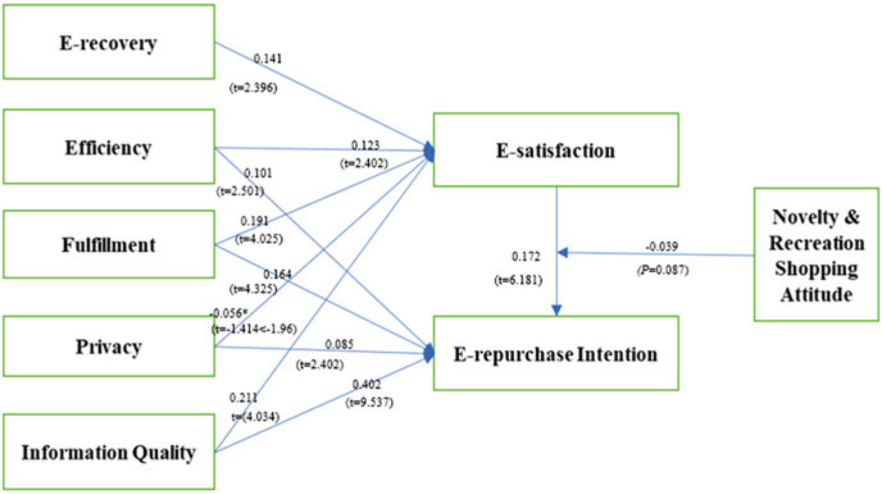


Fig. 2 The moderating effects of shopping attitudes shaped by novelty and recreation

Table 14 Hypothesis test results

Hypothesis	Supported	Not supported
H1a	✓	
H1b	✓	
H1c		✓
H2a	✓	
H2b	✓	
H2c	✓	
H3	✓	
H4	✓	
H5a	✓	
H5b	✓	
H6(partially)	Novelty and recreational characteristics of the shopping attitude	Impulsiveness, perfectionist, brand consciousness, value for money consciousness characteristics of the shopping attitude

relationships between efficiency, fulfillment, privacy, information quality, and e-repurchase intention, as it is explained in Tables 11 and 12. Finally, novelty and recreational shopping attitudes slightly moderate the effect of e-satisfaction on the e-repurchase intention at a 9% significance level, as displayed in Table 13. It can therefore be concluded that the relationship between e-satisfaction and e-repurchase will be slightly weaker for consumers who have a hedonic attitude marked by (a) novelty and (b) recreational aims when they engage in online shopping. Table 14 includes the summary results of the hypothesis test analysis.

5 Discussion

In an examination of the case of a large Turkish e-retailer, this study adopted the multi-item e-SERVQUAL scale developed by Parasuraman et al. (2005), which attempts to capture electronic service quality as a whole within the parameters of E-S-QUAL and E-RecS-QUAL while also measuring e-customers' service quality perceptions, with the ultimate goal of providing e-services companies with insights that they can use to offer superior service quality and thereby enhance both customers' e-satisfaction and e-repurchase intention. Moreover, we adopted the determinants of accuracy and timeliness developed by Doll and Torkzadeh (1988) so that we could better investigate information quality. In that process, we examined the mediating role of customer e-satisfaction and the moderating role of traditional shopping attitudes in a format adopted from Kendal and Sproles (1986). Firstly, the data we obtained provided support for an e-service quality scale with three dimensions efficiency, fulfillment, and privacy—as well as E-Rec-QUAL as unidimensional including all the items of the scale. The e-service quality and e-recovery service dimensions we utilized were developed by Parasuraman et al. (2005). In terms of system availability, the statement "this website uploaded and opened quickly" replaced efficiency, while the remaining two dimensions did not appear statistically. As argued by Blut et al. (2015), the dimensions of e-service quality vary from culture to culture in association with overall service quality and they are also dependent upon the environmental and technical context of the country in question. The findings of this study indicate that Turkish customers perceive e-service quality based on four factors—efficiency, fulfillment, privacy/security, and customer service—instead of the five dimensions proposed by Parasuraman et al. (2005).

Secondly, information quality was also found to be a key feature for the e-satisfaction and e-repurchase intention of consumers. The data supported all the items of accuracy and timeliness proposed by Doll and Torkzadeh (1988) as a unidimensional. Furthermore, that factor was influential on both e-satisfaction and e-repurchase intention. Regarding regression coefficients, we found that the direct effect of information quality on e-repurchase intention, which had a regression coefficient of 0.399, was higher than its effect on e-satisfaction, which had a regression coefficient of 0.205. Hence, it can be concluded that for the case of this study such an approach is suitable for considering information quality as a separate dimension.

Thirdly, the effect of e-service recovery on e-satisfaction was found to be high and direct, whereas that effect proved to be indirect for e-repurchase intention. E-service recovery thus directly influences customer e-satisfaction and has an indirect impact on e-repurchase intention, meaning that customers who were dissatisfied with after-sales e-services would not make use of that website again. This result coincides with the conclusions of Jiang and Rosenbloom (2005), who asserted that "the only truly loyal customers are totally satisfied customers" (p. 152), suggesting that when customer service failures occur, it is unlikely that consumers will engage in repurchases from that site.

Fourthly, only the dimension of a shopping attitude based on novelty and recreation appeared as a uni-dimension among the six traditional dimensions of CIS, and it moderates the effect of e-satisfaction on e-repurchase intention. The significant negative interaction term indicated that when consumers displayed more of a novelty and recreation attitude, e-satisfaction had less of an impact on the e-repurchase intention for consumers dealing with Company XYZ. That represents a critical finding for the company in terms of keeping its customer portfolio loyal. The claim put forward by Jiang and Rosenbloom (2005) about truly loyal customers being totally satisfied customers (p. 152) did not appear to hold for Company XYZ's customers whose shopping attitudes were based on novelty and recreation.

Lastly, our data analysis revealed that the dimensions "price-conscious, habitual, non-perfectionist, and confused by an overabundance of choices" of Sproles and Kendall's (1986) consumer-style index (CSI) did not moderate the effect of e-satisfaction on e-repurchase intention.

6 Conclusion

Based on Parasuraman, Zeithaml, and Malhotra's e-SERVQUAL model integrated with a model of information quality that includes Doll and Torkzadeh's (1988) dimensions of accuracy and timeliness, this study examined and reported on the significant direct and indirect effects of e-service quality and information quality on the e-repurchase intention of consumers through the mediating role of e-satisfaction.

The findings indicated that service quality in Turkey's e-retail industry has a significant positive association with the dimensions of information quality. On the other hand, our mediating analysis demonstrated that other influential variables have an impact on e-repurchase intention besides e-satisfaction, such as the dimensions of information quality and e-service quality. The moderating role of the different shopping style attitudes which are based on the CSI and proposed by Kendall and Sproles was supported by the data analysis. Through the use of data obtained utilizing a questionnaire and structural equation modeling, we found that among the eight shopping style attitudes only the novelty and recreation attitude were found significantly affected the impacts of e-satisfaction on e-repurchase intention, while price-conscious, habitual, non-perfectionist, confused by an overabundance of choices, and novelty and recreation behavior did not significantly moderate those effects.

Furthermore, the relationship between e-satisfaction and the e-repurchase intention was not found to be negatively moderated by novelty and recreation attitudes, meaning that as the degree of novelty and recreation increased, the positive impact of e-satisfaction on e-repurchase intention decreased. The findings of this study shed light on those issues, helping us better understand the dynamics that have contributed to customer e-loyalty and their integrated effects in today's digitalized world and hence also in the e-commerce sector, including the e-retail industry, especially since the Covid-19 pandemic broke out in March of 2020. In that way, our findings have the potential to help guide marketing managers in the e-retail industry to implement effective strategies for maintaining long-lasting relationships with their customers.

It should be noted, however, that this research has some limitations that could be addressed in future studies. First, the study was only conducted with people associated with a particular e-retailer, Company XYZ, which makes the sample demographically homogeneous, thereby limiting the generalizability of the results. Future studies could address that limitation by using a broader sample of online shoppers. Furthermore, replication of the research model in other industries and an examination of other websites that differ from Company XYZ may provide additional insights. Future research could also explore the effects of other variables such as system availability and traditional shopping attitudes, and information quality could be taken up as a new dimension of e-core service quality and be evaluated on a sectoral, regional, and/or culture-specific basis.

References

Abdullah, D., Hamir, N., Nor, N. M., Krishnaswamy, J., & Rostum, A. M. M. (2018). Food quality, service quality, price fairness, and restaurant re-patronage intention: The mediating role of customer satisfaction. *International Journal of Academic Research in Business and Social Sciences, 8*(17), 211–226.

Abels, E. G., White, M. D., & Hahn, K. (1999). A user-based design process for websites. *OCLC Systems and Services, 15*(1), 35–44.

Ahn, T., Ryu, S., & Han, I. (2007). The impact of web quality and playfulness on user acceptance of online retailing. *Information & Management, 44*, 263–275.

Al-Debei, M. (2014). The quality and acceptance of websites: An empirical investigation in the context of higher education. *International Journal of Business Information Systems, 15*(2), 170–188.

Barnes, S. J., & Vidgen, R. T. (2002). An integrative approach to the assessment of e-commerce quality. *Journal of Electronic Commerce Research, 3*(3), 114–127.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182.

Blut, M. (2016). E-service quality: Development of a hierarchical model. *Journal of Retailing, 92*(4), 500–517. https://doi.org/10.1016/j.jretai.2016.09.002

Blut, M., Chowdhry, N., Mittal, V., & Brock, C. (2015). E-service quality: A meta-analytic review. *Journal of Retailing, 91*(4), 679–700. https://doi.org/10.1016/j.jretai.2015.05.004

Brady, M. K., & Robertson, C. J. (2001). Searching for a consensus on the antecedent role of service quality and satisfaction: An exploratory cross-national study. *Journal of Business Research, 51*(1), 53–60. https://doi.org/10.1016/so148.2963(99).00041.7

Ceylan, E., & Alagöz, S. B. (2020). A study to determine the effect of consumers' decision-making styles on the organic food purchasing decision. *Karamanoğlu Mehmetbey University, Journal of Social and Economic Research, 22*(38), 148–163.

Chang, H. J. J. J., Dokko, J., Min, J., & Rakib, M. A. N. (2020). A typology of online shopping consumers and its relation to online shopping perception and obsession. In *International Textile and Apparel Association Annual Conference Proceedings*, 77(1). Iowa State University Digital Press.

Chang, M. Y., Pang, C., Tarn, J. M., Liu, T. S., & Yen, D. C. (2015). Exploring user acceptance of an e-hospital service: An empirical study in Taiwan. *Computer Standards and Interfaces, 38*, 35–43.

Chen, J., Xu, H., & Whinston, A. B. (2011). Moderated online communities and quality of user-generated content. *Journal of Management Information Systems, 28*(2), 237–268.

Civelek, M. E. (2018). *Methodology of structural equation modeling*. Beta Basım. ISBN:978-605-242-090-4.

Cole, S. T., & Illum, S. F. (2006). Examining the mediating role of festival visitors' satisfaction in the relationship between service quality and behavioral intentions. *Journal of Vacation Marketing, 12*(2), 160–173.

Collier, J. E., & Bienstock, C. C. (2006). Measuring service quality in e-retailing. *Journal of Service Research, 8*(3), 260–275.

Connolly, R., Bannister, F., & Kearney, A. (2010). Government website service quality: A study of the Irish revenue online service. *European Journal of Information Systems, 19*(6), 649–667. https://doi.org/10.1057/ejis.2010.45

Demirbaş, E., Salman, G. G., & Uray, N. (2014). An inregrative model on the factors affecting consumers' satisfaction, trust and repurchase intention in online shopping. *Doctoral Thesis*. Bahcesehir University Graduate School of Social Sciences.

Doll, W. J., & Torkzadeh, G. (1988). The measurement of end-user computing satisfaction. *MIS Quarterly, 12*(2), 259–274. https://doi.org/10.2307/248851

Durmuş (Sipahi), B., Yurtkoru, E. S., & Çinko, M. (2010). *Data analysis with SPSS in social sciences* (3rd ed.). Beta. ISBN: 978-975-295-196-8.

Fortes, N., Rita, P., & Pagani, M. (2017). The effects of privacy concerns, perceived risk, and trust on online purchasing behavior. *International Journal of Internet Marketing and Advertising, 11*(4), 307–329. https://doi.org/10.1504/IJIMA.2017.10007887

Fortes, N., & Rita, P. (2016). Privacy concerns and online purchasing behavior: Towards an integrated model. *European Research on Management and Business Economics, 22*(3), 167–176.

Fullerton, G., & Taylor, S. (2002). Mediating, interactive, and non-linear effects in service quality and satisfaction with services research. *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration, 19*(2), 124–136. https://doi.org/10.1111/j.1936-4490.2002.tb00675.x

Gao, Y. (2005). *Web systems design and online consumer behavior*. Idea Group Publishing.

Gounaris, S., Dimitriadis, S., & Stathakopoulos, V. (2010). An examination of the effects of service quality and satisfaction on customers' behavioral intentions in e-shopping. *Journal of Services Marketing, 24*(2), 142–156. https://doi.org/10.1108/08876041011031118

Griffith, D. A., & Krampf, R. F. (1998). An examination of the web-based strategies of the top 100 U.S. retailers. *Journal of Marketing Theory and Practices, 6*(3), 12–23. https://doi.org/10.1080/10696679.1998.11501801

Grönroos, C., Heinonen, F., Isonlemi, K., & Lindholm, M. (2000). The net offer model: A case example from the virtual marketspace. *Management Decision, 8*(4), 243–252.

Schaupp, L. C. (2010). Web site success: Antecedents of web site satisfaction and re-use. *Journal of Internet Commerce, 9*(1), 42–64. https://doi.org/10.1080/15332861.2010.487414

Ha, H., Janda, S., & Muthaly, S. K. (2010). A new understanding of the satisfaction model in an e-re-purchase situation. *European Journal of Marketing, 44*(7/8), 997–1016. https://doi.org/10.1108/03090561011047490

Holloway, B. B., & Beatty, S. E. (2008). Satisfiers and dissatisfiers in the online environment: A critical incident assessment model. *Journal of Service Research, 10*(4), 347–364.

Huang, M.-H. (2003). Modeling virtual expletory and shopping dynamics: An environmental psychology approach. *Information & Management, 41*, 39–47.

Jiang, P., & Rosenbloom, B. (2005). customer intention to return online: Price perception, attribute level performance and satisfaction unfolding over time. *European Journal of Marketing, 39*(1/2), 150–174.

Kau, A. K., Tang, Y. E., & Ghose, S. (2003). Typology of online shoppers. *Journal of Consumer Marketing, 20*, 139–156.

Kaya, B., Behravesh, E., Abubakar, A. M., Kaya, Ö. S., & Orús, C. (2019). The moderating role of website familiarity in the relationships between e-service quality, e-satisfaction and e-loyalty. *Journal of Internet Commerce, 18*(4), 369–394. https://doi.org/10.1080/15332861.2019.1668658

Kendall, E. L., & Sproles, G. B. (1986). A methodology for profiling consumers' decision-making styles. *The Journal of Consumer Affairs, 20*(2), 267–279.

Kesharwani, S. (2020). E-service quality in banking industry-a review. *Global Journal of Enterprise Information System, 12*(2), 111–118.

Kim, M., Kim, J.-H., & Lennon, S. J. (2006). Online service attributes available on apparel retail websites: An E-S-QUAL approach. *Journal of Service Theory and Practice, 16*(1), 51–77.

King, R. C., Schilhavy, R. A., Chowa, C., & Chin, W. W. (2016). Do customers identify with our website? The effects of website identification on repeat purchase intention. *International Journal of Electronic Commerce, 20*(3), 319–354.

Korten, D. C. (2009). *Agenda for a new economy*. BerrettKoehler.

Lestari, V. T., & Ellyawati, J. (2019). Effect of e-service quality on repurchase intention: Testing the role of e-satisfaction as mediator variable. *International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8*(7), 158–162.

Liao, T. H., & Keng, C. J. (2013). Online shopping delivery delay: Finding a psychological recovery. strategy by online consumer experiences. *Computers in Human Behavior, 29*(4), 1849–1861.

Li, H., & Suomi, R. (2007). Evaluating electronic service quality: A transaction process-based evaluation model. *The European Conference on Information Management and Evaluation, September*, 331–340.

Lin, H.-F. (2010). An application of fuzzy AHP for evaluating course website quality. *Computers & Education, 54*, 877–888.

Lin, J., & Lu, H. (2000). Toward an understanding of the behavioral intention to use a website. *International Journal of Information Management, 20*(3), 197–208.

Lynch, J. G., Jr., & Ariely, D. (2000). Wine online: Search cost affects competition on price, quality, and distribution. *Marketing Science, 19*(1), 83–103.

Mai, J.-E. (2012). The quality and qualities of information. *Journal of the American Society for Information Science and Technology, 64*(4), 675–688. https://doi.org/10.1002/asi.22783

Malhotra, N. (2009). *Marketing research: An applied orientation* (6th ed.). Pearson. ISBN-13: 9780133071757.

Mansouri, S. A., Aktas, E., & Besikci, U. (2016). Green scheduling of a two-machine flow shop: Trade-off between makespan and energy consumption. *European Journal of Operational Research, 248*(3), 772–788. https://doi.org/10.1016/j.ejor.2015.08.064

Mun, Y. Y., Yoon, J. J., Davis, J. M., & Lee, T. (2013). Untangling the antecedents of initial trust in web-based health information: the roles of argument quality, source expertise, and user perceptions of information quality and risk. *Decision Support Systems, 55*(1), 284–295.

Öztürk, E., & Şahin, A. (2020). Effects of consumers' decision-making styles on abandoning online shopping bags. *Çukurova University Social Sciences Institute Journal., 29*(3), 70–90.

Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). E-S-QUAL: A multi-item scale for assessing electronic service quality. *Journal of Service Research, 7*(3), 213–233.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing, 49*(3), 41–50.

Rašković, M., Ding, Z., Hirose, M., Žabkar, V., & Fam, K. S. (2020). Segmenting young-adult consumers in East Asia and Central and Eastern Europe–The role of consumer ethnocentrism and decision-making styles. *Journal of Business Research, 108*, 496–507.

Regina Connolly, R., Bannister, F., & Kearney, A. (2010). Government website service quality: A study of the Irish revenue online service. *European Journal of Information Systems, 19*, 649–667.

Reichheld, F. F., & Schefter, P. (2000). E-loyalty: Your secret weapon on the web. *Harvard Business Review, July–August*, 105–113.

Rita, P., Oliveira, T., & Farisa, A. (2019). The impact of e-service quality and customer satisfaction on customer behavior in online shopping. *Heliyon, 5*(10), 1–14. https://doi.org/10.1016/j.heliyon.2019.e02690

Rohm, A. J., & Swaminathan, V. (2002). A typology of online shoppers based on shopping motivations. *Journal of Business Research, 57*, 748–757.

Rosenbaum, M. S., & Losada, M. (2017). How to create a realistic customer journey map. *Business Horizons, 60*(1), 143–150. https://doi.org/10.1016/j.bushor.2016.09.010

Rust, R. T., & Lemon, K. N. (2001). E-service and the consumer. *International Journal of Electronic Commerce, 5*(3), 85–101.

Sahin, A., Kitapçi, H., Altindag, E., & Gok, M. S. (2017). Investigating the impacts of brand experience and service quality. *International Journal of Market Research, 59*(6), 707–724. https://doi.org/10.2501/IJMR-2017-051

Santos, J. (2003). E-service quality: A model of virtual service quality dimensions. *Managing Service Quality, 13*(3), 233–246.

Schmidt, S., Cantallops, A. S., & Santos, C. P. (2008). The characteristics of hotel websites and their Implications for website effectiveness. *International Journal of Hospitality Management, 27*(4), 504–516. https://doi.org/10.1016/j.ijhm.2007.08.002

Smith, D., Menon, S., & Sivakumar, K. (2005). Online peer and editorial recommendations, trust, and choice in virtual markets. *Journal of Interactive Marketing, 19*(3), 15–37.

Srivastava, K., & Sharma, N. K. (2013). Service quality, corporate brand image, and switching behavior: The mediating role of customer satisfaction and repurchase intention. *Services Marketing Quarterly, 34*(4), 274–291.

Szymanski, D. M., & Hise, R. T. (2000). E-satisfaction: An initial examination. *Journal of Retailing, 76*(3), 309–322. https://doi.org/10.1016/S0022-4359(00)00035-X

Taherdoost, H., Sahibuddin, S., Ibrahim, S., Kalantari, A., Jalaliyoon, N., & Ameri, S. (2012). Examination of electronic service definitions. In *2012 International Conference on Advanced Computer Science Applications and Technologies (ACSAT)* (pp. 73–77). IEEE.

Tandon, U., Kiran, R., & Sah, A. N. (2017). Customer satisfaction as a mediator between website service quality and repurchase intention: An emerging economy case. *Service Science, 9*(2), 106–120.

Thompson, S. H. T. (2002). Attitudes towards online shopping and the internet. *Behavior and Information Technology, 21*(4), 259–271.

Vo, N. T., Chovancová, M., & Tri, H. T. (2020). The impact of e-service quality on customer satisfaction and consumer engagement behaviors toward luxury hotels. *Journal of Quality Assurance in Hospitality and Tourism, 21*(5), 499–523. https://doi.org/10.1080/1528008X.2019.1695701

Wang, L., Law, R., Guillet, B. D., Hung, K., & Fong, D. K. C. (2015). Impact of hotel website quality on online booking intentions: E-trust as a mediator. *International Journal of Hospital Management, 47*, 108–115.

Watcher, K. (2002). Longitudinal assessment of web retailers: Issues from a consumer point of view *Journal of Fashion Marketing and Management, 6*(2), 134–145.

Wolfinbarger, M., & Gilly, M. C. (2003). E-TailQ: Dimensionalizing, measuring, and predicting e-tail quality. *Journal of Retailing, 79*(3), 183–198. https://doi.org/10.1016/S0022-4359(03)00034-4

Xiao, Q. (2016). Managing E-commerce platform quality and its performance implication: A multiple-group structural model comparison. *Journal of Internet Commerce, 15*(2), 142–162. https://doi.org/10.1080/15332861.2016.1143214

Yang, Z., Cai, S., Zhou, Z., & Zhou, N. (2005). Development and validation of an instrument to measure user perceived service quality of information presenting web portals. *Information & Management, 42*, 575–589.

Yang, Z., & Fang, X. (2004). Online service quality dimensions and their relationships with satisfaction: A content analysis of customer reviews of securities brokerage services. *International Journal of Service Industry Management, 15*(3), 302–332.

Yang, Z., & Jun, M. (2002). Consumer perception of e-service quality: From internet purchaser and non-purchaser perspectives. *Journal of Business Strategies, 19*(1), 19–40.

Yoo, B., & Donthu, N. (2001). Developing a scale to measure the perceived quality of internet shopping site (SITEQUAL). *Quarterly Journal of Electronic Commerce, 2*(1), 31–45.

Zarei, G., Nuri, B. A., & Noroozi, N. (2019). The effect of Internet service quality on consumers' purchase behavior: The role of satisfaction, attitude, and purchase intention. *Journal of Internet Commerce, 18*(2), 197–220. https://doi.org/10.1080/15332861.2019.1585724

Zeithaml, V. A., Parasuraman, A., & Malhotra, A. (2002). Service quality delivery through websites: A critical review of extant knowledge. *Journal of the Academy of Marketing, 30*(4), 362–375.

The European Survey on OR/MS Education: Statistical Analysis Addressing the Lecturing Modules

Ana Paula Teixeira, João Luís de Miranda, and Margarida Vaz Pato

Abstract This work addresses the European survey conducted with the aim of learning from the most recent developments in the Higher Education (HE) area concerning the Operational Research/Management Science (OR/MS) field. This regards, in particular: (i) the enrolment; (ii) the reduction in first year students' failure rates; (iii) the perceived value of OR/MS courses; (iv) the OR/MS teaching practices; and (v) the transition onto the Labour Market. A statistical analysis is performed to understand if there is a relation of the respondents' perceptions ("*Positive*" and "*Negative*") and the types of OR/MS modules that were lectured at their institutions (type-A, bachelor/master programmes; type-B, separate majors/minors, and type-C, courses). The results analysis of independence tests are presented; most of the queries concerning the restructuring procedures are weakly related with the lecturing modules and the same occurs with half of the queries regarding the first year students; the opposite occurs with the labour market responses, as all of them are independent of the lecturing modules; and, both for the students enrolment and the OR/MS teaching practices, although weakly one-third of the responses are related with the lectured modules. Finally, further developments within the evidence-based approach are discussed too.

A. P. Teixeira
Pole CMAT-UTAD and Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal
e-mail: ateixeir@utad.pt

J. L. de Miranda
Escola Superior de Tecnologia e Gestão, Instituto Politécnico de Portalegre, Portalegre, Portugal

CERENA - Centro de Recursos Naturais e Ambiente, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
e-mail: jlmiranda@ipportalegre.pt; joaoluismiranda@tecnico.ulisboa.pt

M. V. Pato
ISEG, Universidade de Lisboa, Lisbon, Portugal

CMAFcIO - Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal
e-mail: mpato@iseg.ulisboa.pt

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_11

Keywords OR/MS education · Enrolment · Retention · Restructuration · Teaching practices · Transition into labour market

1 Introduction

Education in OR/MS is crucial for the scientific development of the new European generations, as well as for their success in dealing with complex processes, namely in Data Mining, Decision Analysis, Decision Support Systems, Forecasting, Logistics and Scheduling, among others. Recent publications enhance the relevance of this area, for example, Bell et al. (2013), Čibej (2002), Cochran (2012), Dias (2017), Dupuy et al. (2011), Gorman (2018), Johnes (2015), Sharda et al. (2013), Romm (2018), Smilowitz and Keppler (2020). In spite of the significant progress, Education in OR/MS has shown that further developments, that may help to identify key factors related to OR/MS programmes at Higher Education Institutions (HEI), are still required. This work aims to better understand how European HEI are addressing OR/MS education topics, in order to better ponder what has been done in this area, to gain insight on the future directions of this field, and draw some improvements from the most recent developments (e.g., Belien et al. 2018). In particular, this study contributes with a status point of European HEI concerning: (1) the enrolment of students, (2) the reduction in first year students' failure rates and the promotion of continuity, (3) the restructuring and the perceived value of OR/MS courses, (4) the teaching practices and (5) the transition of graduates onto the labour market, relating these topics with the fact that different types of OR/MS modules are or not available at HEI.

2 Literature Review

Note that multi/transdisciplinary topics like the ones studied here are subject of wide dispersion, both on the scientific fields and even geographically. For instance, Johnes (2015) addressed a variety of issues in education, where OR techniques have been applied in order to improve operations and provide solutions, while Mazzaroli (1998) analysed from his survey results the critical success factors for international education marketing within the promotion of Australian education competitiveness in the international landscape. Therefore, the literature revision for this article mainly focuses on the last two decades; however, some older works are selected too due to their relevance.

In the beginning of the twenty-first century, Čibej (2002) focused on the design and implementation of OR/MS education for those called as forgotten populations (namely, users as a whole and sponsors as their subspecies), showing that more than 90% of projects remained unaccomplished without active support of the top management in the firm. Almost a decade later, Dupuy (2011) presented a new

Operations Research/Operational Analysis education software characterised by a strong embedded interactive component. Approximately at the same time, Cochran 2012 was concerned in making the communication between OR professionals (academic or practitioner) and students, clients, subordinates, supervisors, and colleague become more effective. A year later, a book, that offers a tool box of metaphors and associative OR approaches, was published by Bell et al. (2013), being the application of OR approaches to different complex and uncertain problems encountered in higher education management addressed. In 2017, Dias (2017) describes the experience of applying gamification in an OR/MS course taught to undergraduate management students. Challenges, points, personalised feedback, badges, and leader boards were considered and an increase of students' participation in classes and in the percentage of approved students, as well as a better assessment of the course made by the students were reported. Some recommendations on how to implement an OR course for management students are also given. Almost at the same period, a survey of over 20 articles on field-based education in OR/MS and related disciplines was performed by Gorman (2018); the best practices and lessons learned from these articles on the teaching of field-based research are provided. In 2020, Smilowitz and Keppler (2020) analyse a series of questions related to the use of OR/MS to improve public education systems. The history of OR/MS research in education is traced. COVID-19 impact on education provision is discussed and the links between issues in education over time and OR/MS research interests, methods, and technology are explored.

Surveys and questionnaires are applied in many subjects and contexts. For example, in the last decade of the twentieth century, Mingers (1991) presents the results of a survey of practising OR groups in UK organisations, with the aim of elicit the views of OR practitioners as to the desirable content of an OR MSc. course. At the same time, an IFORS survey was prepared to obtain an international picture of the main features of the educational programs in the field of OR/MS; the first one worldwide on this field. Seventeen countries sent information concerning a total of 176 educational programs in OR/MS and the results were published in 1994 by Tavares 1994. The use of a survey in this study is due to the wide documentation required, the high number of attributes to be addressed, and the dynamics and complexity within the study subjects. The *"OR/MS E*ducation" survey (Teixeira et al. 2016, 2018) was directed at HEI boards, deans, and school directors, programmes coordinators, professors, researchers, and other OR/MS professionals, and it was carried out with wide collaboration; namely with the collaboration of EURO (The European Association of Operational Research Societies) and a significant number of OR/MS national societies.

Considering the international cooperation framework, preliminary results were presented and discussed in OR/MS meetings, both at national level (e.g., Ittmann 2018) and at European level (Belien et al. 2016a, b). In a similar framework, under the umbrella of the Institute for Operations Research and Management Science (INFORMS), Sharda et al. (2013) discussed the teaching opportunities in Business and Big Data Analytics, with applications in athletic injuries and healthcare.

Within the Open Science paradigm (e.g., European Commission, COM 2016–178 final), some basic descriptive statistical information was made available both to the scientific community and to the general public. Namely, through the dedicated webpage in the EU-Survey platform, dedicated publications (Belien et al. 2018), and the aggregate report (Teixeira et al. 2016, hereafter the *EURO-report*).

3 Methods

The survey outputs were analysed in order to identify key factors relating to OR/MS education, to propose subfactors, and to plan subsequent phases for the study. Five major factors of interest were addressed, each one with a number of questions related to the factor. Synoptically:

(i) *Enrolment of students*—The first factor concerns to HEI relationships with pre-university institutions, namely, addressing the typical inflow of students originated from secondary schools. The results are analysed in Sect. 4.1, which focuses both the general enrolment measures and the general activities directed to pre-university institutions, while complementing with the OR/MS scope activities targeting the students' inflow.

(ii) *Reduction in first year students' failure rates and the promotion of continuity*—The second factor under analysis is dedicated to students attending to the first year of graduation programs for the very first time. The results are analysed in Sect. 4.2, which focuses the general activities related both with promotion of continuity and the failure's reduction, while again complementing with activities under the scope of OR/MS that are targeting first year students.

(iii) *Restructuring and the perceived value of OR/MS courses*—This factor is dedicated to the redesign of HEI programmes, since the OR/MS field can benefit from programmes restructuration. The results are analysed in Sect. 4.3, looking for a better understanding about the restructuring procedures, as well as better promoting the OR/MS Education area, with regards either to material or human resources.

(iv) *Teaching practices*—The topics within this factor seem important in mapping the education approaches and formats in the European OR/MS education area, namely, to gain an insight that can be used for particular or local comparisons. The results are analysed in Sect. 4.4, which is specifically aimed at classroom activities and the context within such practices are developed.

(v) *Transition of graduates onto the Labour Market*—The last factor under consideration is focusing the most important topics associated with the successful students' outflows from HEI. The results are analysed in Sect. 4.5, which addresses the HEI relationships with Labour Market organisations, both in general and in the specific scope of OR/MS.

With different approaches, the important students' retention factor is also focused by, e.g., Meer (2014), Bennet (2010), and Rowley (2003) while other works address the teaching practices in OR/MS (e.g., Gorman 2018, Dias 2017, Cochran 2012, Dupuy 2011, Cibej 2002).

In the *EURO-report*, graphs were used to present the aggregate information concerning to each one of the queries. Contingency tables and chi-square independent tests were developed in SPSS v25.0 to analyse any possible association among the received answers and the existence of OR/MS lecturing modules (bachelors/masters, majors/minors, or courses).

The answers, originally in the Likert scale 0–5, were conveniently aggregated into three classes:

- "***Positive***" (includes *Satisfactory*, *Good,* and *Excellent*), corresponding to categories "3-4-5" in the Likert scale;
- "***Negative***" (includes *Poor* and *Fair*) corresponding to categories "1–2" in the Likert scale; and
- "***Non-existent/Lacking sufficient information***", also ***NE***, for category "0".

In fact, this analysis is aggregating those three categories and comparing their results with the associated aggregation of two "negative" categories "1–2"; in plus, category "0" is neutrally considered in here, but other interpretations could classify this category differently. Then, in our point of view, the current approach is the most favourable to ponder the positive perceptions of survey respondents. It was decided to use this data format, just making specific mention to the specific components of each one of the first two classes when relevant.

4 Findings and Discussion

During the full period of open consultation, 191 responses were obtained. It is now intended to reflect if the answering is someway related to the types of OR/MS modules (bachelor/master programmes, separate majors/minors, or courses) that are being lectured in the HEI where the respondent is affiliated. With this in mind, the analysis of the possible associations between the received answers and the type of OR/MS modules that are lectured in the respondent's HEI is performed.

Concerning the four options considered in the survey's introductory section (Query $\mathbf{Q_0}$), the purpose was to identify the types of OR/MS modules that were available in the respondent's HEI:

- **A—***"There is a bachelor/master in OR/MS".*
- **B—***"There is a major/minor in OR/MS as part of a bachelor/master".*
- **C—***"OR/MS courses are part of a bachelor/master but do not form a separate major/minor".*
- **D—***"Other; further specification was required if this option was selected".*

The type-C module was indicated by 45% of the respondents ("just" courses, in our perception, because all the other options were excluded when calculating this percentage), about 28% marked type-A, approximately 26% selected type-B, and type-D was referred by about 1%. The distribution of OR/MS type of modules would allow to analyse the existence of possible relationships with the answers received to the different queries, as well as the intensity of those relationships.

Regarding the five topics mentioned before (**i** to **v**), this section also includes: Sect. 4.1, concerns the enrolment of students; Sect. 4.2 addresses the retention of first year students; in Sect. 4.3, the restructuration of OR/MS courses and programmes is analysed; the OR/MS teaching practices are studied in Sect. 4.4, while the transition and retention of graduates onto labour market is addressed in Sect. 4.5; and, finally, the conclusions are presented in Sect. 5.

4.1 Enrolment of Students

The survey's first section concerns to HEI relationships with pre-university institutions. It focuses enrolment planning, joint work projects, the existence of a HEI interlocutor dedicated to promote the collaboration with pre-university teachers, the assessment of enrolment activities and the development of OR/MS materials in cooperation with primary/secondary schools. The following queries were included:

- **A.1—***"In my Higher Education Institution (HEI) there are plans to promote the enrolment of students"*.
- **A.2—***"In my HEI there are joint work projects with pre-university teachers (for example: Mathematics, Physics, Chemistry, Informatics, Economics, etc.), with the aim of providing pedagogic continuity"*.
- **A.3—***"In my HEI there is a dedicated interlocutor to establish and supervise liaisons with primary/secondary schools"*.
- **A.4—***"In my HEI there is assessment of activities related to the enrolment of students"*.
- **A.5—***"My HEI is developing and supporting contents under the scope of OR/MS in association with primary/secondary schools (for example, a website of educational resources, textbooks, ...)"*.
- **A.6—***"In my HEI there are joint work projects under the scope of OR/MS with pre-university teachers"*.

4.1.1 General Enrolment Measures

European HEIs are aware of the importance of attracting new students, and they apply considerable effort in developing and appreciating measures to enrol students. Namely, Fig. 1 shows a very positive perception of the HEI activities to promote students' enrolment; it is directly constructed from the results for Query **A.1** in the *EURO-report*, by aggregating the Likert scale answering in "Positive", "Negative", and "NE".

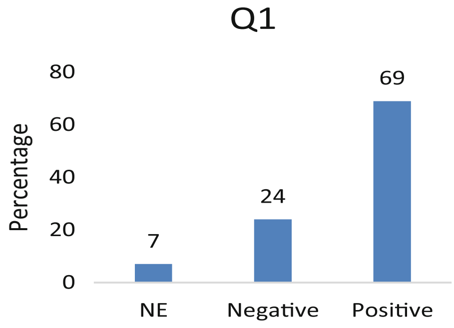


Fig. 1 Query **A.1**

In my Higher Education Institution (HEI) there are plans to promote the enrolment of students

The null, H_o, and alternative, H_1, hypotheses considered to apply the chi-square independent test to Query **A.1** are, respectively:

- **H_o:** There is independence (no relationship) between the existence of plans to promote the enrolment and the type of OR/MS modules existent at the HEIs.
- **H_1:** There is a relationship between the existence of plans to promote the enrolment and the type of OR/MS modules existent at the HEIs.

The independence test is also applied to all the other queries, being the null and alternative hypotheses similar to those stated for Query **A.1**, namely:

- **The null hypotheses** as the non-existence of relationship between the query in analysis and the type of OR/MS modules existent at the HEI (independence); and
- **The alternative hypotheses** as the existence of a relationship between the query answering and the type of OR/MS modules lectured in the HEIs.

The results of all the chi-square tests are presented as suggested in Marôco (2018), and the statements and figures associated with each one of the queries are skipped for reading purposes.

From the results in Table 1—Query **A.1**/Fig. 1, about 69% of the respondents referred their HEI have plans to promote the enrolment; 42% of them rated those plans as *Good* or *Excellent*, and only 7% mentioned that no such plans exist or that they are lacking information on the subject. It is interesting to note that the "*Non-existent/Lacking sufficient information*" answer decreases more than a third when the respondents belong to HEI with at least a bachelor/master in OR/MS (types-A and B), comparing with the answers from the respondents whose HEI "just" have OR/MS courses that are part of a bachelor/master but do not form a separate major/minor (type-C). Additionally, the positive answers also increase considerably (more than 16%) when comparing HEI with type-A modules with HEI with type-C modules ("just" courses). This trend seems to be corroborated by the inferential statistical analysis, which indicates that there is a relationship between the existence of plans to promote the enrolment and the type of OR/MS modules lectured in the

Table 1 Queries **A.1** and **A.4**

% in OR/MS modules		Query **A.1**			Query **A.4**		
		NE/ LSI	Negative	Positive	NE/ LSI	Negative	Positive
OR/MS modules	Courses	12.6%	24.1%	63.2%	18.4%	26.4%	55.2%
	Major/minor	2.0%	30.0%	68.0%	8.0%	42.0%	50.0%
	Bachelor/master	3.7%	16.7%	79.6%	14.8%	29.6%	55.6%
Total		7.3%	23.6%	69.1%	14.7%	31.4%	53.9%

HEI, although a weak one ($\chi^2(4) = 9.55$; p_value $= 0.049$; VCramer $= 0.158$; $N = 191$).

Furthermore, more than half of the respondents (54%) confirmed the assessment of enrolment activities in their HEI (with at least more than a quarter evaluating it as *Good*) and only 15% indicating the non-existence of such activities (see Table 1—Query **A.4**). The chi-square test allows to conclude that the assessment of enrolment activities is strongly related with the types of OR/MS modules existent in the HEI ($\chi^2(4) = 5.059$; p_value $= 0.281$; $N = 191$).

4.1.2 General Activity with Pre-university Institutions

The HEI efforts to establish joint work projects with pre-university education are relatively small (see Table 2—Query **A.2**): just approximately 26% of the respondents rated the existence of those kind of projects positively (with only 7% valuing it as *Excellent*), while more than 40% evaluated it negatively (about a quarter graded it as *Poor*) and almost one third classified these activities as "*Non-existent*". Although the respondents from HEIs with type-B modules (a separate major/minor in OR/MS) correspond to the highest percentage of respondents with negative rating (48%) and the lowest percentage of "*Non-existent/Lacking sufficient information*" answers (24%), inferential statistical analysis confirms that the efforts to establish joint work projects with pre-university education are strong related with the types of OR/MS modules existent in the HEI ($\chi^2(4) = 2.694$; p_value $= 0.61$; $N = 191$).

The perception concerning the existence of a dedicated interlocutor to establish and supervise contacts with primary/secondary schools (see Table 2—Query **A.3**) is slightly better than the efforts to establish joint work projects with pre-

Table 2 Queries **A.2** and **A.3**

% in OR/MS modules		Query A.2			Query A.3		
		NE/ LSI	Negative	Positive	NE/ LSI	Negative	Positive
OR/MS modules	Courses	35.6%	41.4%	23.0%	37.9%	27.6%	34.5%
	Major/minor	24.0%	48.0%	28.0%	20.0%	52.0%	28.0%
	Bachelor/master	35.2%	37.0%	27.8%	46.3%	27.8%	25.9%
Total		32.5%	41.9%	25.7%	35.6%	34.0%	30.4%

university education, with about 30% of the respondents presenting a positive answer (although with just 9% assessing it as *Excellent*), against 34% of negative responses and an additional 36% that referred to it as non-existing. It is interesting to note that almost half (46.3%) of the respondents from HEIs with type-C modules (where OR/MS courses are part of a bachelor/master but do not form a separate major/minor) said that this interlocutor did not exist and more than a quarter (25.9%) rated it negatively. It can also be noticed that more than half (52%) of the respondents of HEIs with type-B of OR/MS modules rated it positively, while the positive answers of the HEIs with other type of OR/MS modules were about 28%. This tendency is supported by the inferential statistical analysis; it can be said that there is no relationship between the existence of a dedicated interlocutor to establish and supervise contacts with primary/secondary schools and the type of OR/MS modules existent in the HEI, although very weak ($\chi^2(4) = 12.64$; p_value $= 0.013$; VCramer $= 0.182$; $N = 191$).

4.1.3 OR/MS Scope Activity with Pre-university Institutions

A small amount of OR/MS material has been developed in cooperation with pre-university schools (see Table 3—Query **A.5**): only slightly more than 10% answered this query positively, while 21% rated it as *Poor* and more than half (60%) of the respondents selected the "*Non-existent*" option. Moreover, although the choice of the "*Non-existent*" option decreases more than 10% when the respondents belong to HEIs with at least a bachelor/master in OR/MS (types-A or B of OR/MS modules), comparing with the ones of the respondents whose HEIs just have OR/MS courses that are part of a bachelor/master but do not form a separate major/minor (type-C), the application of the chi-square test allows to conclude that the development and support of materials in the OR/MS field by HEIs in association with primary/secondary schools is strongly related with the OR/MS type of modules existent in the HEI ($\chi^2(4) = 5.919$; p_value $= 0.205$; $N = 191$).

The situation concerning the OR/MS joint projects with pre-university teachers is even worse (Table 3—Query **A.6**): only about 1 respondent in 12 (8%) evaluated positively this mode of collaboration, with almost a quarter rating it as *Poor* and nearly 2 respondents in 3 (64%) considering it as "*Non-existent*". ~~Similarly to~~ As in the previous item, the existence of joint work projects in the OR/MS field with pre-

Table 3 Queries **A.5** and **A.6**

% in OR/MS modules		Query **A.5**			Query **A.6**		
		NE/ LSI	Negative	Positive	NE/ LSI	Negative	Positive
OR/MS modules	Courses	67.8%	21.8%	10.3%	71.3%	21.8%	6.9%
	Major/minor	52.0%	38.0%	10.0%	62.0%	28.0%	10.0%
	Bachelor/master	55.6%	27.8%	16.7%	55.6%	37.0%	7.4%
Total		60.2%	27.7%	12.0%	64.4%	27.7%	7.9%

university teachers is also modules type dependent in those HEIs ($\chi^2(4) = 4.521$; p_value = 0.34; $N = 191$), although a reduction of the "*Non-existent/Lacking sufficient information*" option can be seen from type-C to type-A.

In summary, the collected results on the independence testing between the *Enrolment* topics in the queries **A1–A6** and the type of modules existing in the HEI follow:

- **A.1, A.3—**The null hypothesis is rejected (dependence, but weak relation);
- **A.2, A.4, A.5, A.6—**The null hypothesis is not rejected (independence).

4.2 Retention of First Year Students

The second section of the survey is dedicated to HEI efforts to promote continuity and reduce failure of students attending to the first year of graduation programs for the very first time. The surveyed aspects cover the planning and assessment of activities to achieve this goal, the existence of a dedicated interlocutor in this subject, and the development of OR/MS material specifically for first year students.

To achieve this goal, the survey included the queries:

- **B.7—***"My HEI has plans to promote continuity and reduce the failure rate of 1st-year students (for example, through tutoring, mentoring, counselling and by supporting the development of specific training programs for faculty staff)."*
- **B.8—***"In my HEI there is a dedicated interlocutor to establish and supervise liaison with 1st-year students."*
- **B.9—***"In my HEI there is assessment of those activities designed to promote continuity and to avoid failure of 1st-year students."*
- **B.10—***"My HEI is developing and supporting contents under the scope of OR/MS for 1st-year students (for example, a website of educational resources, textbooks, . . .)."*

4.2.1 General Promotion of Continuity and Reduce Failure Measures

The majority of the respondents (about 63%) evaluated positively their HEIs efforts to promote study continuity and reduce the failure rate of first year students (Table 4—Query **B.7**). About 41% of respondents evaluated them as *Good* or even *Excellent*. Such efforts might include, but are not limited to, tutoring, mentoring or counselling first year students as well as supporting the development of specific training programmes. The percentage of positive answers increases (from 58.6% to 70.4%) as the respondents belong to HEI with type-A to type-C modules. Inferential statistical analysis corroborates the existence of a relation (although weak) between

Table 4 Queries **B.7**, **B.8**, and **B.9**

% in OR/MS modules		Query **B.7**			Query **B.8**			Query **B.9**		
		NE/ LSI	Negative	Positive	NE/ LSI	Negative	Positive	NE/ LSI	Negative	Positive
OR/MS modules	Courses	9.2%	32.2%	58.6%	19.5%	29.9%	50.6%	16.1%	36.8%	47.1%
	Major/minor	2.0%	36.0%	62.0%	16.0%	26.0%	58.0%	10.0%	44.0%	46.0%
	Bachelor/master	13.0%	16.7%	70.4%	18.5%	29.6%	51.9%	14.8%	25.9%	59.3%
Total		8.4%	28.8%	62.8%	18.3%	28.8%	52.9%	14.1%	35.6%	50.3%

the existence of efforts to promote study continuity and reduce the failure rate of first year students and the modules type in the HEI ($\chi^2(4) = 8.599$; p_value = 0.072; VCramer = 0.15; $N = 191$).

Table 4—Query **B.8** concerns to the existence of a dedicated interlocutor in a HEI that supports and supervises the contacts with first year students. It can be observed that slightly more than half of the respondents (52.9%) classified it positively; in particular, more than one third rated it as *Good* or even *Excellent*. The opinion of the respondents seems similar, strongly dependent of the modules type in their HEI. This is corroborated by inferential statistical analysis ($\chi^2(4) = 0.749$; p_value = 0.945; $N = 191$).

The results relative to the assessment of activities designed to promote continuity and to avoid failure of first year students are presented in Table 4—Query **B.9**. Half of the respondents rated it negatively or reported its non-existence, while the other half classified it positively (with approximately one third evaluated these assessment activities as *Good* or even *Excellent*). Although the positivity is higher among the respondents of HEIs with type-C of OR/MS modules, inferential statistical analysis allows to conclude that the assessment of such activities is strong related with the type of OR/MS modules that exist at the HEI ($\chi^2(4) = 4.523$; p_value = 0.34; $N = 191$).

4.2.2 Activities Under the Scope of OR/MS

One of the purposes of this survey was to find out to what extent European HEIs are developing support material in the OR/MS field specifically for first year students. Examples of such material are websites with educational resources and textbooks. Counterintuitively, the results in Table 5—Query **B.10** may suggest that providing such material is not a common practice in Europe.

Only 25% of the respondents evaluated these initiatives positively; on the contrary, not only about 39% of respondents indicated that either such practice does not exist or they are not aware of any such initiatives in their HEI, but about 36% of them assessed the provision of such materials negatively, while a quarter of them even rated it poorly. The lack of knowledge on these activities decreases from 54% to 24.1% when the OR/MS modules in the HEI goes from type-A to type-C, while the positive evaluations double from 17% to 35%. Inferential statistical

Table 5 Query B.10

% in OR/MS modules		Query **B.10**		
		NE/ LSI	Negative	Positive
OR/MS modules	Courses	54.0%	28.7%	17.2%
	Major/minor	28.0%	44.0%	28.0%
	Bachelor/Master	24.1%	40.7%	35.2%
Total		38.7%	36.1%	25.1%

analysis allows to conclude that there is no relationship among the development and the support of material in the OR/MS fields for first year students and the type of OR/MS modules that exist in those HEIs, although a weak one ($\chi^2(4) = 16.714$; p_value = 0.002; VCramer = 0.209; $N = 191$).

In summary, the collected results on the independence testing between the *Promotion of Continuity* topics in the queries **B7-B10**, and the type of modules existing in the HEI are presented:

- **B.7, B.10—**The null hypothesis is rejected (weak relation);
- **B.8, B.9—**The null hypothesis is not rejected (independence).

4.3 OR/MS Restructuring of Courses and Programs

The topic *Restructuring Procedures* is dedicated to the redesign of HEI programmes, for example, the procedures resulting from the Bologna Agreement that proposed more comparable HE programmes within the signing countries (http://ec.europa.eu/education/policy/higher-education/bologna-process_en.htm). The OR/MS field can benefit from the programmes restructuration; therefore, the opportunity to develop international cooperation programmes, the importance to equip and develop OR/MS laboratories, and to assess the impact of these restructurings on OR/MS activities are surveyed. These factors seem important to better understand such restructuring procedures, as well as to better promote the OR/MS Education area, both with regards to support material and human resources. For that purpose, the results of the following queries were analysed:

- **C.11—***"In my HEI when programmes are restructured, the number of OR/MS courses or modules in Economics/ Engineering/ Exact Sciences/ Management programmes is increased."*
- **C.12—***"In my HEI there are ongoing international cooperation programmes under the scope of OR/MS (for example, MSc/PhD programmes, modules, intensive programmes, curriculum development)."*
- **C.13—***In my HEI there is specific OR/MS equipment or a laboratory (for example: study materials, computer facilities, specific software, technical support, ...).*
- **C.14—***"In my HEI the assessment of 'OR/MS-based' postgraduate programmes (PhD., MSc., advanced courses) or 'OR/MS-based' expertise courses (requested by organisations) has impact on subsequent programmes restructuring."*

Concerning the increase on the number of OR/MS courses or modules in HEI programmes, unfortunately an overall negative perception prevails, as can be observed in Table 6—Query **C.11**: although only 13.6% of the respondents did not have sufficient information, about half of the respondents (49.2%) indicated that in their HEI the number of OR/MS courses decreased after programme restructuring.

Table 6 Queries **C.11** and **C.12**

% in OR/MS modules		Query **C.11**			Query **C.12**		
		NE/ LSI	Negative	Positive	NE/ LSI	Negative	Positive
OR/MS modules	Courses	13.8%	52.9%	33.3%	33.3%	37.9%	28.7%
	Major/minor	10.0%	50.0%	40.0%	22.0%	24.0%	54.0%
	Bachelor/master	16.7%	42.6%	40.7%	11.1%	35.2%	53.7%
Total		13.6%	49.2%	37.2%	24.1%	33.5%	42.4%

More specifically: only about 17% of the respondents indicated the number of OR/MS courses, after programmes restructuration, as *Good* or *Excellent* (while about 29% of the respondents rated it poorly). The chi-square test allows to say that the increase on the number of OR/MS courses or modules in HEI programs is strongly related with the types of OR/MS modules the HEI offers ($\chi^2(4) = 2.213$; p_value $= 0.697$; $N = 191$).

The survey results presented in Table 6—Query **C.12** reveal that about 42% of the respondents evaluated the international cooperation within the OR/MS field positively, with 28% considering it *Good or Excellent*. Nevertheless, 33.5% rated it below the satisfaction level (with about 19% assessing it as *Poor*) and about 24% were not aware of international cooperation with respect to the OR/MS courses or modules in their HEI, most probably because there are none.

Note that a reduction of two thirds of the "*Non-existent*" answers and an increase of the positive answers from 28.7% to about 54% can be observed when comparing HEI with "just" OR/MS courses with HEI with bachelor/master in OR/MS. These differences are corroborated by the chi-square test, that allows to conclude that there is no relationship between the existence of international cooperation programs in the OR/MS field and the type of OR/MS module lectured in the HEI, although weakly ($\chi^2(4) = 15.9$; p_value $= 0.003$; VCramer $= 0.289$; $N = 191$).

The perspective is that HEI offering OR/MS courses or modules have sufficient OR/MS equipment like, e.g., study materials, computer facilities, specific and appropriate software tools, and technical support. As shown in Table 7—Query **C.13**, the majority of the respondents (64.4%) indicated to have sufficient equipment, with 42% rating it as *Good* or *Excellent*. Nonetheless, about 8% of the respondents had either no idea of the existence of specific OR/MS equipment or indicated that there is none and the remaining 27.7% evaluated the availability of specific OR/MS equipment negatively (with 14% rating it poorly). It can also be

Table 7 Queries C.13 and C.14

% in OR/MS modules		Query **C.13**			Query **C.14**		
		NE/LSI	Negative	Positive	NE/LSI	Negative	Positive
OR/MS modules	Courses	13.8%	34.5%	51.7%	37.9%	33.3%	28.7%
	Major/minor	2.0%	28.0%	70.0%	22.0%	38.0%	40.0%
	Bachelor/master	3.7%	16.7%	79.6%	13.0%	20.4%	66.7%
Total		7.9%	27.7%	64.4%	26.7%	30.9%	42.4%

observed; for one side, a reduction of the "*Non-existent*" option from 13.8% to 3.7% and of the negative rating from 34.5% to 16.7%; for other side, an increase of the positive evaluation from 51.7% to 79.6%, as the existing OR/MS modules are no longer "just" courses to be bachelor/master programmes. These results are corroborated by inferential statistical analysis, which allows to conclude that there is no relationship between the existence of specific OR/MS equipment or laboratory and the type of OR/MS modules in the HEI, although weakly ($\chi^2(4) = 15.446$; p_value = 0.004; VCramer = 0.201; $N = 191$).

From Table 7—Query **C.14**, it can be observed that an overall nonpositive perception of the impact that the appraisal of either OR/MS postgraduate programmes (like PhD. and advanced courses) or OR/MS-based expertise courses (requested by organisations) have on subsequent programmes restructuring prevails: 26.7% of the respondents are not aware of any impact of OR/MS follow-up education on the programme restructuring in their HEI, probably because there is none, and an additional 30.9% of the respondents evaluated it negatively. From the remaining 42.4% respondents that rated this impact positively, only 4% classified it as *Excellent*. As in the previous case, it can also be observed a reduction of the "*Non-existent*" option from 37.9% to 13% and an increase of the positive evaluation from 28.7% to 66.7%, as the existing OR/MS modules of in HEI are no longer "just" courses to be bachelor/master programmes. Inferential statistical analysis corroborates the previous results and allows to conclude that there is no relation between the impact of OR/MS-based postgraduate programmes or expertise courses on the subsequent programmes restructuring and the types of OR/MS modules in the HEI, although weakly ($\chi^2(4) = 22.657$; p_value = 0.00; VCramer = 0.244; $N = 191$).

In summary, the collected results on the independence testing between the *Restructuring Procedures* topics in the queries **C.11–C.14** and the type of modules existing in the HEI are as follows:

- **C.11—**The null hypothesis is not rejected (independence);
- **C.12, C.13, C.14—**The null hypothesis is rejected (dependence, but weak relation).

4.4 OR/MS Teaching Practices

The section *Teaching Practices* is specifically aimed at classroom activities and the context within which these are developed; it comprises the queries

- **D.15**—*"In my HEI syllabus comparisons with other HEI occur often, OR/MS techniques are up-to-date, companies needs are regularly checked."*

Table 8 Query **D.15**

% in OR/MS modules		Query **D.15**		
		NE/LSI	Negative	Positive
OR/MS modules	Courses	13.8%	35.6%	50.6%
	Major/minor	8.0%	38.0%	54.0%
	Bachelor/master	7.4%	27.8%	64.8%
Total		10.5%	34.0%	55.5%

- **D.16**—*"In my HEI OR/MS courses or modules have become less theoretical and more practically oriented (real-life cases)."*
- **D.17**—*"In my HEI OR/MS courses or modules are mainly be taught through lectures instead of by using more active teaching formats (for example, case studies, group assignments, company projects)."*

The comparison and transparency of OR/MS syllabus, the balance between theoretical and practical approaches of OR/MS courses and modules, as well as the learning formats are evaluated in this sector. These factors seem important in mapping the education approaches and formats in the European OR/MS education area, to gain an insight that can be used to perform some comparisons.

One of the surveyed aspects intends to let one know if European HEIs are properly monitoring the relevance of their OR/MS education (e.g., by comparing their programs with other HEI, regularly updating OR/MS techniques, or checking materials alignment with company needs).

The results in Table 8—Query **D.15** suggest that there is still significant scope for improvement in this regard; namely, more than half of the respondents classified the monitoring positively (55.5%), while only 9% indicated it as *Excellent*; for other side, 10.5% referred that this activity is "*Non-existent*" or that they lacked information about it, while 34% evaluated it negatively. Furthermore, reductions both on the "*Non-existent*" option (from 13.8% to 7.4%) and the negative assessment (from 35.6% to 27.8%) can be seen; it can be also observed an increase of the positive evaluation (from 50.6% to 64.8%), as the existing OR/MS modules are no longer "just" courses to be bachelor/master programmes. Nonetheless, the syllabi comparisons with other HEI, the use of up-to-date OR/MS techniques and the regular checking of the companies' needs are strongly related with the type of OR/MS modules that exist at HEIs ($\chi^2(4) = 3.855$; p_value = 0.426; $N = 191$).

Another surveyed issue concerns the fact that OR/MS courses in European HEI have become less theoretical and more practically oriented and involving real-life case studies. From the 191 respondents, just 4% of the respondents indicated they lack information on this issue. As this is a negligible percentage and the conditions necessary to apply the chi-square test were not fulfilled, only the 184 positive/negative responses will be analysed. According to Table 9—Query **D.16**, more than two thirds of the 184 respondents said that this happened in their HEI, with approximately 40% rating it as *Good* or *Excellent*, and about 29% of them did not see a pattern towards more practically oriented OR/MS education in their HEI.

Table 9 Queries **D.16** and **D.17**

% in OR/MS modules		Query **D.16**		Query **D.17**	
		Negative	Positive	Negative	Positive
OR/MS modules	Courses	32.1%	67.9%	48.2%	51.8%
	Major/minor	22.9%	77.1%	52.1%	47.9%
	Bachelor/master	30.8%	69.2%	28.3%	71.7%
Total		29.3%	70.7%	43.5%	56.5%

As in the previous issue, the opinions about OR/MS courses or modules becoming less theoretical and more practically oriented are strongly related with the type of OR/MS modules existent at HEIs ($\chi^2(2) = 1.325$; p_value = 0.516; $N = 184$).

When compared with traditional lectures, modern teaching approaches increasingly value active teaching formats like case studies, group assignments or company projects. Does OR/MS education within European HEIs follow the new trend? From the 191 respondents, only a very small percentage (about 3%) indicated lack of information on this issue. As this is a negligible percentage and the conditions necessary to apply the chi-square test were not fulfilled, the analysis only focused on the 186 negative / positive responses.

The results in Table 9—Query **D.17** show that less than half (43.5%) of the analysed HEI still mainly use traditional lectures, while 56.5% use more active teaching formats. It is also possible to observe a 20% decrease for the negative rating and a 20% increase for the positive evaluation, as the existing OR/MS modules are no longer "just" courses to be bachelor/master programmes. Furthermore, using inferential statistics it can be concluded that a relation between the type of OR/MS modules in the HEI and the use of more practical formats for teaching OR/MS modules exists, although weakly ($\chi^2(3) = 7.193$; p_value = 0.027; VCramer = 0.197; $N = 186$).

In summary, the collected results on the independence testing between the topics under analysis in the queries **D15–D17** and the type of modules existing in the HEI are:

- **D.15, D.16**—The null hypothesis is not rejected (independence);
- **D.17**—The null hypothesis is rejected (dependence, but weak relation).

4.5 Transition and Retention of Graduates onto Labour Market

The sector *Labour Market* focuses on HEI relationships with Labour Market organisations, both in general and in the specific OR/MS scope. The surveyed aspects cover important considerations related to curricula content, to HEI boards and decision-makers, employability concerns and the transition of graduates into the Labour Market; these are all important factors for both program accreditation and students' enrolment. For that, the results of the following queries are analysed.

- **E.18**—*"My HEI promotes the transition of graduates onto the Labour Market (for example, by providing: a website, a support office, business materials, professional computer/facilities, professional software, dedicated activities/events, and consultation with employers/organisations when designing new programmes)."*
- **E.19**—*"In my HEI, there is a dedicated interlocutor to establish and supervise liaisons with Labour Market organisations."*
- **E.20**—*"In my HEI there is assessment of activities related to the Labour Market."*
- **E.21**—*"My HEI specifically promotes entry into the Labour Market in the OR/MS field."*
- **E.22**—*"My HEI provides scientific and technical updates into organisations and institutions, particularly within the OR/MS field."*

4.5.1 General Measures

In this subsection the general relationships of HEI with Labour Market organisations are studied.

Information in Table 10—Query **E.18** seems to support that HEIs are effectively concerned about the transition of graduates onto the Labour Market. In fact, although 30.4% of the respondents evaluated the efforts made by their HEIs to promote this transition negatively and 6.8% indicated a lack of information in their case, almost two thirds of the respondents evaluated it positively, with 42% of them rating those efforts as *Good* or *Excellent*. Additionally, the chi-square test shows that the promotion of the transition of graduates onto the Labour Market is independent of the type of OR/MS modules that the HEI offers ($\chi^2(4) = 2.888$; p_value = 0.577; $N = 191$).

Furthermore, (see Table 10—Query **E.19**), though about 20% of the respondents indicated lack of information concerning the existence in their HEI of a dedicated interlocutor to establish and supervise contacts with Labour Market organisations and 29.3% of them assessed negatively its existence, more than half of the respondents (50.8%) evaluated it positively, with 36% of them rating this initiative as *Good* or *Excellent*. As it happens with the previous item, inferential statistical analysis also allows to conclude that the existence of a dedicated interlocutor to establish and supervise liaisons with Labour Market organisations is strong related with the types of OR/MS modules the HEI offers ($\chi^2(4) = 0.644$; p_value = 0.958; $N = 191$). In a similar way, despite the assessment of Labour Market activities received about 31% of negative responses and 14.1% of the respondents indicated a lack of information on this issue, more than half of the respondents (55%) evaluated this statement positively (see Table 10—Query **E.20**). Once more, the chi-square

Table 10 Queries E.18, E.19, and E.20

% in OR/MS modules		Query **E.18**			Query **E.19**			Query **E.20**		
		NE/LSI	Negative	Positive	NE/LSI	Negative	Positive	NE/LSI	Negative	Positive
OR/MS modules	Courses	5.7%	35.6%	58.6%	20.7%	31.0%	48.3%	13.8%	32.2%	54.0%
	Major/minor	8.0%	22.0%	70.0%	18.0%	30.0%	52.0%	12.0%	32.0%	56.0%
	Bachelor/ master	7.4%	29.6%	63.0%	20.4%	25.9%	53.7%	16.7%	27.8%	55.6%
Total		6.8%	30.4%	62.8%	19.9%	29.3%	50.8%	14.1%	30.9%	55.0%

test indicates that the assessment of activities related to the Labour Market is strong related with the type of OR/MS modules that the HEI offers ($\chi^2(4) = 0.677$; p_value = 0.954; $N = 191$).

4.5.2 OR/MS Scope Measures

Now, the relationships of HEIs with Labour Market organisations within the OR/MS field will be focused specifically. In this case, an opposing trend to the one recorded in the previous subsection can be observed. In Table 11—Query **E.21**, a nonpositive pattern is clearly shown for HEI efforts in promoting entry of graduates into the OR/MS field Labour Market. This issue received positive ratings from about 35% of the respondents (with just 7% of *Excellent*) against 41% negative answers (with 22% of the participants assessing it poorly), and 24% of respondents selected the non-existing option.

Moreover, although an increase of 20% on the positive evaluations and a 14% decrease of the "*Non-existent*" option as the OR/MS modules existent in the HEI are not "just" courses to be bachelor/master programmes, statistical inference allows to conclude that the promotion into the OR/MS specific Labour Market is strong related with the type of OR/MS modules in the HEIs ($\chi^2(4) = 7.504$; p_value = 0.112; $N = 191$).

When specifically addressing scientific and technical updating on the OR/MS field, the results are worse (Table 11—Query **E.22**): the positive responses just reach 30% of the respondents (with only 5% of *Excellent*), in contrast with 46.8% negative evaluations (with 26% of the participants assessing it as *Poor*) and 23.2% of the respondents selecting the "*Non-existent"* option. The strong relationship towards the type of OR/MS modules that exist at HEIs is also verified for scientific and technical updates provided by HEIs into organisations and institutions. Particularly within the OR/MS field ($\chi^2(4) = 3.739$; p_value = 0.443; $N = 191$), though a considerable decrease of the negative assessment (from 52.3% to 40.7%) and an increase of the positive evaluation (from 24.4% to 38.9%) can be observed as the OR/MS modules existent in the HEI are not "just" courses to be bachelor/master programmes.

In summary, for all the collected results on the independence testing between the *Labour Market* topics under analysis in the queries **E18–E22**, and the type of

Table 11 Queries **E.21** and **E.22**

% in OR/MS modules		Query **E.21**			Query **E.22**		
		NE/LSI	Negative	Positive	NE/LSI	Negative	Positive
OR/MS modules	Courses	31.0%	42.5%	26.4%	23.3%	52.3%	24.4%
	Major/minor	20.0%	42.0%	38.0%	26.0%	44.0%	30.0%
	Bachelor/master	16.7%	37.0%	46.3%	20.4%	40.7%	38.9%
Total		24.1%	40.8%	35.1%	23.2%	46.8%	30.0%

modules existing in the HEI, the null hypothesis is not rejected (independence is substantiated).

5 Conclusions

Within the Open Science paradigm, this work allows a first inferential analysis on the results published at the *EURO-Report* on OR/MS Education. It could be observed that all the answers received are either independent or weakly related to the lectured types of OR/MS modules at the respondent HEI.

A. Concerning the enrolment of students, the existing types of OR/MS modules are related, although weakly, not only with the existence of plans to promote the enrolment of students, but also with the existence of a dedicated interlocutor to establish and supervise liaisons with primary/secondary schools; the independence checks on all the other topics related to this subject.
B. In relation to the approaches within first year students, both the existence of plans to promote continuity and reduce the failure rate and the development/ support of contents under the scope of OR/MS are weakly related to the existent types of OR/MS modules.
C. Regarding the restructuring procedures, there is a weak relation between the types of OR/MS modules and each one of the following topics:
   - the existence of ongoing international cooperation programmes under the scope of OR/MS;
   - the existence of specific OR/MS equipment or a laboratory; and
   - the assessment of "OR/MS-based" postgraduate programmes or "OR/MS-based" expertise courses, while independence is verified for the increasing number of OR/MS courses or modules in Economics/ Engineering/ Exact Sciences/ Management.
D. For the teaching practices, only the predominant use of lectures instead of more active teaching formats to teach OR/MS courses or modules is not independent of the types of OR/MS modules existent at the respondent HEI and the relation is weak.
E. Finally, there is no relation between each one of the topics on labour market and the existent types of OR/MS modules.

In this way, further developments will include statistical analysis with other factors and/or subfactors, dedicated interviews and meetings, and grounded theory techniques to adequately support country comparisons.

Acknowledgements The authors thank *Universidade de Trás-os-Montes e Alto Douro*, *Instituto Politécnico de Portalegre*, and *Instituto Superior de Economia e Gestão*. They thank CMAT-*Centro de Matemática*, CERENA-*Centro de Recursos Naturais e Ambiente,* and CMAFcIO, and the support from national funding of *FCT- Fundação para a Ciência e a Tecnologia* under the strategic projects UIDB/00013/2020, UIDB/04028/2020, and UID/MAT/04561/2020. The authors also thank the survey respondents for their helpful collaboration, without them this study could not have been carried out, and thank two anonymous reviewers for their valuable comments.

References

Beliën, J., Ilchenko, K., Miranda, J., Oliveira, J. F., Sörensen, K., Sofer, A., & Zekic-Susac, M. (2016a). Panel discussion on OR/MS education: Countries' differences within a shared European view? In *European Conference on Operational Research (EURO),* 2016/07/03–2016/07/06, Poznań. www.euro-online.org/conf/admin/tmp/program-euro28.pdf and https://ec.europa.eu/eusurvey/files/c45158f9-3b6b-4dda-836c-4e20537d5fe5

Beliën, J., Ittmann, H., Laumanns, M., Miranda, J., Teixeira, A. P., & Vaz Pato, M. (2016b). European study on OR/MS education: Brief results and preliminary trends. In *European conference on operational research (EURO),* 2016/07/03–2016/07/06, Poznań. www.euro-online.org/conf/admin/tmp/program-euro28.pdf

Beliën, J., Teixeira, A. P., & Ittmann, H. W. (2018). In J. L. De Miranda, M. Laumanns, & M. V. Pato (Eds.), *Advances in operations research education: European studies*. Springer. https://doi.org/10.1007/978-3-319-74104-8. isbn:978-3-319-74103-1. http://www.springer.com/us/book/9783319741031

Bell, G., Warwick, J., & Galbraith, P. (Eds.). (2013). *Higher education management and operational research: Demonstrating new practices and metaphors* (Vol. 54). Springer Science & Business Media.

Bennett, R., & Kane, S. (2010). Factors associated with high first year undergraduate retention rates in business departments with non-traditional student intakes. *The International Journal of Management Education, 8*(2), 53–66.

Čibej, J. A. (2002). Operations research education for forgotten populations. *European Journal of Operational Research, 140*(2), 225–231.

Cochran, J. J. (2012). You want them to remember? Then make it memorable! Means for enhancing operations research education. *European Journal of Operational Research, 219*(3), 659–670.

COM. (2016). 178 final, *European Cloud Initiative - Building a competitive data and knowledge economy in Europe* (EC, Brussels, April-2016). https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=15266

Dias, J. (2017). Teaching operations research to undergraduate management students: The role of gamification. *The International Journal of Management Education, 15*(1), 98–111.

Dupuy, A., Ishutkin, V., Pickl, S., & Tschiedel, R. (2011). Judgment based analysis via an interactive learning software for modern operations research education and training. In *Operations research proceedings 2010* (pp. 623–628). Springer.

Gorman, M. F. (2018). A survey of research in field-based education: A summary of process, best practices, and lessons learned. *INFORMS Transactions on Education, 18*(3), 145–161.

Ittmann, H. W. (2018). A south African perspective on OR/MS education. In *Advances in operations research education* (pp. 59–77). Springer. https://doi.org/10.1007/978-3-319-74104-8_6. https://link.springer.com/chapter/10.1007/978-3-319-74104-8_6

Johnes, J. (2015). Operational research in education. *European Journal of Operational Research, 243*(3), 683–696.

Marôco, J. (2018) Análise Estatística com o SPSS Statistics, 7th edition, ReportNumber, Análise e Gestão de Informaão, Lda.

Mazzarol, T. (1998). Critical success factors for international education marketing. *International Journal of Educational Management, 12*, 163–175.

Meer, N. M., & Chapman, A. (2014). Assessment for confidence: Exploring the impact that low-stakes assessment design has on student retention. *The International Journal of Management Education, 12*(2), 186–192.

Mingers, J. (1991). The content of MSc operational research courses: Results of a questionnaire to OR groups. *Journal of the Operational Research Society, 42*(5), 375–382.
Rowley, J. (2003). Retention: Rhetoric or realistic agendas for the future of higher education. *International Journal of Educational Management, 17*(6), 248–253.
Romm, N. R. (2018). Reflections on a multi-layered intervention in the South African public education system: Some ethical implications for Community Operational Research. *European Journal of Operational Research, 268*(3), 971–983.
Sharda, R., Asamoah, D. A., & Ponna, N. (2013). Business analytics: Research and teaching perspectives. In *Proceedings of the ITI 2013 35th international conference on information technology interfaces* (pp. 19–27). IEEE. https://doi.org/10.2498/iti.2013.0589
Smilowitz, K., & Keppler, S. (2020). On the use of operations research and management in public education systems. *INFORMS TutORials in Operations Research null(null):* 84–105. https://doi.org/10.1287/educ.2020.0216
Tavares, L. V. (1994). Education in OR/MS at a glance: Results of the first IFORS survey. *International Transactions in Operational Research, 1*(4), 389–394.
Teixeira, A. P., Ittmann, H. W., Beliën, J., Miranda, J. L., Laumanns, M., & Pato, M. V. (2016). European study on OR/MS education. *Aggregate Report–Preliminary Results*, Portalegre Polytechnics Institute, Portalegre, Portugal. isbn:978-989-8806-09-3. https://ec.europa.eu/eusurvey/files/ab582c30-c2d0-42e6-a911-338354d6f4b0 and https://ec.europa.eu/eusurvey/files/cfadefb6-a182-4704-aba4-cb49670c9998
Teixeira, A. P., Ittmann, H. W., Beliën, J., de Miranda, J. L., Laumanns, M., & Pato, M. V. (2018). European study on OR/MS education. In *Advances in operations research education* (pp. 1–12). Springer. https://doi.org/10.1007/978-3-319-74104-8_1. https://link.springer.com/chapter/10.1007/978-3-319-74104-8_1

The Skill Wage Premium in the Turkish Labor Market from *The Race Between Technology and Education* Perspective

Bilge Eriş Dereli, Narod Erkol, Haluk Levent, and Benan Zeki Orbay

Abstract In Turkey, a dramatic surge in the number of university graduates has increased the share of skilled labor in the workforce over the last two decades. This study aims to analyze the effects of this change on the skill wage premium by using Goldin and Katz's (*The race between education and technology*. Belknap for Harvard University Press, 2008) framework in *The Race between Education and Technology*. In contrast with previous studies on developed economies, our results show that the skill wage premium is declining in Turkey, with the primary reasons being supply-related factors and the significant rise in the real minimum wage. This indicates that no significant change in the Turkish labor market's skill demand in the 2004–2019 period originated from skill-biased technological change.

Keywords Wage gap · Skill wage premium · Skilled labor supply · Turkish labor market · Skill-biased technological change

B. Eriş Dereli
Department of Economics, Marmara University, İstanbul, Turkey
e-mail: bilge.eris@marmara.edu.tr

N. Erkol
School of Advanced Vocational Studies, Banking and Insurance, İstanbul Bilgi University, Beyoğlu, İstanbul, Turkey
e-mail: narod.erkol@bilgi.edu.tr

H. Levent · B. Zeki Orbay (✉)
Business Administration Department, İstanbul Bilgi University, İstanbul, Turkey
e-mail: haluk.levent@bilgi.edu.tr; benan.orbay@bilgi.edu.tr

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_12

1 Introduction

Since the early 2000s, the number of Turkish universities has increased dramatically,[1] and university enrolment has risen from 1.5 million to 7.9 million students. Whether the Turkish economy has the potential to create sufficient growth in employment to accommodate these university graduates remains a critical economic issue that needs to be examined. If the demand for skilled labor is not sufficient, most of these graduates will be overeducated for their jobs. This study aims to analyze changes in the wage premium for university education before and after this upsurge in the number of university graduates and understand the structure of the Turkish labor market (TLM) by examining the factors affecting the wage gap across levels of different education.

Between 1990 and 2020, the TLM experienced a significant structural change resulting from substantial demographic and social transitions. Life expectancy, mean age, and the elderly population all increased remarkably, and the average education level improved as well. In the same period, illiteracy rates fell drastically, while the share of workers with a university degree increased remarkably. The total number of students enrolled in universities in Turkey grew to reach 8.24 million in the academic year of 2020–2021 (including vocational schools and Open University students) according to Council of Higher Education statistics; in other words, 9.85 out of every 100 people in Turkey are university students. The employment rate for workers with university degrees rose from 16% to 23% for females and from 17% to 27% for males.

Nonetheless, the quality of both secondary and university education is highly problematic. According to the Program for International Student Assessment (PISA), which measures an education system's success according to certain educational outcomes, Turkey ranked 50th among 72 countries in 2015. In Turkey, about three out of every five students fail to meet international baseline educational standards in reading, mathematics, and science. The Adult Skills Measurement Program (ASMP), which assesses at the international level the problem-solving skills of adults between the ages of 16 and 65, indicated that Turkey was among the lowest ranking Organization for Economic Co-operation and Development (OECD) countries in terms of reading, digital literacy, and technology in 2015. This indicates that Turkish labor market participants' soft skills fall far below the average OECD level.

In the last three decades, agricultural dominance in the Turkish labor force dropped, whereas the number of wage earners more than doubled. The female participation rate, the average education level of the labor force, and university graduates' share of employment augmented due to this drastic socio-economic

[1] The main motivation for the newly established state universities was related to underdeveloped regions' economic problems rather than (and/or alongside) educational demand. For further information, see Polat (2017).

transformation. Moreover, the real minimum wage doubled, demonstrating a secular trend over the same period.

Gaining a better understanding of the TLM's structural changes requires a detailed long-term analysis of the wage premium changes for university graduates. The wage premium of university graduates is closely related to the labor market's supply and demand dynamics. If the increase in demand for university graduates lags behind the increase in supply, a significant change in their wage premium is unlikely to be observed. A recent World Bank report (Acar & Del Carpio, 2019) pointed to the decline in university graduates' wage premium in the TLM after 2012 and related it to the disequilibrium in the labor market due to the substantial increase in better-educated workers. Eren (2018) also indicated that the decline in the wage premium for university graduates is highly related to the relative excess supply of university graduates. Using Household Labor Force Surveys released by the Turkish Statistical Institute (TURKSTAT) between 2004 and 2019, we aim to examine the relationship between the changes in the wage premium for university education and the structural changes in the labor market. In the descriptive part of our study, we observed a sharp increase in the number of university students after 2008, reflecting the so-called a university in every city policy launched in Turkey in 2006. The number of students in Turkish universities has quadrupled in the last two decades. Following this dramatic increase, the ratio of skilled[2] workers (workers with tertiary education) to unskilled workers (workers with high school, secondary school, or primary school education, and those who are literate but uneducated) in employment also increased sharply in this period. However, an analysis of the wage changes for workers with different education levels in the last 15 years indicates that tertiary-educated skilled workers' wages grew the slowest among workers from all education levels. Besides, there has been a substantial increase in the minimum real wage over the last 15 years, particularly in 2015. All of these changes have resulted in a substantial decrease in skilled workers' wage premium in the last decade.

This paper aims to understand the reasons for the changes in the wage gap between skilled and unskilled workers in the Turkish Labor Market (TLM). Using Autor et al.'s (2008) conceptual two-factor constant elasticity of substitution (CES) production function model, we try to explain the effects of market-driven changes in supply and demand on wage differentials in the TLM. In other words, we discuss whether the TLM's skill wage premium changes can be explained using Goldin and Katz's (2008) framework in *The Race between Technology and Education*.

[2] In this context, "skill" indicates holding a degree and does not necessarily reflect ability. Here, it is important to mention that the Turkish secondary and tertiary education system is not entirely capable of giving students the skills (such as artificial intelligence, big data, and automation) they need to succeed in the twenty-first century.

2 Literature Review

There is a sizeable body of literature analyzing wage differentials from different perspectives. The rising wage gap in the U.S. economy in particular has attracted the attention of many economists. Most of the studies in this area have tried to understand the dynamics of supply and demand shifts for skills and their impact on the rising inequality in wages. Katz and Murphy (1992), Goldin and Katz (2008), and Autor et al. (2008) have published the leading studies modeling the wage differentials using the conceptual framework of the race between the supply and demand for skills. According to their model, technological developments increase the demand for skills and create incentives for human capital investments. If the demand for skills rises faster than the supply, wage differentials arise. Autor et al. (2008) showed that Katz and Murphy's (1992) model, which used a CES production function with two types of factors—university equivalent and high school equivalent—does an excellent job of explaining the growth in the skill wage premium in the U.S. economy for the 1963–1992 period. After 1992, the relative supply growth slowed. Thus, the model overpredicts the skill wage premium. Autor et al. (2008) pointed out that the simple CES model with two factors is insufficient to explain the labor market changes that started in the early 1990s. They showed that significant wage differentials existed within education groups and that experience also played an important role in the wage differentials.

Other studies have concluded that wage inequality in the U.S. economy after the 1980s was primarily based on the erosion of the minimum wage (e.g., Card & DiNardo, 2002; Lee, 1999; Lemieux, 2006). Acemoglu and Autor (2012) indicated that Goldin and Katz's race between technology and education framework successfully explained the distribution of wages in the United States based on the supply and demand factors for human capital in the twentieth century. They also pointed out that this model's shortcomings in terms of explaining the wage differentials in the last two decades could be resolved by including a richer set of interactions between skills and job tasks. Autor et al. (2020) extended Goldin and Katz's (2008) model estimations to the nineteenth and twenty-first centuries. They showed that although educational wage differentials explained 75% of wage inequality between 1980 and 2000, only 38% of the wage inequality could be explained by educational wage differentials between 2000 and 2017. They also specified that an important part of the wage inequality came from the differentials within rather than between education groups.

Few studies have focused on wage inequality in the TLM. Bakış and Polat (2015) analyzed the evolution of wage inequality in Turkey in two sub-periods (2002–2004 and 2004–2010) using Household Labor Force Survey (HLFS) data between 2002 and 2010. Their principal findings showed that the relative supply of highly educated employees to less educated ones remained constant in the first period but relative wages improved in favor of less educated workers. However, in the second period, relative supply increased, and relative wages stayed stable or continued to rise, favoring more educated workers. Their results showed that a sharply increasing minimum wage and significant institutional changes create an asymmetrical welfare improvement (asymmetrical in terms of there being relatively more considerable progress for the lower-level wage deciles). Eren (2018) also showed that there was a significant decline in the wage premium of university graduates in the period from 2004 to 2015. In order to capture the effects of demand- and supply-side dynamics on the wage premium, Eren (2018) used a relative excess supply measure, which is simply the ratio of the number of unemployed university graduates to the number of unemployed high school graduates.

In this paper, using the canonical two-skills model developed by Katz and Murphy (1992), Goldin and Katz (2008), and Autor et al. (2008), we aim to understand the effects of skill supply-driven and skill demand-driven forces on the skill wage premium in the TLM. In contrast with the studies that have been carried out for the U.S. labor market, our results show that the wage gap in the TLM between skilled and unskilled workers has been declining in the last two decades and that the primary reason for this decline is supply-related factors.

Another important finding in our study indicates that the skill demand in the TLM did not significantly change between 2004 and 2019. This result may also point to there being no significant skill-biased technological change (SBTC) in this period. The estimation results show that a 10% increase in the relative supply of skilled/unskilled labor equivalent only decreases the skill wage premium by 2.7% ($1/\sigma = 0.27$). According to this result, the elasticity of substitution between skilled and unskilled workers in the TLM is relatively high compared to the U.S. economy. It is much easier to substitute unskilled workers with skilled ones in Turkey; thus, the changes in the relative supply of skilled/unskilled labor equivalent have less of an impact on the skill wage premium than in the USA. It is also possible to say that the rapid surge in university graduates in the last decade pushed an important segment of the unskilled workers out of the labor market. One can conclude from here that a significant proportion of university graduates work in jobs that only require a high school education.

Our estimation results also show that the real minimum wage is another essential determinant in the decline in the skill wage premium. This result is an important indicator showing that Goldin and Katz's hypothesized race between technology and education is not valid in the TLM. Human capital investments in Turkey are not driven by demand-side factors. The upsurge in the real minimum wage, the rapid increase in university graduates, and the insufficient demand increase for university graduates are the most likely sources of the decline in the skill wage premium.

3 Data and Methodology

3.1 Summary Statistics and Sample Selection

We use cross-sectional data from the HLFS collected between 2004 and 2019. The surveys are representative for the entire country because they cover all geographical regions. Table 1 presents summary statistics for some demographics and labor force indicators for the entire population in 2004, 2012, and 2019.

Table 1 Demographics and labor force indicators

	Female			Male		
	2004	2012	2019	2004	2012	2019
Demographics						
Age						
15–24	28%	23%	21%	27%	23%	22%
25–34	24%	23%	20%	23%	22%	20%
35–44	22%	22%	22%	22%	21%	22%
45–54	16%	19%	20%	17%	19%	20%
55–64	10%	14%	17%	10%	14%	16%
Education level						
Below high school	82%	77%	72%	76%	71%	62%
High school	13%	14%	15%	17%	17%	21%
University	5%	9%	13%	7%	12%	17%
Labor force						
Labor force participation						
Below university	84%	75%	72%	90%	83%	79%
University	16%	25%	28%	10%	17%	21%
Unemployment						
Below university	77%	70%	66%	91%	87%	83%
University	23%	30%	34%	9%	13%	17%
Employment						
Below university	85%	75%	73%	90%	83%	79%
University	15%	25%	27%	10%	17%	21%
Firm size						
Less than 10	72%	63%	61%	64%	57%	57%
10–49	13%	18%	16%	16%	21%	17%
More than 50	15%	19%	23%	20%	22%	26%
Labor force						
Labor force participation rate						
Below university	23%	26%	28%	71%	70%	68%
University	70%	69%	70%	84%	83%	84%
Unemployment rate						
Non-agricultural						
Youth (15–24)	20%	20%	30%	21%	16%	22%
Below university	11%	11%	14%	11%	9%	12%
University	17%	14%	18%	9%	6%	9%
Employment rate						
Below university	20%	23%	24%	64%	64%	60%
University	58%	60%	57%	76%	78%	76%

Source: Authors' calculations based on the HLFS

Turkey is experiencing a rapid social transition, and the demographic results of this substantial societal transformation continue to affect the TLM significantly. Table 1 shows the basic labor market statistics for females and males in 2004, 2012, and 2019. The share of young people (15–24) decreased by approximately 20% (from 52% to 41% and from 50% to 42% in the female and male populations, respectively). However, the number of university graduates exploded (increased around 2.5 times) during this period, including females and males.

The female labor force participation rate is closely related to education levels, particularly to university graduation, because of relatively high reservation wages for women. Drastically increasing their education level stimulated female participation, with the non-agricultural female labor force participation rate doubling over two decades from 16% to around 30%. The male participation rate was relatively stable at approximately 70% throughout the entire period. New job creation has been stimulated mainly by the growth performance of the economy. The growth elasticity of employment in Turkey is around 0.5 (Gürsel et al., 2015), which means that if the economy grows by 10%, the expectation of new job creation will be 5%.

In Turkey, the mean yearly growth rate was relatively high in the first decade of the 2000s; however, following a significant deterioration in the second decade, the unemployment rate jumped to reach double digits toward the end of the period. The composition of the jobless workforce also changed significantly during the period; due to the upsurge in the number of university graduates, their share in employment and unemployment increased remarkably. These shares continue to rise.

The study sample is restricted to wage earners aged between 25 and 64 years old with 0–39 years of potential experience who were working in a non-agriculture private sector at the time of the survey. Table 2 summarizes the basic employment statistics for females and males in the sample in 2004, 2012, and 2019. Reflecting changes in the entire population, the share of workers with university degrees in our sample increased remarkably throughout the period, both for females and males. The employment rate for workers with university degrees rose from 16% to 23% for females and from 17% to 27% for males. As expected, the share of workers with university degrees across non-registered employment is very tiny compared to the share of workers with an education level below a university degree.

Another striking observation in the TLM in the 2000s is the substantial increase in the real minimum wage. The minimum wage-earning worker's share in the total wage earners is over 40%, and Turkey ranks high among OECD countries with this share. A more than doubling of the real minimum wage caused an expansion of minimum wage earners in the bottom deciles of the real hourly wage distribution. Industry oversensitivity to wage costs (resulting from many different structural problems) is the fundamental reason for this expansion. Industry's inability to pay higher salaries creates an accumulation at the bottom rather than pushing all wages up.

Table 2 Employment statistics

	Female			Male		
	2004	2012	2019	2004	2012	2019
Employment						
Below university	85%	74%	68%	94%	87%	81%
University	15%	26%	32%	6%	13%	19%
Non-registered employment						
Below university	95%	94%	92%	98%	95%	93%
University	5%	6%	8%	2%	5%	7%
Occupation						
High skilled	19%	22%	21%	12%	15%	17%
Skilled	64%	59%	63%	74%	69%	70%
Low skilled	17%	19%	16%	14%	16%	13%
Job tenure in years						
Below university	4	3	4	6	4	5
University	4	4	4	5	4	5
Employment rate						
Below university	7%	9%	9%	28%	34%	27%
University	16%	23%	23%	17%	27%	27%
Non-registered employment rate						
Below university	3%	3%	2%	12%	8%	5%
University	2%	1%	1%	3%	2%	2%

Source: Authors' calculations based on the HLFS

3.2 Composition-Adjusted Wage and Relative Supply of Skilled/Unskilled Labor Equivalents Series

The most critical steps in estimating the effect of the relative labor supply on the wage gap are calculating composition-adjusted wage and equivalent supply series. Composition-adjusted wage series correspond to average wages within specified groups holding the composition effect fixed by eliminating the impact of primary factors on wages in each group. Equivalent supply series represent the labor supply augmented by efficiency wages. The steps to calculate composition-adjusted wages and the relative supply of skilled/unskilled labor equivalent series from 2004 to 2019 are adapted from Autor et al. (2008). Considering the characteristics of our original data, these steps are described below.

Composition-Adjusted Wage Series

The surveys provide information on monthly earnings, and by making use of this information, hourly wages are estimated using Eq. (1)

$$hw = \frac{me * 12}{hpw * 52}, \tag{1}$$

where *hw*, *me*, and *hpw* represent the hourly wages, monthly earnings, and usual hours worked per week in the primary job, respectively. Then, hourly wages are converted to 2004 constant prices using suitable gross domestic product deflators.

Below-Mincerian (see Mincer, 1974) wage function is estimated with suitable weights in each year using Ordinary Least Squares (OLS) to control the effects of gender, education level, years of experience, tenure, full-time/part-time division, and region on log real hourly wages:

$$\begin{aligned}\log(w_i) = {} & \beta_0 + \beta_1 male_i + \beta_2 ps_i + \beta_3 hs_i + \beta_4 uni_i + \beta_5 \log\left(\exp_i\right) \\ & + \beta_6 \log\left(\exp_i\right)^2 + \beta_7\left[ps_i * \log\left(\exp_i\right)\right] + \beta_8\left[hs_i * \log\left(\exp_i\right)\right] \\ & + \beta_9\left[uni_i * \log\left(\exp_i\right)\right] + \beta_{10}\log(ten_i) + \beta_{11} ft_i + \beta_{12} reg_i + u_i,\end{aligned} \tag{2}$$

where log(*w*) is the logarithm of real hourly wage, *male* is a dummy variable that takes the value of 1 if the individual is male and 0 otherwise, *ps* is a dummy variable that takes the value of 1 if the individual graduated from primary school and 0 otherwise, *hs* a is a dummy variable that takes the value of 1 if the individual graduated from either secondary school or high school and 0 otherwise, *uni* a is a dummy variable that takes the value of 1 if the individual graduated from university and 0 otherwise, log(*exp*) is the logarithm of the years of potential experience of the worker (calculated as $age - cumulative\ years\ of\ schooling - 6$), log(*ten*) is the logarithm of the years spent in the worker's current job, *ft* is a dummy variable that takes the value of 1 if the worker is employed in a full-time job (full-time workers usually work between 35 and 60 h each week) and 0 otherwise, and *reg* is a categorical regional variable consisting of 12 regions.[3]

Using the coefficients obtained from the above regression, predicted log real hourly wage series are calculated as follows:

$$\begin{aligned}\widehat{\log(w_i)} = {} & b_0 + b_1 male_i + b_2 ps_i + b_3 hs_i + b_4 uni_i + b_5 \log\left(exp\_gr_i\right) \\ & + b_6 \log\left(exp\_gr_i\right)^2 + b_7\left[ps_i * \log\left(exp_i\right)\right] + b_8\left[hs_i * \log\left(exp_i\right)\right] \\ & + b_9\left[uni_i * \log\left(exp_i\right)\right] + b_{10}\log(av\_ten_i) + b_{11} ft_i,\end{aligned} \tag{3}$$

where *exp_gr* is a categorical variable with four categories (0–9, 10–19, 20–29, and 30–39 years) and *av_ten* is the sample average of the tenure variable within a specific year.

Composition-adjusted log real hourly wages are calculated by sorting data into gender-education-experience groups. This sorting results in 32 groups based on

[3] The 12 geographical regions are taken from TURKSTAT's Nomenclature of Territorial Units for Statistics-1 (NUTS1) classification system and are as follows: İstanbul, West Marmara, Aegean, East Marmara, West Anatolia, Mediterranean, Central Anatolia, West Black Sea, East Black Sea, Northeast Anatolia, Central East Anatolia, and Southeast Anatolia.

two genders, four education levels, and four experience categories. Annual log real hourly wages are the average of these composition-adjusted log real hourly wages in each year, weighted by the usual number of hours worked per week.

Figure 1 presents the trends in real log hourly wages in our employment sample and composition-adjusted real log hourly wages at different education levels. It explains wage inequality between education levels. As shown in Panel A, the change in the real log hourly wage for university graduates increases until 2010 and then follows a more stable pattern. The wage gap between university graduates and graduates with a degree below the university level decreases monotonically, especially after 2010. The change in composition-adjusted wages shown in Panel B helps us identify the ceteris paribus effect of education level on the real log hourly wage. As can readily be observed, when we look at the ceteris paribus effect of education, the increase in the log hourly wage for university graduates up to 2010 also vanishes, and the increase in the composition-adjusted wage of university graduates is smallest across all education levels. On the other hand, the increases in all other education levels become more pronounced, especially after 2011. These changes lead to a decrease in the wage gap between university graduates and with less educated graduates.

As previously stated, there is a drastic increase in the real minimum wage in the TLM in the 2000s. This important change can certainly explain the decreasing wage gap between skilled and unskilled workers.

Equivalent Supply Series

Other important variables in our analysis are the series of skilled and unskilled labor supply. Educational attainment serves as an indirect measure of an individual's skill level in the absence of direct measures of skill supply, assuming that qualifications can be acquired through education. Despite the fact that formal education alone is not a complete measure of skill, the availability of educational attainment information in many datasets makes it a widely used measure of skill level. Especially in the earnings inequality literature, skilled and unskilled workers are determined such that their acquired education level is compared with a chosen threshold education level. Workers with an education level above (below) that threshold education level are classified as skilled (unskilled) workers. Juhn et al. (1993), Goldin and Katz (2008), Acemoglu and Autor (2012), and Balakumar and Chattopadhyay (2020) are some examples. In our classification, skilled workers comprise workers with a university degree, and unskilled workers include workers with a degree below university education.

Following Autor et al. (2008), we converted raw supply series to supplies calculated in efficiency units to adjust for labor force composition changes by gender, education level, and experience groups. Similar to composition-adjusted wage calculation, we generated 32 groups based on two genders, four education levels, and four experience categories. For each of these 32 groups, the real hourly wages are normalized as dividing the group average wage by the base wage, which is the weighted predicted real hourly wage of skilled males each year. Efficiency wages are calculated as the average of the normalized wages for all years and for

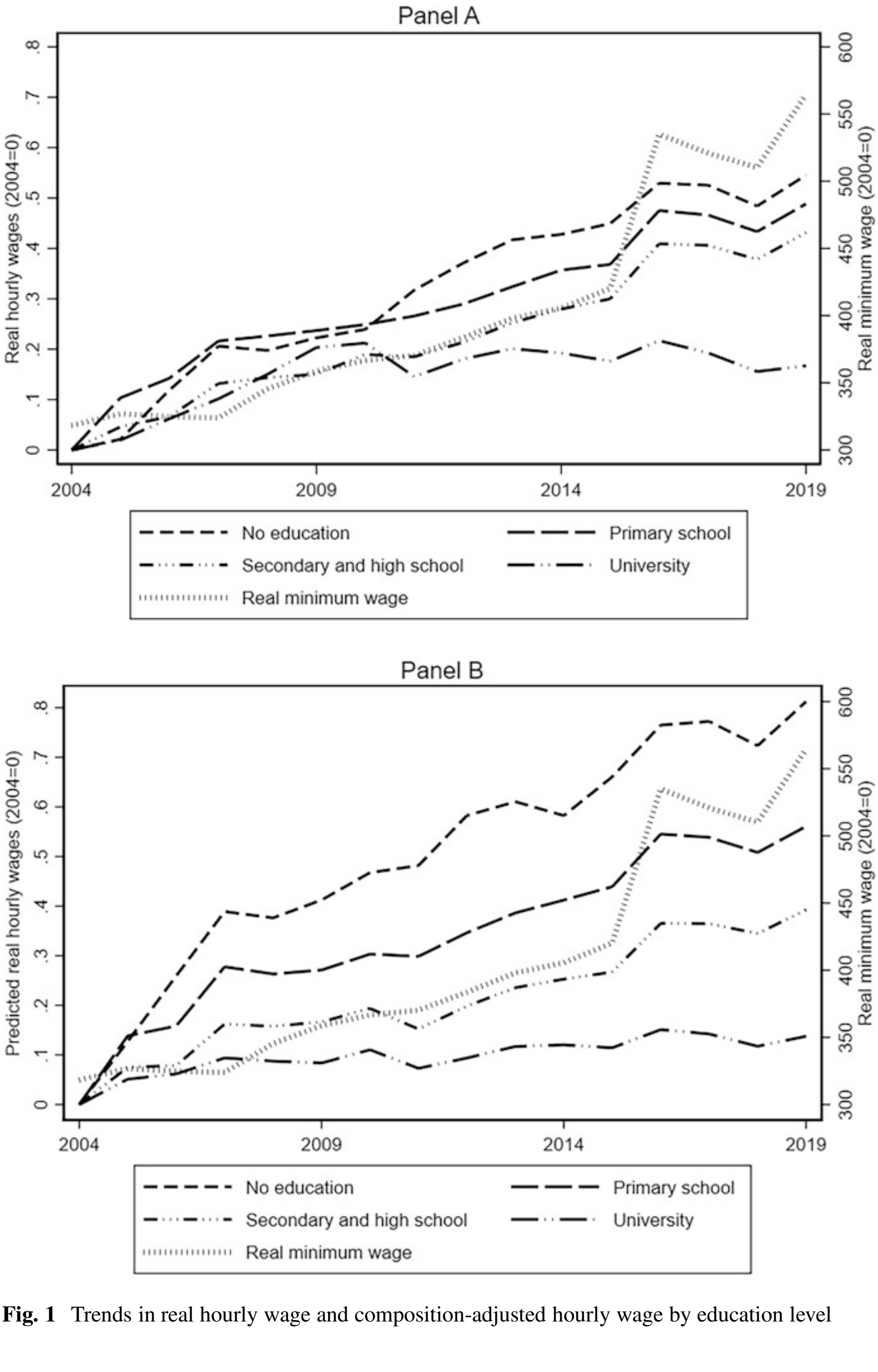


Fig. 1 Trends in real hourly wage and composition-adjusted hourly wage by education level

each skill level. Then, efficiency labor supply is calculated as the multiplication of the efficiency wages and the total number of workers in that group. The annual aggregate skilled-labor supply equivalent is calculated as the total efficiency units of labor supplied by workers with a university degree in a particular year. Similarly, the annual aggregate unskilled labor supply equivalent is calculated as the total efficiency units of labor supplied by workers with a degree below university in a particular year.

Skill Wage Premium and Relative Supply of Skilled/Unskilled Labor Equivalents

Using the two series described above, skill wage premium and relative supply of skilled/unskilled labor equivalents are calculated. The annual relative supply of skilled/unskilled labor equivalent is calculated based on the logarithm of the proportion of the skilled-equivalent labor supply to the unskilled-equivalent labor supply. Skill wage premium is the logarithm of the proportion of real hourly composition-adjusted wages of skilled workers to the real hourly composition-adjusted wages of unskilled workers.

Figure 2 presents the evolution of the skill wage premium and the relative supply of skilled/unskilled labor equivalents. As this figure shows, the skill wage premium is decreasing while the relative supply of skilled/unskilled labor equivalents accompanies the skill wage premium with a gradual increase, generating a negative relationship between the two over the entire period. This observation is

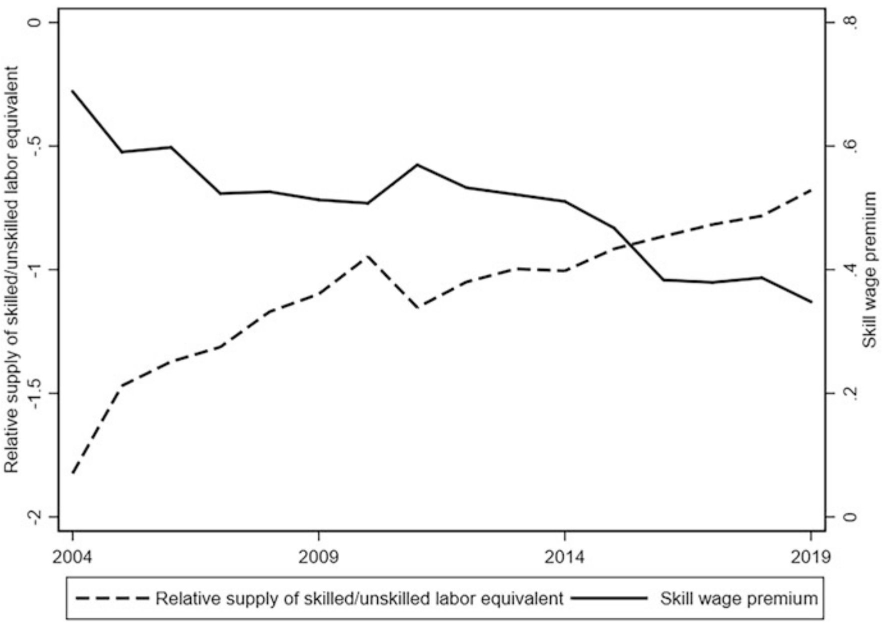


Fig. 2 Relative supply of skilled/unskilled labor equivalents and skill wage premium

quite interesting because many of the studies mentioned above have claimed that there is a positive relationship between these two series in developed countries due to the relatively faster increase in skill demand due to SBTC. One exception is the work of Blundell et al. (2021), which indicated an increasing trend in the number of people with a university degree while the university wage premium remained relatively flat. Figure 2 implies that the story is different in the TLM. According to a recent World Bank report (Acar & Del Carpio, 2019), one potential explanation of this negative relationship is the pressure that rapid growth in the relative supply of better-educated workers places on earnings premiums.

4 Conceptual Framework: The Race Between Education and Technology

Following Tinbergen's (1974) leading study, skill demand increases with the increasing skill supply. The relative demand for skills is related to the SBTC. The canonical model used in the literature has successfully explained this race between education and technology (see Autor et al., 2008; Katz & Murphy, 1992; Goldin & Katz, 2008). Technological changes are assumed to take a factor-augmenting form in the canonical model and constitute two imperfectly substitutable skill groups: skilled and unskilled. The production function for the aggregate economy postulates a CES, as given below:

$$Q_t = \left[\alpha_t (a_t S_t)^\rho + (1-\alpha_t)(b_t U_t)^\rho\right]^{1/\rho}, \tag{4}$$

where Q_t is the aggregate output, S_t and U_t are the quantities of skilled and unskilled labor in period t, a_t and b_t represent the labor-augmenting technological change at time t, α_t is a distribution factor determining the relative importance of skilled labor in the production process, and ρ is the production parameter which can be written as a function of the elasticity of substitution (σ) between skilled and unskilled labor such that $\sigma = \frac{1}{1-\rho}$. The main reason for an increase in the ratio of a_t/b_t or α_t is the SBTC, which results in an increase in skill demand. Following the assumption that skilled and unskilled labor wages are equal to their marginal products, one can easily derive the following relationship between relative wages and relative skill supply from Eq. (4).

$$\frac{w_{S_t}}{w_{U_t}} = \frac{\alpha_t}{1-\alpha_t}\left(\frac{a_t}{b_t}\right)^{\frac{\sigma-1}{\sigma}}\left(\frac{S_t}{U_t}\right)^{-\frac{1}{\sigma}}, \tag{5}$$

Taking the logs of both sides, we obtain:

$$\ln\left(\frac{w_{S_t}}{w_{U_t}}\right) = \frac{1}{\sigma}\left[D_t - \ln\left(\frac{S_t}{U_t}\right)\right], \tag{6}$$

where $D_t = \sigma\left[\ln\left(\frac{\alpha_t}{1-\alpha_t}\right)\right] + (\sigma - 1)\ln\left(\frac{a_t}{b_t}\right)$. Here, D_t represents the relative skill demand, which depends on the SBTC parameters and the relative importance parameters of skill. As Autor et al. (2020) stated, the terms in brackets in Eq. (3) show how the demand and supply factors for skill affect the skill wage premium. It must be pointed out here that the aggregate elasticity of the substitution between skilled and unskilled labor is the main determinant of how relative skill supply and skill demand parameters affect the skill wage premium. As indicated by Acemoglu and Autor (2012), if there is no SBTC and the relative importance of skill does not change over time (i.e., a_t/b_t and α_t is constant), the increase in skill supply leads to a decline in the skill wage premium. In the literature, it has generally been assumed that there is a log-linear trend increase in the demand for skills, depending on the technological changes. Starting with Katz and Murphy (1992), a simple time trend variable was used to capture these demand changes. In addition to a simple time trend, Autor et al. (2008) also used the unemployment rate of males between the ages of 25 and 54 to capture unobserved demand shifts. They also included the log real minimum wage as a control variable in their model.

In the following sections, we also estimate a version of Eq. (6) using Autor et al.'s (2008) specifications.

5 Estimation and Findings

Based on the conceptual relationship given in Eq. (6), different specifications of the below model are estimated using an OLS methodology:

$$\log\left(swp_t\right) = \beta_0 + \beta_1 \log\left(rssue_t\right) + \beta_2 mw_t + \beta_3 u_t + \beta_4 \text{time} + \varepsilon_t, \tag{7}$$

where *swp* is the skill wage premium, *rssue* is the relative supply of skilled/unskilled labor equivalent, *mw* is real minimum wages, *u* is the unemployment rate of individuals aged between 15 and 64 years old, *time* is time-fixed effects, and *t* represents time periods. The real minimum wage and unemployment rate are incorporated to control the demand- and supply-side factors.

Table 3 presents Katz and Murphy's (1992) model estimation results with different control variables. The first column shows the simple regression results of the skill wage premium, including the only relative supply of skilled/unskilled labor equivalent. We found a relatively high substitution effect (the absolute value of the coefficient of *rssue* is $1/\sigma = 0.27$, and thus $\sigma = 3.7$) when compared with developed countries. This result indicates that it is relatively easier to substitute unskilled workers with skilled ones in the TLM. In other words, the impact of the shifts in the relative skill supply has smaller effects on relative wages. For example, in Turkey, a 10% increase in the relative skill supply only causes a 2.7% decline in the skill wage premium; however, according to Autor et al. (2008), this causes a 6% decline in the U.S. economy.

Table 3 Regression models for the skill wage premium

	(1)	(2)	(3)	(4)	(5)
Relative supply of skilled/unskilled labor equivalent	−0.27*** (0.02)	−0.24*** (0.08)	−0.18*** (0.05)	−0.24*** (0.03)	−0.17*** (0.05)
Year		−0.00 (0.01)			
Real minimum wage			−0.17** (0.08)		−0.15* (0.08)
Unemployment rate				−0.01 (0.01)	−0.01 (0.01)
Constant	0.13*** (0.03)	3.95 (10.6)	1.26** (0.51)	0.25** (0.09)	1.22** (0.5)
Observations	16	16	16	16	16
R^2	0.891	0.892	0.921	0.908	0.931

Standard errors in parentheses
*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

The second column shows the estimation results, including an additional time trend variable to capture the skill demand shifts originating from SBTC in time. We also know from previous studies that there is a log-linear trend increase in skill demand that originates from technological changes in developed countries (see examples in Acemoglu & Autor, 2012; Autor et al., 2008). This trend increase captures the unobserved skill demand shifts in the U.S. economy, which D_t denotes in the conceptual framework [see Eq. (6)]. As demonstrated in Column 2, the linear trend coefficient is insignificant in the TLM. This result may point to the lack of SBTC in time, and thus it can be considered evidence of an insufficient increase in the demand for skills in the TLM during our sample period. A closer look at the data explains why this insignificance may have occurred. Panel A of Fig. 1 displays stagnation beginning in 2010 for university graduates' real hourly wages in contrast with the secular increase during the entire period in the real wages of workers with lower education levels. We know from Fig. 2 that there is a negative relationship between the skill wage premium and the relative supply of skilled/unskilled labor equivalent. As Acemoglu and Autor (2012) also stated, this negative relationship implies that technology, more specifically a_t/b_t, remains constant over time in this period in the TLM. In other words, if there is an improvement in skill augmenting technological change (which corresponds to an increase in a_t/b_t in the above conceptual framework), the skill premium must increase due to the higher demand for skills. However, the lack of such an increase in the TLM indicates that there was no significant technological change favoring skills—or that the current simple model and gross data are insufficient to explain the complicated labor market dynamics.

In Column 3 and Column 4, we added real minimum wages and unemployment rate variables to better control labor market cyclical conditions in Turkey. Lastly, we estimated the full model in Column 5. The unemployment rate coefficient is insignificant and surprisingly indicates that unemployment does not affect the wage gap.

However, the inclusion of the real minimum wage as an explanatory variable diminishes the effect of relative supply of skilled/unskilled labor equivalent on wage gap. The coefficients of real minimum wage and relative supply of skilled/unskilled labor equivalent slightly differentiate between model (3) and the full model (5). We observe from Column 3 that the increase in the real minimum wage can explain a portion of the decline in the skill wage premium. The rise in the relative supply of skilled/unskilled labor equivalent still significantly contributes to this decline, but the effect is less pronounced. In other words, the inverse elasticity of substitution is smaller: a 10% increase in the relative supply of skilled/unskilled labor equivalent reduces the skill wage premium by 1.8%. These results show that there is not currently a race between technology and education in the TLM. The upsurge in the number of university graduates is not demand driven, as suggested by Goldin and Katz's model. Thus, in contrast to the observations made for the U.S. labor market, the skill wage premium has been a declining trend in the TLM for the last two decades.

6 Discussion and Conclusions

As a result of the substantial increase in the number of universities established in Turkey in the last two decades, university enrolment has gone from 1.5 million to 7.9 million students. As Polat (2017) stated, this increase has not been directly related to the needs and capacity of the Turkish economy. A critical economic issue that needs to be clarified is the economy's potential to create sufficient employment growth to accommodate Turkish university graduates. Using Autor et al.'s (2008) conceptual two-factor CES production function model, we try to decompose the effects of market-driven changes in supply and demand on wage differentials in the TLM.

In developed countries, increases in the supply of skilled labor generally occur due to skill-favoring technological changes. A substantial increase in skill demand creates a skill supply surge and causes an increase in the skill wage premium. Previous studies using data for developed countries confirmed the simple canonical model's explanatory power, which analyzes this relationship between technology and education. In contrast to these studies, we observe that as the skilled labor share has increased in the workforce, there has been a substantial decline in the skill wage premium in Turkey over the last two decades. This negative relationship implies that there has been no significant skill-favoring technological change in this period in the TLM. The lack of sufficient skill demand is the culprit behind the decline in the skill wage premium. This result indicates that the Turkish economy does not yet have enough potential to create jobs to accommodate the ever-increasing number of university graduates it is producing. In other words, the increase in the skilled labor supply in Turkey is not driven by demand-side factors, which explains why Goldin and Katz's race between technology and education hypothesis is not valid in the TLM.

One of the extraordinary characteristics of the TLM is its high share of minimum wage earners. This share has risen over the last two decades and is now at 40%, which is exceptionally high compared to the rest of the OECD countries. Our estimation results show that the real minimum wage is also a significant variable reducing the skill wage premium. The drastic increase in the real minimum wage in the 2000s was also an important economic factor reducing industrial capability to pay high wages for skilled labor. As a result, a crowding-out effect occurred in the labor market. The less educated labor force started to be replaced by university graduates, and the rapid expansion in the pool of university graduates simultaneously deteriorated the mean skill level of educated laborers.

In conclusion, based on the quality of Turkish university education, a hierarchy has implicitly appeared in the job application process. Most low-skilled jobs are now being taken by lower-quality university graduates. This educational mismatch is also likely to be responsible for the wage stagnation of university graduates. Whether Turkey's higher education system has sufficient capacity to produce skilled laborers with the necessary skills is another critical question that still needs to be answered. In other words, Turkey needs to identify whether it faces a skills shortage or a skills surplus. Further research using microdata, including information about the labor force's various skills levels, is undoubtedly necessary to shed light on these questions.

References

Acar, A., & Del Carpio, X. (2019). *Turkey jobs diagnostic*. World Bank. License: Creative Commons Attribution CC BY 3.0 IGO.

Acemoglu, D., & Autor, D. (2012). What does human capital do? A review of Goldin and Katz's the race between education and technology. *Journal of Economic Literature, 50*(2), 426–463.

Autor, D., Goldin, C., & Katz, L. F. (2020). *Extending the race between education and technology*. NBER Working Paper, No.26705.

Autor, D., Katz, L. F., & Kearny. (2008). Trends in U.S. wage inequality: Revising the revisionists. *The Review of Economics and Statistics, 90*(2), 300–323.

Bakış, O., & Polat, S. (2015). Wage inequality in Turkey, 2002–2010. *Economics of Transition, 23*(1), 169–212.

Balakumar, S., & Chattopadhyay, S. (2020). Skilled-unskilled wage gap without discrimination-the case of India. *Applied Economics Letters*. https://doi.org/10.1080/13504851.2020.1851644

Blundell, R., Green, D., & Wenchao, J. (2021). The UK as a technological follower: Higher education expansion and the college wage premium. *The Review of Economic Studies.* Forthcoming.https://www.ucl.ac.uk/~uctp39a/BGJ_MS24921manuscript.pdf

Card, D. E., & DiNardo, J. E. (2002). Skill-biased technological change and rising wage inequality: Some problems and puzzles. *Journal of Labor Economics, 20*(4), 733–783.

Eren, O. (2018). *Evolution of the university wage premium in Turkey 2004-2015*. The Central Bank of the Republic of Turkey, Working Paper, No: 18/08.

Goldin, C., & Katz, L. F. (2008). *The race between education and technology*. Belknap for Harvard University Press.

Gürsel, S., Uysal, G., & Durmaz, M. (2015). *Düşük büyüme olağanüstü istihdam artışı*. BETAM, Bahçeşehir University Research Brief No.184.

Juhn, C., Murphy, K. M., & Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy, 101*, 410–442.

Katz, L. F., & Murphy, K. M. (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *Quarterly Journal of Economics, 107*, 35–78.

Lee, D. S. (1999). Wage inequality in the U.S. during the 1980s: Rising dispersion or falling minimum wage? *Quarterly Journal of Economics, 114*(3), 977–1023.

Lemieux, T. (2006). Increased residual wage inequality: Composition effects, noisy data, or rising demand for skill. *American Economic Review, 96*(3), 461–498.

Mincer, J. (1974). *Schooling, experience, and earnings*. Columbia University Press.

Polat, S. (2017). The expansion of higher education in Turkey: Access, equality and regional returns to education. *Structural Change and Economic Dynamics, 43*, 1–14.

Tinbergen, J. (1974). Substitution of graduate by other labour. *Kyklos, 27*(2), 217–226.

Turkish Statistical Institute (TURKSTAT). (2004–2019). *Household labor force survey*. Turkish Statistical Institute.

The Status Granted to Organizational Identity Characteristics: An Application of the Best-Worst Method and Regression Analysis

Başak Topaler and Nur Ayvaz Çavdaroğlu

Abstract According to the traditional conceptualization in management literature, organizations gain status through demonstrations of superior performance. Criticizing this narrow focus on "achieved" status, this study identifies the status value ascribed to organizations within socially constructed systems of norms and values. Utilizing concepts offered by institutional research in organization theory, we propose historical legacy, endorsement, and prominence in the field as sources of ascribed status for organizational identity characteristics. The Turkish higher education field constitutes the empirical setting of our research. We conducted interviews with key stakeholders in this field (i.e., students, academicians, managers in industry, and high school counselors) to elicit the organizational characteristics they perceive as relevant for defining university identity. The status value ascribed to these identity characteristics was documented via surveys with separate samples of these four stakeholder groups using a relatively new but widely recognized MCDM method, the best-worst method (BWM). Our findings provide evidence for the view that institutions have a broad influence on social hierarchy systems, yet with some nuances. We discuss theoretical implications for the research on status in markets and practical implications for higher education institutions.

Keywords Status · Institutional theory · Higher education field · Multi-criteria decision-making · Best-worst method

Authors are equal contributors; their names are listed in alphabetical order.

B. Topaler (✉) · N. A. Çavdaroğlu
Kadir Has University Department of Business Administration, Cibali Fatih İstanbul,, Turkey
e-mail: basak.topaler@khas.edu.tr; nur.cavdaroglu@khas.edu.tr

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_13

1 Introduction

Status hierarchies form in almost any social setting, where actors occupying higher-ranking positions are afforded greater prestige and deference (Goode, 1978; Sauder et al., 2012; Weber, 1978). Research has extensively addressed the benefits of high status in various contexts (Chae et al., 2020; Chen et al., 2012; Piazza & Castellucci, 2014; Sauder et al., 2012). Given important benefits of high status in the context of markets, such as improved access to valuable resources and lower transaction costs, organizations have a strong motivation to advance their status ranking and actively respond in case they experience a loss of status (Askin & Bothner, 2016; Breznik & Law, 2019; Chen et al., 2012; Rider & Negro, 2015; Sauder & Espeland, 2009). Yet, we still know little about the origins of status hierarchies and the underlying bases of status differences among organizations (Pearce, 2011).

According to the traditional conceptualization of status in markets, organizational status is merit-based, achieved through mechanisms such as demonstrations of superior performance or affiliation with high-status others (Phillips & Podolny, 1996; Podolny, 1993; Sauder, 2006). Yet, this conceptualization of status is highly overlapping with the concept of reputation, which refers more closely to economists' notion of perceived quality and captures differences in merit (Fombrun & Shanley, 1990; Pfarrer et al., 2010). Indeed, scholars in the field of sociology distinguish between *achieved* status that actors receive based on their individual accomplishments and *ascribed status* that is generally beyond actors' control and inherited from the status of the social group with which they share some identity characteristics such as gender or race (Foladare, 1969; Jasso, 2001; Thomas-Hunt & Phillips, 2011). Status is a distinct and useful concept to the extent that it refers to such unearned differences in social rank that generate privileges (Barron & Rolfe, 2012; Ertug & Castellucci, 2013; Park et al., 2020; Pollock et al., 2015; Washington & Zajac, 2005).

Despite the lack of attention in the literature on organization and management theory, the idea of ascribed status is relevant in the context of organizational fields. Similar to individuals in social settings, organizations in markets are perceived and evaluated by audiences based on their identity characteristics such as structural properties, products, and technologies (Hannan et al., 2007; Hsu & Hannan, 2005; Peteraf & Shanley, 1997; Porac et al., 1995; Rosa et al., 1999). Studies also note that some identity characteristics of organizations might be perceived as more worthy or respectable, yet without providing any reference to the underlying bases of these perceptions (Park & Podolny, 2000; Phillips & Zuckerman, 2001; Sharkey, 2014). Indeed, there is no systematic research on the factors that affect this status value ascribed to organizational identity characteristics. Thus, existing literature provides an incomplete picture of status dynamics in organizational fields.

This study aims to fill this gap by investigating the institutional influences that shape the status value of identity characteristics in an organizational field. We study these dynamics in the context of higher education, which is widely characterized as a "prestige market," where universities try to attain and protect this asset (Askin

& Bothner, 2016; Breznik & Law, 2019). We conducted interviews with key stakeholders in the Turkish higher education field (i.e., students, academicians, managers in industry, and high school counselors) to elicit the organizational characteristics that they perceive as relevant for defining university identity. The status value ascribed to these identity characteristics was documented via surveys with separate samples of these four stakeholder groups using a relatively new but widely recognized MCDM method, the best-worst method (BWM).

Our findings provide significant theoretical implications for the research on status in markets, as they identify the status value ascribed to organizational identity characteristics within socially constructed systems of norms and values. This conceptualization of organizational status helps to differentiate from related but distinct concepts such as reputation or image (Barron & Rolfe, 2012; Ertug & Castellucci, 2013; Park et al., 2020; Pollock et al., 2015; Washington & Zajac, 2005). Our findings also have important implications for strategic management theory since the idea of status granted to organizational identity characteristics indicates strategic trade-offs with respect to internal allocation of organizational resources (Rosenzweig & Easton, 2010). As a final contribution, our study adds to the recent trend in research applying MCDM techniques by introducing the opinions of several stakeholder groups. This multi-stakeholder perspective may be especially relevant in contexts like higher education, where the stakeholder community is broad and potentially fragmented.

The remaining parts of the chapter are organized as follows. Section 2 presents a literature review on the status concept and status dynamics in organizational fields. Section 3 introduces the empirical setting of our analysis and the factors that have shaped stakeholder perceptions of status value in this context. Section 4 explains the methodological framework that is followed in this study. The results of the empirical analyses are presented in Sect. 5, and Sect. 6 discusses their interpretations and implications. Finally, Section 7 concludes the chapter by stating the main contributions and the limitations of the study.

2 Literature Review

According to a widely accepted definition, status is "a socially constructed, inter-subjectively agreed-upon and accepted ordering or ranking of individuals, groups, organizations, or activities in a social system" (Washington & Zajac, 2005; p. 284). Research, in the context of markets, documents important advantages enjoyed by organizations that are granted a high status. High-status firms can charge higher prices for the same quality output and experience greater growth in sales for a given level of quality (Benjamin & Podolny, 1999). By enhancing visibility and trustworthiness, organizational status lowers barriers of entry into markets and improves access to critical resources (Fombrun & Shanley, 1990; Stuart et al., 1999).

Despite volumnious research on how status affects organizations' actions and market outcomes (Chae et al., 2020; Chen et al., 2012; Piazza & Castellucci, 2014; Pollock et al., 2019; Sauder et al., 2012), we still do not know much about the origins of status hierarchies and the underlying bases of status distinctions (Pearce, 2011). Studies on status in markets posit that organizations gain status through demonstrations of superior performance or affiliation with high-status others (Phillips & Podolny, 1996; Podolny, 1993; Sauder, 2006). Yet, this conceptualization is highly overlapping with the notion of reputation, which has its roots in the economics discipline and captures the idea that stakeholders rely on a history of past performance to predict future performance (Fombrun & Shanley, 1990; Pfarrer et al., 2010). Indeed, the long-standing intellectual tradition of status in sociology establishes that actors gain status not only based on their accomplishments but also through inheritance from the status of the social group with which they share some identity characteristics such as gender or race. These two sources of social standing are differentiated as *achieved* vs. *ascribed* status (Correll & Ridgeway, 2003; Foladare, 1969; Jasso, 2001; Prato et al., 2019; Ridgeway, 1991).

The idea of ascribed status due to identity characteristics is largely neglected in the literature on organization and management theory (Barron & Rolfe, 2012; Ertug & Castellucci, 2013; Pollock et al., 2015; Washington & Zajac, 2005). Indeed, it is well established that organizations are perceived and evaluated by audiences based on their identity characteristics such as structural properties, products, and technologies (Hannan et al., 2007; Hsu & Hannan, 2005; Peteraf & Shanley, 1997; Porac et al., 1995; Rosa et al., 1999). Further, a few studies recognize that organizations may gain (or lose) status on the basis of their identity characteristics such as niche width (Park & Podolny, 2000), target customer segment (Phillips & Zuckerman, 2001), or affiliated industry categories (Sharkey, 2014). Unfortunately, though, these studies are silent about the factors that lead to these status judgments.

Stakeholders' assessments of status value in an organizational field are likely to be framed by a socially constructed system of norms and values in that particular context (Bitektine, 2011; Campbell et al., 2019; Kakkar et al., 2020; Pollock et al., 2019). Institutional research in organization theory provides a useful path to study such cognitive and normative frames in organizational fields. We specifically focus on three concepts proposed by this stream of research to build a collective sense of approval in an organizational field; and therefore, may be influential in shaping the status value of organizational identity characteristics: a) historical legacy, b) endorsement, and c) prominence.

The first of these concepts refers to organizational characteristics that are established in the early history of an organizational field and serve as legitimate identity templates (Greenwood & Hinings, 1993; Haack et al., 2014; King et al., 2011; Pólos et al., 2002; Zamparini & Lurati, 2017). Beyond this cognitive legitimacy, these ancestral identity characteristics are likely to inherit a historical legacy and be perceived by audiences in the field as more prestigious (Heugens & Lander, 2009; Sharkey, 2014).

The second concept refers to the endorsement of some organizational identity characteristics through various external and often publicized evaluations such as rankings, ratings, and critiques. Organizational identity characteristics that receive this endorsement are in general perceived as "desirable standards" or "the way things should be done" (Bromley & Powell, 2012; Lee et al., 2017). Beyond this normative legitimacy, organizational identity characteristics endorsed by such intermediaries are likely to be granted symbolic capital or prestige (Bourdieu & Wacquant, 2013; Holland & Ford, 2020; Rao, 1994).

The third concept refers to the (positive) visibility of certain organizational identity characteristics due to being adopted by successful or prominent organizations in the field (DiMaggio & Powell, 1983; Fombrun & Shanley, 1990). Such organizational characteristics are likely to be associated with successful outcomes and thus gain a higher status value in the eyes of evaluating stakeholders (Baum et al., 2000; Haunschild & Miner, 1997; Kraatz, 1998).

As an important caveat, the three influences mentioned above (historical legacy, endorsement, and prominence) are defined and determined in the particular context of an organizational field (Bitektine, 2011; Rao et al., 2005; Vigneron & Johnson, 1999; Zuckerman & Kim, 2003). The following section identifies the manifestations of each institutional influence in the context of the empirical setting of this research (i.e., the Turkish higher education field).

3 Empirical Setting

The Turkish higher education field with a history dating back to the early 1930s constitutes the empirical setting of our research. Reaching the size of 18 universities (all public) in the early 1980s, the field was re-designed with an overhaul in legislation. Field governance was centralized under the authority of a national-level board (YÖK—the Turkish abbreviation for "higher education council"). The new legislation permitted philanthropic foundations to establish not-for-profit private universities. The establishment of new universities (state or private) required both YÖK and government approval. Since then, the number of universities has increased more than tenfold, reaching 209 in 2020 (131 public and 78 private).

In line with our theoretical framework on the factors that affect stakeholder perceptions of status value, below we identify the university characteristics with a historical legacy, those that have received endorsement, and those that have become prominent in the Turkish higher education field.

3.1 Organizational Identity Characteristics with a Historical Legacy

The University Reform that took place in 1933 in the Turkish higher education field reconfigured *Darülfünun* (The House of Sciences, inherited from the Ottoman Empire) under the name of Istanbul University. This first university, as well as a replica of it established in Ankara, was modeled after the Continental European "classical" university. Another university model emerged in 1944 with the establishment of Istanbul Technical University (ITU), which specialized in engineering and architecture (Okyar, 1967). These initial universities (Istanbul, Ankara, and ITU) were provided with extensive autonomy according to the university law in 1946 and gained a distinguished position compared to non-university institutions in the Turkish higher education field (Gürüz, 2008; Tekeli, 2010; Üsdiken et al., 2013).

The classical universities (Istanbul and Ankara) came to be characterized by their large sizes and faculty of medicine. This classical university identity attracted many followers from newly established universities. The technical university identity, however, stayed with its sole exemplar (ITU) and as a distinctive identity (Topaler et al., 2021). Accordingly, the organizational identity characteristics that are granted historical legacy in the context of the Turkish higher education field have been the ones that characterize the classical university identity: i.e., the presence of the faculty of medicine and large size.

3.2 Organizational Identity Characteristics That Have Received an Endorsement

Like many other organizational fields, contexts of higher education have recently been under the influence of various forms of external, and often publicized, evaluations via rankings, ratings, and critics (Bloch & Mitterle, 2017; Bromley & Powell, 2012; Dearden et al., 2019; Rindova et al., 2018). The criteria applied by these evaluation systems are increasingly accepted as measures of success or appropriateness (Dill & Soo, 2005; Espeland et al., 2016).

As a centrally governed field, statements and recommendations of the higher education council (YÖK) have traditionally been influential in defining norms and appropriate behavior in the context of the Turkish higher education field. Since the early 2000s, though, the field has been subject to the penetrating influence of the global trend in higher education, which emphasizes evaluation, standardization, and benchmarking. Global rankings in higher education endorse university features such as research output, faculty excellence, internationalization, and funding (Altbach & Salmi, 2011; Bloch & Mitterle, 2017; Campbell et al., 2019; Horta, 2009). Among these features, internationalization (in terms of students and academicians) has been very limited in the context of the Turkish higher education field (https://istatistik.yok.gov.tr/). Financial resources (or endowments) have not been a significant feature that differentiates universities either, perhaps because the Turkish higher education field has traditionally been characterized by publicly owned universities whose budgets are determined by the state. Private universities, in this context, can only be established by non-profit foundations and cannot be for-profit.

University characteristics that have received endorsement in the context of the Turkish higher education field are research orientation and quality orientation in teaching. With respect to the former, the past two decades have witnessed increasing pressures from YÖK to produce international publications. With respect to the latter, a higher faculty-to-student ratio (occasionally labeled as being a "boutique" university) has recently emerged as a mark of quality orientation in teaching. These two university characteristics (research orientation and quality orientation in teaching) also constitute evaluation criteria applied by URAP (https://www.urapcenter.org/), a non-profit organization that has been publishing yearly rankings of Turkish universities based on their academic performance since 2009.

3.3 Organizational Identity Characteristics That Have Become Prominent in the Field

As indicated above, identity characteristics adopted by successful or prominent members of an organizational field enjoy (positive) visibility or prominence. The most visible indicator of university success in the Turkish higher education field is selectivity in admissions. All universities admit students through a centralized examination administered by the Student Selection and Placement Center. Since this is the only mechanism of matching students with universities, admittance scores are highly visible, and they are closely monitored by audiences in the Turkish higher education field.[1]

Accordingly, university characteristics that have become prominent in the context of the Turkish higher education field are those associated with higher selectivity in admissions. As will be explained in the subsequent section, we determine the exact list of prominent university characteristics in this field via regression analysis using archival data on the entire set of organizational characteristics considered by stakeholders as identity-relevant (see Sect. 4.3). According to the results of this regression analysis, two organizational characteristics of universities are revealed to have a significantly greater contribution to higher selectivity in admissions: English-medium instruction and location (i.e., establishment in larger cities).

4 Methods

The first step of our methodological framework is to identify the set of organizational characteristics perceived by audiences in the Turkish higher education field as relevant for defining university identity. We then apply multi-criteria decision-making techniques to document the stakeholders' prioritization of these characteristics in terms of status value. The final methodological step that we present

[1] Selectivity in admissions is recognized as a proxy for prestige in many other contexts of higher education (Askin & Bothner, 2016; Davies & Zarifa, 2012; Holland & Ford, 2020; Kraatz et al., 2010; Massey et al., 2003).

in this section aims to determine the prominence of university characteristics in this field, which we have theorized as an antecedent of their status value (noted above as the third institutional influence).

4.1 Determining the Characteristics That the Stakeholders Apply to Evaluate Universities

Key stakeholders of universities in the context of the Turkish higher education field are: a) students, b) academicians, c) managers in the industry (who constitute potential recruiters of university graduates), and d) high school counselors (who advise pupils in university selection). In order to identify the organizational identity characteristics that these stakeholders apply to evaluate universities, 10 interviews were conducted with each of the four stakeholder groups (Table S1 in Appendix presents sample statistics).

In these interviews, we applied card-sorting techniques (Budhwar, 2000; Daniels et al., 2002) where we presented a sample of 30 universities[2] (to ensure feasibility in the exercise) with names on laminated cards. We asked the respondents to classify these universities, using any criteria that they find relevant. They were then asked to name each cluster and describe its characteristics. Once the informants were done with the first round of clustering, they were invited to try different cluster solutions until they could not come up with anymore. Interviews were taped with the informants' consent and lasted between 1 hour and one and a half hours. The informants mentioned six distinct university characteristics for distinguishing universities, which are operationalized as below:

1. *Quality orientation in teaching*: ratio of the total number of full and associate professors to the total number of students
2. *Research orientation*: publications in journals covered by the Web of Science database per member of full-time academic staff
3. *Size*: total student intake of a university
4. *Location*: 3 for the three largest cities (i.e., Istanbul, Ankara, and Izmir), 2 for other large cities, and 1 for the remaining cities
5. *Faculty of medicine*: 1 if the faculty exists in the university, 0 otherwise
6. *English-medium instruction*: proportion of departments in the university in which instruction is in English

[2] This set is randomly selected among the whole set of universities in the Turkish higher education field. It is available from the first author upon request.

4.2 Documenting Stakeholders' Prioritization of University Characteristics in Terms of Status Value (Best-Worst Method)

Separate samples of the four stakeholder groups were surveyed to document the way the above-mentioned six university characteristics are prioritized in terms of status value. The questionnaire[3] that we designed were used to apply multiple-criteria decision-making (MCDM) techniques. Multi-criteria decision-making (MCDM) techniques create a ranking among several solution alternatives that need to be evaluated with respect to multiple criteria (Triantaphyllou, 2000). They have been widely applied in the context of higher education to solve various decision problems involving resource allocation, performance measurement, financial and operations management (Caballero et al., 2004; Hein et al., 2015; Ho et al., 2006; Mustafa & Goh, 1996). Over the years, several methodologies have been developed to use the main principles of MCDM. In this study, we utilize a relatively newer technique, namely Best-Worst Method (BWM).

BWM was initially developed by Rezaei (2015) in order to make pairwise comparisons more consistent and efficient. The properties of this method are further illustrated in Rezaei (2016); and an extension of the original BWM for capturing the opinions of several decision makers is introduced by Mohammadi and Rezaei (2020). We particularly apply the group BWM developed in this latest work of Mohammadi and Rezaei (2020).

According to BWM, two vectors of pairwise comparisons are utilized to determine the weights of the evaluation criteria. The first vector identifies how significant each criterion is with respect to the "most important" criterion of all, while the second vector displays the relative importance of all criteria with respect to the "least important" of all. Next, a non-linear minmax model (which can be transformed into a LP formulation with a unique solution) is used to compute the weights of all criteria.

Although the analysis could be conducted with several other MCDM techniques traditionally used to rank decision alternatives in various contexts, we chose to proceed with BWM due to two factors. First, this method has not been explored extensively in the literature, mainly because it is a relatively new decision-making technique. This study will contribute to the extant BWM literature by applying the group BWM technique in an unconventional context, namely measuring the status perceptions in higher education. Secondly, the BWM method presents important advantages compared to similar MCDM methodologies. The advantages of using BWM can be depicted as follows: First, BWM requires the decision maker (DM) to identify the best and the worst criteria before conducting the pairwise comparisons among the criteria. The DM having a clear understanding of the range of evaluation could lead to more consistent and reliable comparisons. As a second advantage,

[3] Available from the first author upon request.

the anchoring bias that the DM might have while conducting pairwise comparisons is mitigated through the use of two pairwise comparison vectors formed based on two opposite references (best and worst) in a single optimization model. Using a single vector for the input data in a pairwise comparison-based method is more efficient, yet it does not permit to check the consistency of the provided pairwise comparisons in this manner. Using a full matrix allows checking the consistency of the provided pairwise comparisons, yet it is not very efficient in terms of data (and time) and might still induce inconsistencies. BWM combines the advantages of both methodologies by letting the DM make two sets of pairwise comparisons, enabling him/her to check consistency without creating extra confusion.

The procedure for the original BWM follows the following steps.

Step 1: The DM establishes the set of decision criteria $C = \{c_1, c_2, \ldots, c_n\}$.

Step 2: The DM determines the best (c_B) and the worst (c_W) criteria from C which was established in the first step. The DM does not perform any pairwise comparison in this stage. This selection is subjective; the best criterion is the most significant or the most desirable one according to the DM, while the worst criterion is the least significant or the least desirable criterion among others.

Step 3: The DM performs the pairwise comparison between the best (c_B) and the other criteria in C. At this stage, the DM adjusts his/her evaluation of the best criterion with respect to the other criteria by choosing a number between one and nine. One means both criteria are equally significant and nine means the best criterion is extremely more significant than the other criterion under consideration. The pairwise comparisons altogether produce the "Best-to-Others" vector A_B as

$$A_B = \{a_{B1}, a_{B2}, \ldots, a_{Bn}\}$$

where a_{Bj} displays the preference ratio of the best criterion c_B with respect to criterion c_j.

Step 4: Similar to Step 3, the DM now performs the pairwise comparison between the worst (c_W) and the other criteria in the set C. At this stage, the DM again selects a number between one and nine in order to reflect his/her preferences of the other criteria with respect to the worst criterion. One means equally important and nine means extremely more important. The pairwise comparisons would finally produce the "Others-to-Worst" vector A_W as

$$A_W = \{a_{W1}, a_{W2}, \ldots, a_{Wn}\}$$

where a_{Wj} shows how much criterion c_j is preferred with respect to the worst criterion c_W.

Step 5: One obtains the criteria weights $w^* = \left\{w_1^*, w_2^*, \ldots, w_n^*\right\}$by solving the following non-linear programming model:

$$\min_w \max_j \left\{ \left| \frac{w_B}{w_j} - a_{Bj} \right|, \left| \frac{w_j}{w_W} - a_{jW} \right| \right\} \tag{1}$$

$$s.t. \sum_j w_j = 1, w_j \geq 0, \ \ \forall j = 1, 2, \ldots n. \tag{2}$$

The non-linear programming model (1)–(2) aims to minimize the absolute differences $\left|\frac{w_B}{w_j} - a_{Bj}\right|, \left|\frac{w_j}{w_W} - a_{jW}\right|$ for all criteria $j = 1, 2, \ldots n$, by setting the optimal weight vector such that all weights are non-negative and add up to 1. Although BWM has various advantages, it was originally designed for evaluating the decisions of a single DM. However, the opinions of several experts are often required to avoid bias in real-life situations. BWM can still be used to create a weighted average (arithmetic or geometric) of expert opinions, especially if there were sufficiently many experts whose total number would allow utilizing the law of large numbers. However, for smaller groups (i.e., groups of fewer than 30 individuals involved), a more appropriate and precise approach would be to use the Bayesian BWM, which assumes a joint probability distribution for all decision makers. In particular, instead of using a precise point, a distribution as in the Maximum Likelihood Estimation is utilized to estimate the parameters. The Dirichlet distribution that is used in the Bayesian inference as the prior to the multinomial can perfectly be substituted for the weight vector since both of them produce non-negative values and all components add up to one in both. The steps of Bayesian Group BWM are similar to those of the original BWM method, with the exception of Step 5. The only difference is the optimization problem being substituted with a probabilistic model. The proposed Bayesian model, yet, is more informative about the confidence of the relation between each pair of criteria. In particular, given the preference vectors A_B^k and A_W^k of all decision makers $k = 1, \ldots$ K, the optimal weight $w*$ is computed by applying Bayes rule to the set of probabilities of observing each vector A_B^k and A_W^k provided that each decision maker is independent in his/her decisions and the optimal weight vector is $w*$. Since these probabilities are expressed in a chain affecting each other, the model is called hierarchical. We refer the reader to Mohammadi and Rezaei (2020) for the details of this procedure and the associated formulations.

The Bayesian BWM method further produces a "credal ranking" among the criteria, which computes the degree to which a criterion is superior to another. This notion is important since the precedence relations among the criteria are not very clear for a weight vector with values very close to each other. In particular, credal ranking provides the probability that criterion c_i being superior to c_j as follows:

$$P\left(c_i > c_j\right) = \int 1_{\left\{w_i^* > w_j^*\right\}} P\left(w^*\right) \tag{3}$$

where $w*$ is the optimal weight vector, $P(w*)$ its distribution, and $1_{\{.\}}$ is the indicator function which takes the value 1 if the condition is satisfied and 0 otherwise. Clearly, $P(c_i > c_j) + P(c_j > c_i) = 1$. Therefore, c_i is more important than c_j if and only if $P(c_i > c_j) > 0.5$ in Eq. (3). As a result, a threshold of 0.5 can be applied to the credal ranking to obtain the traditional ranking of criteria.

4.3 Estimating the Effects of University Characteristics on University Selectivity (Regression Analysis)

As identified above, the prominence of university characteristics in the context of the Turkish higher education field corresponds to their association with higher selectivity scores. In order to estimate the effects of the six university characteristics mentioned above on university selectivity, we applied regression analysis using archival data.

We recorded student entrance scores in the centralized university examination for each program in a university and in each score type (verbal, quantitative, language, and equally weighted). We then z-standardized these scores within score types. A university's overall selectivity in admissions equaled the mean z-score of its programs. The data for university characteristics were gathered from four sources: (a) the Law on the Organization of Higher Education Institutions, (b) the annual central university examination manuals (Student Selection and Placement Center, 2014), (c) annual higher education statistics (Council of Higher Education, 2014), and (d) the Web of Science database.

For detrending, we included year as a covariate to account for the linear effect of time by Curran et al. (2012); Wang and Maxwell (2015). The ordinary least squares (OLS) regression model that we estimated is formally represented as

$$\begin{aligned} Y = {} & \beta_0 + \beta_1 \text{ (quality in teaching)} + \beta_2 \text{ (research orientation)} + \beta_3 \text{ (size)} \\ & + \beta_4 \text{ (faculty of medicine)} + \beta_5 \text{ (English-medium instruction)} \\ & + \beta_6 \text{ (location)} + \text{time} + \varepsilon \end{aligned} \tag{1}$$

Z-score standardization is applied to all independent variables before they are included in the regression model. Quantile-quantile plots showed no violation of normality. Table 1 presents the descriptive statistics and pairwise correlations between the study variables. There are no issues with multicollinearity since the variance inflation factors for all variables are less than 2.5.

Table 1 Descriptive statistics and correlations (regression analysis)[a]

	Variables	Mean	SD	1	2	3	4	5	6
1	Quality in teaching	0.19	0.13						
2	Size	0.46	0.42	−0.19					
3	Faculty of medicine	0.49	0.50	0.23	0.53				
4	English-medium instruction	0.24	0.33	0.03	−0.39	−0.28			
5	Location (larger cities)	1.09	0.84	0.35	−0.26	−0.02	0.57		
6	Research orientation	0.41	0.56	0.37	−0.07	0.04	0.37	0.33	
7	Time (year divided by 1000)	2.01	0.05	0.28	0.02	0.03	0.37	−0.02	0.24

[a] $n = 2263$. Correlations greater than 0.04 are significant at 0.05 (two-tailed test)

Table 2 Regression analysis results[a]

Variable	β coefficient	*Effect size*
Quality in teaching	0.09* (0.03)	*0.033*
Research orientation	0.08* (0.03)	*0.030*
Size	−0.03 (0.04)	*0.003*
Faculty of medicine	0.06* (0.03)	*0.022*
English-medium instruction	0.11** (0.04)	*0.089*
Location (larger cities)	0.23*** (0.04)	*0.161*
Time (years)	0.01 (0.54)	
Constant	−0.06* (0.03)	
Model F	25.85***	
R^2	0.72	

[a]$N = 2263$. Standard errors are in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Effect sizes show Eta-squared values

According to the regression results (Table 2), university selectivity is positively affected by faculty of medicine, quality orientation in teaching, English-medium instruction, location (establishment in larger cities), and research orientation. Effect sizes (calculated as eta-squared values) show that location and English-medium instruction have the strongest effect on university selectivity, followed by quality orientation in teaching, research orientation, and faculty of medicine.

5 Findings

The questionnaire that applies the Best-Worst Method (BWM) to document stakeholder judgments of university characteristics was sent in the form of a web-based survey. Responses were received from 17 students, 21 academicians, 24 managers, and 11 counselors. The respondents' answers were transformed into the best-to-others and others-to-worst vectors, which were then used as inputs for the Bayesian BWM analysis. An example of best-to-others and others-to-worst vectors (i.e., A_B^k and A_W^k of all counselors $k = 1, \ldots\ 11$) is provided in Table 3 for the 11 counselors who have participated in our survey. The preference vectors of other groups are not presented here for the sake of brevity.

After the preference vectors of the four groups are identified, the following joint probability distribution is sought for each group j:

$$P\left(w^{*j}, w^{j,1:K_j} | A_B^{1:K_j},\ \ A_W^{1:K_j}\right) \tag{5}$$

where w^{*j} is the overall optimal weight for group j, the superscript indicates the total of all vectors in the base, $w^{j,1:K_j}$ is the total of all weight vectors of the decision makers in group j, and $A_B^{1:K_j},\ \ A_W^{1:K_j}$ indicate the total of all preference vectors in group j ($j = 1, 2, 3, 4$).

Table 3 Best-to-others and Others-to-Worst vectors of counselors

	Best	Best-medicine	Best-English	Best-research	Best-location	Best-teaching	Best-size
Coun. 1	Research or.	5	4	1	6	3	8
Coun. 2	Qual. In teaching	5	5	2	5	1	9
Coun. 3	Research or.	3	2	1	2	2	2
Coun. 4	Research or.	7	7	1	9	7	7
Coun. 5	Research or.	8	7	1	8	7	9
Coun. 6	Research or.	3	6	1	2	5	2
Coun. 7	Qual. In teaching	4	4	2	4	1	5
Coun. 8	Research or.	3	3	1	6	2	6
Coun. 9	Research or.	3	3	1	3	3	3
Coun. 10	Research or.	3	7	1	5	3	5
Coun. 11	Research or.	2	3	1	3	3	4
	Worst	**Worst-medicine**	**Worst-English**	**Worst-research**	**Worst-location**	**Worst-teaching**	**Worst-size**
Coun. 1	Size	2	6	8	3	7	1
Coun. 2	Size	2	2	8	2	9	1
Coun. 3	Medicine	1	2	5	2	2	2
Coun. 4	Location	2	2	9	1	2	2
Coun. 5	Size	2	7	9	6	7	1
Coun. 6	English	2	1	6	5	2	2
Coun. 7	Size	2	2	5	4	5	1
Coun. 8	Location	3	3	8	1	7	2
Coun. 9	Location	3	3	4	1	2	3
Coun. 10	English	2	1	7	3	2	2
Coun. 11	Size	2	2	4	2	3	1

Note that the value of $w^{j,k}$ is dependent on $A_W^{j,k}$ and $A_B^{j,k}$, and the value of w^{*j} is also dependent on $w^{j,k}$. That is, the individual preference of k^{th} decision maker in group j is independent of w^{*j} given $w^{j,k}$:

$$P\left(A_W^{j,k}|w^{*j}, w^{j,k}\right) = P\left(A_W^{j,k}|w^{j,k}\right) \tag{6}$$

The application of the Bayes rule to the joint probability expression (6) with a consideration of all independence among the variables leads to the following expression:

$$P\left(w^{*j}, w^{j,1:K_j}|A_B^{1:K_j},\ \ A_W^{1:K_j}\right) \ltimes P\left(A_B^{1:K_j},\ \ A_W^{1:K_j}|w^{*j}, w^{j,1:K_j}\right)$$
$$\times P\left(w^{*j}, w^{j,1:K_j}\right) = P\left(w^{*j}\right)\prod_{k=1}^{K_j} P\left(A_B^{j,k}|w^{j,k}\right) P\left(A_W^{j,k}|w^{j,k}\right) P\left(w^{j,k}|w^{*j}\right) \tag{7}$$

where $\ltimes$indicates proportionality. The last equality is obtained by using (a) the probability chain rule, (b) the conditional independence of different variables, and (c) the independent decision making of each decision maker. This is a hierarchical model where the estimation of the parameters in Eq. (7) relies on the estimation of other variables. The underlying Matlab code accounts for this hierarchical structure by using Markov-chain Monte Carlo technique. The final output is the posterior distribution of weights for every single decision maker and the aggregated optimal weights w^{*j} for the group j. We refer the reader to Mohammadi and Rezaei (2020) for further detail.

We used the Matlab implementation on the BWM website to apply the analysis (https://bestworstmethod.com/software/). The runs take approximately 15–25 minutes for six university characteristics ("criteria" in the language of BWM) and the indicated sizes of the stakeholder groups on a computer with 16 GB RAM and Matlab R2019b version. According to the analysis results, the optimal weights computed for each group (i.e., w^{*j}, $j = 1,2,3,4$) are displayed in Table 4.

We also computed the credal rankings of university characteristics for each stakeholder group as expressed in Eq. (3). Figure 1 shows these values. The number combining two university characteristics identifies the credal ranking of

Table 4 Optimal weights of university characteristics across the stakeholder groups

	Quality in teaching	Research orientation	Size	Faculty of medicine	English medium	Location (larger cities)
Students	0.1910	0.2940	0.0845	0.0905	0.2038	0.1362
Academicians	0.1940	0.2845	0.0890	0.1539	0.1536	0.1250
Managers	0.1865	0.2436	0.1229	0.1295	0.1991	0.1183
Counselors	0.1878	0.3441	0.0974	0.1228	0.1256	0.1222

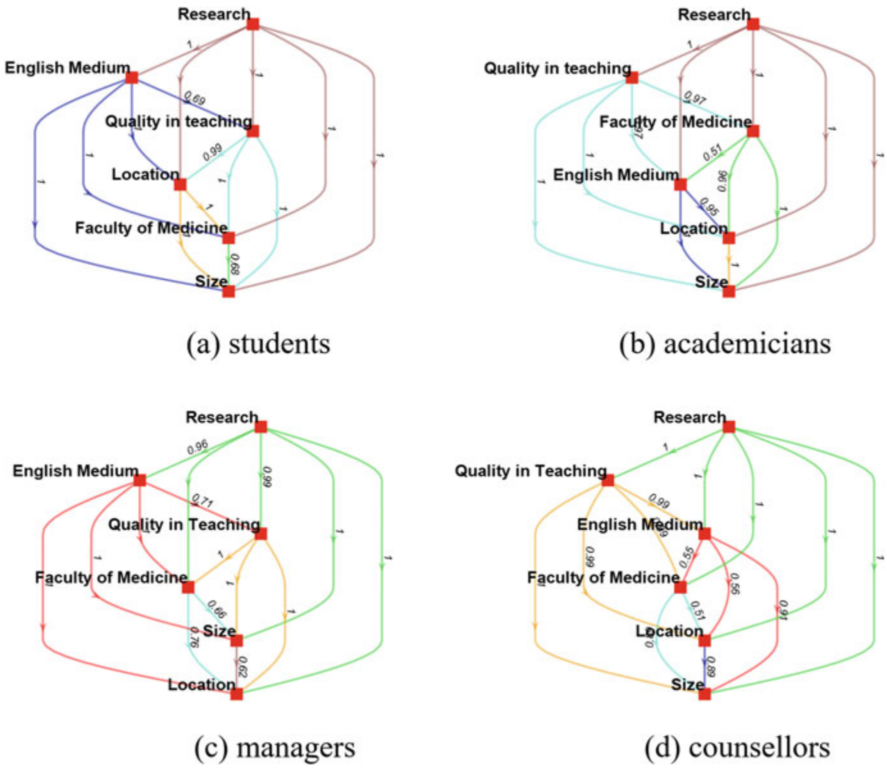


Fig. 1 Credal rankings of university characteristics across the stakeholder groups. (**a**) students, (**b**) academicians, (**c**) managers, (**d**) counselors

the characteristic above compared to the one below (i.e., the degree to which the characteristic above is superior to the one below). For instance, students regarded research orientation as the most appreciated university characteristic. At the other extreme, size and faculty of medicine are perceived as the least valued university characteristics in the eyes of this stakeholder group. English-medium instruction is certainly perceived as more valuable than the presence of the faculty of medicine (ranking = 1), whereas it is more valued than quality orientation in teaching only with some confidence (ranking = 0.69).

According to the results displayed in Table 4 and Fig. 1, overall, the evaluations are consistent among the four groups of stakeholders. In particular, all groups regard the research orientation as the most valuable criterion to define the prestige of a higher education institution. Quality orientation in teaching follows as the second valuable criterion according to academicians and counselors, while English-medium instruction is regarded as the second most valuable according to the students and managers. Students appreciate the location (establishment in larger cities) higher than the other groups, while they rank faculty of medicine at a lower position compared to other stakeholders. Academicians value the presence of a faculty of medicine in an institution more than the other groups. Size is, in general, the least appreciated criterion although managers think that location (establishment in larger cities) is less valuable than university size.

Overall, BWM analyses suggest that university characteristics that have received endorsement (i.e., research orientation and quality orientation in teaching) are ascribed with the greatest status value. University characteristics that have prominence in the field take the second place, as English-medium instruction receives middle to high status value, and location (establishment in larger cities) receives low to middle status value. The university characteristics that have historical legacy are perceived to rank lower since both faculty of medicine and (large) size receive low to middle status value.

As identified in the Empirical Setting section, historical legacy, endorsement, and prominence in the context of the Turkish higher education field represent past, recent, and time-independent institutional influences, respectively. In order to account for this variation in temporal distance, we did additional analyses where we compared the opinions across several stakeholder groups regarding their age. To this end, we categorized participants into four different age groups, namely 20–30, 30–40, 40–50, and above-50. After dropping the participants who did not provide age information, the numbers of participants in the age groups 20–30, 30–40, and 40–50 were 18, 21, and 10. Table 5 displays the optimal weights computed for the given age groups. Since there were only three participants in the above-50 group, our inferences below exclude this age group.

As can be seen from the results displayed in Table 5, the status value that the participants attribute to the presence of the faculty of medicine increases with age, while the value of English-medium instruction decreases. In a sense, the status value ascribed to university characteristics with historical legacy (faculty of medicine and size) and the status value ascribed to university characteristics that have prominence in the field (English-medium instruction and location) become similar when we control for participant age. There does not seem a large discrepancy among groups regarding the remaining university characteristics. All groups regard research orientation as the most significant feature, while size becomes the least appreciated attribute.

We conducted an additional analysis to see if there exist any discrepancies between the stakeholders in state versus private universities. Ownership structure represents an important divide among universities in the context of the Turkish

Table 5 Optimal weights of university characteristics across age groups

	Quality in teaching	Research orientation	Size	Faculty of medicine	English medium	Location (larger cities)
20–30	0.1894	0.2954	0.0861	0.0906	0.2019	0.1367
30–40	0.1898	0.2530	0.1087	0.1306	0.1958	0.1221
40–50	0.1711	0.2915	0.1177	0.1557	0.1452	0.1188
50+	0.1556	0.2961	0.1099	0.1648	0.0965	0.1771

Table 6 Optimal weights of university characteristics according to stakeholders affiliated with state vs. private universities

	Quality in teaching	Research orientation	Size	Faculty of medicine	English medium	Location (larger cities)
Stakeholders affiliated with state univ.	0.2038	0.2895	0.0864	0.1119	0.1693	0.1392
Stakeholders affiliated with private univ.	0.1850	0.2899	0.0893	0.1292	0.1826	0.1240

higher education field. Despite being subject to the same higher education law in the country, private universities are distinct from state universities in terms of their governance structure, revenue generation, and employment relationship with their staff (Barblan et al., 2008). In our sample, 11 academicians and 8 students are affiliated with state universities (Group 1), whereas 10 academicians and 9 students are affiliated with private universities (Group 2). Table 6 displays the optimal weights computed for these two groups.

According to Table 6, status judgments of stakeholders affiliated with state and private universities are mostly consistent. Research orientation is ascribed with the greatest status value, and quality orientation in teaching follows next. Participants from private universities regard English-medium instruction as almost equally valuable as quality orientation in teaching, possibly because instruction in English has been a distinguishing feature of private universities in the context of the Turkish higher education field. Stakeholders in private universities value the faculty of medicine slightly more than the stakeholders affiliated with state universities. On the other hand, stakeholders affiliated with state universities prioritize universities' location more than the faculty of medicine. Rankings of the remaining criteria are largely the same in the state and private groups.

6 Discussion

In this research, we propose three institutional influences that serve as sources of ascribed status for organizational identity characteristics. Our empirical investigation in the context of Turkish higher education shows that stakeholders attach the highest status value to those university characteristics that have received endorsement in the field. Those university characteristics that have historical legacy (due to characterizing archetypal identities in the field) or those with prominence in the field (due to characterizing successful exemplars) are not valued to the same extent.

Our findings provide evidence for the view that institutions have a broad influence on the social hierarchy systems (O'Brien & Dietz, 2011; Pratto et al., 2006), yet with some nuances. We observe, for instance, the more recent institutional source of status (i.e., endorsement) makes a more significant influence on the status value of organizational characteristics than the one with the historical roots. Indeed, one might expect the latter influence to be stronger since early emerging status beliefs in an organizational field are likely to be taken-for-granted, and therefore stable over time (Prato et al., 2019). This expectation is also in line with the idea of "past anchoring," which posits that templates introduced in the past will provide normative prescriptions (Chan et al., 2021; King et al., 2011). Yet, this has not been the case in our empirical setting.

An immediate explanation for this may be the "temporal myopia" of audiences, that is, a tendency to focus on the short run (Greve, 2008; Levinthal & March, 1993; Miller, 2002). Yet, the additional analyses where we differentiated between stakeholder age groups provide more nuanced reasoning. We find that stakeholders at older age groups attach greater status value to university characteristics with a historical legacy, whereas they attach lower status value to university characteristics that have recently received endorsement in the field. This observation makes sense since the older members of a field are more likely to be influenced by the status dynamics in its early history (see Marquis & Tilcsik, 2013 for a more detailed discussion of this imprinting argument). Overall, our findings suggest that there is value in taking a comprehensive approach to institutional sources of status in an organizational field, which accounts for past, present, as well as time-independent phenomena.

Our study extends existing research recognizing that organizations may gain (or lose) status based on some identity characteristics without a focus on how these characteristics are associated with such status value (Park & Podolny, 2000; Phillips & Zuckerman, 2001; Sharkey, 2014). Here, we introduce the institutional theory perspective to propose factors that shape stakeholders' status judgments with respect to organizational identity characteristics. Future research in other empirical contexts may extend our insights further by demonstrating other cognitive and normative underpinnings of ascribed status in organizational fields.

7 Conclusions

Criticizing the narrow focus in management literature on organizational status based on demonstrations of superior performance (i.e., achieved status), this study identifies the status value ascribed to organizational identity characteristics within socially constructed systems of norms and values. In this way, we provide further evidence for the idea of loose coupling between organizational status and actual performance or quality (Elsbach & Cable, 2019). Our empirical analyses in the context of the higher education field show that a wide variety of university characteristics—such as location or language of instruction—have status value, despite their lack of a direct association with higher quality or excellence. These insights help differentiate the concept of organizational status from related but distinct concepts such as reputation or image (Barron & Rolfe, 2012; Ertug &

Castellucci, 2013; Park et al., 2020; Pollock et al., 2015; Washington & Zajac, 2005). Further, our study suggests that the dynamics of social hierarchies may be highly contextual, shaped by the norms and values in a particular setting.

Our findings also have important implications for strategic management theory. Intangible assets like status have strategic importance for organizations, as they provide sustainable competitive advantage (Barney, 1991). The idea of status value ascribed to organizational identity characteristics indicates strategic trade-offs with respect to internal allocation of organizational resources, which was not recognized in the prior literature (see Rosenzweig & Easton, 2010, for more on this topic). Organizations that aim to improve their social standing in the field may prioritize their investments into various identity characteristics such as structural properties or technologies based on the status value attached to these characteristics by evaluating audiences.

An additional contribution of our study is to the literature on higher education. Concerns about status position have been an ever-present issue in higher education fields. Despite this significance of high status for university outcomes, little has been done to explore stakeholders' perceptions of status value. Existing research instead focuses on published rankings as the main indicator of university status (Dearden et al., 2019; Dill & Soo, 2005; Espeland et al., 2016; Rindova et al., 2018; Torres-Olave et al., 2020). Our demonstration of the status value attached to a wider and deeper set of university characteristics suggests that studies of status in higher education fields should take a more comprehensive perspective.

Our findings are of great value for decision making in universities and for developing state policies regarding higher education. In particular, university administrators should recognize the significance of improving the quality orientation in teaching and the research infrastructure to improve their prestige and ranking in admissions. Similarly, governmental bodies responsible for developing the budget plans for state universities might take these two factors into account in allocating the resources. In extending the investments in higher education, the government or a private body should be aware that language of instruction and faculty composition are other significant factors that may contribute to the status of a new university.

As a final contribution, our study adds to the recent trend in research applying MCDM techniques by introducing the opinions of several stakeholder groups. There are earlier examples of such research in various contexts. For instance, Garcia et al. (2016) consider alternative management options for sustainable corporate development from a stakeholder versus shareholder perspective. Soltani et al. (2015) present a review of the MCDM studies that have considered multiple stakeholders such as government, municipalities, and the public in the Municipal Solid Waste Management problem. Scott et al. (2015) develop a decision support model to aid with supplier selection and order allocation process of companies in a multi-stakeholder setting. Our insights indeed suggest that this multi-stakeholder perspective may be especially relevant in contexts like higher education, where the stakeholder community is broad and potentially fragmented. Although we observe some heterogeneity in the opinions of different categories of stakeholders (academicians, students, counselors, and managers), age groups, and public/private affiliation, our conclusions in this respect are limited since we did not conduct a systematic investigation of stakeholder heterogeneity. Future research may study this fragmentation in the stakeholder community and its implications on decision making.

The fact that our research is conducted in a single empirical context may raise concerns about whether our findings are specific to the Turkish higher education field. In fact, studies in other contexts of higher education observed that audiences differentiate between universities based on their language of instruction (Lau & Lin, 2017) or ownership structure (Marginson, 2007). Still, the perceived value of organizational characteristics may be shaped by various contextual influences (Collins & Park, 2016). In this respect, future studies on social evaluation dynamics in other organizational fields would increase confidence in the generalizability of our findings and conclusions.

Methodologically, our use of the BWM technique to compare the worth of organizational characteristics in the eyes of stakeholders introduces an important novelty. This method may be superior to the direct ratings of organizations, which are increasingly common in organizational fields, yet suffer from various types of rater bias (Bromley & Powell, 2012; Salmi & Saroyan, 2007). Other MCDM techniques such as the analytical hierarchy process (AHP) can also be applied to document the relative valence of organizational characteristics, which will then serve to validate published ratings and rankings of organizations.

Appendix

Table S1 Sample statistics for interview participants

Stakeholder group	Sample size	Gender	Residence
Students	10	5 females, 5 males	4 in small cities, 6 in large cities
Academicians	10	6 females, 4 males	3 in small cities, 7 in large cities
Counselors	10	3 females, 7 males	2 in small cities, 8 in large cities
Managers	10	8 females, 2 males	3 in small cities, 7 in large cities

References

Altbach, P. G., & Salmi, J. (Eds.). (2011). *The road to academic excellence: The making of world-class research universities*. The World Bank.

Askin, N., & Bothner, M. S. (2016). Status-aspirational pricing: The “Chivas Regal” strategy in US higher education, 2006–2012. *Administrative Science Quarterly, 61*(2), 217–253.

Barblan, A., Ergüder, Ü., & Gürüz, K. (2008). *Higher education in Turkey: Institutional autonomy and responsibility in a modernising society*. Bononia University Press.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120.

Barron, D. N., & Rolfe, M. (2012). It ain't what you do, it's who you do it with: Distinguishing reputation and status. In M. L. Barnett & T. G. Pollock (Eds.), *Oxford handbook of corporate reputation* (pp. 160–178). Oxford University Press.

Baum, J. A., Calabrese, T., & Silverman, B. S. (2000). Don't go it alone: Alliance network composition and startups' performance in Canadian biotechnology. *Strategic Management Journal, 21*(3), 267–294.

Benjamin, B. A., & Podolny, J. M. (1999). Status, quality and social order in the California wine industry. *Administrative Science Quarterly, 44*, 563–589.

Bitektine, A. (2011). Toward a theory of social judgments of organizations: The case of legitimacy, reputation, and status. *Academy of Management Review, 36*(1), 151–179.

Bloch, R., & Mitterle, A. (2017). On stratification in changing higher education: The "analysis of status" revisited. *Higher Education, 73*(6), 929–946.

Bourdieu, P., & Wacquant, L. (2013). Symbolic capital and social classes. *Journal of Classical Sociology, 13*(2), 292–302.

Breznik, K., & Law, K. M. (2019). What do mission statements reveal about the values of top universities in the world? *International Journal of Organizational Analysis, 27*(5), 1362–1375.

Bromley, P., & Powell, W. W. (2012). From smoke and mirrors to walking the talk: Decoupling in the contemporary world. *Academy of Management Annals, 6*(1), 483–530.

Budhwar, P. (2000). The use of visual card sorting technique to study managers' belief structure–An international comparative study. *Journal of Managerial Psychology, 15*(5), 440–455.

Caballero, R., Galache, T., Gómez, T., Molina, J., & Torrico, A. (2004). Budgetary allocations and efficiency in the human resources policy of a university following multiple criteria. *Economics of Education Review, 23*(1), 67–74.

Campbell, C. M., Jimenez, M., & Arrozal, C. A. N. (2019). Prestige or education: College teaching and rigor of courses in prestigious and non-prestigious institutions in the US. *Higher Education, 77*(4), 717–738.

Chae, H., Song, J., & Lange, D. (2020). Basking in reflected glory: Reverse status transfer from foreign to home markets. *Strategic Management Journal, 42*, 802.

Chan, T., Lee, G. Y., & Jung, H. (2021). Anchored differentiation: The role of temporal distance in the comparison and evaluation of new product designs. *Organization Science, 32*, 1523.

Chen, Y. R., Peterson, R. S., Phillips, D. J., Podolny, J. M., & Ridgeway, C. L. (2012). Introduction to the special issue: Bringing status to the table—Attaining, maintaining, and experiencing status in organizations and markets. *Organization Science, 23*, 299–307.

Collins, F. L., & Park, G. S. (2016). Ranking and the multiplication of reputation: Reflections from the frontier of globalizing higher education. *Higher Education, 72*(1), 115–129.

Correll, S. J., & Ridgeway, C. (2003). Expectation states theory. In J. Delamater (Ed.), *Handbook of social psychology* (pp. 29–51). Springer.

Council of Higher Education. (2014). *Annual Higher Education Statistics (Yükseköğretim İstatistikleri)*. Available at: www.istatistik.yok.gov.tr

Curran, P. J., Lee, T., Howard, A. L., Lane, S., & MacCallum, R. A. (2012). Disaggregating within-person and between-person effects in multilevel and structural equation growth models. In J. R. Harring & G. R. Hancock (Eds.), Advances in longitudinal methods in the social and behavioral sciences (pp. 217–253). Charlotte, NC: Information Age.

Daniels, K., Johnson, G., & De Chernatony, L. (2002). Task and institutional influences on managers' mental models of competition. *Organization Studies, 23*, 31–62.

Davies, S., & Zarifa, D. (2012). The stratification of universities: Structural inequality in Canada and the United States. *Research in Social Stratification and Mobility, 30*, 143–158.

Dearden, J. A., Grewal, R., & Lilien, G. L. (2019). Strategic manipulation of university rankings, the prestige effect, and student university choice. *Journal of Marketing Research, 56*(4), 691–707.

Dill, D. D., & Soo, M. (2005). Academic quality, league tables, and public policy: A cross-national analysis of university ranking systems. *Higher Education, 49*(4), 495–533.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review, 48*(2), 147–160.

Elsbach, K. D., & Cable, D. M. (2019). Explaining stakeholder identification with moderate prestige collectives: A study of Nascar fans. *Organization Studies, 40*(9), 1279–1305.

Ertug, G., & Castellucci, F. (2013). Getting what you need: How reputation and status affect team performance, hiring, and salaries in the NBA. *Academy of Management Journal, 56*, 407–431.

Espeland, W. N., Sauder, M., & Espeland, W. (2016). *Engines of anxiety: Academic rankings, reputation, and accountability*. Russell Sage Foundation.

Foladare, I. S. (1969). A clarification of "ascribed status" and "achieved status". *The Sociological Quarterly, 10*(1), 53–61.

Fombrun, C., & Shanley, M. (1990). What's in a name? Reputation building and corporate strategy. *Academy of Management Journal, 33*(2), 233–258.

Garcia, S., Cintra, Y., Rita de Cássia, S. R., & Lima, F. G. (2016). Corporate sustainability management: A proposed multi-criteria model to support balanced decision-making. *Journal of Cleaner Production, 136*, 181–196.

Goode, W. J. (1978). *The celebration of heroes: Prestige as a control system*. University of California Press.

Greenwood, R., & Hinings, C. R. (1993). Understanding strategic change: The contribution of archetypes. *Academy of Management Journal, 36*, 1052–1081.

Greve, H. R. (2008). A behavioral theory of firm growth: Sequential attention to size and performance goals. *Academy of Management Journal, 51*(3), 476–494.

Gürüz, K. (2008). *Yirmi Birinci Yüzyılın Başında Türk Milli Eğitim Sistemi*. Türkiye İş Bankası Kültür Yayınları.

Haack, P., Pfarrer, M. D., & Scherer, A. G. (2014). Legitimacy-as-feeling: How affect leads to vertical legitimacy spillovers in transnational governance. *Journal of Management Studies, 51*(4), 634–666.

Hannan, M. T., Pólos, L., & Carroll, G. R. (2007). *Logics of organization theory: Audiences, codes, and ecologies*. Princeton University Press.

Haunschild, P. R., & Miner, A. S. (1997). Modes of interorganizational imitation: The effects of outcome salience and uncertainty. *Administrative Science Quarterly, 42*(3), 472–500.

Hein, N., Kroenke, A., & Júnior, M. M. R. (2015). Professor assessment using multi-criteria decision analysis. *Procedia Computer Science, 55*, 539–548.

Heugens, P. P., & Lander, M. W. (2009). Structure! Agency!(and other quarrels): A meta-analysis of institutional theories of organization. *Academy of Management Journal, 52*(1), 61–85.

Ho, W., Dey, P. K., & Higson, H. E. (2006). Multiple criteria decision-making techniques in higher education. *International Journal of Educational Management, 20*, 319–337.

Holland, M. M., & Ford, K. S. (2020). Legitimating prestige through diversity: How higher education institutions represent ethno-racial diversity across levels of selectivity. *The Journal of Higher Education, 92*(1), 1–30.

Horta, H. (2009). Global and national prominent universities: Internationalization, competitiveness and the role of the state. *Higher Education, 58*, 387–405.

Hsu, G., & Hannan, M. T. (2005). Identities, genres, and organizational forms. *Organization Science, 16*(5), 474–490.

Jasso, G. (2001). Studying status: An integrated framework. *American Sociological Review, 66*(1), 96–124.

Kakkar, H., Sivanathan, N., & Gobel, M. S. (2020). Fall from grace: The role of dominance and prestige in the punishment of high-status actors. *Academy of Management Journal, 63*(2), 530–553.

King, B. G., Clemens, E. S., & Fry, M. (2011). Identity realization and organizational forms: Differentiation and consolidation of identities among Arizona's charter schools. *Organization Science, 22*(3), 554–572.

Kraatz, M. S. (1998). Learning by association? Interorganizational networks and adaptation to environmental change. *Academy of Management Journal, 41*(6), 621–643.

Kraatz, M. S., Ventresca, M. J., & Deng, L. (2010). Precarious values and mundane innovations: Enrollment management in American liberal arts colleges. *Academy of Management Journal, 53*, 1521–1545.

Lau, K., & Lin, C. Y. (2017). Internationalization of higher education and language policy: The case of a bilingual university in Taiwan. *Higher Education, 74*(3), 437–454.

Lee, B. H., Hiatt, S. R., & Lounsbury, M. (2017). Market mediators and the trade-offs of legitimacy-seeking behaviors in a nascent category. *Organization Science, 28*(3), 447–470.

Levinthal, D. A., & March, J. G. (1993). The myopia of learning. *Strategic Management Journal, 14*(2), 95–112.

Marginson, S. (2007). The public/private divide in higher education: A global revision. *Higher Education, 53*(3), 307–333.

Marquis, C., & Tilcsik, A. (2013). Imprinting: Toward a multilevel theory. *Academy of Management Annals, 7*(1), 195–245.

Massey, D., Charles, C., Lundy, G., & Fischer, M. (2003). *The source of the river: The social origins of freshmen at America's Selective Colleges and Universities*. Princeton University Press.

Miller, K. D. (2002). Knowledge inventories and managerial myopia. *Strategic Management Journal, 23*(8), 689–706.

Mohammadi, M., & Rezaei, J. (2020). Bayesian best-worst method: A probabilistic group decision making model. *Omega, 96*, 102075.

Mustafa, A., & Goh, M. (1996). Multi-criterion models for higher education administration. *Omega, 24*(2), 167–178.

O'Brien, J. A. M. E. S., & Dietz, J. O. E. R. G. (2011). Maintaining but also changing hierarchies: What social dominance theory has to say. *Status in management and organizations*, 55–84.

Okyar, O. (1967). Universities in Turkey. *Minerva, 6*(2), 213–243.

Park, D. Y., & Podolny, J. M. (2000). The competitive dynamics of status and niche width: US investment banking, 1920–1949. *Industrial and Corporate Change, 9*(3), 377–414.

Park, S., Yang, D., Cha, H., & Pyeon, S. (2020). The halo effect and social evaluation: How organizational status shapes audience perceptions on corporate environmental reputation. *Organization & Environment, 33*(3), 464–482.

Pearce, J. L. (2011). Introduction: The power of status. In J. L. Pearce (Ed.), *Status in management and organizations* (pp. 1–22). Cambridge Univ. Press.

Peteraf, M., & Shanley, M. (1997). Getting to know you: A theory of strategic group identity. *Strategic Management Journal, 18*, 165–186.

Pfarrer, M. D., Pollock, T. G., & Rindova, V. P. (2010). A tale of two assets: The effects of firm reputation and celebrity on earnings surprises and investors' reactions. *Academy of Management Journal, 53*, 1131–1152.

Phillips, D. J., & Zuckerman, E. W. (2001). Middle-status conformity: Theoretical restatement and empirical demonstration in two markets. *American Journal of Sociology, 107*, 379–429.

Phillips, D., & Podolny, J. M. (1996). The dynamics of organizational status. *Industrial and Corporate Change, 5*, 453–471.

Piazza, A., & Castellucci, F. (2014). Status in organization and management theory. *Journal of Management, 40*(1), 287–315.

Podolny, J. M. (1993). A status-based model of market competition. *American Journal of Sociology, 98*, 829–872.

Pollock, T. G., Lashley, K., Rindova, V. P., & Han, J. H. (2019). Which of these things are not like the others? Comparing the rational, emotional, and moral aspects of reputation, status, celebrity, and stigma. *Academy of Management Annals, 13*(2), 444–478.

Pollock, T. G., Lee, P. M., Jin, K., & Lashley, K. (2015). (Un) tangled: Exploring the asymmetric coevolution of new venture capital firms' reputation and status. *Administrative Science Quarterly, 60*(3), 482–517.

Pólos, L., Hannan, M. T., & Carroll, G. R. (2002). Foundations of a theory of social forms. *Industrial and Corporate Change, 11*(1), 85–115.

Porac, J. F., Thomas, H., Wilson, F., Paton, D., & Kanfer, A. (1995). Rivalry and the industry model of Scottish knitwear producers. *Administrative Science Quarterly, 40*, 203–227.

Prato, M., Kypraios, E., Ertug, G., & Lee, Y. G. (2019). Middle-status conformity revisited: The interplay between achieved and ascribed status. *Academy of Management Journal, 62*(4), 1003–1027.

Pratto, F., Sidanius, J., & Levin, S. (2006). Social dominance theory and the dynamics of intergroup relations: Taking stock and looking forward. *European Review of Social Psychology, 17*(1), 271–320.

Rao, H. (1994). The social construction of reputation: Certification contests, legitimation, and the survival of organizations in the American automobile industry: 1895–1912. *Strategic Management Journal, 15*(1), 29–44.

Rao, H., Monin, P., & Durand, R. (2005). Border crossing: Bricolage and the erosion of categorical boundaries in French gastronomy. *American Sociological Review, 70*(6), 968–991.

Rezaei, J. (2015). Best-worst multi-criteria decision-making method. *Omega, 53*, 49–57.

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

Rider, C. I., & Negro, G. (2015). Organizational failure and intraprofessional status loss. *Organization Science, 26*, 633–649.

Ridgeway, C. L. (1991). The social construction of status value: Gender and other nominal characteristics. *Social Forces, 70*(2), 367–386.

Rindova, V. P., Martins, L. L., Srinivas, S. B., & Chandler, D. (2018). The good, the bad, and the ugly of organizational rankings: A multidisciplinary review of the literature and directions for future research. *Journal of Management, 44*(6), 2175–2208.

Rosa, J. A., Porac, J. F., Runser-Spanjol, J., & Saxon, M. S. (1999). Sociocognitive dynamics in a product market. *Journal of Marketing, 63*, 64–77.

Rosenzweig, E. D., & Easton, G. S. (2010). Trade-offs in manufacturing? A meta-analysis and critique of the literature. *Production and Operations Management, 19*(2), 127–141.

Salmi, J. & Saroyan, A. (2007). League tables as policy instruments. *Higher Education Management and Policy, 19*(2), 1–38.

Sauder, M. (2006). Third parties and status systems: How the structures of status systems matter. *Theory and Society, 35*, 299–321.

Sauder, M., & Espeland, W. N. (2009). The discipline of rankings: Tight coupling and organizational change. *American Sociological Review, 74*, 63–82.

Sauder, M., Lynn, F., & Podolny, J. M. (2012). Status: Insights from organizational sociology. *Annual Review of Sociology, 38*, 267–283.

Scott, J., Ho, W., Dey, P. K., & Talluri, S. (2015). A decision support system for supplier selection and order allocation in stochastic, multi-stakeholder and multi-criteria environments. *International Journal of Production Economics, 166*, 226–237.

Sharkey, A. J. (2014). Categories and organizational status: The role of industry status in the response to organizational deviance. *American Journal of Sociology, 119*(5), 1380–1433.

Soltani, A., Hewage, K., Reza, B., & Sadiq, R. (2015). Multiple stakeholders in multi-criteria decision-making in the context of municipal solid waste management: A review. *Waste Management, 35*, 318–328.

Stuart, T. E., Hoang, H., & Hybels, R. C. (1999). Interorganizational endorsements and the performance of entrepreneurial ventures. *Administrative Science Quarterly, 44*(2), 315–349.

Student Selection and Placement Center. (2014). *Annual Central University Examination Manuals (Yükseköğretim Programları ve Kontenjanları Kılavuzu)*. Available at: https://www.osym.gov.tr/TR,851/2014.html

Tekeli, İ. (2010). *Tarihsel Bağlamı içinde Türkiye'de Yükseköğretimin ve YÖK'ün Tarihi*. Tarih Vakfı Yurt Yayınları.

Thomas-Hunt, M. C., & Phillips, K. W. (2011). The malleability of race in organizational teams: A theory of racial status activation. In J. L. Pearce (Ed.), *Status in management and organizations* (pp. 238–266). Cambridge University Press.

Torres-Olave, B., Brown, A. M., Franco Carrera, L., & Ballinas, C. (2020). Not waving but striving: Research collaboration in the context of stratification, segmentation, and the quest for prestige. *The Journal of Higher Education, 91*(2), 275–299.

Topaler, B., Koçak, Ö., & Üsdiken, B. (2021). Positioning new identities for appeal: Configurations of optimal distinctiveness amid ancestral identities. *Strategic Organization*. https://doi.org/10.1177/1476127021999966

Triantaphyllou, E. (2000). Multi-criteria decision making methods. In *Multi-criteria decision making methods: A comparative study* (pp. 5–21). Springer.

Üsdiken, B., Topaler, B., & Koçak, Ö. (2013). Yasa, Piyasa ve Örgüt Tiplerinde Çeşitlilik: 1981 Sonrasında Türkiye'de Üniversiteler. *Ankara Üniversitesi Siyasal Bilgiler Fakültesi Dergisi, 68*(3), 187–223.

Vigneron, F., & Johnson, L. W. (1999). A review and a conceptual framework of prestige-seeking consumer behavior. *Academy of Marketing Science Review, 1*(1), 1–15.

Wang, L., & Maxwell, S. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods 20*(1), 63–83.

Washington, M., & Zajac, E. J. (2005). Status evolution and competition: Theory and evidence. *Academy of Management Journal, 48*, 282–296.

Weber, M. (1978). In G. Roth & G. Wittich (Eds.), *Economy and society: An outline of interpretive sociology* (Vol. II: 956–975, pp. 980–989). University of California Press.

Zamparini, A., & Lurati, F. (2017). Being different and being the same: Multimodal image projection strategies for a legitimate distinctive identity. *Strategic Organization, 15*(1), 6–39.

Zuckerman, E. W., & Kim, T. Y. (2003). The critical trade-off: Identity assignment and box-office success in the feature film industry. *Industrial and Corporate Change, 12*(1), 27–67.

A Multi-Criteria Decision-Making Approach for Hazmat Transportation

Zafer Yilmaz and Vedat Verter

Abstract Hazardous materials (hazmat) transportation is a niche segment of the transport industry, whereby the cargo imposes risk on the public health, the environment, and the property. A plethora of methodologies have been developed to find optimal routes for hazmat vehicles, which consider hazmat accident risk, population exposure, and cost. While everyone recognizes the relevance of these factors, we are unaware of studies that do not take into consideration all the factors together.

In this study, we consider three main factors, and nine sub-factors together for hazmat transportation and propose a practical methodology to find optimal routes. First, we propose finding the factor weights using AHP methodology. In our case study based in Istanbul, Turkey, where we elicited the views of eight international experts on hazmat transportation, the most important main and sub-factors are found as "Consequences" and "Population exposure," respectively. Next, we propose finding the arcs one composite score for each arc on a road network by combining the data at the sub-factor level using TOPSIS methodology and factor weights found in the first step. Finally, optimal routes between origin–destination (OD) pairs can be identified using ArcGIS network analysis tool, in which total route score is minimized. We compare the optimal routes found using our methodology and the methods used in previous studies. The results are encouraging from the perspective of practical applicability of the three-step procedure we propose in this chapter.

Keywords ArcGIS · AHP · Hazardous Materials Transportation · Multi-Criteria Decision-Making · TOPSIS

Z. Yilmaz (✉)
Business Administration Department, TED University, Kolej Çankaya – Ankara, Turkey
e-mail: zafer.yilmaz@tedu.edu.tr

V. Verter
Supply Chain Management Department, Michigan State University, East Lansing, MI, USA
e-mail: verter@broad.msu.edu

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_14

1 Introduction

Hazardous materials (hazmat) transportation constitutes a niche segment of the transport sector due to the risk caused by dangerous goods carried. If a hazmat accident causes an explosion or hazmat release, its consequences (fatalities, injuries, and property–environment damages) are often more spectacular than an ordinary traffic accident. Therefore, hazmat accidents are accepted as "low risk high consequence" events. For example, 2700 fatalities were reported in 1982 due to a gasoline truck explosion in a tunnel in Afghanistan. A train incident in Quebec/Canada caused 47 fatalities in 2013 and hence the responsible company Irving Oil has been ordered to pay $four million. Considering the possibility of significant loss of human life and cost, no one dares to neglect the very low incident risks on road segments while estimating the hazmat risk of a road segment. The hazmat risk on a road segment is associated with the past hazmat accidents data in traditional hazmat transportation models.

There are two primary stakeholders in hazmat transportation: the *regulators* (legal authorities) who try to decrease the risk on the population, property, and environment by determining the available links on the road network of hazmat transportation, and the *carriers* who focus on cost minimization by entailing routing decisions (Kara & Verter, 2004). The average travel cost for a hazmat vehicle is around $ 250 per hour including the estimated hourly fuel cost (Verma & Verter, 2010). 2.5 billion tons of hazmat were shipped in the USA in 2012 (Ditta et al., 2019), which cost billions of dollars. To increase the buy-in from industry, the transportation cost needs to be a central concern, since any solution causing travel time extensions or delays will cause important amount of cost increases. The prevailing research show that there are three most important factors that affect hazmat transportation decisions:

- Hazmat transportation risk due to hazmat accidents causing explosion or hazmat release.
- Transportation cost (driver cost, travel cost of the hazmat vehicle, etc.).
- Consequences (fatalities, injuries, property and environmental damage, evacuation or clean-up costs, etc.).

Most of the hazmat transportation studies focus on hazmat risk assessment, routing, scheduling, and consequences analyses (Yilmaz et al., 2016). As we discuss in more detail in the next section, while some of the researchers only focus on finding optimal solutions with respect to the hazmat transportation risk, there are also some researchers who propose bi-objective solutions by both considering risk and cost. Some researchers only focus on the consequences since the hazmat accident probabilities are too low and consequences are too high. Recent studies focus on value-at-risk models since consequences involve dramatic losses although the probability of a hazmat accident is too low. Complex mathematical models are developed and exact or heuristic solution procedures are proposed. It is often hard to implement these complex methods in practice, when the decision maker does not have deep methodology knowledge. In addition, we are not aware of any prevailing studies, which focus on *all relevant factors* together in proposing solutions for hazmat transportation.

In this study, in an effort to fulfill the gap mentioned above, we develop a methodology based on multi-criteria decision-making approach, which can be easily used in hazmat transportation practice. Our methodology consists of three steps. First, we determine the *criteria* (factors) that affect hazmat transportation and find their weights using Analytic Hierarchy Process (AHP) methodology. Next, the road segments (arcs) are conceived as *alternatives* and the scores for the road segments are obtained using Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) methodology, which combines the data of all factors into one unique score for each road segment by also including the criteria weights found in the first step. Finally, those scores for the road segments are imported and adopted to be used in ArcGIS network analysis tool and optimal routes between origin destination (OD) pairs are found in which the scores of road segments are minimized. Hence, the proposed methodology includes all the factors together and gives practical solutions for hazmat transportation problems. We present a Case Study in Istanbul, Turkey to find the optimal routes for hazmat vehicles in order to show the efficiency of the proposed methodology.

The organization of the chapter is as follows. The next section describes the key literature on which this chapter has been developed. The methodology for the proposed model is defined in the third section. The case study is presented, and its findings are highlighted in Sect. 4. We conclude with a discussion in Sect. 5 and the opportunities for future research in Sect. 6.

2 Literature Review

There are many studies that focus on different aspects of hazmat transportation in the literature. Despite very low incident risks, hazmat transportation has been a very popular topic in many studies due to the value at risk. Different kinds of hazmat transportation related topics can be found in the literature nevertheless the most popular ones are risk assessment and routing. Erkut and Verter (1998) define the traditional risk as the risk of transporting hazmat B over a unit road segment A (such as a one mile stretch) and they formulate the risk as the multiplication of p_{AB} and CAB where p_{AB} = probability of an incident on the unit road segment A for hazmat B, and CAB is population along the unit road segment A within the neighborhood associated with hazmat B. They claim that estimates of incident probabilities are between 0.1 and 0.8 per million miles. Later studies focus on CAB and they extend the definition of CAB by including cost of damage on nature, evacuation cost, property damage cost, etc. However, the risk p_{AB} is always considered same which is named as traditional hazmat risk.

Hazmat accidents are considered as low probability high consequences events. Kang et al. (2014) propose value at risk model to generate route choices for a hazmat shipment based on a specified risk confidence level. Toumazis and Kwon (2015), in their paper, apply an advanced risk measure, called conditional value-at-risk (CVaR), for routing hazmat trucks which offers a flexible, risk-averse, and computationally tractable routing method that is appropriate for hazmat accident mitigation strategies.

Kang et al. (2014) summarize the risk assessment formulas in different studies as follows (For all min $l \in$P);

Expected risk: $\sum_{(i,j)} \in \mathrm{A}^l\ p_{ij}C_{ij}$
Incident consequence: $\sum_{(i,j)} \in \mathrm{A}^l\ C_{ij}$
Incident probability: $\sum_{(i,j)} \in \mathrm{A}^l\ p_{ij}$
Perceived risk: $\sum_{(i,j)} \in \mathrm{A}^l\ p_{ij}\ (C_{ij})^{\mathrm{q}}$
Mean-variance: $\sum_{(i,j)} \in \mathrm{A}^l\ (p_{ij}C_{ij} + kp_{ij}\ (C_{ij})^2)$
Disutility: $\sum_{(i,j)} \in \mathrm{A}^l\ p_{ij}(\exp(kC_{ij} - 1))$
Maximum risk: $\max_{(i,j)} \in \mathrm{A}^l\ C_{ij}$
Minimax (Uncertain probabilities): $\min_w \max_p \sum_{(i,j)} \in \mathrm{A}^l\ w_{ij}\ (p_{ij}C_{ij} + c_{ij})$
Conditional probability, $\sum_{(i,j)} \in \mathrm{A}^l\ p_{ij}C_{ij} \sum_{(i,j)} \in \mathrm{A}^l\ p_{ij}$

where l is the number of links, p_{ij} is accident probability on link $(i, j) \in A$, c_{ij} is cost on the link $(i, j) \in A$, C_{ij} is accident consequence on the link $(i, j) \in$A, A is the number of road segments, P is the set of available paths for shipment s.

The studies summarized above focus on accident risk, the consequences, and sometimes both. Relatively in all studies, the incident risk remains constant along an arc which is p_{ij}: accident probability on arc (i, j).

In some of the researches, in addition to focusing on hazmat risk assessment and consequences, cost is also included. Kara and Verter (2004) find bi-level solutions by considering both risk and travel costs on road segments to meet the carriers' travel cost concerns. The literature surveys; Erkut et al. (2007), Yilmaz et al. (2016), and Ditta et al. (2019) are advised for detailed information about hazmat transportation risk assessment studies.

Eventually, hazmat accident risk, consequences, and costs are the main factors on a road segment that researchers concentrate on. The researchers propose very complex algorithms to find optimal or heuristic solutions while finding the best routes between OD pairs. They usually develop complex math models and use some solvers (i.e., GAMS and CPLEX) to find the optimal routes. In most of the studies, as it is summarized by Kang et al. (2014), multiplication of hazmat risk and possible consequences or either only hazmat risk or consequences are minimized in the objective functions of the math models.

In this study, we believe that the factors that affect the hazmat transportation decisions (i.e., risk, consequences, and cost) will have different weights on the decisions that should be taken into consideration while finding optimal routes. So, AHP methodology is used in our study to find the weights of the main and sub-factors. The experts who work on hazmat transportation and published articles in Web of Science indexed journals are asked to fill out a questionnaire. Their replies are analyzed by using AHP methodology in order to find the importance weights of the factors. Next, the road segments are conceived as alternatives in TOPSIS methodology and the risk, cost and consequences related data for each alternative (road segment) are combined into one score for each alternative by using TOPSIS methodology. Finally, ArcGIS network analysis tool which uses Dijkstra algorithm is used to find optimal routes between OD pairs in which the scores obtained in TOPSIS methodology are minimized. The details of the methodology are given in the next section.

There are a very limited number of hazmat transportation related studies in which multi-criteria decision-making methodologies (i.e., AHP and TOPSIS) are included. For example, Sattayaprasert et al. (2008) propose a method to create a risk-based network for hazmat logistics by route prioritization with AHP. This research has been conducted with the information of short-range freight transportation mainly for gasoline movement and for a specific case and location only. The researchers define main and sub-criteria and high-medium and low-level risks for those criteria. Our study differs from this research since in addition to the criteria proposed in Sattayaprasert et al.'s (2008) research, we include hazmat vehicles and other types of vehicle accident risks in our main and sub-criteria. The other difference is that we assign scores for each road segment by using TOPSIS methodology rather than prioritizing the candidate routes which makes our study to focus on all possible routes. Li et al. (2019), propose a decision support model for risk management of hazardous materials road transportation. They use Fuzzy-AHP to build a hierarchical risk assessment system and determine the importance rating of each risk factor. They focus on direct and indirect risk factors. In our study, by using AHP, in addition to the importance rating of risk factors, we find the importance ratings of cost and consequences factors. Jun and Wei (2010), in their study which is presented in a conference just find the weights of safety, time, and profit by using AHP without including any sub-factors or consequences. Huang (2006) in his study, considers safety, costs, and security. GIS is used to quantify the factors on each link in the network that contributes to each of the evaluation criteria for a possible route. AHP is used to assign weights to the factors; exposure, socio-economic, risks of terrorism, traffic conditions, and emergency response. Each route can then be quantified by a cost function and the suitability of the routes for HAZMAT transportation can be compared. We focus on three main and nine sub-factors which makes our study more comprehensive compared to Huang's study in which only five main factors are considered. Chen et al. (2019) propose a PHFLTS- and TOPSIS-Integrated Multi-Perspective Approach to evaluate and select HazMat Transportation Companies. A case study is applied in China and five candidate companies are sorted by using TOPSIS. TOPSIS is used in a totally different way in our study to find the scores of each road segment.

Geographic Information Systems (GIS) is now widely used in hazmat transportation studies (Ditta et al., 2019) since GIS makes it is possible to attain perfect information about some attributes of road networks (i.e., distances, times, and traffic). For example, Zografos and Androutsopoulos (2008) estimate the total population within a specific selected area by using GIS. Kawprasert and Barkan (2008) use GIS to compute the distance and the type of traffic control system on a HAZMAT transportation network. Rashid et al. (2010) develop a GIS application to build a spatial model for the assessment on the consequences of liquefied petroleum gas release accidents in road transportation. Kim et al. (2011) offer a GIS tool that provides online routing instructions for HAZMAT vehicles given the vehicle's current location and updated information concerning traffic and weather conditions. Samanlioglu (2013) proposes a multi-objective mathematical model for the industrial hazardous waste location-routing problem in which the data is obtained by a combination of GIS software and the regional geographical database. Readers may refer to Holeczek (2019)'s review study for all researches in which GIS is used. Different from the above studies in which GIS is used, we use GIS to find the optimal routes between OD pairs since GIS network analysis tool provides a wide range of network analysis methodologies such as a Dijkstra-based methodology to find shortest paths.

3 Methods

Referring to the past hazmat transportation related researches explained in the introduction and literature review sections, there are three main factors considered in hazmat transportation which are: Risk, Cost, and Consequences.

In almost all hazmat transportation models, as it is summarized by Kang et al. (2014), "$p_{ij}C_{ij}$" is the objective of the model where p_{ij} and C_{ij} are risk and consequences on the road segment (arc) *ij*. A general road network for the hazmat transportation is given in Fig. 1 to explain the objective function of the traditional hazmat transportation models. It is accepted that a hazmat vehicle is going to travel from Origin (node 1) to its Destination (node 7). Let p_{ij} and C_{ij} be the hazmat accident risk and possible consequences on arc *ij*, respectively. x_{ij} is the binary variable and becomes 1 when the arc *ij* is included in the solution.

The objective of the traditional hazmat risk assessment model is to minimize total hazmat accident risks and consequences which is formulated as:

$$\sum \left(p_{ij} C_{ij}\right) x_{ij} \tag{1}$$

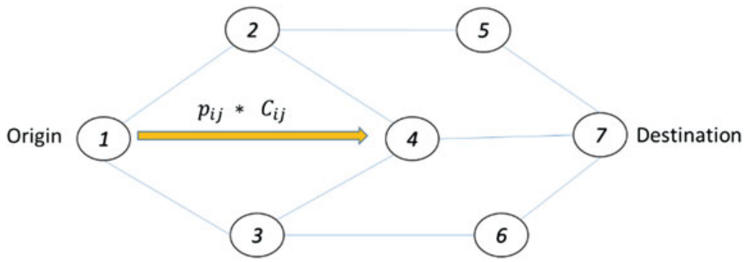


Fig. 1 A general road network for hazmat transportation

Most of the models consider hazmat risk and its consequences as past hazmat accidents and the number of population who will be affected from the accident, respectively. However, we accept that in addition to past hazmat accidents, other risks (i.e., the past accident data of other vehicles and types of the roads) should also be considered in hazmat risk assessments in order to find more realistic solutions. The probability of a hazmat accident is very low (1×10^{-6} on the average on a 1 mile arc in the USA), and there are many arcs on which there are no past hazmat accidents so that the probability becomes zero on those arcs. However, it will not be realistic to accept a zero hazmat risk on arcs since there is always an accident risk if a vehicle is traveling on an arc with a traffic flow. Thus, the probabilities of accidents caused by other vehicle types should also be included in risk assessments on arcs to have a more realistic risk assessment. We also believe that the road types are effective on the occurrence of accidents. For example, the accident probabilities will not be the same in one-way and two-way roads. Eventually, for the hazmat risk assessment main factor we define three sub-factors; past hazmat accidents, past other types of accidents, and road types.

We assign two sub-factors which are driver cost and fuel usage cost for the "Cost" main factor. The last main factor considered in this study is "Consequences" and we assign four sub-factors (referring to past researches), which are population living along the arcs (which is important for fatality and injury estimations), property damage (damage on the hazmat vehicle, other vehicles, and surrounding properties after a hazmat vehicle explosion), environmental damage, and evacuation and clean-up costs (if there is a necessity to evacuate the accident region or clean up the region after a hazmat release). Past researches consider only one factor (usually population) or two factors for consequence analysis. However, our study, to the best of our knowledge is the first study that consider all the factors together. We include all possible main and sub-factors (criteria) in our study which are depicted in Table 1.

The traditional hazmat transportation model (Eq. 1) should be updated with the following formula to include all factors in the objective function:

$$\text{Min} \sum \left[\left(p_{ij}^{T}\, C_{ij}^{T} \right) x_{ij} + \text{Min} \sum \left(\text{Cost}_{ij} \right) x_{ij} \right] \tag{2}$$

where p_{ij}^{T} is the total risk (including R1, R2, and R3 factors), C_{ij}^{T} is all possible consequences (including CN1, CN2, CN3, and CN4), Cost_{ij} is the cost of traveling on arc ij, and x_{ij} is the binary variable and becomes 1 when the arc ij is included in

Table 1 Main and sub-criteria

	Main criteria		
	Risk	Cost	Consequences
Sub-criteria	Past hazmat accidents (R1)	Fuel usage (C1)	Population (CN1)
	Past other accidents (R2)	Driver cost C2)	Property damage (CN2)
	Road type (R3)		Environmental damage (CN3)
			Evacuation and clean-up (CN4)

the solution. We realize that as more factors and objectives are included, the problem becomes more complicated to be solved in a reasonable time. Therefore, in order to reduce the complexity of the problem and find solutions in a reasonable time, we propose a new approach for the hazmat transportation problems which is explained in the following paragraphs.

It is obvious that the aforementioned factors will affect the optimal solution with different weights. So, in the first step of our study, we find the criteria weights by using *AHP methodology* which is developed by Saaty, (Saaty, 1977) for solving decision-making problems. It is one of the most effective multi-criteria decision-making methodology (MDCM) used for finding criteria weights. AHP is described in the following steps:

Decision-Making Problem: In this step, decision points and the factors affecting the decision points are determined.

Creating a Cross-Factor Comparison Matrix: The inter-factor comparison matrix is a dimensional square matrix which is given below. The values of matrix components on the diagonals are 1 since the relevant factor is compared with itself.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

The comparison of the factors is made one to one and mutually according to their importance values. Saaty's factor scale (see Table 2) is used for one-to-one comparison of factors.

Comparisons are made for values that lie above the diagonal of all values of the comparison matrix. For the components under the diagonal, Eq. 3 is used.

$$a_{ji} = \frac{1}{a_{ij}} \tag{3}$$

Table 2 Saaty's 1–9 comparison scale

Level	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
3	Moderate importance	Experience and judgement slightly favor one activity over another
5	Strong importance	Experience and judgement strongly favor one activity over another
7	Very strong importance	An activity is favored very strongly over another; its dominance 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

Determining Percentage Importance of Factors: In order to determine the weights of the factors, column vectors forming the comparison matrix are used and column B with nxn components is formed. Equation 4 is used to calculate column B vector values.

$$b_{ij} = \frac{a_{ij}}{\sum_{i=1}^{n} a_{ij}} \tag{4}$$

C matrix shown below is created when B column vectors are combined in a matrix format.

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

The importance values of the factors relative to each other can be obtained by using Eq. 5.

$$W_i = \frac{\sum_{j=1}^{n} c_{ij}}{n} \tag{5}$$

Measuring Consistency in Factor Benchmarking: Consistency Ratio (CR) and the priority vector provide the possibility to test the consistency of the comparisons. AHP is based on the essence of the CR calculation by comparing the number of factors with a coefficient λ called the Basic Value. After calculating λ, the Consistency Index (CI) can be calculated using the Eq. 6.

$$CI = \frac{\lambda - n}{n - 1} \tag{6}$$

CR is obtained using Eq. 7 and Random Indicator (RI) is shown in Table 3.

$$CR = \frac{CI}{RI} \tag{7}$$

Table 3 Random consistency index

Random consistency index					
N	RI	N	RI	N	RI
1	0	6	1.24	7	1.32
2	0	4	0.9	8	1.41
3	0.58	5	1.12	9	1.45

In addition, a calculated CR value of less than 0.10 indicates that the comparisons made by the decision maker are consistent. Eventually, the weights of the factors (criteria) determined in this study are found by using AHP methodology.

In the second step of this study, the arcs are considered as alternatives (a_{ij}). For all alternatives, we assign the values considering the sub-criteria. In order to assign the risk values; number of the past hazmat accidents and other types of accidents on the arcs (alternatives) are used. Types of the roads that the hazmat vehicles travel are also importantly affect the accident risks and road type values for the arcs are assigned with respect to number of lanes. The related studies, reports, and open-source data are used to assign four consequences (CN1, CN2, CN3, and CN4) values for the arcs. Cost values (C1 and C2) for the arc are assigned considering unit cost for the drivers and fuel cost with respect to distance traveled.

In third step, we use TOPSIS methodology, which is proposed by Hwang and Yoon (1981). It is a very commonly used MCDM methodology for selecting the best alternatives or sorting the alternatives.

In general, the process for the TOPSIS algorithm starts with forming the decision matrix (D) representing the satisfaction value of each criterion with each alternative. The rows show alternatives while columns show criteria in the matrix below:

$$D = \begin{bmatrix} \mathrm{a}_{11} & a_{12} & \cdots & \mathrm{a}_{1\mathrm{n}} \\ \mathrm{a}_{21} & a_{22} & \cdots & \mathrm{a}_{2\mathrm{n}} \\ \vdots & \vdots & & \vdots \\ \mathrm{a}_{\mathrm{m}1} & \mathrm{a}_{\mathrm{m}2} & \cdots & \mathrm{a}_{\mathrm{nm}} \end{bmatrix}$$

Next, the matrix is normalized with a desired normalizing scheme. In the basic matrix $i = 1,2,\ldots..,\mathrm{m}$ and $j = 1,2,\ldots\ldots.,\mathrm{n}$ values are normalized by using vector normalization below:

$$r_{ij} = \frac{a_{ij}}{\sqrt{\sum_{i=1}^{m} \mathrm{a}_{ij}^2}}$$

Next, the normalized values (r_{ij}) are multiplied by the criteria weights (we use criteria weights obtained using AHP methodology) to find $\mathrm{V_{ij}}$ values.

Subsequently, the positive-ideal A^+ (associated with the criteria having a positive impact) and negative-ideal A^- (associated with the criteria having a negative impact) values are calculated among $\mathrm{V_{ij}}$ values for each criterion.

The distance of each alternative (S_i^+ and S_i^-) from positive and negative ideal values (A^+ and A^-) is calculated with a distance measure.

Later, the similarity to the worst condition ($C_i^{+)}$ is calculated:

$$C_i^+ = \frac{S_i^-}{S_i^+ + S_i^-}$$

Finally, the alternatives are ranked based on their relative closeness to the ideal solution.

Eventually, after applying TOPSIS, we find unique values for each alternative a_{ij} in which all sub-criteria values are considered. So, we propose Eq. 8 to be used instead of Eq. 2.

$$\sum \left(a_{ij}\right) x_{ij} \tag{8}$$

where a_{ij} is the value (that is obtained from AHP-TOPSIS methodologies) for arc ij which include all the aforementioned factors.

In the last step, rather than writing a linear model which also includes Eq. 8 in the objective function, we use ArcGIS to find the optimal routes between the OD pairs in which those unique values of alternatives are minimized. ArcGIS usage is a very practical way of finding optimal routes. The changes in the related data may require writing new linear models in the studies in which linear models are proposed. However, those changes can easily be adopted for ArcGIS which makes it a more practical tool for finding optimal routes.

Hazmat transportation problems are generally large-scale problems and excessive computational requirements are needed so that many researchers propose heuristic solutions. The bi-objective (risk and cost) hazmat problems are usually solved by the proposed heuristic models. We propose a new approach to overcome the large-scale hazmat transportation problems. Another important contribution of the study is that all factors (including risk and cost) of hazmat transportation are considered together in this study so as to find optimal solutions and hence both the concerns of carries (main concern is cost) and the legal authorities (main concern is risk) are satisfied. In addition, our methodology gives the opportunity to ban some of the arcs on the road network and find optimal solutions without including those banned arcs.

4 Case Study Findings

In this section, we present a *Case Study* in Istanbul, Turkey to find the optimal routes for hazmat vehicles in order to show the efficiency of the proposed methodology. In the first step of our analysis, the experts who work on hazmat transportation and recently published articles in Web of Science indexed journals are asked to fill out a questionnaire that includes comparisons for main and sub-criteria. Eight experts from the USA, Canada, Brazil, Italy, China, Iran, and Turkey filled out our questionnaire in helping this study. The experts are chosen from the countries in which hazmat transportation has an important share. Note that those counties are chosen from the most crowded continents (Asia, North and South America, Europe), which are located in different regions of the world. Those experts are aimed to bring a multinational and cross-continental perspective while assigning

Table 4 Main and sub-criteria weights

Main criteria	Weight	Sub-criteria	Weight-1	Weight-2
Risk	0.249	Past hazmat accidents	0.3415	0.0852
		Past other type vehicle accidents	0.1780	0.0444
		Road type	0.4805	0.1198
Cost	0.109	Fuel usage	0.5000	0.0543
		Driver cost	0.5000	0.0543
Consequences	0.642	Population	0.6025	0.3868
		Property damage	0.0743	0.0477
		Environmental damage	0.2334	0.1499
		Evacuation and clean-up cost	0.0898	0.0576
Total	1.000		Total	1.000

importance weights for main and sub-criteria. Their replies are analyzed using AHP methodology steps explained in the third section in order to find the importance weights of the factors. The importance weights for main and sub-criteria are given in Table 4.

We find that the most important main criteria for hazmat transportation is "Consequences" with a score 0.642. "Risk" and "Cost" have 0.249 and 0.109 criteria weighs. Weight-1 column shows the sub-criteria weights for each main criterion. For example, the sub-criteria weights for "Risk" main criteria are found as 0.3415, 0.1780, and 0.4805 for past hazmat accidents, past other type vehicle accidents, and road type sub-criteria, respectively. Main criteria weights (scores in Weight column) and sub-criteria weights (scores in Weight-1 column) are multiplied to find the sub-criteria scores in general (Scores in Weight 2 column). For example, among 9 sub-criteria, "Population" has the greatest weight score which is 0.3868. Second most important sub criteria is found as "environmental damage" with a score 0.1499. All CR values in our AHP analysis are below 0.06, which is less than the edge value 0.1 that makes our analysis consistent.

In the second step, we obtain and adopt the case study data for TOPSIS analysis which will be focused in the third step. We select Istanbul-Bahçelievler for the case study. Istanbul is the most crowded city in Turkey with a population of over 15 million. Thousands of tankers carry fuels to meet the demand in fuel stations. Many fuel stations unfortunately are located in the crowded parts of the urban areas in Istanbul. Hazmat transportation is very important in terms of both cost and risk. An explosion of a hazmat vehicle may cause hundreds of casualties, thousands on injuries, and million dollars of cost. Istanbul has a great road network and we prefer to focus on Bahçelievler which is a town of Istanbul for our case analysis. Bahçelievler road network includes 8508 arcs. Those arcs are considered to be alternatives in TOPSIS analysis. We need the scores of those arcs with respect to nine sub-criteria in order to apply TOPSIS analysis. The explanation of how the related data for nine sub-criteria considered in this study is adopted for our case study is given in Table 5.

Table 5 DATA collection and adaptation for the case study

Factor (Criteria)	DATA collection and adaptation
Past hazmat accident	The real data for 2018 are obtained from the Istanbul police headquarters. There is only yearly accidents data for towns of Istanbul. We distributed Bahçelievler's hazmat accidents for the year 2018 to its districts evenly with respect to a total distance of road networks of the districts.
Past other types of accidents	The accidents data for all districts of Bahçelievler for 2018 are obtained from the Istanbul police headquarters. The district accidents data are evenly distributed to the arcs with respect to the arc lengths.
Road type	We use Istanbul road network data in terms of "shape file" format to be used in ArcGIS tool. Road types are available in this file.
Driver cost	The average driver cost for Istanbul is accepted to be ₺ 4000 (₺: Turkish Lira) per month. The drivers work 40 hour a week and 160 hours per month. The hourly rate is accepted to be ₺25. We first assigned average hazmat vehicle speeds on the arcs with respect to road types. Next, we find the time to travel on each arc. The hourly driver cost is adopted for the arcs according to the travel times on those arcs.
Fuel cost	Fuel costs for the arcs are assigned with respect to arc lengths. So, the arcs with greater lengths are assigned greater fuel costs.
Population	The population of the districts of Bahçelievler for the year 2019 is available on the official website of Bahçelievler government given below: http://www.bahcelievler.gov.tr/bahcelievler-in-nufus-durumu The district populations are evenly distributed to the arcs with respect to the arc lengths. So, the longer arcs are assigned greater population values.
Property damage	The property damage values for the arcs are assigned with respect to road widths. We consider that the traffic density will be higher in the arcs with wider widths causing greater property damage due to greater presence of vehicles around hazmat accident point.
Environmental damage	We assign "1" for the arcs closer to the parks, green areas are and "0" for the arcs away from the parks and green areas. So, we force the ArcGIS tool to select the arcs with "0" values.
Evacuation and clean-up	We consider that evacuation costs will be higher for crowded areas. We assign greater values for the arcs closer to schools, hospitals, and libraries.

In the third step, we apply TOPSIS analysis explained in Sect. 3. We use the sub-criteria weights found in first step while applying TOPSIS. We combine nine sub-criteria scores for each arc to one unique score by applying TOPSIS. Eventually, we find scores for 8508 arcs. Those scores are between 0 and 1. Since the objectives in all criteria are minimization, the arcs with the scores closest to zero are considered to be best options for finding optimal routes.

In the last step of this study, we use ArcGIS network analysis tool which minimizes those scores while finding optimal routes between OD pairs. TOSIS scores of all arcs are imported and ArcGIS network analysis road networks are constructed to find optimal routes. We first find optimal routes for four OD pairs with respect to TOPSIS score, which includes and represents all nine criteria scores. Next, we also find optimal routes with respect to the objectives; minimizing past hazmat accidents, past all vehicle accidents, travel cost, and population encounter

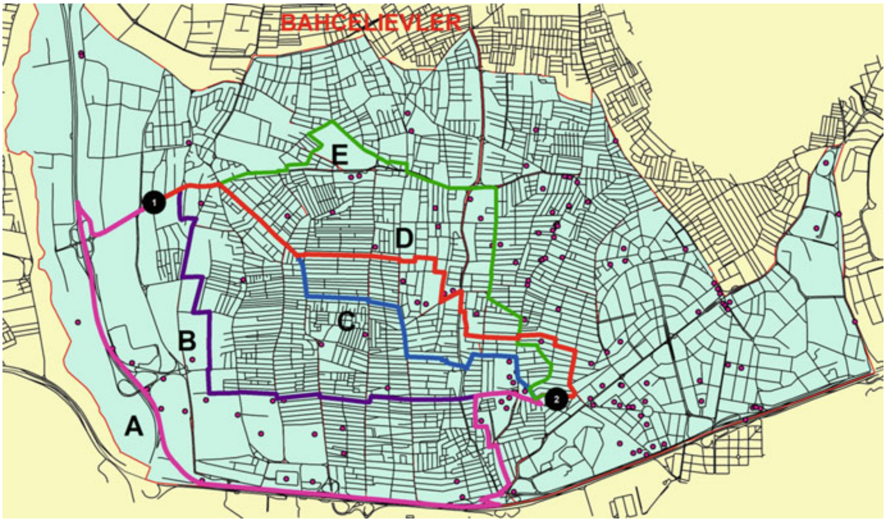


Fig. 2 Routes found for first OD pair

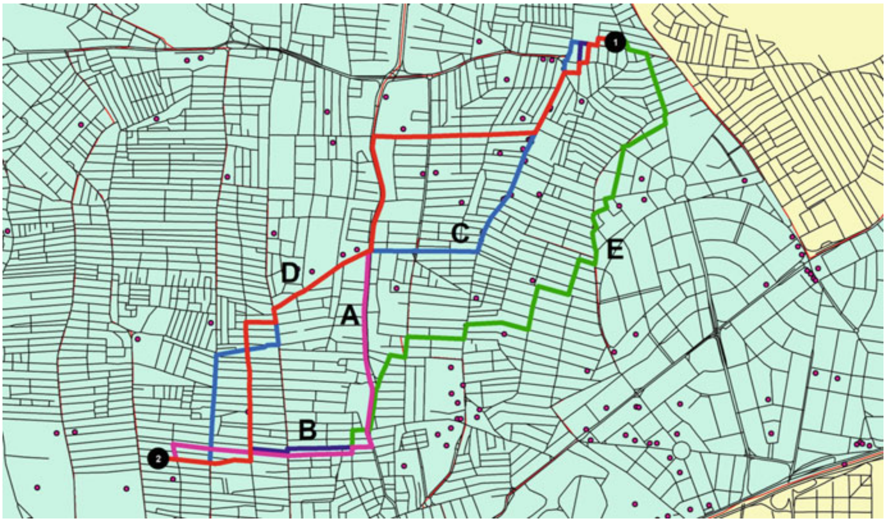


Fig. 3 Routes found for second OD pair

the hazmat risk in order to apply a comparison analysis. Figures 2, 3, 4, and 5 show the routes found from origin 1 to destination 2. We select different origins and destinations for each trial. The optimal routes found for 5 different objectives are represented with different colors. The routes A, B, C, D, and E given in red, pink, dark blue, blue, and green colors are the optimal routes found for the objectives which minimize TOSIS score, past total accidents (all vehicle accidents),

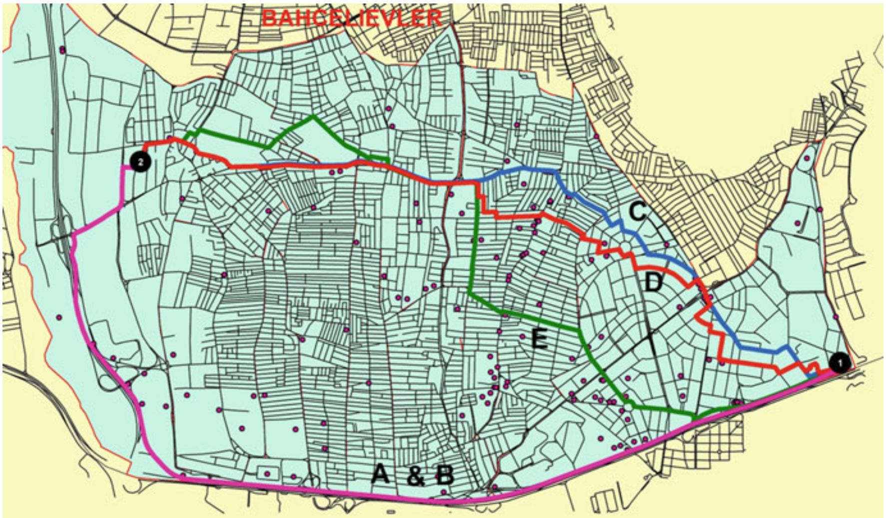


Fig. 4 Routes found for third OD pair

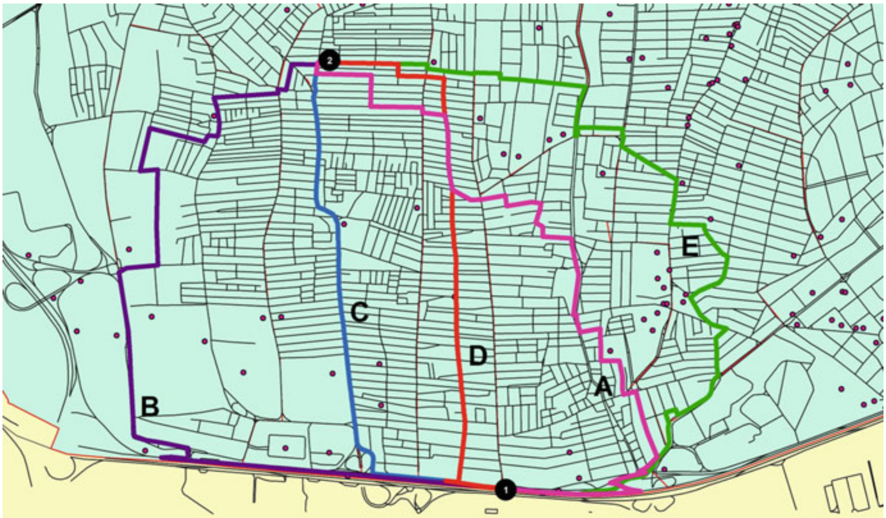


Fig. 5 Routes found for fourth OD pair

past hazmat vehicle accidents, cost, and population respectively. Bahçelievler town region is given in blue shaded area in the figures, in which black lines show the road network. Color codes of objectives and scores of routes are given in Table 6.

We find different optimal routes considering 5 different objectives for the first, second, and fourth OD pairs. However, the routes found for the objectives which minimize total accidents and hazmat accidents are same in third OD pair (the route with pink color in Fig. 4).

Table 6 shows route names and color codes, 5 different objectives, and the scores found for those objectives for 4 OD pairs.

For each OD pair, there are 5 routes found for 5 objectives and their corresponding scores are depicted in Table 6. When we focus on the routes found for the first OD pair; first row shows the scores of route A with pink color considering 5 criteria which are given in "Route Scores Considering 5 Criteria" columns. The objective of first row is minimizing past total accidents and the corresponding past total accident score for route A is found as 582.63. The others scores given in the other columns of route A are the scores of route A considering other 4 criteria and distance. The scores of 4 other routes (B, C, D, and E) are given in the following rows of Table 6 for first OD pair. When the scores for "Total Ac." column for first OD pair is compared, we see that best score belongs to route A (marked with *) since the objective is minimizing past total accidents. The *-marked scores in the columns of Table 6 show the best routes for the corresponding objectives.

In this study, we propose to use TOPSIS scores for finding the optimal routes which take into account all criteria. We compare scores of red routes (in which TOPSIS scores are minimized) with the scores of other routes so as to show if our proposal is realistic. For example, when we focus on "Population" column, of course, the best routes are green color routes in which the only objective is to minimize hazmat risk on the population. We realize that the second best scores for 4 OD pairs considering population belong to red routes. In addition, the red routes again have the second best scores for "Cost" and "Distance." When the accidents are considered (total and only hazmat accidents), the red routes give the third best scores after the total accidents and hazmat accidents routes. Eventually, if we focus only on one objective while finding the optimal route for a hazmat vehicle, we should prefer the routes found for the corresponding objective. However, if we aim to find optimal routes by considering all criteria, our proposed methodology which uses the scores obtained using AHP-TOPSIS methodologies gives the best solutions for finding optimal routes.

We conduct a sensitivity analysis and examine the effects of the changes in the weights of the main criteria (risk, cost, and consequences) focused on this study. Table 7 depicts 5 different weight compositions of the main criteria. The weights given in TOPSIS 1 column are the originally found weights (recall Table 4). We assign the same weights (0.333 for each criterion) for the main criteria in TOPSIS 2 column.

We find the optimal routes (Fig. 6) which minimize the TOPSIS scores with respect to 5 different weight compositions.

We realize that optimal routes may change when the main criteria weights are changed since we found 5 different optimal routes for 5 different weight compositions. Table 8 shows route names and color codes, 5 different objectives in which TOPSIS scores (T1, T2, T3, T4, and T5) are minimized with respect to 5 different weight compositions and the scores found for those objectives (accidents,

Table 6 Case study scores for 4 trials

				Route scores considering 5 criteria					
OD pair	Route name	Route color	Route objective (Min. of)	Total Ac.	Hazmat Ac.	Cost (TRY)	TOPSIS	Population proportion	Distance (meters)
1	A	Pink	Total Ac.	582.63*	80.04	20.54	4.71	494.84	7653.77
	B	D. Blue	Hazmat Ac.	703.63	72.31*	13.98	3.50	351.19	5158.67
	C	Blue	Cost	1138.63	138.87	11.58*	3.18	265.92	4321.44
	D	Red	TOPSIS	1265.63	132.04	13.27	2.64*	248.07	4918.36
	E	Green	Population	2490.63	235.87	14.89	4.18	211.33*	5574.11
2	A	Pink	Total Ac.	440.63*	62.13	10.87	1.84	174.23	4026.68
	B	D. Blue	Hazmat Ac.	488.63	61.75*	10.79	1.84	173.85	3999.27
	C	Blue	Cost	1088.63	115.79	9.42*	2.40	172.07	3539.14
	D	Red	TOPSIS	665.63	67.49	10.27	1.76*	162.45	3801.38
	E	Green	Population	1821.78	146.86	11.58	2.13	131.49*	4260.89
3	A	Pink	Total Ac.	1053.50*	67.62	22.27	3.88	452.15	8342.36
	B	D. Blue	Hazmat Ac.	1053.50	67.62*	22.27	3.88	452.15	8342.36
	C	Blue	Cost	3241.57	186.75	17.36*	3.83	286.22	6538.47
	D	Red	TOPSIS	2707.57	161.88	19.14	3.32*	280.66	7150.97
	E	Green	Population	3709.57	248.36	19.95	4.47	245.45*	7405.78
4	A	Pink	Total Ac.	442.30*	64.22	11.91	2.24	252.66	4408.84
	B	D. Blue	Hazmat Ac.	640.73	40.35*	13.05	3.25	380.86	4892.25
	C	Blue	Cost	1076.73	87.83	7.29*	2.01	205.45	2678.07
	D	Red	TOPSIS	693.70	61.44	7.65	1.61*	180.62	2808.46
	E	Green	Population	2391.27	229.02	12.90	3.21	141.50*	4791.12

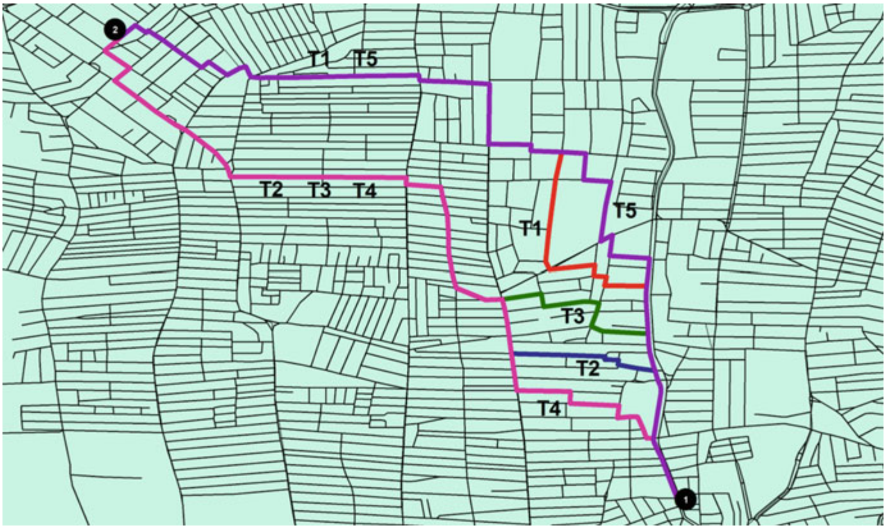


Fig. 6 Routes found for different criteria weights

hazmat accidents, cost, population exposure, and distance traveled). Optimal scores for each column are marked with * and ♦.

Green route (T3) gives the best scores if “Total Accidents” and “Hazmat Accidents” columns are considered since the criterion weight assigned for risk is the highest (0.6) in this route. Pink route (T4) gives the best scores if “cost” and “distance” traveled are considered since we assign 0.6 for cost criterion in this route so as to mostly minimize the cost of travel. The criterion score that we assign for “Consequences” is 0.6 in Purple route (T5). Hence, we obtain best score of “Population” column in route T5 since the main focus is on consequences in this route. Eventually, the sensitivity analysis proves that the optimal routes may change when the weights of the criteria are changed. In this study, we first find the criteria weights with respect to the opinions of experts who study on hazmat transportation and next find optimal routes by taking into account those weights. However, the decision makers may put more emphasis on some criteria which causes changes in the criteria weights and optimal routes can be found accordingly using our methodology.

Table 7 Different weight compositions of the main criteria

AHP Weight	TOPSIS 1 (T1)	TOPSIS 2 (T2)	TOPSIS 3 (T3)	TOPSIS 4 (T4)	TOPSIS 5 (T5)
Risk	0.249	0.333	0.600	0.200	0.200
Cost	0.109	0.333	0.200	0.600	0.200
Consequences	0.642	0.333	0.200	0.200	0.600

Table 8 Scores of the routes found in Fig. 6 for different objectives

Route name	Route color	Obj. (Min.of)	T1	T2	T3	T4	T5	Total Ac.	Haz. Ac.	Cost TRY	Pop.	Dist. (m)
T1	Red	T1	1.806*	1.969	2.222	1.987	1.808	705.3	73.8	9.185	192.7	3403
T2	D.Blue	T2	1.840	1.928*	2.179	1.894	1.823	513.3	65.2	8.433	207.6	3131
T3	Green	T3	1.845	1.943	2.176*	1.927	1.834	485.3 ◆	63.4◆	8.651	210.1	3205
T4	Pink	T4	1.864	1.947	2.234	1.893*	1.840	574.3	69.9	8.389◆	208.2	3115◆
T5	Purple	T5	1.808	1.972	2.245	1.985	1.807*	788.3	77.0	9.176	191.6◆	3404

5 Discussion

In this study, we propose a practical methodology to find optimal routes for hazmat vehicles considering three main and nine sub-criteria. The main focuses in the studies which are introduced in the literature survey are hazmat accident risk and the consequences. Most of the studies consider population exposure as consequence and propose solutions in which the hazmat risk on the population living along arcs is decreased. Some studies also consider risk and cost together by proposing bi-level models to find optimal routes. Different from the other studies, we include all those risks and consequences and additionally the other vehicle's accident risk in our proposed model which we believe shows the strength of our study. The results of the case study are important to explain the contribution of our study. The score change proportions between optimal routes for 4 OD pairs are given in Table 9. Most important factors in hazmat transportation are considered as hazmat accident risk, cost, and population exposure referring to the past studies. Hence, we compare our proposed study results with the results in which only hazmat accident risk or cost or population exposure are considered. For example, if the results for OD pair 4 are compared we see that when the optimal route (red route) found using our proposed methodology (TOPSIS row in OD Pair 4 in Table 9) is used hazmat accident risk will be increased by 1.52 times rather than using the optimal route found in which hazmat accident risk is minimized. The red route will also increase the cost and population exposure 1.05 and 1.28 times, respectively. However, if the route which only minimizes population exposure (Population row in OD Pair 4) is used, hazmat accident risk and cost will be increased 5.68 and 1.67 times. The increases are higher

Table 9 Comparison of optimal routes for 4 OD pairs

OD Pair	Route	Hazmat Ac.	Cost (TRY)	Population
1	Hazmat Ac.	1.00	1.21	1.66
	Cost	1.92	1.00	1.26
	TOPSIS	1.83	1.15	1.17
	Population	3.26	1.29	1.00
2	Hazmat Ac.	1.00	1.15	1.32
	Cost	1.88	1.00	1.31
	TOPSIS	1.09	1.09	1.24
	Population	2.38	1.23	1.00
3	Hazmat Ac.	1.00	1.28	1.84
	Cost	2.76	1.00	1.17
	TOPSIS	2.39	1.10	1.14
	Population	3.67	1.15	1.00
4	Hazmat Ac.	1.00	1.79	2.69
	Cost	2.18	1.00	1.45
	TOPSIS	1.52	1.05	1.28
	Population	5.68	1.77	1.00

than TOPSIS route if the routes which minimize hazmat accident risk or cost are used (see the scores in OD Pair 4). So, using our proposed methodology will give better solutions than focusing only on the important factors separately as in the previous studies.

Our methodology also considers other factors such as other types of vehicle's accident risk, property damage, environmental damage, evacuation, and clean-up cost while finding optimal routes for hazmat vehicles. All nine sub-criteria focused in this study affect the optimal route with different weights and those weights are found using AHP methodology. We consult the researchers who focus hazmat transportation in their studies rather than consulting legal authorities (who prefer to decrease hazmat risk on the population) or carriers (who prefer to decrease the cost of transportation). To the authors' best of knowledge, this is the first study in which weights of three main and nine sub-factors are found.

6 Conclusions

In this study, we determine three main and nine sub-factors together for hazmat transportation referring to previous studies. Next, the researchers who publish hazmat transportation related studies are asked to compare main and sub-factors and their comparison scores are used to find importance weights of main and sub-factors using AHP methodology. We find that the most important main criteria for hazmat transportation are "Consequences" with a score 0.642. "Risk" and "Cost" have 0.249 and 0.109 criteria weighs. Regarding nine sub-criteria, "Population" has the greatest weight score which is 0.3868 and second most important sub-criteria are found as "environmental damage" with a score 0.1499. To the best of our knowledge, our study is the first study in which all criteria weights are found.

Next, the arcs on a road network are considered as alternatives and one score for each arc is found by combining the data which belong to nine factors using TOPSIS methodology and criteria weights found in the first step. This is the first study that TOPSIS is used in this way to combine the data for nine factors into one score for arcs on the road network.

Finally, optimal routes between OD pairs are found using ArcGIS network analysis tool in which total route score is minimized. We compare the optimal routes found using our methodology and the methodology used in previous studies. We also conduct a sensitivity analysis and examine the effects of the changes in the priorities of the criteria weights on the results. The results encourage us to propose our methodology in which all hazmat transportation related factors are included while finding optimal routes for hazmat vehicles. Our study proposes a more practical methodology for finding optimal routes comparing with studies in the literature which propose very complicated mathematical models and solution ways. Another important advantage of our methodology is that when some of the values related to constraints or parameters are changed, this can easily be adopted in our model to find optimal routes between any OD pairs on the road network in a very short time. In addition, our study which uses ArcGIS to find optimal routes is available to ban the arcs with higher scores (which include the scores of nine factors) while finding optimal routes.

For future studies, our methodology can be adopted to find optimal routes for other vehicles carrying different kinds of cargoes since there are also other related criteria that affect the other transportation types and our methodology can be used to find first the criteria weights and next the optimal routes between OD pairs.

Our study is the first study in which criteria weights are assigned for hazmat transportation. The researchers can use main and sub-criteria weights found in this study in their researches in which other types of mathematical models and solutions are used.

We aim to include a multinational perspective while assigning criteria weights. Our questionnaire can be applied to the local experts in further studies if the main focus is on hazmat transportation inside the country.

We use AHP methodology to find the criteria weights. Other MDCM methodologies can be used in future studies to find the criteria weights and the results obtained from those methodologies can be compared.

References

Chen, Z. S., Li, M., Kong, W. T., & Chin, K. S. (2019). Evaluation and selection of HazMat transportation alternatives: A PHFLTS- and TOPSIS-integrated multi-perspective approach. *International Journal of Environmental Research and Public Health, 16*, 4116. https://doi.org/10.3390/ijerph16214116

Ditta, A., Figueroa, O., Galindo, G., & Yie-Pinedo, R. (2019). A review on research in transportation of hazardous materials. *Socio-Economic Planning Sciences, 68*(2019), 100665.

Erkut, E., Tjandra, S.A., & Verter, V., (2007). Chapter 9 hazardous materials transportation. In: *Transportation* (Vol. 14, pp. 539–621). Elsevier.

Erkut, E., & Verter, V. (1998). Modeling of transport risk for hazardous materials. *Operations Research, 46*(5), 625–642.

Holeczek, N. (2019). Hazardous materials truck transportation problems: A classification and state of the art literature review. *Transportation Research Part D, 69*, 305–328.

Huang, B. (2006). GIS-based route planning for hazardous material transportation. *Journal of Environmental Informatics, 8*(1), 49–57.

Hwang, C. L., & Yoon, P. (1981). Multiple attribute decision making. In *Lecture notes in economics and mathematical systems* (p. 7). Springer-Verlag.

Jun, Y., & Wei, C., (2010). AHP application in selection of routes for hazardous materials transportation. *Proceedings of international conference on optoelectronics and image processing, IEEE Computer Society*, ISBN: 978-0-7695-4252-2.

Kang, Y., Batta, R., & Kwon, C. (2014). Generalized route planning model for hazardous material transportation with VaR and equity considerations. *Computers & Operations Research., 43*, 237–247.

Kara, B. Y., & Verter, V. (2004). Designing a road network for hazardous materials transportation. *Transportation Science, 38*(2), 188–196.

Kawprasert, A., & Barkan, C. P. (2008). Effects of route rationalization on hazardous materials transportation risk. *Transportation Research Record: Journal of the Transportation Research Board, 2043*(1), 65–72.

Kim, M., Miller-Hooks, E., & Nair, R. A. (2011). Geographic information system-based real-time decision support framework for routing vehicles carrying hazardous materials. *Journal of Intelligent Transportation Systems, 15*(1), 28–41.

Li, Y. L., Yanga, Q., & Chin, K. S. (2019). A decision support model for risk management of hazardous materials road transportation based on quality function deployment. *Transportation Research Part D*, 74, 154–173

Rashid, Z., El-Harbawi, M., & Shariff, A. (2010). Assessment on the consequences of liquefied petroleum gas release accident in the road transportation via GIS approaches. *Journal of Applied Sciences, 10*, 1157–1165.

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology, 15*(1), 234–281.

Samanlioglu, F. (2013). A multi-objective mathematical model for the industrial hazardous waste location-routing problem. *European Journal of Operational Research, 226*(2), 332–340.

Sattayaprasert, W., Hanaoka, S., Taneerananon, P., & Pradhananga, R. (2008). Creating a risk-based network for hazmat logistics by route prioritization with AHP: Case study: Gasoline logistics in Rayong, Thailand. *IATSS Research, 32*(1), 74–87.

Toumazis, I., & Kwon, C. (2015). Worst-case conditional value-at-risk minimization for hazardous materials transportation. *Transportation Science, 50*(4), 1174–1187.

Verma, M., & Verter, V. (2010). A lead-time based approach for planning rail–truck intermodal transportation of dangerous goods. *European Journal of Operational Research., 202*(3), 696–706.

Yilmaz, Z., Erol, S., & Aplak, H. S. (2016). Transportation of hazardous materials (hazmat) a literature survey. *Pamukkale University Journal of Engineering Sciences, 22*(1), 39–53.

Zografos, K. G., & Androutsopoulos, K. N. (2008). A decision support system for integrated hazardous materials routing and emergency response decisions. *Transport Research Part C Emerging Technologies, 16*(6), 684–703.

Strategic Multi-criteria Decision-Making Against Pandemics Using Picture and Spherical Fuzzy AHP and TOPSIS

Cengiz Kahraman, Irem Ucal Sari, and Sezi Çevik Onar

Abstract With the coronavirus that emerged suddenly and affected the masses, it became clear how unprepared we were for the pandemic situation both individually and at the state level. Over the past period, many different transmission prevention strategies have been implemented in different countries, and their results have been observed. Unfortunately, a complete preventive strategy has not yet been developed, and the disease continues to spread. There is a long and arduous process ahead of us in the fight against coronavirus. In order to better manage this process, this study aims to compare the strategic decisions to be taken by the states, especially during the periods when the disease spreads. The intervention strategies have a strong impact on economics, social life, health systems, and environment. As a result of intervention methods, the income of individuals, health sector, tourism sector and companies are affected economically. Yet, modeling strategic decisions under pandemic conditions is complicated since multiple factors should be considered. The uncertainties and imprecision in the evaluation process increase the complexity of the decision-making process. In this study, the decision-making procedure criteria are defined using multi-criteria decision-making methods, and the alternative strategies are compared using TOPSIS method. Picture fuzzy sets and spherical fuzzy sets are used to handle uncertainty in the decision-making process, and the results are compared. These two methods are compared based on the results and the applicability.

Keywords COVID-19 · Intervention strategies · Picture fuzzy sets · Spherical fuzzy sets · AHP · TOPSIS

C. Kahraman (✉) · I. Ucal Sari · S. Çevik Onar
Industrial Engineering Department, Istanbul Technical University, Macka, Istanbul, Turkey
e-mail: kahramanc@itu.edu.tr; ucal@itu.edu.tr; cevikse@itu.edu.tr

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_15

1 Introduction

Defining the appropriate intervention strategy is crucial for minimizing the impacts of the pandemics such as COVID-19. This study focuses on defining the intervention strategies under pandemics such as COVID-19 and defining the best intervention strategy under different scenarios. Strict measures such as total lockdown can be more effective for eliminating the pandemics, yet sustaining these measures is very tough since they cause severe problems in the populations' physical, psychological, and economic wealth. Selecting the right time with an appropriate magnitude and duration becomes the key for intervention strategies' overall success. However, modeling strategic decisions under pandemic conditions is complicated since multiple factors should be considered. The uncertainties and imprecision in the evaluation process increase the complexity of the decision-making process. This study defines a hierarchical decision model for evaluating the best COVID-19 intervention strategy under different scenarios. Based on the literature review, health-related costs, public prevention-related costs, and social impacts are considered as the three main criteria. Health-related costs can be defined as the hospital intensive and non-intensive care costs, labor costs, and supply chain costs. Public prevention-related costs are the loss in net domestic business closure, labor costs, incentives, and supply of preventive materials. Social impacts are the future educational problems (such as non-return to school, inequality in education and social services), social and psychological well-being and physical well-being.

In literature, many studies focus on evaluating the impact of intervention strategies. These studies can be grouped as mathematical modeling-based approaches, data-driven approaches, and hybrid approaches. These studies highlight that although stricter intervention strategies are more effective, applying and maintaining these strategies is very hard. Selecting the timing, magnitude, and duration of the intervention strategies is crucial for eliminating pandemics such as COVID-19.

Kennedy et al. (2020) model the effects of intervention strategies for COVID-19 by using a mathematical modeling approach. In this compartmental model, three different strategies, namely, continuous, intermittent, and stepping-down social distancing strategies, are evaluated. The results indicate that the stepping-down strategy every 80 days in 2 years is the most effective long-term strategy for minimizing the peak number of COVID-19 patients and deaths. The study also highlights that personal precautions such as applying proper hygiene precautions, using masks, and social distancing are vital in reducing COVID-19. Reddy et al. (2021) focus on the cost-effectiveness of COVID-19 intervention strategies in South Africa. This study highlights that in the low- and middle-income countries, financial measures are crucial since these countries have limited reserves, maintaining strict intervention strategies dramatically affects the overall wealth of their nations. The authors develop a microsimulation model to evaluate the health and financial impacts of the interventions, namely, the healthcare testing, contact tracing, isolation and guarantee centers, mass symptom screening, and molecular testing. The study shows that a combination of these approaches can achieve a low-cost and effective intervention strategy. Li et al. (2020) develop a mathematical model with a climate indicator for understanding the spreading of COVID-19. This study shows that an intensive 8 weeks intervention strategy that limits local and international is effective. In the high-risk areas continuing these intensive intervention strategies claimed to be necessary for suppressing the spread of COVID-19.

Imai et al. (2020) evaluate the impact of intervention strategies for COVID-19. In this study, social distancing interventions such as school closures, remote working, and quarantine are examined through a detailed literature review and a data-driven approach. The impact of intervention strategies is collected from government websites and media. The results indicate that travel restrictions are slowing the spread of COVID-19, but it is not easy to prove the impact of intervention strategies due to the complex nature of the COVID-19 outbreak. The study also highlights the importance of these intervention strategies' timing, the applicability, and the impact of these strategies can be optimized by defining the optimal time and duration for these strategies. Ngonghala et al. (2020) conducted a mathematical model for evaluating the impact of intervention strategies. Intervention strategies that are taken into account in this study are social distancing, contact tracing, quarantine, isolation, and the use of face masks. The results indicate that when the reproduction number is less than unity, the spread of disease can be stopped. Both the social distancing and usage of face masks are very important for decreasing the spread and overall impact of COVID-19. In this data-driven approach, several simulations are conducted using data collected from New York and the entire USA. Similar to Imai et al. (2020), this study also claims that the intervention strategy's timing and duration are crucial. Pinto Neto et al. (2021) develop a multi-objective genetic algorithm design-based optimization model for evaluating the impact of strategic interventions in Sao Paulo. In this epidemiological compartmental model, different intervention strategy scenarios are evaluated based on the magnitude and timing of social distancing and personal protection. The study indicates that social distancing with stepping down at every 80 days is the optimal intervention strategy. Kantor and Kantor (2020) survey revealing the public perception through the intervention strategies. The study shows that attributes such as age, gender, and race impact applying the intervention strategies. Thus, while considering the intervention strategies, we should take people's perceptions.

Dighe et al. (2020) develop an established Bayesian framework for modeling the COVID-19 outbreak in South Korea. The results show that enhanced testing and contact tracing strategies and strong social distancing are critical for South Korean success. Panovska-Griffiths et al. (2020) try to define the optimal strategy for reopening schools in the UK by using a stochastic individual-based model. The model shows that large-scale testing is crucial for avoiding the further COVID-19 waves. Kneale et al. (2020) focus on school closures in COVID-19. In this model, children's health and well-being, children's education, impacts on teachers and other school staff, the school organization, considerations for parents and families, public health considerations, and broader economic impacts are considered the main criteria for evaluating the closure decision.

COVID-19 intervention strategies have economic, social, health and environment effects. As a result of intervention methods, the income of individuals, health sector, tourism sector and companies are affected economically. Depression can be seen in individuals due to the feeling of loneliness, family relationships, communication and social skills are affected, and inequality in education can occur. Interventions also affect various health problems (Rauschenberg et al., 2021; Roberts et al., 2021; Robinson et al., 2021; Rozhnova et al., 2021; Russell et al. 2020). Analytical hierarchy process (AHP) and TOPSIS methods are commonly used for modeling complex decision problems. Combining AHP and TOPSIS approaches for evaluating a multi-criteria decision model is a widely employed approach since it enables defining the weights of criteria and score of alternatives with a higher sensitivity without increasing the complexity. In order to model, imprecision and uncertainty picture fuzzy and spherical fuzzy extensions of fuzzy AHP and fuzzy TOPSIS methods are utilized. This study uses scenario-based picture fuzzy and spherical fuzzy extensions of fuzzy AHP and fuzzy TOPSIS methods for evaluating the intervention strategies. In MCDM problems, many alternatives are evaluated with many criteria. In these evaluations, the uncertainty in human thoughts can best be represented by fuzzy sets. The general purpose of all fuzzy set extensions is how to better express this uncertainty in human thought. This uncertainty, which is generally found in linguistic evaluations, can be handled with two-dimensional fuzzy sets such as intuitionistic, Pythagorean, Fermatean, q-rung orthopair fuzzy sets, or it can be represented by three-dimensional fuzzy sets such as picture fuzzy, spherical fuzzy, or neutrosophic. The motivation in this study is to explore how uncertainty in human thought can be captured by three-dimensional fuzzy sets. We need three-dimensional fuzzy sets because the decision maker's degree of hesitancy in his linguistic judgments should be assigned by himself, not as a value arising from the other two parameters. Therefore, picture fuzzy or spherical fuzzy sets should be preferred. The contribution of this study is to show in detail how vagueness and impreciseness in the evaluations of decision makers can be handled in an MCDM problem with picture fuzzy and spherical fuzzy sets.

The rest of the paper is organized as follows: Section 2 briefly explains the preliminaries on picture fuzzy sets and spherical fuzzy sets. Section 3 explains the proposed scenario-based hybrid multi-criteria decision-making (MCDM) methodology. In Sect. 4, the proposed model is applied to define intervention strategy. The last section concludes and gives further suggestions.

2 Preliminaries on Picture Fuzzy Sets and Spherical Fuzzy Sets

Both picture fuzzy sets and spherical fuzzy sets are extensions defined to represent vagueness with three parameters. However, spherical fuzzy sets are designed to offer a larger domain area to the decision maker. In other words, the decision maker has greater freedom in assigning membership degrees, and it ensures that the error is eliminated when the sum of the degrees exceeds one. If less restriction is desired during the assignment of membership degree, spherical fuzzy sets should be preferred to picture fuzzy sets. However, if it can be guaranteed by the decision maker that the rule of not exceeding 1 will be provided, picture fuzzy sets may be preferred due to the ease of calculation.

2.1 Picture Fuzzy Sets

IFS has been extended to picture fuzzy sets (PFS) in order to consider the refusal degree of decision makers (Cuong & Kreinovich, 2014). PFS based approaches are effective methods to meet different expert judgments such as yes, abstain, no, and refusal. PFS based models are successful in symbolizing uncertain information in different processes such as cluster analysis and pattern recognition.

A picture fuzzy set Ã on the universe X is an object of the form

$$\widetilde{\mathrm{A}} = \left\{\left\langle x;\, \mu_A(x), \nu_A(x), \pi_A(x)\right\rangle | x \in X\right\} \tag{1}$$

where $\mu_{\widetilde{\mathrm{A}}}(x) \in [0, 1]$ is called the "degree of positive membership of Ã," $\nu_{\widetilde{\mathrm{A}}}(x) \in [0, 1]$ is called the "degree of negative membership of $\widetilde{\mathrm{A}}$," and $\pi_{\widetilde{\mathrm{A}}}(x) \in [0, 1]$ is called the "degree of neutral membership of $\widetilde{\mathrm{A}}$" and $\mu_{\widetilde{\mathrm{A}}}(x)$, $\nu_{\widetilde{\mathrm{A}}}(x)$, and $\pi_{\widetilde{\mathrm{A}}}(x)$ satisfy the following condition: $0 \le \mu_{\widetilde{\mathrm{A}}}(x) + \nu_{\widetilde{\mathrm{A}}}(x) + \pi_{\widetilde{\mathrm{A}}}(x) \le 1, \forall x \in X$. Then for $x \in X$, $\rho_{\widetilde{\mathrm{A}}}(x) = 1 - \mu_{\widetilde{\mathrm{A}}}(x) - \nu_{\widetilde{\mathrm{A}}}(x) - \pi_{\widetilde{\mathrm{A}}}(x)\big)$could be called the degree of refusal membership of x in $\widetilde{\mathrm{A}}$. Thereafter, $\left\langle x;\, \mu_{\widetilde{\mathrm{A}}}(x), \nu_{\widetilde{\mathrm{A}}}(x), \pi_{\widetilde{\mathrm{A}}}(x)\right\rangle$ will be given as $\left(x;\, \mu_{\widetilde{\mathrm{A}}}, \nu_{\widetilde{\mathrm{A}}}, \pi_{\widetilde{\mathrm{A}}}\right)$.

Voting can be a good illustration of such a situation as the human voters may be divided into four groups of those who: vote for, hesitant, and vote against, refusal of the voting (Cuong & Kreinovich, 2013). Some definitions and theorems of PFS are given in the following (Wei, 2017).

Definition 1. Let $\widetilde{\alpha} = \left(\mu_a, \nu_\alpha, \pi_\alpha\right)$be a picture fuzzy number (PFN). Then, the score function S of a picture fuzzy number can be given as follows (Cuong & Kreinovich, 2014):

$$S\left(\widetilde{\alpha}\right) = \mu_a - \pi_\alpha,\; S\left(\widetilde{\alpha}\right) \in [-1, 1] \tag{2}$$

Definition 2: Basic operators of Single-valued PFSs;

$$\widetilde{A}_p \oplus \widetilde{B}_p = \left(\mu_{\widetilde{A}_p} + \mu_{\widetilde{B}_p} - \mu_{\widetilde{A}_p}\mu_{\widetilde{B}_p}, \nu_{\widetilde{A}_p}\nu_{\widetilde{B}_p}, \pi_{\widetilde{A}_p}\pi_{\widetilde{B}_p}\right) \tag{3}$$

$$\widetilde{A}_p \otimes \widetilde{B}_p = \left(\mu_{\widetilde{A}_p}\mu_{\widetilde{B}_p}, \nu_{\widetilde{A}_p} + \nu_{\widetilde{B}_p} - \nu_{\widetilde{A}_p}\nu_{\widetilde{B}_p}, \pi_{\widetilde{A}_p}\pi_{\widetilde{B}_p} - \pi_{\widetilde{A}_p}\pi_{\widetilde{B}_p}\right) \tag{4}$$

$$\lambda \widetilde{A}_p = \left\{ \left(1 - \left(1 - \mu_{\widetilde{A}_p} \right)^{\lambda}, \nu_{\widetilde{A}_p}{}^{\lambda}, \pi_{\widetilde{A}_p}{}^{\lambda} \right) \right\} \text{ for } \lambda > 0 \quad (5)$$

$$\widetilde{A}_p^{\lambda} = \left\{ \left(\mu_{\widetilde{A}_p}{}^{\lambda}, 1 - \left(1 - \nu_{\widetilde{A}_p} \right)^{\lambda}, 1 - \left(1 - \pi_{\widetilde{A}_p} \right)^{\lambda} \right) \right\} \text{ for } \lambda > 0 \quad (6)$$

Picture fuzzy arithmetic aggregation operators are used for aggregating the different evaluations of multiexperts. Picture fuzzy weighted averaging (PFWA) operator and picture fuzzy weighted geometric (PFWG) operator are aggregation operators of PFS as arithmetic and geometric aggregation operators which are developed by Guiwu Wei (2017).

Definition 3. Let α_j ($j = 1, 2, \cdots, n$) be a collection of PFNs. The picture fuzzy weighted averaging (PFWA) operator is a mapping $P^n \rightarrow P$ such that

$$\text{PFWA}_{\omega} (\alpha_1, \alpha_2, \ldots, \alpha_n) = \oplus_{j=1}^{n} \left(\omega_j \alpha_j \right) \quad (7)$$

where $\omega = (\omega_1, \omega_2, \ldots, \omega_n)^{\mathrm{T}}$ be the weight vector of α_j ($j = 1, 2, \cdots, n$), and $\omega_j > 0$, $\sum_{j=1}^{n} \omega_j = 1$ (Wei, 2017).

Theorem 1. The aggregated value by using PFWA operator is also a PFN, where

$$\text{PFWA}_{\omega} (\alpha_1, \alpha_2, \ldots, \alpha_n) = \left(1 - \prod_{j=1}^{n} \left(1 - \mu_{\alpha_j} \right)^{\omega_j}, \prod_{j=1}^{n} \left(\nu_{\alpha_j} \right)^{\omega_j}, \prod_{j=1}^{n} \left(\pi_{\alpha_j} \right)^{\omega_j} \right) \quad (8)$$

where $\omega = (\omega_1, \omega_2, \ldots, \omega_n)^{\mathrm{T}}$ be the weight vector of α_j ($j = 1, 2, \ldots, n$), and $\omega_j > 0$, $\sum_{j=1}^{n} \omega_j = 1$ (Wei, 2017).

Definition 4. Let α_j ($j = 1, 2, \ldots, n$) be a collection of PFNs. The picture fuzzy weighted geometric (PFWG) operator is a mapping $P^n \rightarrow P$ such that

$$\text{PFWG}_{\omega} (\alpha_1, \alpha_2, \ldots, \alpha_n) = \otimes_{j=1}^{n} \left(\alpha_j{}^{\omega_j} \right) \quad (9)$$

where $\omega = (\omega_1, \omega_2, \ldots, \omega_n)^{\mathrm{T}}$ be the weight vector of α_j ($j = 1, 2, \ldots, n$), and $\omega_j > 0$, $\sum_{j=1}^{n} \omega_j = 1$ (Wei, 2017).

Theorem 2. The aggregated value by using PFWG operator is also a PFN, where

$$\text{PFWG}_{\omega} (\alpha_1, \alpha_2, \ldots, \alpha_n) = \left(\prod_{j=1}^{n} \left(\mu_{\alpha_j} \right)^{\omega_j}, 1 - \prod_{j=1}^{n} \left(1 - \nu_{\alpha_j} \right)^{\omega_j}, 1 - \prod_{j=1}^{n} \left(1 - \pi_{\alpha_j} \right)^{\omega_j} \right) \quad (10)$$

where $\omega = (\omega_1, \omega_2, \ldots, \omega_n)^{\mathrm{T}}$ be the weight vector of α_j ($j = 1, 2, \ldots, n$), and $\omega_j > 0$, $\sum_{j=1}^{n} \omega_j = 1$ (Wei, 2017).

Definition 5: Single-valued Picture Fuzzy Weighted Averaging operator (PFWA) with respect to $w = (w_1, w_2 \ldots\ldots, w_n)$; $w_i \in [0, 1]$; $\sum_{i=1}^{n} w_i = 1$, is defined as (Wei, 2017)

$$\begin{aligned}\mathrm{PFWA}_w\left(\widetilde{A}_1, \widetilde{A}_2, \ldots, \widetilde{A}_n\right) &= w_1\widetilde{A}_1 + w_2\widetilde{A}_2 + \cdots + w_n\widetilde{A}_n \\ &= \left(1 - \prod_{i=1}^{n}\left(1 - \mu_{\widetilde{A}_i}\right)^{w_i}, \prod_{i=1}^{n}\left(\nu_{\widetilde{A}_i}\right)^{w_i}, \prod_{i=1}^{n}\left(\pi_{\widetilde{A}_i}\right)^{w_i}\right)\end{aligned} \tag{11}$$

Definition 6: Score functions and Accuracy functions of sorting picture fuzzy numbers are defined by Wei (2017);

$$\text{Score}\ \left(\widetilde{A}_p\right) = \frac{1}{2}\left(1 + 2\mu_{\widetilde{A}_p} - \nu_{\widetilde{A}_p} - \frac{\pi_{\widetilde{A}_p}}{2}\right) \tag{12}$$

$$\text{Accuracy}\ \left(\widetilde{A}_p\right) = \mu_{\widetilde{A}_p} + \nu_{\widetilde{A}_p} + \pi_{\widetilde{A}_p} \tag{13}$$

Note that: $\widetilde{A}_p < \widetilde{B}_p$ if and only if

(a) Score $\left(\widetilde{A}_p\right) <$ Score $\left(\widetilde{B}_p\right)$ or
(b) Score $\left(\widetilde{A}_p\right) =$ Score $\left(\widetilde{B}_p\right)$ and Accuracy $\left(\widetilde{A}_p\right) <$ Accuracy $\left(\widetilde{A}_p\right)$

Definition 7: Distance formulas for picture fuzzy numbers $\left(\tilde{A}_p, \tilde{B}_p\right)$ are defined as follows (Cuong & Kreinovich, 2014):

(a) The normalized Hamming Distance;

$$D_H\left(\widetilde{A}_p, \widetilde{B}_p\right) = \frac{1}{n}\sum_{i=1}^{n}\left(\left|\mu_{\widetilde{A}_p}(x_i) - \mu_{\widetilde{B}_p}(x_i)\right| + \left|\nu_{\widetilde{A}_p}(x_i) - \nu_{\widetilde{B}_p}(x_i)\right| + \left|\pi_{\widetilde{A}_p}(x_i) - \pi_{\widetilde{B}_p}(x_i)\right|\right) \tag{14}$$

(b) The normalized Euclidean Distance;

$$D_E\left(\widetilde{A}_p, \widetilde{B}_p\right) = \left(\frac{1}{n}\sum_{i=1}^{n}\left(\left(\mu_{\widetilde{A}_p}(x_i) - \mu_{\widetilde{B}_p}(x_i)\right)^2 + \left(\nu_{\widetilde{A}_p}(x_i) - \nu_{\widetilde{B}_p}(x_i)\right)^2 + \left(\pi_{\widetilde{A}_p}(x_i) - \pi_{\widetilde{B}_p}(x_i)\right)^2\right)\right)^{1/2} \tag{15}$$

2.2 *Spherical Fuzzy Sets*

Spherical fuzzy sets (SFSs) are one of the latest extensions of intuitionistic fuzzy sets. However, it is in the class of three parameter fuzzy set extensions such as picture fuzzy sets and neutrosophic sets. These parameters are membership degree, non-membership degree, and hesitancy (indeterminacy) degree. All these parameters must be determined to satisfy the condition that their squared sum is at most equal to 1. The difference of the squared sum from 1 is called the squared refusal degree. SFSs are also in the class of fuzzy set extensions having a larger domain for the assignment of the mentioned degrees such as Pythagorean fuzzy sets, Fermatean fuzzy sets, and q-rung orthopair fuzzy sets. The preliminaries of SFSs are presented in the following:

Definition 8. A single-valued SFS, $\widetilde{A}_S$ of the universe of discourse X is given by (Gündoğdu & Kahraman, 2019a):

$$\widetilde{A}_s = \left\{x, \mu_{\widetilde{A}_s}(x), \vartheta_{\widetilde{A}_s}(x), I_{\widetilde{A}_s}(x) \;\middle|\; x \in X\right\} \tag{16}$$

where $\mu_{\widetilde{A}_s}(u), \vartheta_{\widetilde{A}_s}(u), I_{\widetilde{A}_s}(u) : U \to [0, 1]$ are the degree of membership, non-membership, and indeterminacy of x to $\widetilde{A}_S$, respectively, and

$$0 \le \mu^2_{\widetilde{A}_s}(x) + \vartheta^2_{\widetilde{A}_s}(x) + I^2_{\widetilde{A}_s}(x) \le 1 \tag{17}$$

Then, $1 - \sqrt{\mu^2_{\widetilde{A}_s}(x) + \vartheta^2_{\widetilde{A}_s}(x) + I^2_{\widetilde{A}_s}(x)}$ is defined as refusal degree of x in X.

Definition 9. Assume that $\widetilde{A}_s$ and $\widetilde{B}_s$ are any two Spherical Fuzzy sets. So the basic operations of SFSs can be defined as follows (Gündoğdu & Kahraman, 2019a):

$$\widetilde{A}_s \oplus \widetilde{B}_s = \left\{\sqrt{\mu^2_{\widetilde{A}_s} + \mu^2_{\widetilde{B}_s} - \mu^2_{\widetilde{A}_s}\mu^2_{\widetilde{B}_s}},\ \vartheta_{\widetilde{A}_s}\vartheta_{\widetilde{B}_s}, \sqrt{\left(1-\mu^2_{\widetilde{B}_s}\right)I^2_{\widetilde{A}_s} + \left(1-\mu^2_{\widetilde{A}_s}\right)I^2_{\widetilde{B}_s} - I^2_{\widetilde{A}_s}I^2_{\widetilde{B}_s}}\right\} \tag{18}$$

$$\widetilde{A}_s \otimes \widetilde{B}_s = \left\{\mu_{\widetilde{A}_s}\mu_{\widetilde{B}_s}, \sqrt{\vartheta^2_{\widetilde{A}_s} + \vartheta^2_{\widetilde{B}_s} - \vartheta^2_{\widetilde{A}_s}\vartheta^2_{\widetilde{B}_s}}, \sqrt{\left(1-\vartheta^2_{\widetilde{B}_s}\right)I^2_{\widetilde{A}_s} + \left(1-\vartheta^2_{\widetilde{A}_s}\right)I^2_{\widetilde{B}_s} - I^2_{\widetilde{A}_s}I^2_{\widetilde{B}_s}}\right\} \tag{19}$$

$$k\widetilde{A}_s = \left\{\sqrt{1-\left(1-\mu^2_{\widetilde{A}_s}\right)^k}, \vartheta^k_{\widetilde{A}_s}, \sqrt{\left(1-\mu^2_{\widetilde{A}_s}\right)^k - \left(1-\mu^2_{\widetilde{A}_s} - I^2_{\widetilde{A}_s}\right)^k}\right\};\quad k \ge 0 \tag{20}$$

$$\widetilde{A}_s^k = \left\{\mu^k_{\widetilde{A}_s}, \sqrt{1-\left(1-\vartheta^2_{\widetilde{A}_s}\right)^k}, \sqrt{\left(1-\vartheta^2_{\widetilde{A}_s}\right)^k - \left(1-\vartheta^2_{\widetilde{A}_s} - I^2_{\widetilde{A}_s}\right)^k}\right\};\quad k \ge 0 \tag{21}$$

Definition 10. Spherical Fuzzy Weighted Arithmetic Mean (SFWAM) with respect to $w = (w_1, w_2, \ldots, w_n)$; $w_i \in [0, 1]$; $\sum_{i=1}^{n} w_i = 1$, SFWAM is defined as

(Gündoğdu & Kahraman, 2019a)

$$\mathrm{SFWAM}_w\left(\widetilde{A}_{S1},\widetilde{A}_{S2},\ldots,\widetilde{A}_{Sn}\right)=w_1\widetilde{A}_{S1}+w_2\widetilde{A}_{S2}+\cdots+w_n\widetilde{A}_{Sn}$$
$$=\left\{\sqrt{1-\prod_{i=1}^{n}\left(1-\mu_{A_s}^2\right)^{w_i}},\prod_{i=1}^{n}\vartheta_{A_s}^{w_i},\sqrt{\prod_{i=1}^{n}\left(1-\mu_{A_s}^2\right)^{w_i}-\prod_{i=1}^{n}\left(1-\mu_{A_s}^2-I_{A_s}^2\right)^{w_i}}\right\} \tag{22}$$

Definition 11. Spherical Fuzzy Weighted Geometric Mean (SFWGM) with respect to $w=(w_1,w_2,\ldots,w_n)$; $w_i\in[0,1]$; $\sum_{i=1}^{n}w_i=1$, SFWGM is defined as (Gündoğdu & Kahraman, 2019a)

$$\mathrm{SFWGM}_w\left(\widetilde{A}_{S1},\widetilde{A}_{S2},\ldots,\widetilde{A}_{Sn}\right)=\widetilde{A}_{S1}^{w_i}+\widetilde{A}_{S2}^{w_i}+\cdots+\widetilde{A}_{Sn}^{w_i}$$
$$=\left\{\prod_{i=1}^{n}\mu_{A_s}^{w_i},\sqrt{1-\prod_{i=1}^{n}\left(1-\vartheta_{A_s}^2\right)^{w_i}},\sqrt{\prod_{i=1}^{n}\left(1-\vartheta_{A_s}^2\right)^{w_i}-\prod_{i=1}^{n}\left(1-\vartheta_{A_s}^2-I_{A_s}^2\right)^{w_i}}\right\} \tag{23}$$

In the following, we give the definition of interval-valued spherical fuzzy sets (IVSFS) and summarize distance measurement, arithmetic operations, and aggregation and defuzzification operations.

Definition 12: An Interval-Valued Spherical Fuzzy Set (IVSFS) $\widetilde{A}_S$ of the universe of discourse U is defined as in Eq. (24) (Gündoğdu & Kahraman, 2021a)

$$\widetilde{A}_s=\left\{u,\left(\left[\mu_{\widetilde{A}_s}^L(u),\mu_{\widetilde{A}_s}^U(u)\right],\left[\upsilon_{\widetilde{A}_s}^L(u),\upsilon_{\widetilde{A}_s}^U(u)\right],\left[\pi_{\widetilde{A}_s}^L(u),\pi_{\widetilde{A}_s}^U(u)\right]\middle|\,u\in U\right)\right\} \tag{24}$$

where $0\le\mu_{\widetilde{A}_s}^L(u)\le\mu_{\widetilde{A}_s}^U(u)\le1$, $0\le\upsilon_{\widetilde{A}_s}^L(u)\le\upsilon_{\widetilde{A}_s}^U(u)\le1$ and $0\le\left(\mu_{\widetilde{A}_s}^U(u)\right)^2+\left(\upsilon_{\widetilde{A}_s}^U(u)\right)^2+\left(\pi_{\widetilde{A}_s}^U(u)\right)^2\le1$. For each $u\in U$, $\mu_{\widetilde{A}_s}^U(u)$, $\upsilon_{\widetilde{A}_s}^U(u)$, and $\pi_{\widetilde{A}_s}^U(u)$ are the upper degrees of membership, non-membership, and hesitancy of u to $\widetilde{A}_s$, respectively. For each $u\in U$, if $\mu_{\widetilde{A}_s}^L(u)=\mu_{\widetilde{A}_s}^U(u)$, $\upsilon_{\widetilde{A}_s}^L(u)=\upsilon_{\widetilde{A}_s}^U(u)$, and $\pi_{\widetilde{A}_s}^L(u)=\pi_{\widetilde{A}_s}^U(u)$, then IVSFS $\widetilde{A}_s$ reduces to a single-valued SFS.

For an IVSFS $\widetilde{A}_S$, the pair $\left\langle\left[\mu_{\widetilde{A}_s}^L(u),\mu_{\widetilde{A}_s}^U(u)\right],\left[\upsilon_{\widetilde{A}_s}^L(u),\upsilon_{\widetilde{A}_s}^U(u)\right],\left[\pi_{\widetilde{A}_s}^L(u),\pi_{\widetilde{A}_s}^U(u)\right]\right\rangle$ is called an interval-valued spherical fuzzy number (IVSFS). For convenience, the pair $\left\langle\left[\mu_{\widetilde{A}_s}^L(u),\mu_{\widetilde{A}_s}^U(u)\right],\left[\upsilon_{\widetilde{A}_s}^L(u),\upsilon_{\widetilde{A}_s}^U(u)\right],\left[\pi_{\widetilde{A}_s}^L(u),\pi_{\widetilde{A}_s}^U(u)\right]\right\rangle$ is denoted by $\widetilde{\alpha}=\langle[a,b],[c,d],[e,f]\rangle$ where $[a,b]\subset[0,1]$, $[c,d]\subset[0,1]$, $[e,f]\subset[0,1]$, and $b^2+d^2+f^2\le1$.

Obviously, $\widetilde{\alpha}^*=\langle[1,1],[0,0],[0,0]\rangle$ is the largest IVSFS, $\widetilde{\alpha}^-=\langle[0,0],[1,1],[0,0]\rangle$ is the smallest IVSFS, and $\widetilde{\alpha}^{*/-}=\langle[0,0],[0,0],[1,1]\rangle$ is between largest and smallest IVSFS number.

Some operations are defined over IVSFS as below:

Definition 13: Let $\widetilde{\alpha} = \langle [a, b], [c, d], [e, f] \rangle$, $\widetilde{\alpha}_1 = \langle [a_1, b_1], [c_1, d_1], [e_1, f_1] \rangle$, and $\widetilde{\alpha}_2 = \langle [a_2, b_2], [c_2, d_2], [e_2, f_2] \rangle$ be IVSFS then (Gündoğdu & Kahraman, 2021a)

$$\widetilde{\alpha}_1 \cup \widetilde{\alpha}_2 = \{[\max\{a_1, a_2\}, \max\{b_1, b_2\}], \ [\min\{c_1, c_2\}, \min\{d_1, d_2\}], \ [\min\{e_1, e_2\}, \min\{f_1, f_2\}]\} \tag{25}$$

$$\widetilde{\alpha}_1 \cap \widetilde{\alpha}_2 = \{[\min\{a_1, a_2\}, \min\{b_1, b_2\}], [\max\{c_1, c_2\}, \max\{d_1, d_2\}], [\min\{e_1, e_2\}, \min\{f_1, f_2\}]\} \tag{26}$$

$$\begin{aligned} \widetilde{\alpha}_1 \oplus \widetilde{\alpha}_2 = \Big\{ & \Big[\left((a_1)^2 + (a_2)^2 - (a_1)^2(a_2)^2\right)^{1/2}, \left((b_1)^2 + (b_2)^2 - (b_1)^2(b_2)^2\right)^{1/2}\Big], \\ & [c_1 c_2, d_1 d_2], \Big[\left(\left(1-(a_2)^2\right)(e_1)^2 + \left(1-(a_1)^2\right)(e_2)^2 - (e_1)^2(e_2)^2\right)^{1/2}, \\ & \left(\left(1-(b_2)^2\right)(f_1)^2 + \left(1-(b_1)^2\right)(f_2)^2 - (f_1)^2(f_2)^2\right)^{1/2}\Big]\Big\} \end{aligned} \tag{27}$$

$$\begin{aligned} \widetilde{\alpha}_1 \otimes \widetilde{\alpha}_2 = \Big\{ & [a_1 a_2, b_1 b_2], \Big[\left((c_1)^2 + (c_2)^2 - (c_1)^2(c_2)^2\right)^{1/2}, \left((d_1)^2 + (d_2)^2 - (d_1)^2(d_2)^2\right)^{1/2}\Big], \\ & \Big[\left(\left(1-(c_2)^2\right)(e_1)^2 + \left(1-(c_1)^2\right)(e_2)^2 - (e_1)^2(e_2)^2\right)^{1/2}, \\ & \left(\left(1-(d_2)^2\right)(f_1)^2 + \left(1-(d_1)^2\right)(f_2)^2 - (f_1)^2(f_2)^2\right)^{1/2}\Big]\Big\} \end{aligned} \tag{28}$$

Multiplication by a scalar; λ > 0 (Gündoğdu & Kahraman, 2021a)

$$\begin{aligned} \lambda\widetilde{\alpha} = \Big\{ & \Big[\left(1-\left(1-a^2\right)^{\lambda}\right)^{1/2}, \left(1-\left(1-b^2\right)^{\lambda}\right)^{1/2}\Big], \\ & \left[c^{\lambda}, d^{\lambda}\right], \Big[\left(\left(1-a^2\right)^{\lambda} - \left(1-a^2-e^2\right)^{\lambda}\right)^{1/2}, \left(\left(1-b^2\right)^{\lambda} - \left(1-b^2-f^2\right)^{\lambda}\right)^{1/2},\Big]\Big\} \end{aligned} \tag{29}$$

λ^{th} Power of $\widetilde{\alpha}$; λ > 0 (Gündoğdu & Kahraman, 2021a)

$$\begin{aligned} \widetilde{\alpha}^{\lambda} = \Big\{ & \left[a^{\lambda}, b^{\lambda}\right], \Big[\left(1-\left(1-c^2\right)^{\lambda}\right)^{1/2}, \left(1-\left(1-d^2\right)^{\lambda}\right)^{1/2}\Big], \\ & \Big[\left(\left(1-c^2\right)^{\lambda} - \left(1-c^2-e^2\right)^{\lambda}\right)^{1/2}, \left(\left(1-d^2\right)^{\lambda} - \left(1-d^2-f^2\right)^{\lambda}\right)^{1/2}\Big]\Big\} \end{aligned} \tag{30}$$

Definition 14: Let $\lambda, \lambda_1, \lambda_2 \geq 0$, then (Gündoğdu & Kahraman, 2021a)

(a) $\widetilde{\alpha}_1 \oplus \widetilde{\alpha}_2 = \widetilde{\alpha}_2 \oplus \widetilde{\alpha}_1$
(b) $\widetilde{\alpha}_1 \otimes \widetilde{\alpha}_2 = \widetilde{\alpha}_2 \otimes \widetilde{\alpha}_1$
(c) $\lambda(\widetilde{\alpha}_1 \oplus \widetilde{\alpha}_2) = \lambda.\widetilde{\alpha}_1 \oplus \lambda.\widetilde{\alpha}_2$

(d) $(\widetilde{\alpha}_1 \otimes \widetilde{\alpha}_2)^\lambda = \widetilde{\alpha}_1^\lambda \otimes \widetilde{\alpha}_2^\lambda$
(e) $\lambda_1.\widetilde{\alpha} \oplus \lambda_2.\widetilde{\alpha} = (\lambda_1 + \lambda_2).\widetilde{\alpha}$
(f) $\widetilde{\alpha}^{\lambda_1} \otimes \widetilde{\alpha}^{\lambda_2} = \widetilde{\alpha}^{\lambda_1+\lambda_2}$

Definition 15: Let $\widetilde{\alpha}_j = \langle [a_j, b_j], [c_j, d_j], [e_j, f_j] \rangle$ be a collection of Interval-valued Spherical Weighted Arithmetic Mean (IVSWAM) with respect to $w_j = (w_1, w_2, \ldots, w_n)$; $w_j \in [0, 1]$ and $\sum_{j=1}^n w_j = 1$, IVSWAM is defined as (Gündoğdu & Kahraman, 2021a)

$$\begin{aligned} &\mathrm{IVSWAM}_w\left(\widetilde{\alpha}_1, \widetilde{\alpha}_2, \ldots, \widetilde{\alpha}_n\right) \\ &= w_1.\widetilde{\alpha}_1 \oplus w_2.\widetilde{\alpha}_2 \oplus \cdots \oplus w_n.\widetilde{\alpha}_n \\ &= \Bigg\{ \left[\left(1 - \prod_{j=1}^n \left(1 - a_j^2\right)^{w_j}\right)^{1/2}, \left(1 - \prod_{j=1}^n \left(1 - b_j^2\right)^{w_j}\right)^{1/2} \right], \left[\prod_{j=1}^n c_j^{w_j}, \prod_{j=1}^n d_j^{w_j} \right], \\ &\quad \left[\left(\prod_{j=1}^n \left(1 - a_j^2\right)^{w_j} - \prod_{j=1}^n \left(1 - a_j^2 - e_j^2\right)^{w_j} \right)^{1/2}, \right. \\ &\quad \left. \left(\prod_{j=1}^n \left(1 - b_j^2\right)^{w_j} - \prod_{j=1}^n \left(1 - b_j^2 - f_j^2\right)^{w_j} \right)^{1/2} \right] \Bigg\} \end{aligned} \tag{31}$$

Definition 16: Let $\widetilde{\alpha}_j = \langle [a_j, b_j], [c_j, d_j], [e_j, f_j] \rangle$ be a collection of Interval-valued Spherical Geometric Mean (IVSWGM) with respect to $w_j = (w_1, w_2, \ldots, w_n)$; $w_j \in [0, 1]$ and $\sum_{j=1}^n w_j = 1$, IVSWGM is defined as (Gündoğdu & Kahraman, 2021a)

$$\begin{aligned} &\mathrm{IVSWGM}_w\left(\widetilde{\alpha}_1, \widetilde{\alpha}_2, \ldots, \widetilde{\alpha}_n\right) \\ &= \widetilde{\alpha}_1^{w_1} \oplus \widetilde{\alpha}_2^{w_2} \oplus \cdots \oplus \widetilde{\alpha}_n^{w_n} = \Bigg\{ \left[\prod_{j=1}^n a_j^{w_j}, \prod_{j=1}^n b_j^{w_j} \right], \left[\left(1 - \prod_{j=1}^n \left(1 - c_j^2\right)^{w_j}\right)^{1/2}, \right. \\ &\quad \left. \left(1 - \prod_{j=1}^n \left(1 - d_j^2\right)^{w_j}\right)^{1/2} \right], \left[\left(\prod_{j=1}^n \left(1 - c_j^2\right)^{w_j} - \prod_{j=1}^n \left(1 - c_j^2 - e_j^2\right)^{w_j} \right)^{1/2}, \right. \\ &\quad \left. \left(\prod_{j=1}^n \left(1 - d_j^2\right)^{w_j} - \prod_{j=1}^n \left(1 - d_j^2 - f_j^2\right)^{w_j} \right)^{1/2} \right] \Bigg\} \end{aligned} \tag{32}$$

Definition 17: The score function of IVSFS number α is defined as

$$\mathrm{Score}\,(\widetilde{\alpha}) = S\,(\widetilde{\alpha}) = \frac{a^2 + b^2 - c^2 - d^2 - \left(e/2\right)^2 - \left(f/2\right)^2}{2} \tag{33}$$

whereScore $(\widetilde{\alpha}) = S(\widetilde{\alpha}) \in [-1, +1]$. Clearly, the greater the$S(\widetilde{\alpha})$, the larger the α. In particular, when $S(\widetilde{\alpha}) = 1$ then $\widetilde{\alpha} = \langle [1, 1], [0, 0], [0, 0] \rangle$; when $S(\widetilde{\alpha}) = -1$ then α is the smallest IVSFS number $\widetilde{\alpha} = \langle [0, 0], [1, 1], [0, 0] \rangle$ (Gündoğdu & Kahraman, 2021a).

Definition 18: The accuracy function of IVSFS number α is defined as

$$\text{Accuracy } (\widetilde{\alpha}) = H(\widetilde{\alpha}) = \frac{a^2 + b^2 + c^2 + d^2 + e^2 + f^2}{2} \tag{34}$$

where$H(\widetilde{\alpha}) \in [0, 1]$.

Note that: $\widetilde{\alpha}_1 < \widetilde{\alpha}_2$ if and only if $S(\widetilde{\alpha}_1) < S(\widetilde{\alpha}_2)$ or $S(\widetilde{\alpha}_1) = S(\widetilde{\alpha}_2)$ and $H(\widetilde{\alpha}_1) < H(\widetilde{\alpha}_2)$ (Gündoğdu & Kahraman, 2021a).

Definition 19: Let $\widetilde{\alpha}_1 = \langle [a_1, b_1], [c_1, d_1], [e_1, f_1] \rangle$, and $\widetilde{\alpha}_2 = \left\langle [a_2, b_2], [c_2, d_2], [e_2, f_2] \right\rangle$ be two IVSFS numbers, then we define the distance between $\widetilde{\alpha}_1$ and $\widetilde{\alpha}_2$ as follows (Peng & Yang, 2016a, 2016b):

$$d(\widetilde{\alpha}_1, \widetilde{\alpha}_2) = \frac{1}{4}\left(\left|a_1^2 - a_2^2\right| + \left|b_1^2 - b_2^2\right| + \left|c_1^2 - c_2^2\right| + \left|d_1^2 - d_2^2\right| + \left|e_1^2 - e_2^2\right| + \left|f_1^2 - f_2^2\right|\right) \tag{35}$$

3 Scenario-Based Hybrid MCDM Methodology

In cases where the current situation, which directly affects the decision to be made, changes rapidly, it may be necessary to constantly repeat the decision models. To avoid this situation, a scenario-based decision model is proposed in this study. This proposed methodology utilizes from picture fuzzy and spherical fuzzy extensions of fuzzy AHP and fuzzy TOPSIS methods. Fuzzy methods with interval values are preferred when there is more uncertainty in assigning degrees of membership and non-membership. Therefore, IVSF AHP-TOPSIS methodology is preferred to PF AHP-TOPSIS methodology. If point membership degrees can be assigned and we are sure of these values, there is no need to use interval value fuzzy set methodologies. In the literature, there are lots of MCDM methods which are extended using picture fuzzy sets such as AHP (Gündoğdu et al., 2021; Tey et al., 2019), TOPSIS (Jin et al., 2021; Qiyas et al., 2021; Ambrin et al., 2021), CODAS (Simic et al., 2021; Zhang et al., 2021), COPRAS (Lu et al., 2021), VIKOR (Joshi & Kumar, 2021), and GRA (Ullah et al., 2021). A much newer extension, spherical fuzzy sets, has also been used in many MCDM methods such as AHP (Oztaysi et al. 2020 Gündoğdu & Kahraman, 2021c), TOPSIS (Gündoğdu & Kahraman, 2019a, 2019b, 2021b), CODAS (G*ü*ndogdu & Kahraman, 2019d), and VIKOR (Gündoğdu et al., 2019; Akram et al., 2021).

Steps of the proposed methodology and the basics of the utilized methods are detailed as follows:

Table 1 Linguistic terms and picture fuzzy and interval-valued spherical fuzzy scales

Linguistic terms	Picture fuzzy numbers	Spherical fuzzy numbers
Strongly agree	(0.9, 0, 0.1)	([0.85, 0.95], [0.10, 0.15], [0.05, 0.15])
Agree	(0.7, 0.1, 0.2)	([0.65, 0.75], [0.20, 0.25], [0.20, 0.25])
Moderately agree	(0.55, 0.15, 0.30)	([0.50, 0.55], [0.45, 0.55], [0.30, 0.40])
Disagree	(0.2, 0.1, 0.7)	([0.20, 0.25], [0.65, 0.75], [0.20, 0.25])
Strongly disagree	(0.1, 0, 0.9)	([0.10, 0.15], [0.85, 0.95], [0.05, 0.15])

3.1 Determination of the Criteria and Scenarios

In the first step of the proposed methodology, the criteria and the situation scenarios which directly change the weights of the criteria are determined.

Then, the decision makers are asked to determine the current situations membership degrees for the determined situation scenarios using the linguistic scale that is given in Table 1. While determining the scales, attention was paid to ensure that the score values of picture fuzzy and spherical fuzzy numbers were compatible with the scale used in classical methods.

The membership functions of the current situation to each situation scenario can be calculated using fuzzy weighted arithmetic mean operator that is given in Eqs. (7) and (22) for picture fuzzy sets and spherical fuzzy sets, respectively. The membership values of each situation can be defuzzified using score functions and normalized to obtained weights that will be used in the aggregation of criteria weights of different scenarios. Then criteria weights are calculated for each of the scenario using fuzzy extensions of AHP.

3.2 Picture Fuzzy Extensions of AHP and TOPSIS

3.2.1 Picture Fuzzy AHP

Picture fuzzy AHP method utilizes picture fuzzy sets in the determination of linguistic statements. Picture fuzzy AHP method is proposed by Gündoğdu et al. (2021) and Ju et al. (2019).

The following proposed methodology differs from the previous picture fuzzy AHP methods by the used aggregation operators and defuzzification method. In the aggregation of different expert opinions, geometric mean is used as in the classical method. The defuzzification method used in the previous proposed methods can lead to negative weights. This problem has been avoided by modifying the score function used in the defuzzification step.

Step 1. Define problem and construct hierarchical structure of the problem.

Step 2. Establish pairwise comparison matrices using linguistic statements that are determined in Table 2 where a picture fuzzy number is denoted as $\widetilde{A}_p = \left(\mu_{\widetilde{A}_p}(u), \nu_{\widetilde{A}_p}(u), \pi_{\widetilde{A}_p}(u)\right)$.

Step 3. Construct picture fuzzy comparison matrices.

Pairwise comparison matrices are formed for each decision maker using picture fuzzy scale given in Table 2 as in the form of Eq. (36) and $k = 1, \ldots s$.

$$\widetilde{A}_p^{(k)} = \begin{bmatrix} 1 & \tilde{a}_{p12}^{(k)} & \ldots & \tilde{a}_{p1n}^{(k)} \\ \tilde{a}_{p21}^{(k)} & 1 & \ldots & \tilde{a}_{p2n}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{a}_{pn1}^{(k)} & \tilde{a}_{pn2}^{(k)} & \cdots & 1 \end{bmatrix} \tag{36}$$

Step 4. Check the consistency of each fuzzy pairwise comparison matrix.

Consistency ratio of each pairwise comparison matrix can be calculated using score indices that are given in Table 2 and classical consistency check formula (Saaty, 1980). Pairwise comparison matrices are considered as consistent when consistency ratio (CR) is less than 0.1.

Step 5. Compute aggregated picture fuzzy comparison matrix.

When there are several decision makers, evaluations of each criterion can be aggregated using picture fuzzy weighted geometric mean operator by Eq. (37) where w_k represents weight of the decision maker that satisfies $\sum_{k=1}^{s} w_k = 1$.

$$\tilde{w}_{p_{ij}} = \left[\prod_{k=1}^{s}\left(\mu_{ij^{(k)}}\right)^{w_k}, \prod_{k=1}^{s}\left(\nu_{ij^{(k)}}\right)^{w_k}, 1 - \prod_{k=1}^{s}\left(1 - \pi_{ij^{(k)}}\right)^{w_k}\right] \tag{37}$$

Step 6. Calculate the local and global weights of criteria.

Table 2 Linguistic terms and picture fuzzy scale for pairwise comparisons

Linguistic terms	Picture fuzzy numbers	Saaty's Scale
Absolutely More Importance (AMI)	(0.9, 0.05, 0)	9
Very High Importance (VHI)	(0.8, 0.15, 0.05)	7
High Importance (HI)	(0.7, 0.2, 0.1)	5
Slightly More Importance (SMI)	(0.6, 0.35, 0.05)	3
Equally Importance (EI)	(0.5, 0.5, 0)	1
Slightly Low Importance (SLI)	(0.35, 0.6, 0.05)	1/3
Low Importance (LI)	(0.2, 0.7, 0.1)	1/5
Very Low Importance (VLI)	(0.15, 0.8, 0.05)	1/7
Absolutely Low Importance (ALI)	(0.05, 0.9, 0)	1/9

The picture fuzzy local weights of each criterion are obtained by using picture fuzzy weighted average mean operator and given in Eq. (38).

$$\tilde{w}_{p_j} = \left[\left(1 - \prod_{j=1}^{n}\left(1-\mu_j\right)^{w_j}\right), \quad \prod_{k=1}^{s}\left(\nu_j\right)^{w_j}. \prod_{k=1}^{s}\left(\pi_j\right)^{w_j}\right] \tag{38}$$

where $w_j = 1/n$.

To compute the global weights, the local weight of each criterion is defuzzified by a modified score function $[S\left(\tilde{w}_{p_j}\right)]$ using Eq. (39) and normalized using Eq. (40).

$$S\left(\tilde{w}_{p_j}\right) = \frac{1 + 2\mu_j - \nu_j - \frac{\pi_j}{2}}{2} \tag{39}$$

$$\overline{w}_{p_j} = \frac{S\left(\tilde{w}_{p_j}\right)}{\sum_{j=1}^{n} S\left(\tilde{w}_{p_j}\right)} \tag{40}$$

Step 7. Compute final picture fuzzy weights

Final picture fuzzy weights for each criterion and sub-criterion are obtained by multiplying the picture fuzzy global weights at first level by each related sub-criterion global weight for each sub-criterion.

3.2.2 Picture Fuzzy TOPSIS

Extension of fuzzy TOPSIS method using picture fuzzy sets is proposed by Torun and Gordebil (2019). Picture fuzzy TOPSIS consists of the following steps:

Step 1. Construct the decision matrix of each decision maker using picture fuzzy scale given in Table 2.

Step 2. Calculate the aggregated decision matrix using Picture Fuzzy Weighted Arithmetic Mean (PFWAM) operator in the form of Eq. (41):

$$\widetilde{A}_p = \begin{array}{c} \\ f_1 \\ f_2 \\ \vdots \\ f_m \end{array} \begin{array}{c} \begin{array}{cccc} X_1 & X_2 & \dots & X_n \end{array} \\ \begin{bmatrix} \tilde{f}_p 11 & \tilde{f}_p 12 & \dots & \tilde{f}_p 1n \\ \tilde{f}_p 21 & \tilde{f}_p 22 & \dots & \tilde{f}_p 2n \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{f}_p m1 & \tilde{f}_p m2 & \dots & \tilde{f}_p mn \end{bmatrix} \end{array} \tag{41}$$

Step 3. Construct the weighted decision matrix $\overline{Y}_{p_w}$

$$\overline{Y}_{p_w} = \begin{array}{c} \\ f_1 \\ f_2 \\ \vdots \\ f_m \end{array} \begin{array}{c} \begin{array}{cccc} X_1 & X_2 & \dots & X_n \end{array} \\ \begin{bmatrix} \tilde{v}_p 11 & \tilde{v}_p 12 & \dots & \tilde{v}_p 1n \\ \tilde{v}_{p21} & \tilde{v}_p 22 & \dots & \tilde{v}_p 2n \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{v}_p m1 & \tilde{v}_p m2 & \dots & \tilde{v}_p mn \end{bmatrix} \end{array} \tag{42}$$

where $\tilde{v}_{p_{ij}} = \tilde{w}_{p_i} \otimes \tilde{f}_{p_{ij}}$.
Step 4. Defuzzify the weighted picture fuzzy decision matrix by using Eq. (43)

$$S\left(\overline{Y}_{p_w}\right) = \frac{1 + 2\mu_j - \nu_j - \frac{\pi_j}{2}}{2} \tag{43}$$

Step 5. Determine the positive ideal solution (PIS) and negative ideal solution (NIS).

PF_PIS can be calculated using Eq. (44) or Eq. (45)

$$\text{PFPIS} = \left\{ C_j, \max_i < S\left(\overline{Y}_{p_w}\right) > \middle| j = 1, 2, \dots, n \right\} \tag{44}$$

$$\text{PFPIS} = \left\{ \left\langle C_1, \left(\mu_1^+, \nu_1^+, \pi_1^+\right)\right\rangle, \ \dots, \left\langle C_n, \left(\mu_n^+, \nu_n^+, \pi_n^+\right)\right\rangle \right\} \tag{45}$$

PF_NIS can be calculated using Eq. (46) or Eq. (47)

$$\text{PFNIS} = \left\{ C_j, \min_i < S\left(\overline{Y}_{p_w}\right) > \middle| j = 1, 2, \dots, n \right\} \tag{46}$$

$$\text{PFPIS} = \left\{ \left\langle C_1, \left(\mu_1^-, \nu_1^-, \pi_1^-\right)\right\rangle, \ \dots, \left\langle C_n, \left(\mu_n^-, \nu_n^-, \pi_n^-\right)\right\rangle \right\} \tag{47}$$

Step 6. Calculate the distance of each alternative from PIS and NIS, respectively, using normalized Euclidian distances as follows:

$$D_i''^{+} = \sqrt{\frac{1}{2n} \sum_{i=1}^{n} \left(\left(\mu_{ij} - \mu_j^+\right)^2 + \left(\nu_{ij} - \nu_j^+\right)^2 + \left(\pi_{ij} - \pi_j^+\right)^2 \right)} \tag{48}$$

$$D_i''^{-} = \sqrt{\frac{1}{2n} \sum_{i=1}^{n} \left(\left(\mu_{ij} - \mu_j^-\right)^2 + \left(\nu_{ij} - \nu_j^-\right)^2 + \left(\pi_{ij} - \pi_j^-\right)^2 \right)} \tag{49}$$

Table 3 Linguistic terms and their corresponding interval-valued spherical fuzzy numbers

Linguistic terms	Interval-valued spherical fuzzy numbers	Saaty's Scale
Absolutely More Importance (AMI)	([0.85, 0.95], [0.10, 0.15], [0.05, 0.15])	9
Very High Importance (VHI)	([0.75, 0.85], [0.15, 0.20], [0.15, 0.20])	7
High Importance (HI)	([0.65, 0.75], [0.20, 0.25], [0.20, 0.25])	5
Slightly More Importance (SMI)	([0.55, 0.65], [0.25, 0.30], [0.25, 0.30])	3
Equally Importance (EI)	([0.50, 0.55], [0.45, 0.55], [0.30, 0.40])	1
Slightly Low Importance (SLI)	([0.25, 0.30], [0.55, 0.65], [0.25, 0.30])	1/3
Low Importance (LI)	([0.20, 0.25], [0.65, 0.75], [0.20, 0.25])	1/5
Very Low Importance (VLI)	([0.15, 0.20], [0.75, 0.85], [0.15, 0.20])	1/7
Absolutely Low Importance (ALI)	([0.10, 0.15], [0.85, 0.95], [0.05, 0.15])	1/9

Step 7. Calculate the relative degree of closeness with respect to the positive ideal solution

$$C_i = \frac{D_i^-}{D_i^- + D_i^+} \tag{50}$$

Step 8. Rank the alternatives according to C_i values where the best alternative has biggest C_i value.

3.3 Interval-Valued Spherical Fuzzy Extensions of AHP and TOPSIS

3.3.1 Interval-Valued Spherical AHP

Interval-valued spherical fuzzy AHP (IVS-AHP) method consists of the following steps (Gündoğdu & Kahraman, 2021c):

Step 1. Define problem and construct hierarchical structure of the problem

Step 2. Establish pairwise comparison matrices using linguistic statements that are determined in Table 3 where an interval-valued spherical number is denoted as $\widetilde{A}_s = \left(\left[\mu^L_{\widetilde{A}_s}(u), \mu^U_{\widetilde{A}_s}(u)\right], \left[\nu^L_{\widetilde{A}_s}(u), \nu^U_{\widetilde{A}_s}(u)\right], \left[\pi^L_{\widetilde{A}_s}(u), \pi^U_{\widetilde{A}_s}(u)\right]\right)$.

Step 3. Construct interval-valued spherical fuzzy comparison matrices.

Pairwise comparison matrices are formed for each decision maker using interval-valued spherical fuzzy scale given in Table 3 as in the form of Eq. (51) and $k = 1, \ldots s$.

$$\widetilde{A}_s^{(k)} = \begin{bmatrix} 1 \; \tilde{a}_{s_{12}}^{(k)} & \ldots \; \tilde{a}_{s_{1n}}^{(k)} \\ \tilde{a}_{s_{21}}^{(k)} \; 1 & \ldots \; \tilde{a}_{s_{2n}}^{(k)} \\ \vdots \quad \vdots & \ddots \quad \vdots \\ \tilde{a}_{s_{n1}}^{(k)} \; \tilde{a}_{s_{n2}}^{(k)} & \cdots \quad 1 \end{bmatrix} \tag{51}$$

Step 4. Check the consistency of each fuzzy pairwise comparison matrix.

Consistency ratio of each pairwise comparison matrix can be calculated using score indices that are given in Table 3 and classical consistency check formula (Saaty, 1980). Pairwise comparison matrices are considered as consistent when consistency ratio (CR) is less than 0.1.

Step 5. Compute aggregated interval-valued spherical fuzzy comparison matrix.

When there are several decision makers, evaluations of each criterion can be aggregated using interval-valued spherical fuzzy weighted geometric mean operator by Eq. (52) where w_k represents weight of the decision maker that satisfies

$$\begin{aligned} \textstyle\sum_{k=1}^{s} w_k = 1 \tilde{w}_{s_{ij}} = \Bigg[& \prod_{k=1}^{s} \mu_{ij}^{L(k)\,w_k}, \prod_{k=1}^{s} \mu_{ij}^{U(k)\,w_k}, \left[1 - \prod_{k=1}^{s}\left(1 - \nu_{ij}^{L(k)\,2}\right)^{w_k}\right]^{1/2}, \\ & \left[1 - \prod_{k=1}^{s}\left(1 - \nu_{ij}^{U(k)\,2}\right)^{w_k}\right]^{1/2}, \left(\prod_{k=1}^{s}\left(1 - \nu_{ij}^{L(k)\,2}\right)^{w_k} - \prod_{k=1}^{s}\left(1 - \nu_{ij}^{L(k)\,2} - \pi_{ij}^{L(k)\,2}\right)^{w_k}\right)^{1/2}, \\ & \left(\prod_{k=1}^{s}\left(1 - \nu_{ij}^{U(k)\,2}\right)^{w_k} - \prod_{k=1}^{s}\left(1 - \nu_{ij}^{U(k)\,2} - \pi_{ij}^{U(k)\,2}\right)^{w_k}\right)\Bigg] \end{aligned} \tag{52}$$

Step 6. Calculate the local and global weights of criteria.

The spherical fuzzy local weights of each criterion are obtained by using IVSWAM operator and given in Eq. (53). The weighted arithmetic mean is used to compute the interval-valued spherical fuzzy weights.

$$\begin{aligned} \tilde{w}_{s_j} = \Bigg[& \left(1 - \prod_{j=1}^{n}\left(1 - \mu_j^{L^2}\right)^{w_j}\right)^{1/2}, \left(1 - \prod_{j=1}^{n}\left(1 - \mu_j^{U^2}\right)^{w_j}\right)^{1/2}, \prod_{k=1}^{n}\left(\nu_j^{L}\right)^{w_j} \cdot \prod_{k=1}^{n}\left(\nu_j^{U}\right)^{w_j}, \\ & \left(\prod_{j=1}^{n}\left(1 - \mu_j^{L^2}\right)^{w_j} - \prod_{j=1}^{n}\left(1 - \mu_j^{L^2} - \pi_j^{L^2}\right)^{w_j}\right)^{1/2}, \\ & \left(\prod_{j=1}^{n}\left(1 - \mu_j^{U^2}\right)^{w_j} - \prod_{j=1}^{n}\left(1 - \mu_j^{U^2} - \pi_j^{U^2}\right)^{j}\right)^{1/2}\Bigg] \end{aligned} \tag{53}$$

where $w_j = 1/n$.

To compute the global weights, the local weights of each criterion are defuzzified by a modified score function [$S\left(\tilde{w}_{s_j}\right)$] using Eq. (54) and normalized using Eq. (55).

$$S\left(\tilde{w}_{s_j}\right)=\frac{\left(\mu_j^L\right)^2+\left(\mu_j^U\right)^2-\left(\nu_j^L\right)^2-\left(\nu_j^U\right)^2-\left(\frac{\left(\pi_j^L\right)}{2}\right)^2-\left(\frac{\left(\pi_j^U\right)}{2}\right)^2}{2}+1 \tag{54}$$

$$\overline{w}_{s_j}=\frac{S\left(\tilde{w}_{s_j}\right)}{\sum_{j=1}^{n} S\left(\tilde{w}_{s_j}\right)} \tag{55}$$

Step 6. Compute final interval-valued spherical fuzzy weights.

The final interval-valued spherical fuzzy weights for each criterion and sub-criterion are calculated by multiplying the fuzzy global weights at first level by each related sub-criterion global weight.

3.3.2 Interval-Valued Spherical: TOPSIS

Gündoğdu and Kahraman (2019b) proposed one of the newest extensions of TOPSIS method using interval-valued spherical fuzzy sets. The method consists of the following steps:

Step 1. Construct the decision matrix of each decision maker using picture fuzzy scale given in Table 3.

Step 2. Calculate the aggregated decision matrix using Interval-valued Spherical Weighted Arithmetic Mean (IVSWAM) operator Eq. (56):

$$\widetilde{A}_s=\begin{array}{c} \\ f_1 \\ f_2 \\ \vdots \\ f_m \end{array}\begin{array}{c} \begin{array}{cccc} X_1 & X_2 & \ldots & X_n \end{array} \\ \begin{bmatrix} \tilde{f}_s 11 & \tilde{f}_s 12 & \ldots & \tilde{f}_s 1n \\ \tilde{f}_s 21 & \tilde{f}_s 22 & \ldots & \tilde{f}_s 2n \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{f}_s m1 & \tilde{f}_s m2 & \ldots & \tilde{f}_s mn \end{bmatrix} \end{array} \tag{56}$$

where $\tilde{f}_s ij=\left(\left[\mu_{ij}^L(u),\mu_{ij}^U(u)\right],\left[\nu_{ij}^L(u),\nu_{ij}^U(u)\right],\left[\pi_{ij}^L(u),\pi_{ij}^U(u)\right]\right)$

Step 3. Construct the weighted interval-valued spherical fuzzy decision matrix $\overline{Y}_{s_w}$:

$$\overline{Y}_{s_w}=\begin{array}{c} \\ f_1 \\ f_2 \\ \vdots \\ f_m \end{array}\begin{array}{c} \begin{array}{cccc} X_1 & X_2 & \ldots & X_n \end{array} \\ \begin{bmatrix} \tilde{v}_s 11 & \tilde{v}_s 12 & \ldots & \tilde{v}_s 1n \\ \tilde{v}_{s21} & \tilde{v}_s 22 & \ldots & \tilde{v}_s 2n \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{v}_s m1 & \tilde{v}_s m2 & \ldots & \tilde{v}_s mn \end{bmatrix} \end{array} \tag{57}$$

where

$$\tilde{v}_{s_{ij}} = \tilde{w}_{s_i} \otimes \tilde{f}_{s_{ij}} = \left[\mu^L_{ijw}(u), \mu^U_{ijw}(u)\right], \left[\nu^L_{ijw}(u), \nu^U_{ijw}(u)\right], \left[\pi^L_{ijw}(u), \pi^U_{ijw}(u)\right].$$

Step 4. Defuzzify the weighted interval-valued spherical fuzzy decision matrix by using Eq. (58)

$$S\left(\overline{Y}_{s_w}\right) = \frac{\left(\mu^L_j\right)^2 + \left(\mu^U_j\right)^2 - \left(\nu^L_j\right)^2 - \left(\nu^U_j\right)^2 - \left(\frac{\left(\pi^L_j\right)}{2}\right)^2 - \left(\frac{\left(\pi^U_j\right)}{2}\right)^2}{2} + 1 \tag{58}$$

Step 5. Determine the Interval-valued Spherical Fuzzy Positive Ideal Solution (IVSF-PIS) and the Interval-valued Spherical Fuzzy Negative Ideal Solution (IVSF-NIS).

IVSF_PIS can be calculated using Eq. (59) or Eq. (60)

$$\text{IVSFPIS} = \left\{C_j, \max_i < S\left(\overline{Y}_{s_w}\right) > \middle| j = 1, 2, \ldots, n\right\} \tag{59}$$

$$\begin{aligned} \text{IVSFPIS} = \Big\{&\Big\langle C_1, \Big(\left[\mu^{L+}_1, \mu^{U+}_1\right], \left[\nu^{L+}_1, \nu^{U+}_1\right], \left[\pi^{L+}_1, \pi^{U+}_1\right]\Big)\Big\rangle, \ \ldots, \\ &\Big\langle C_n, \Big(\left[\mu^{L+}_n, \mu^{U+}_n\right], \left[\nu^{L+}_n, \nu^{U+}_n\right], \left[\pi^{L+}_n, \pi^{U+}_n\right]\Big)\Big\rangle\Big\} \end{aligned} \tag{60}$$

IVSF_NIS can be calculated using Eq. (61) or Eq. (62)

$$\text{IVSFNIS} = \left\{C_j, \min_i < S\left(\overline{Y}_{s_w}\right) > \middle| j = 1, 2, \ldots, n\right\} \tag{61}$$

$$\begin{aligned} \text{IVSFNIS} = \Big\{&\Big\langle C_1, \Big(\left[\mu^{L-}_1, \mu^{U-}_1\right], \left[\nu^{L-}_1, \nu^{U-}_1\right], \left[\pi^{L-}_1, \pi^{U-}_1\right]\Big)\Big\rangle, \ \ldots, \\ &\Big\langle C_n, \Big(\left[\mu^{L-}_n, \mu^{U-}_n\right], \left[\nu^{L-}_n, \nu^{U-}_n\right], \left[\pi^{L-}_n, \pi^{U-}_n\right]\Big)\Big\rangle\Big\} \end{aligned} \tag{62}$$

Step 6. Calculate the distance of each alternative from IVSF-PIS and the IVSF-NIS, respectively, using normalized distance formula (Peng & Yang, 2016a, 2016b):

$$\begin{aligned} D''^{+}_i = \tfrac{1}{4\text{n}} \textstyle\sum_{\text{j}=1}^{\text{n}} \Big(&\left|\left(\mu^L_{ij}\right)^2 - \left(\mu^+_j\right)^2\right| + \left|\left(\mu^U_{ij}\right)^2 - \left(\mu^+_j\right)^2\right| + \left|\left(\nu^L_{ij}\right)^2 - \left(\nu^+_j\right)^2\right| + \\ &\left|\left(\nu^U_{ij}\right)^2 - \left(\nu^+_j\right)^2\right| + \left|\left(\pi^L_{ij}\right)^2 - \left(\pi^+_j\right)^2\right| + \left|\left(\pi^U_{ij}\right)^2 - \left(\pi^+_j\right)^2\right|\Big) \end{aligned} \tag{63}$$

$$D_i^{\prime\prime -} = \frac{1}{4\mathrm{n}} \sum_{\mathrm{j}=1}^{\mathrm{n}} \left(\left| \left(\mu_{ij}^L\right)^2 - \left(\mu_j^-\right)^2 \right| + \left| \left(\mu_{ij}^U\right)^2 - \left(\mu_j^-\right)^2 \right| + \left| \left(\nu_{ij}^L\right)^2 - \left(\nu_j^-\right)^2 \right| \right.$$
$$\left. + \left| \left(\nu_{ij}^U\right)^2 - \left(\nu_j^-\right)^2 \right| + \left| \left(\pi_{ij}^L\right)^2 - \left(\pi_j^-\right)^2 \right| + \left| \left(\pi_{ij}^U\right)^2 - \left(\pi_j^-\right)^2 \right| \right) \tag{64}$$

Step 7. Calculate the relative degree of closeness with respect to the positive ideal solution

$$C_i = \frac{D_i^-}{D_i^- + D_i^+} \tag{65}$$

Step 8. Rank the alternatives according to C_i values where the best alternative has biggest C_i value.

4 Application

Based on the literature review the hierarchy of the decision model that is given in Fig. 1 is developed for defining the COVID-19 Intervention Strategies. In this hierarchy health-related costs, public prevention-related costs, and social impacts are the main criteria. Health-related costs involve all the costs occurred due to intensive and non-intensive care costs such as the machines, electricity, etc., the associated labor costs which include the cost of medical workers, and supply chain costs which include the cost of medical materials used (Bavli et al., 2020; Roberts et al., 2021). The public prevention-related costs involve the net loss in the total domestic business, the labor costs which involve the costs on the government that are related with unemployment, direct and indirect incentive costs that are provided by the governments, and the supply of preventive materials such as surgical masks and disinfectants (Danielli et al., 2021; Harling et al., 2021). The last criteria are the social impacts which involve the future educational problems (such as non-return to school, inequality in education and social services, etc.), the social and psychological well-being of citizens and physical well-being of citizens (Pedrozo-Pupo and Campo-Arias 2020; Piltch-Loeb et al., 2021).

Based on the literature, nine different intervention strategies by combining different measures such as lockdown, border restrictions, closure of schools and businesses, curfew and social distancing are defined. These nine intervention strategy alternatives can be given as follows:

A1: Total Lockdown with internal and external border restrictions, schools are closed and restriction of all nonessential businesses including delivery services

A2: Total Lockdown with internal and external border restrictions, schools are open and restriction of all nonessential businesses including delivery services

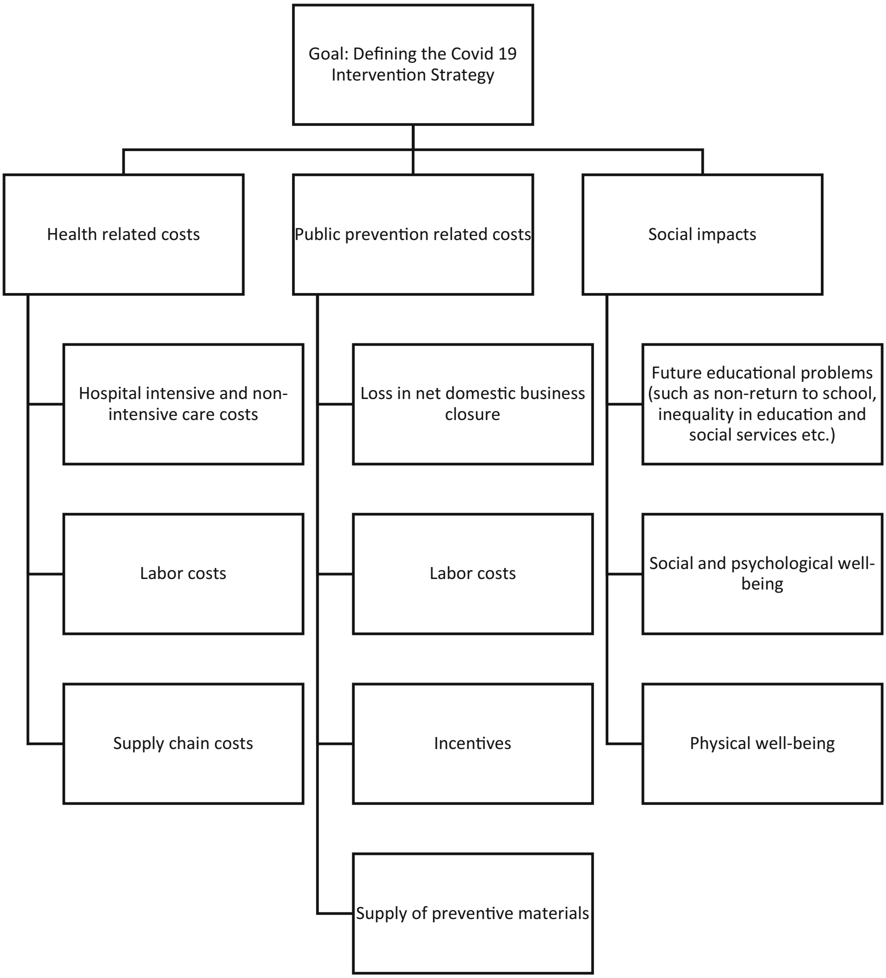


Fig. 1 Defining COVID-19 intervention strategies

A3: Total Lockdown with internal and external border restrictions, schools are closed and going out for business purposes is possible; only restaurants are closed
A4: Internal border restrictions reducing the ability to move freely (transportation) within a country, external border restrictions, Quarantine/lockdown of patients and those suspected of infection, schools are closed, businesses are open
A5: Internal border restrictions reducing the ability to move freely (transportation) within a country, quarantine of patients and those suspected of infection, external border restrictions, schools are open, businesses are open
A6: Curfew, External border precautions but no restrictions, quarantine of patients and those suspected of infection, schools are closed, businesses are open

Table 4 Assessments of the DMs for the current situation in terms of resource usage rate

	Low	Medium	High
DM 1	Disagree	Moderately agree	Agree
DM 2	Moderately agree	Agree	Moderately agree
DM 3	Disagree	Moderately agree	Moderately agree

A7: Curfew, External border precautions but no restrictions, Quarantine/lockdown of patients and those suspected of infection schools are open, businesses are open
A8: Quarantine/lockdown of patients and those suspected of infection, social distancing and Restrictions of mass gatherings no further restrictions
A9: Quarantine/lockdown of patients, no other restrictions

In order to determine the most appropriate strategy according to the current pandemic conditions, the proposed three-stage model is applied. The experts whose opinions were taken in the study were selected from among the managers who closely observed the economic, social, and health effects of the alternative strategies applied before, and are authorized to take COVID-19 measures. At first, three managers from different levels determined linguistically the degree to which the current situation belongs to low, medium, and high resource usage rate scenarios using the scale given in Table 1. Resource utilization rate is a data that includes hospital bed occupancy, intensive care bed occupancy, use of medical equipment and drug use, and can be an indicator for the number of COVID-19 cases and severe patients and expresses the severity of the situation. The questions are asked as “Can the resource usage rate be described as “low” in the current situation?” for three situations (low, medium, and high). The statements of the decision makers (DMs) are given in Table 4:

Using PFWAM operator membership value of current situation to “low resource usage rate” scenario is calculated as (0.3396, 0.1145, 0.5277) where it is calculated as (0.6069, 0.1310, 0.2621) for medium and high resource usage rates. Scores of the membership values are calculated using Eq. (12) and normalized. The weights of the situations are calculated as 0.25, 0.375, and 0.375 for low, medium, and high resource usage rates, respectively.

Using SFWAM operator membership value of current situation to “low resource usage rate” scenario is calculated as ([0.4771, 0.5324], [0.575, 0.6763], [0.2461, 0.3049]) where it is calculated as ([0.7234, 0.7776], [0.3434, 0.4228], [0.2666, 0.3469]) for medium and high resource usage rates. Scores of the membership values are calculated using Eq. (54) and normalized. The weights of the situations are calculated as 0.2324, 0.3838, and 0.3838 for low, medium, and high resource usage rates, respectively.

Then decision makers made pairwise comparisons of the criteria for the determined three pandemic situations. For convenience, all criteria are defined as benefit criteria as shown below:

C1: Decrease on health-related costs

C11: Decrease in hospital intensive and non-intensive care costs
C12: Decrease in healthcare professionals' costs
C13: Decrease in supply chain costs

C2: Decrease in public prevention-related costs

C21: Decrease in the loss in net domestic business closure
C22: Decrease in labor costs
C23: Decrease in the amount of incentives
C24: Decrease in procurement costs of preventive materials

C3: Decrease in social impact

C31: Reduction in possible future educational problems
C32: Increase in social and psychological well-being
C33: Increase in physical well-being

As an example, linguistic judgements of one decision maker are given in Table 5. Consistency ratios of the comparisons are all calculated below 0.10.

Fuzzy AHP method based on picture fuzzy and spherical fuzzy scales is applied to each of the pandemic situation, separately. As a result of the calculations, three different criteria weights are calculated, each for a different pandemic situation. The results of Picture fuzzy AHP are given in Table 6.

Local and global weights of criteria calculated using Spherical fuzzy AHP method for the situations when resource usage rate is low, medium, and high are shown in Tables 7, 8, and 9, respectively.

To determine current situation's criteria weights, fuzzy weighted arithmetic mean operator is used for the criteria weights of three different situations and normalized score values of current situation's membership functions for each situation. The obtained final criteria weights are given in Table 10 both using picture fuzzy sets and spherical fuzzy sets.

As it is seen in Table 10, normalized score functions of picture fuzzy weights and spherical fuzzy weights are close to each other but different. This difference may come from the definition of spherical fuzzy sets' membership functions whose squared sums of their parameters can be at most equal to 1. The ranking of the criteria in picture fuzzy sets C11>C12>C13>C21>C24>C23>C22>C31>C33>C32, whereas it is C11>C24>C12>C21>C23>C13>>C22>C33>C31>C32 in spherical fuzzy sets. The rank of the criteria C24 and C13 is changed significantly. C24 gained more importance while C13 lost importance.

At the last stage of the proposed methodology, nine alternative strategies which are detailed in Sect. 2 are examined and managers are asked to evaluate the alternative strategies according to the criteria.

Decision makers' linguistic assessments for the alternatives are given in Table 11.

Table 5 Example of pairwise comparisons of criteria

	When resource usage rate is low				When resource usage rate is medium				When resource usage rate is high			
	C1	**C2**	**C3**		**C1**	**C2**	**C3**		**C1**	**C2**	**C3**	
C1	EI	EI	HI		EI	SLI	HI		EI	HI	VHI	
C2		EI	SMI			EI	VHI			EI	SMI	
C3			EI				EI				EI	
	C11	**C12**	**C13**		**C11**	**C12**	**C13**		**C11**	**C12**	**C13**	
C11	EI	SLI	ALI		EI	SMI	VHI		EI	EI	AMI	
C12		EI	SLI			EI	HI			EI	HI	
C13			EI				EI				EI	
	C21	**C22**	**C23**	**C24**	**C21**	**C22**	**C23**	**C24**	**C21**	**C22**	**C23**	**C24**
C21	EI	SMI	LI	VLI	EI	SMI	LI	LI	EI	EI	SMI	SMI
C22		EI	ALI	ALI		EI	LI	ALI		EI	SMI	HI
C23			EI	EI			EI	SLI			EI	EI
C24				EI				EI				EI
	C31	**C32**	**C33**		**C31**	**C32**	**C33**		**C31**	**C32**	**C33**	
C31	EI	SMI	LI		EI	SMI	SLI		EI	LI	SLI	
C32		EI	VLI			EI	LI			EI	SMI	
C33			EI				EI				EI	

Table 6 Criteria weights obtained from PF-AHP

	When resource usage rate is low		When resource usage rate is medium		When resource usage rate is high	
	Local weights	Global weights	Local weights	Global weights	Local weights	Global weights
C1	(0.45, 0.48, 0)		(0.53, 0.41, 0)		(0.72, 0.19, 0)	
C2	(0.53, 0.41, 0)		(0.63, 0.3, 0)		(0.48, 0.44, 0)	
C3	(0.41, 0.51, 0)		(0.32, 0.62, 0)		(0.31, 0.63, 0)	
C11	(0.37, 0.59, 0)	(0.16, 0.8, 0)	(0.63, 0.31, 0)	(0.25, 0.71, 0)	(0.71, 0.25, 0)	(0.4, 0.51, 0)
C12	(0.5, 0.46, 0)	(0.18, 0.78, 0)	(0.42, 0.53, 0)	(0.21, 0.75, 0)	(0.4, 0.56, 0)	(0.32, 0.6, 0)
C13	(0.52, 0.45, 0)	(0.19, 0.78, 0)	(0.38, 0.58, 0)	(0.21, 0.76, 0)	(0.35, 0.61, 0)	(0.31, 0.61, 0)
C21	(0.41, 0.55, 0)	(0.17, 0.8, 0)	(0.42, 0.54, 0)	(0.22, 0.74, 0)	(0.56, 0.43, 0)	(0.16, 0.8, 0)
C22	(0.36, 0.6, 0)	(0.16, 0.81, 0)	(0.38, 0.58, 0)	(0.21, 0.75, 0)	(0.37, 0.59, 0)	(0.14, 0.83, 0)
C23	(0.48, 0.5, 0)	(0.18, 0.79, 0)	(0.44, 0.52, 0)	(0.22, 0.74, 0)	(0.45, 0.53, 0)	(0.15, 0.81, 0)
C24	(0.46, 0.53, 0.07)	(0.17, 0.8, 0)	(0.54, 0.42, 0.11)	(0.23, 0.73, 0)	(0.46, 0.49, 0.09)	(0.15, 0.82, 0)
C31	(0.55, 0.4, 0)	(0.17, 0.78, 0)	(0.43, 0.51, 0)	(0.12, 0.86, 0)	(0.42, 0.52, 0)	(0.11, 0.86, 0)
C32	(0.37, 0.61, 0)	(0.15, 0.82, 0)	(0.41, 0.52, 0)	(0.12, 0.86, 0)	(0.48, 0.45, 0)	(0.12, 0.85, 0)
C33	(0.46, 0.51, 0)	(0.16, 0.8, 0)	(0.48, 0.45, 0)	(0.12, 0.85, 0)	(0.44, 0.51, 0)	(0.12, 0.86, 0)

Table 7 Criteria weights when resource usage rate is low are obtained from IVSF-AHP

	Local weights	Global weights
C1	([0.42, 0.48], [0.47, 0.57], [0.27, 0.34])	
C2	([0.5, 0.57], [0.38, 0.47], [0.27, 0.35])	
C3	([0.39, 0.45], [0.5, 0.6], [0.26, 0.34])	
C11	([0.37, 0.43], [0.54, 0.65], [0.25, 0.33])	([0.23, 0.26], [0.81, 0.85], [0.15, 0.21])
C12	([0.45, 0.52], [0.42, 0.51], [0.27, 0.35])	([0.26, 0.3], [0.77, 0.82], [0.17, 0.23])
C13	([0.47, 0.53], [0.42, 0.51], [0.26, 0.34])	([0.26, 0.3], [0.76, 0.82], [0.17, 0.23])
C21	([0.4, 0.46], [0.49, 0.59], [0.25, 0.33])	([0.29, 0.33], [0.75, 0.79], [0.17, 0.23])
C22	([0.36, 0.42], [0.55, 0.66], [0.23, 0.31])	([0.27, 0.31], [0.77, 0.82], [0.16, 0.22])
C23	([0.44, 0.51], [0.46, 0.55], [0.26, 0.34])	([0.3, 0.35], [0.73, 0.78], [0.18, 0.24])
C24	([0.43, 0.5], [0.47, 0.57], [0.25, 0.33])	([0.3, 0.34], [0.73, 0.78], [0.17, 0.24])
C31	([0.52, 0.61], [0.35, 0.43], [0.25, 0.32])	([0.25, 0.29], [0.75, 0.81], [0.17, 0.23])
C32	([0.36, 0.41], [0.55, 0.65], [0.25, 0.33])	([0.21, 0.24], [0.83, 0.87], [0.14, 0.19])
C33	([0.43, 0.5], [0.45, 0.55], [0.25, 0.33])	([0.23, 0.27], [0.79, 0.84], [0.16, 0.21])

Table 8 Criteria weights when resource usage rate is medium are obtained from IVSF-AHP

	Local weights	Global weights
C1	([0.5, 0.58], [0.37, 0.45], [0.26, 0.34])	
C2	([0.6, 0.69], [0.29, 0.37], [0.25, 0.32])	
C3	([0.34, 0.38], [0.59, 0.69], [0.24, 0.32])	
C11	([0.59, 0.68], [0.27, 0.33], [0.24, 0.31])	([0.34, 0.4], [0.66, 0.72], [0.19, 0.25])
C12	([0.41, 0.46], [0.5, 0.59], [0.26, 0.35])	([0.28, 0.33], [0.76, 0.8], [0.16, 0.21])
C13	([0.39, 0.44], [0.53, 0.63], [0.26, 0.34])	([0.27, 0.32], [0.77, 0.81], [0.15, 0.21])
C21	([0.41, 0.47], [0.48, 0.57], [0.26, 0.34])	([0.35, 0.41], [0.7, 0.75], [0.16, 0.22])
C22	([0.36, 0.42], [0.53, 0.63], [0.24, 0.32])	([0.33, 0.39], [0.73, 0.77], [0.15, 0.21])
C23	([0.41, 0.47], [0.47, 0.57], [0.26, 0.33])	([0.35, 0.41], [0.7, 0.75], [0.16, 0.22])
C24	([0.49, 0.57], [0.37, 0.45], [0.24, 0.31])	([0.38, 0.45], [0.65, 0.7], [0.17, 0.24])
C31	([0.42, 0.48], [0.49, 0.58], [0.25, 0.32])	([0.19, 0.21], [0.86, 0.9], [0.14, 0.19])
C32	([0.39, 0.45], [0.5, 0.6], [0.26, 0.34])	([0.18, 0.21], [0.86, 0.9], [0.13, 0.18])
C33	([0.45, 0.52], [0.44, 0.54], [0.25, 0.32])	([0.19, 0.22], [0.85, 0.89], [0.14, 0.2])

Fuzzy TOPSIS extensions using picture fuzzy sets and spherical fuzzy sets are applied after aggregation of the decision makers' evaluation using PFSWAM and SFWAM operators.

Positive and negative idea solutions obtained using PF-TOPSIS and IVSF-TOPSIS are given in Tables 12 and 13, respectively.

The results of the calculations for the current situation are summarized in Table 14.

The ranking results indicate that the first two alternatives are same whereas A3 goes from third rank in PF-TOPSIS to sixth rank in IVSF-TOPSIS and the other alternatives change their rank one lower or one upper positions.

Table 9 Criteria weights when resource usage rate is high are obtained from IVSF-AHP

	Local weights	Global weights
C1	([0.68, 0.79], [0.23, 0.29], [0.2, 0.27])	
C2	([0.47, 0.55], [0.4, 0.48], [0.26, 0.33])	
C3	([0.33, 0.38], [0.6, 0.71], [0.24, 0.32])	
C11	([0.68, 0.77], [0.26, 0.33], [0.21, 0.28])	([0.49, 0.59], [0.51, 0.58], [0.16, 0.24])
C12	([0.4, 0.46], [0.52, 0.63], [0.25, 0.34])	([0.39, 0.47], [0.68, 0.72], [0.13, 0.2])
C13	([0.38, 0.43], [0.56, 0.66], [0.25, 0.33])	([0.37, 0.46], [0.7, 0.74], [0.13, 0.2])
C21	([0.52, 0.6], [0.37, 0.45], [0.27, 0.35])	([0.29, 0.34], [0.73, 0.77], [0.17, 0.23])
C22	([0.36, 0.41], [0.54, 0.64], [0.25, 0.32])	([0.24, 0.29], [0.8, 0.84], [0.14, 0.19])
C23	([0.41, 0.47], [0.47, 0.57], [0.26, 0.34])	([0.26, 0.31], [0.77, 0.81], [0.15, 0.21])
C24	([0.43, 0.5], [0.46, 0.55], [0.25, 0.33])	([0.27, 0.31], [0.77, 0.81], [0.15, 0.21])
C31	([0.38, 0.44], [0.5, 0.6], [0.26, 0.34])	([0.18, 0.2], [0.87, 0.91], [0.13, 0.18])
C32	([0.45, 0.52], [0.42, 0.51], [0.27, 0.35])	([0.19, 0.22], [0.85, 0.9], [0.14, 0.19])
C33	([0.45, 0.52], [0.42, 0.51], [0.27, 0.35])	([0.19, 0.22], [0.85, 0.9], [0.14, 0.19])

Additionally, the ranks of the alternatives for the determined resource usage rate levels without taking into account membership function of current situation are calculated and the results are shown in Table 15.

In Table 15, the ranking of the alternatives with respect to the sum of the rank columns in PF-TOPSIS is A2>A1>A3=A5>A4>A6>A7>A9>A8, whereas it is A2>A5>A1>A4>A6>A3>A7>A8>A9 in SF-TOPSIS. These results indicate that A2 is emerging as the most viable alternative strategy in all circumstances. The strategy A2 includes total lockdown with internal and external border restrictions and restriction of all nonessential businesses including delivery services but schools are open. It is also seen that, the proposed method, which is observed to highlight different strategies at different resource utilization rates, will be very useful in determining the appropriate strategy according to the dynamically changing resource utilization rate due to emerging new variants.

When the results are examined in terms of different resource utilization rates, it is observed that as the resource utilization rate increases, in other words, as the number of cases increases, alternatives with more strict closure are more prioritized, as expected. The proposed method will be very useful in determining the appropriate strategy according to the dynamically changing resource usage rate due to new variants.

In addition, the evaluations are also analyzed using crisp AHP and TOPSIS methods. The weights and rankings obtained by the classical method are given in Tables 16 and 17.

In crisp case, although the criteria weights differ considerably according to resource usage rates, it has been observed that there is almost no differentiation in the ranking of the of alternative strategies for different resource use rates considered in the evaluation. This shows that the fuzzy methodologies used are very sensitive to the uncertainty in the data.

Table 10 Criteria global weights for current situation

	Picture fuzzy weights (PFW)	Normalized score of PFW	Spherical fuzzy weights (SFW)	Normalized score of SFW
C11	(0.2909, 0.6464, 0)	0.1265	([0.3906, 0.4693], [0.6288, 0.6856], [0.1704, 0.2388]	0.1538
C12	(0.2495, 0.6956, 0)	0.1157	([0.3213, 0.3868], [0.7283, 0.7742], [0.1505, 0.2117]	0.1145
C13	(0.2417, 0.7049, 0)	0.1137	([0.3121, 0.3758], [0.7408, 0.7851], [0.1477, 0.2078]	0.1094
C21	(0.1865, 0.7782, 0)	0.0989	([0.3134, 0.3694], [0.7209, 0.7678], [0.164, 0.2249]	0.1145
C22	(0.173, 0.7942, 0)	0.0954	([0.2855, 0.3372], [0.7666, 0.8067], [0.1489, 0.2058]	0.0974
C23	(0.185, 0.7803, 0)	0.0985	([0.3067, 0.3617], [0.7336, 0.7786], [0.1599, 0.22]	0.1099
C24	(0.1864, 0.7787, 0)	0.0989	([0.3229, 0.3807], [0.7106, 0.7586], [0.1659, 0.228]	0.1189
C31	(0.1307, 0.8391, 0)	0.0846	([0.2003, 0.2306], [0.8378, 0.8818], [0.1445, 0.1964]	0.062
C32	(0.1256, 0.8455, 0)	0.0833	([0.1912, 0.2199], [0.8513, 0.8926], [0.1387, 0.1892]	0.0569
C33	(0.1303, 0.8398, 0)	0.0845	([0.2017, 0.232], [0.8356, 0.8807], [0.146, 0.1986]	0.0627

Table 11 Linguistic assessments of alternatives

	C11	C12	C13	C21	C22	C23	C24	C31	C32	C33
	Decision maker 1									
A1	AMI	VHI	VHI	SMI	VHI	LI	VLI	ALI	ALI	ALI
A2	VHI	VHI	AMI	SMI	HI	LI	VLI	VLI	VLI	SLI
A3	HI	SMI	VHI	SLI	HI	SLI	LI	SLI	VLI	SLI
A4	HI	EI	HI	SLI	HI	SLI	LI	SLI	SLI	EI
A5	SMI	SLI	SMI	EI	SMI	EI	EI	EI	EI	EI
A6	EI	LI	EI	SMI	SMI	HI	HI	HI	SMI	HI
A7	SLI	LI	EI	HI	EI	HI	VLI	VLI	HI	HI
A8	SLI	VLI	LI	VHI	LI	VHI	VLI	VLI	HI	AMI
A9	LI	VLI	VLI	AMI	LI	AMI	VLI	ALI	VHI	AMI
	Decision maker 2									
A1	SMI	HI	VHI	AMI	AMI	AMI	AMI	AMI	ALI	ALI
A2	SMI	HI	VHI	AMI	AMI	AMI	AMI	LI	VLI	ALI
A3	SMI	HI	VHI	AMI	VHI	VHI	VHI	LI	VLI	ALI
A4	SMI	SMI	HI	HI	HI	HI	HI	VHI	ALI	VLI
A5	SMI	SMI	HI	HI	HI	HI	HI	LI	LI	SLI
A6	HI	SLI	SMI	SMI	SMI	SMI	SMI	VHI	ALI	VLI
A7	HI	SLI	SMI	SMI	SMI	SMI	SLI	LI	SMI	SMI
A8	ALI	LI	SLI	SMI	SLI	SMI	SLI	VLI	SMI	SMI
A9	ALI	LI	SLI	SMI	SLI	SMI	SLI	VLI	SMI	SMI
	Decision maker 3									
A1	AMI	VHI	AMI	ALI	VHI	VLI	AMI	ALI	ALI	ALI
A2	VHI	VHI	VHI	ALI	VHI	VLI	VHI	SMI	LI	VLI
A3	VHI	HI	VHI	VLI	HI	VLI	VHI	ALI	VLI	VLI
A4	HI	HI	HI	SMI	SLI	HI	HI	AMI	VLI	VLI
A5	SMI	SMI	SMI	HI	SLI	HI	SLI	AMI	SMI	VLI
A6	HI	SLI	HI	HI	SLI	SMI	SMI	ALI	SLI	LI
A7	SMI	SLI	SLI	HI	SLI	SMI	SLI	SLI	HI	SLI
A8	LI	LI	ALI	HI	LI	HI	VLI	ALI	VHI	HI
A9	ALI	VLI	ALI	VHI	VLI	VHI	ALI	ALI	AMI	HI

5 Conclusion

Strategic decision-making is one of the multi-criteria decision-making types whose results are observed in the long term and which contain significant uncertainty. In this chapter, strategic decision-making models for different scenarios under pandemic conditions are discussed and modeled in an uncertain environment. The uncertainty in models has been addressed with the help of fuzzy sets theory, which has been used frequently in recent years. Picture fuzzy sets and spherical fuzzy sets have been used in modeling uncertainty. In the evaluation of strategies, vagueness and impreciseness in the linguistic expressions used by decision makers are captured with picture and spherical fuzzy sets.

Table 12 Positive and negative ideal solutions in PF-TOPSIS

	Positive ideal solution	Negative ideal solution
C11	(0.2447, 0.6802, 0)	(0.0299, 0.9391, 0)
C12	(0.1924, 0.7459, 0.063)	(0.0458, 0.9184, 0.0794)
C13	(0.2033, 0.7356, 0)	(0.0467, 0.928, 0)
C21	(0.1492, 0.8088, 0)	(0.1068, 0.8553, 0.063)
C22	(0.1455, 0.8156, 0)	(0.0439, 0.931, 0.0794)
C23	(0.148, 0.8106, 0)	(0.0962, 0.8717, 0.05)
C24	(0.1472, 0.808, 0)	(0.0412, 0.94, 0.05)
C31	(0.1, 0.8657, 0)	(0.0311, 0.951, 0.063)
C32	(0.1005, 0.8668, 0)	(0.021, 0.9637, 0.063)
C33	(0.1005, 0.8641, 0)	(0.0065, 0.984, 0)

The contribution of this chapter can be given as follows:

1. Modeling the uncertainty in human thoughts through picture fuzzy sets and spherical fuzzy sets
2. Application of the proposed methodologies to the prioritization of intervention strategies
3. Scenario-based evaluation under different conditions.

The originality of this study is that the membership values are determined by experts according to the current resource usage rate being low, medium, and high. Afterwards, the membership degrees of the current situation are multiplied by the weights of the criteria and importance of alternatives found for each situation. By this way, the alternative strategies are evaluated specifically for the current situation. The methods used are based on two bases, namely, picture fuzzy sets and spherical fuzzy sets. The results obtained are similar but different. Still, the top-ranked alternative strategy remains the same under all circumstances. The proposed models are run successfully and the results are validated. As a result of the analysis, it is seen that the best strategy is A2 under all conditions. This strategy says that total lockdown with internal and external border restrictions and restriction of all nonessential businesses including delivery services but schools are open.

For future work, we recommend using other fuzzy set extensions. These fuzzy sets extensions may include hesitant fuzzy sets, type-2 fuzzy sets, Pythagorean fuzzy sets, or Fermatean fuzzy sets. In addition, normalization strategies can be compared with the models used and the most appropriate strategy specific to the period can be determined. Other multi-criteria decision-making methods such as ELECTRE, WASPAS, FUCOM, EDAS, CODAS, or COPRAS can be used instead of AHP-TOPSIS methods in this study.

Table 13 Positive and negative ideal solutions in IVSF-TOPSIS

	Positive ideal solution	Negative ideal solution
C11	([0.3082, 0.4259], [0.6376, 0.6993], [0.1909, 0.2687])	([0.0554, 0.0891], [0.8722, 0.9282], [0.1427, 0.1741])
C12	([0.2317, 0.3183], [0.737, 0.7861], [0.1853, 0.2434])	([0.0541, 0.0844], [0.8778, 0.9304], [0.1542, 0.1789])
C13	([0.2466, 0.337], [0.746, 0.7931], [0.1658, 0.231])	([0.0558, 0.0851], [0.8793, 0.9306], [0.1546, 0.1813])
C21	([0.2347, 0.3199], [0.7289, 0.7793], [0.1906, 0.2538])	([0.165, 0.2299], [0.7506, 0.8027], [0.2216, 0.2704])
C22	([0.2255, 0.3024], [0.7712, 0.8138], [0.1652, 0.2267])	([0.0584, 0.0856], [0.8713, 0.9192], [0.1714, 0.1968])
C23	([0.2297, 0.3132], [0.7412, 0.7895], [0.1861, 0.2483])	([0.1572, 0.2219], [0.7813, 0.8348], [0.1917, 0.2406])
C24	([0.2454, 0.3386], [0.7238, 0.7784], [0.1708, 0.2483])	([0.0577, 0.0862], [0.8666, 0.9229], [0.1671, 0.194])
C31	([0.1438, 0.1942], [0.845, 0.8909], [0.1605, 0.2105])	([0.0239, 0.0388], [0.9487, 0.9818], [0.0978, 0.1072])
C32	([0.1432, 0.1904], [0.8552, 0.8975], [0.1564, 0.2042])	([0.0191, 0.033], [0.961, 0.99], [0.0773, 0.0852])
C33	([0.1452, 0.1953], [0.8409, 0.8871], [0.1683, 0.2168])	([0.0202, 0.0348], [0.9572, 0.989], [0.0813, 0.0895])

Table 14 Results of fuzzy TOPSIS extensions for current situation

	PF-TOPSIS				IVSF-TOPSIS			
	Distance to PIS	Distance to NIS	Ci	Rank	Distance to PIS	Distance to NIS	Ci	Rank
A1	0.0718	0.1757	0.7099	2	0.0342	0.0999	0.745	2
A2	0.0657	0.1651	0.7153	1	0.0334	0.1052	0.759	1
A3	0.0842	0.1508	0.6417	3	0.046	0.0961	0.6763	6
A4	0.0935	0.1482	0.6132	5	0.0411	0.1018	0.7124	4
A5	0.077	0.1275	0.6235	4	0.0412	0.1052	0.7186	3
A6	0.0979	0.1197	0.5501	6	0.0439	0.1001	0.6951	5
A7	0.1212	0.1028	0.4589	7	0.0595	0.0833	0.5833	7
A8	0.1686	0.0685	0.2889	9	0.0913	0.0434	0.3222	8
A9	0.1768	0.0825	0.3182	8	0.0992	0.0345	0.258	9

Table 15 Results of fuzzy TOPSIS extensions for determined resource usage rate levels

	Low resource usage rate		Medium resource usage rate		High resource usage rate		Sum of the ranks in PF-TOPSIS	Sum of the ranks in SF-TOPSIS
	Rank							
	PF-TOPSIS	IVSF-TOPSIS	PF-TOPSIS	IVSF-TOPSIS	PF-TOPSIS	IVSF-TOPSIS		
A1	3	5	2	2	1	2	6	9
A2	1	2	1	1	2	1	4	4
A3	5	6	3	6	3	5	11	17
A4	4	3	5	4	4	3	13	10
A5	2	1	4	3	5	4	11	8
A6	6	4	6	5	6	6	18	15
A7	7	7	7	7	7	7	21	21
A8	9	8	9	8	9	8	27	24
A9	8	9	8	9	8	9	24	27

Table 16 Results of crisp AHP

	Local weights of low resource usage rate	Local weights of medium resource usage rate	Local weights of high resource usage rate
C11	0.06	0.22	0.52
C12	0.11	0.05	0.11
C13	0.11	0.04	0.08
C21	0.11	0.12	0.09
C22	0.09	0.10	0.03
C23	0.15	0.13	0.04
C24	0.14	0.25	0.05
C31	0.13	0.03	0.02
C32	0.03	0.02	0.03
C33	0.07	0.04	0.02

Table 17 Results of crisp TOPSIS

	Low resource usage rate	Medium resource usage rate	High resource usage rate
A1	1	1	1
A2	2	2	2
A3	3	3	3
A4	4	4	4
A5	6	6	6
A6	5	5	5
A7	9	7	7
A8	8	9	9
A9	7	8	8

References

Akram, M., Kahraman, C., & Zahid, K. (2021c). Group decision-making based on complex spherical fuzzy VIKOR approach. *Knowledge-Based Systems, 216*, 106793.

Ambrin, R., Ibrar, M., De La Sen, M., Rabbi, I., & Khan, A. (2021). Extended TOPSIS method for supplier selection under picture hesitant fuzzy environment using linguistic variables. *Journal of Mathematics, 2021*.

Bavli, I., Sutton, B., & Galea, S. (2020). Harms of public health interventions against covid-19 must not be ignored. *BMJ, 371*. https://doi.org/10.1136/bmj.m4074

Cuong, B. C., & Kreinovich, V. (2013, December). Picture Fuzzy Sets-a new concept for computational intelligence problems. In *2013 Third World Congress on Information and Communication Technologies (WICT 2013)* (pp. 1–6). IEEE.

Cuong, B. C., & Kreinovich, V. (2014). Picture fuzzy sets. *Journal of Computer Science and Cybernetics, 30*(4), 409–420.

Danielli, S., Patria, R., Donnelly, P., Ashrafian, H., & Darzi, A. (2021). Economic interventions to ameliorate the impact of COVID-19 on the economy and health: An international comparison. *Journal of Public Health, 43*(1), 42–46. https://doi.org/10.1093/pubmed/fdaa104

Dighe, A., Cattarino, L., Cuomo-Dannenburg, G., Skarp, J., Imai, N., Bhatia, S., Gaythorpe, K. A. M., Ainslie, K. E. C., Baguelin, M., Bhatt, S., Boonyasiri, A., Brazeau, N. F., Cooper, L. V., Coupland, H., Cucunuba, Z., Dorigatti, I., Eales, O. D., Van Elsland, S. L., Fitzjohn, R. G., . . .

Riley, S. (2020). Response to COVID-19 in South Korea and implications for lifting stringent interventions. *BMC Medicine, 18*(1), 1–12.

Gündoğdu, F. K., & Kahraman, C. (2019a). Spherical fuzzy sets and spherical fuzzy TOPSIS method. *Journal of Intelligent & Fuzzy Systems, 36*(1), 337–352.

Gündoğdu, F. K., & Kahraman, C. (2019b). A novel fuzzy TOPSIS method using emerging interval-valued spherical fuzzy sets. *Engineering Applications of Artificial Intelligence, 85*, 307–323.

Gündoğdu, F. K., & Kahraman, C. (2021a). Properties and arithmetic operations of spherical fuzzy sets. In *Decision making with spherical fuzzy sets* (pp. 3–25). Springer.

Gündoğdu, F. K., & Kahraman, C. (2021b). Optimal site selection of electric vehicle charging station by using spherical fuzzy TOPSIS method. In *Decision making with spherical fuzzy sets* (pp. 201–216). Springer.

Gündoğdu, F. K., & Kahraman, C. (2021c). Hospital performance assessment using interval-valued spherical fuzzy analytic hierarchy process. In C. Kahraman & F. Kutlu Gündoğdu (Eds.), *Decision making with spherical fuzzy sets. Studies in fuzziness and soft computing* (Vol. 392). Springer. https://doi.org/10.1007/978-3-030-45461-6_15

G*ü*ndogdu, F. K., & Kahraman, C. (2019d). Extension of CODAS with spherical fuzzy sets. *Journal of Multiple-Valued Logic & Soft Computing, 33*.

Gündoğdu, F. K., Duleba, S., Moslem, S., & Aydın, S. (2021). Evaluating public transport service quality using picture fuzzy analytic hierarchy process and linear assignment model. *Applied Soft Computing, 100*, 106920.

Gündoğdu, F. K., Kahraman, C., & Karaşan, A. (2019, July). Spherical fuzzy VIKOR method and its application to waste management. In *International Conference on Intelligent and Fuzzy Systems* (pp. 997–1005). Springer.

Harling, G., Gómez-Olivé, F. X., Tlouyamma, J., Mutevedzi, T., Kabudula, C. W., Mahlako, R., & Herbst, K. (2021). Protective behaviors and secondary harms resulting from nonpharmaceutical interventions during the COVID-19 epidemic in South Africa: Multisite, prospective longitudinal study. *JMIR Public Health and Surveillance, 7*(5), e26073. https://doi.org/10.2196/26073

Imai, N., Gaythorpe, K. A., Abbott, S., Bhatia, S., van Elsland, S., Prem, K., … Ferguson, N. M. (2020). Adoption and impact of non-pharmaceutical interventions for COVID-19. *Wellcome Open Research, 5*.

Jin, J., Zhao, P., & You, T. (2021). Picture fuzzy TOPSIS method based on CPFRS model: An application to risk management problems. *Scientific Programming*, 2021.

Joshi, R., & Kumar, S. (2021). A novel VIKOR approach based on weighted correlation coefficients and picture fuzzy information for multicriteria decision making. *Granular Computing*, 1–14.

Ju, Y., Ju, D., Gonzalez, E. D. S., Giannakis, M., & Wang, A. (2019). Study of site selection of electric vehicle charging station based on extended GRP method under picture fuzzy environment. *Computers & Industrial Engineering, 135*, 1271–1285.

Kantor, B. N., & Kantor, J. (2020). Non-pharmaceutical interventions for pandemic COVID-19: A cross-sectional investigation of US general public beliefs. *Frontiers in Medicine, 7*, 384.

Kennedy, D. M., Zambrano, G. J., Wang, Y., & Neto, O. P. (2020). Modelling the effects of intervention strategies on COVID-19 transmission dynamics. *Journal of Clinical Virology, 128*.

Kneale, D., O'Mara-Eves, A., Rees, R., & Thomas, J. (2020). School closure in response to epidemic outbreaks: Systems-based logic model of downstream impacts. *F1000Research, 9*.

Li, R., Chen, B., Zhang, T., Ren, Z., Song, Y., Xiao, Y., Hou, L., Cai, J., Xu, B., Li, M., Chan, K. K. Y., Tu, Y., Yang, M., Yang, J., Liu, Z., Shen, C., Wang, C., Xu, L., Liu, Q., … Xu, B. (2020). Global COVID-19 pandemic demands joint interventions for the suppression of future waves. *Proceedings of the National Academy of Sciences, 117*(42), 26151–26157.

Lu, J., Zhang, S., Wu, J., & Wei, Y. (2021). COPRAS method for multiple attribute group decision making under picture fuzzy environment and their application to green supplier selection. *Technological and Economic Development of Economy, 27*(2), 369–385.

Ngonghala, C. N., Iboi, E., Eikenberry, S., Scotch, M., MacIntyre, C. R., Bonds, M. H., & Gumel, A. B. (2020). Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel Coronavirus. *Mathematical Biosciences, 325*, 108364.

Oztaysi, B., Onar, S. C., & Kahraman, C. (2020, July). A fuzzy pricing model for mobile advertisements by using spherical fuzzy AHP scoring. In *International Conference on Intelligent and Fuzzy Systems* (pp. 142–150). Springer.

Panovska-Griffiths, J., Kerr, C. C., Stuart, R. M., Mistry, D., Klein, D. J., Viner, R. M., & Bonell, C. (2020). Determining the optimal strategy for reopening schools, the impact of test and trace interventions, and the risk of occurrence of a second COVID-19 epidemic wave in the UK: A modelling study. *The Lancet Child & Adolescent Health, 4*(11), 817–827.

Pedrozo-Pupo, J. C., & Campo-Arias, A. (2020). Depression, perceived stress related to COVID, post-traumatic stress, and insomnia among asthma and COPD patients during the COVID-19 pandemic. *Chronic Respiratory Disease, 17*, 1479973120962800. https://doi.org/10.1177/1479973120962800

Peng, X., & Yang, Y. (2016a). Fundamental properties of interval-valued Pythagorean fuzzy aggregation operators. *International Journal of Intelligent Systems, 31*(5), 444–487.

Peng, X., & Yang, Y. (2016b). Pythagorean fuzzy Choquet integral based MABAC method for multiple attribute group decision making. *International Journal of Intelligent Systems, 31*(10), 989–1020.

Piltch-Loeb, R., Merdjanoff, A., & Meltzer, G. (2021). Anticipated mental health consequences of COVID-19 in a nationally-representative sample: Context, coverage, and economic consequences. *Preventive Medicine, 145*, 106441. https://doi.org/10.1016/j.ypmed.2021.106441

Pinto, N. O., Kennedy, D. M., Reis, J. C., Wang, Y., Brizzi, A. C. B., Zambrano, G. J., de Souza, J. M., Pedroso, W., de Mello Pedreiro, R. C., de Matos Brizzi, B., Abinader, E. O., & Zângaro, R. A. (2021). Mathematical model of COVID-19 intervention scenarios for São Paulo—Brazil. *Nature Communications, 12*(1), 1–13.

Qiyas, M., Abdullah, S., Al-Otaibi, Y. D., & Aslam, M. (2021). Generalized interval-valued picture fuzzy linguistic induced hybrid operator and TOPSIS method for linguistic group decision-making. *Soft Computing, 25*(7), 5037–5054.

Rauschenberg, C., Schick, A., Hirjak, D., Seidler, A., Paetzold, I., Apfelbacher, C., Riedel-Heller S. G., Reininghaus, U. (2021). Evidence synthesis of digital interventions to mitigate the negative impact of the COVID-19 pandemic on public mental health: Rapid meta-review. *Journal of Medical Internet Research, 23*(3), e23365. https://doi.org/10.2196/23365

Reddy, K. P., Shebl, F. M., Foote, J. H., Harling, G., Scott, J. A., Panella, C., Fitzmaurice, K. P., Flanagan, C., Hyle, E. P., Neilan, A. M., Mohareb, A. M., Bekker, L. G., Lessells, R. J., Ciaranello, A. L., Wood, R., Losina, E., Freedberg, K. A., Kazemian, P., & Siedner, M. J. (2021). Cost-effectiveness of public health strategies for COVID-19 epidemic control in South Africa: A microsimulation modelling study. *The Lancet Global Health, 9*(2), e120–e129.

Roberts, N. J., McAloney-Kocaman, K., Lippiett, K., Ray, E., Welch, L., & Kelly, C. (2021). Levels of resilience, anxiety and depression in nurses working in respiratory clinical areas during the COVID pandemic. *Respiratory Medicine, 176*, 106219. https://doi.org/10.1016/j.rmed.2020.106219

Robinson, E., Boyland, E., Chisholm, A., Harrold, J., Maloney, N. G., Marty, L., Mead, B. R., Noonan, R., & Hardman, C. A. (2021). Obesity, eating behavior and physical activity during COVID-19 lockdown: A study of UK adults. *Appetite, 156*, 104853. https://doi.org/10.1016/j.appet.2020.104853

Rozhnova, G., van Dorp, C. H., Bruijning-Verhagen, P., Bootsma, M. C., van de Wijgert, J. H., Bonten, M. J., & Kretzschmar, M. E. (2021). Model-based evaluation of school-and non-school-related measures to control the COVID-19 pandemic. *Nature Communications, 12*(1), 1–11. https://doi.org/10.1038/s41467-021-21899-6

Russell, T. W., Wu, J., Clifford, S., Edmunds, J., Kucharski, A. J., & Jit, M. (2020). The effect of international travel restrictions on internal spread of COVID-19. medRxiv. https://doi.org/10.1101/2020.07.12.20152298.

Saaty, T. (1980). *The analytic hierarchy process (AHP) for decision making*. Kobe, Japan.

Simic, V., Karagoz, S., Deveci, M., & Aydin, N. (2021). Picture fuzzy extension of the CODAS method for multi-criteria vehicle shredding facility location. *Expert Systems with Applications, 175*, 114644.

Tey, D. J. Y., Gan, Y. F., Selvachandran, G., Quek, S. G., Smarandache, F., Abdel-Basset, M., & Long, H. V. (2019). A novel neutrosophic data analytic hierarchy process for multi-criteria decision making method: A case study in Kuala Lumpur stock exchange. *IEEE Access, 7*, 53687–53697.

Torun, H., & Gördebil, M. (2019, July). Multi criteria decision making based on TOPSIS method with extended fuzzy sets. In *International Conference on Intelligent and Fuzzy Systems* (pp. 558–566). Springer.

Ullah, W., Ibrar, M., Khan, A., & Khan, M. (2021). Multiple attribute decision making problem using GRA method with incomplete weight information based on picture hesitant fuzzy setting. *International Journal of Intelligent Systems, 36*(2), 866–889.

Wei, G. (2017). Some cosine similarity measures for picture fuzzy sets and their applications to strategic decision making. *Informatica, 28*(3), 547–564.

Zhang, S., Wei, G., Wang, R., Wu, J., Wei, C., Guo, Y., & Wei, Y. (2021). Improved CODAS method under picture 2-tuple linguistic environment and its application for a green supplier selection. *Informatica, 32*(1), 195–216.

A Multi-Attribute Decision-Making Model for Hospital Location Selection

Gül İmamoğlu and Y. Ilker Topcu

Abstract The demand for the healthcare sector is increasing day by day due to the population growth, the prolongation of the average human life, changing eating habits, and the rapid spread of epidemics such as COVID-19 in large areas. New hospitals are being built in Turkey, as in other countries, to meet the increase in demand in the health sector and reach high health standards. Therefore, the selection of hospital location arises as a vital decision problem. In the case of an incorrect hospital location, loss of life can occur, as well as a significant financial burden. For this reason, decision makers should handle hospital location selection with analytical and rational methods. For this purpose, in this study, a hybrid Multi-Attribute Decision-Making model, in which the importance of the attributes is determined using the Analytic Network Process method, and evaluating the alternatives is carried out by the PROMETHEE method was proposed. As a case study, the proposed model was utilized to select a suitable location for a new hospital in Trabzon, Turkey.

Keywords Healthcare · Hospital Location Selection · Multi-Attribute Decision-Making · Analytic Network Process · PROMETHEE

1 Introduction

Due to the environmental factors brought about by modern life, changing eating habits, and the prolongation of average life expectancy, the proportion of the elderly population and diseases is increasing. In addition, pandemics such as the COVID-19 outbreak increase the number of patients dramatically in a short time.

G. İmamoğlu (✉)
Industrial Engineering Department, Karadeniz Technical University, Trabzon, Türkiye
e-mail: gulimamoglu@ktu.edu.tr

Y. I. Topcu
Industrial Engineering Department, Istanbul Technical University,İstanbul, Türkiye
e-mail: ilker.topcu@itu.edu.tr

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_16

All these factors lead to an increase in the demand for the health sector. Accordingly, new hospitals have to be built as hospitals are essential components of health systems and critical points of reforms in the health sector. They provide constant nursing care, beds, and meals for their patients who undergo medical therapy at the hands of professional physicians.

Health expenditures are increasing both in Turkey and other countries. For this reason, governments attach great importance to the more effective use of hospitals (Atılgan, 2016). The efficiency of hospitals is affected by many factors. These factors can be grouped under clinical effectiveness and managerial effectiveness (Fragkiadakis et al., 2016).

Hospital location selection is one of the critical elements of managerial effectiveness. An appropriate selection of the location will attract many potential patients and affect the hospital's business success (Soltani and Marandi, 2011). On the other hand, an improper location selection for a hospital can result in the loss of life. Furthermore, it will also create a sizeable monetary load, likewise any wrong location selection for any purpose. For these reasons, the hospital location selection should be focused on analytical and rational methods.

Location selection problems are often addressed under a "set-covering" framework as the goal is to achieve maximum coverage at minimum cost. In addition to coverage and cost, there should be other evaluation factors when choosing a hospital location. Demand, centrality, proximity to major roads, distance to noise centers, and other similar factors need to be considered. Therefore, Multi-Attribute Decision-Making (MADM) methods, which consider many factors and offer an integrated assessment that evaluates alternatives with respect to these factors, can be utilized to treat hospital location decision problems.

This study differs from other hospital location selection studies, which will be discussed in the following section, regarding a careful consideration of the evaluation attributes and usage of actual data not considered in previous studies. Besides, PROMETHEE (Preference Ranking Organization Method for Enrichment Evaluations), an effective outranking MADM method, was utilized for the first time in the hospital location selection.

Another contribution of this study is the analysis of a local hospital location selection problem. A case study was conducted in Trabzon Province. Trabzon is an attractive destination to become a leading international medical tourism center by providing treatment and rehabilitation services to international patients, especially patients from Azerbaijan, Georgia, Iran, and Arab countries. In Trabzon, there is a hospital shortage for both locals and those who come for health tourism.

This chapter aims to identify the factors that should be taken into account when choosing a hospital location. Additionally, it also aims to reveal the importance of these factors and recommend policymakers the appropriate locations for a hospital to be built in Trabzon. For this purpose, we used an integrated decision approach utilizing Analytic Network Process (ANP) and PROMETHEE methods.

The rest of the chapter is organized as follows. In section "Literature Review," a literature review of hospital location selection is presented. The theoretical background of the methods utilized in the proposed approach is provided in section “Methods.” Section “Application of the Proposed Approach” presents the application of the proposed approach for the hospital location selection in Trabzon province and the revealed results. Finally, some concluding remarks and further suggestions are specified in section “Conclusions and Further Suggestions.”

2 Literature Review

The studies on location selection are prevalent in the literature. When we searched the Web of Science Core Collection with the topic “location selection” and “multi-attribute decision-making,” we have found 37 articles published in 2020 and 2021.

Some of them are analyzing the following location selection problems in the energy sector:

- Thermal power plant (Milovanovic et al., 2021)
- Offshore wind power station/farm (Abdel-Basset et al., 2021; Tercan et al., 2020)
- Marine current energy production plant (Yucenur and Ipekci, 2021)
- Artificial recharge site with treated wastewater (Mahmoudi et al., 2021)
- Solar site / photovoltaic plants (Kannan et al., 2021; Furtado and Sola, 2020)
- Distributed photovoltaic power stations for high-speed railway (Li et al., 2020)

There are also some studies addressing location selection problems in logistics and supply chain management:

- Inland terminal for shipping lines (Liang et al., 2021)
- Warehouse location or distribution center (Ehsanifar et al., 2020; Agrebi and Abed 2021; Liu and Li, 2020)
- Emergency shelter in preparation for immediate, short-, and long-term floods (Lee et al., 2020)

On the other hand, in an earlier study, Imamoglu (2015) conducted an extensive literature survey on location selection. She found studies in various areas such as logistics (Yang et al., 2007; Demirel et al., 2010; Kampf et al., 2011; Dey et al., 2013; Wey, 2015), energy (Magaji and Mustafa, 2011; Yunna and Geng, 2014; Kumar and Srikanth, 2015), environment (Meng and Lu, 2011; Kabir and Sumi, 2014; Gang et al., 2015), manufacturing (Guneri et al., 2009; Shen and Yu, 2009; Mokhtarian and Hadi-Vencheh, 2012; Rahmaniani and Ghaderi, 2015), informatics (Lu et al., 2011; Xie et al., 2013), retailing (Cheng et al., 2007), tourism and hospitality (Chou et al., 2008; Ishizaka et al., 2013; Dock et al., 2015), and disaster management (Kılcı et al., 2015).

The field of application of the studies on location selection has recently shifted to the energy sector. This outcome is not surprising as the climate crisis, affordable and clean energy, and environmental sustainability are among the top global challenges the world is facing. However, since 2020, humanity has been paying close attention to the worldwide pandemic that has brought about a health crisis as well as an economic crisis, an inequality crisis, an education crisis, and other crises. So health care management will be again one of the promising fields. For instance, in the recent literature, the following location selection studies address healthcare management in a multi-criteria decision-making environment. Zeferino et al. (2021) proposed an integrated approach for selecting a suitable quarantine facility for COVID-19 patients. Mirzahossein et al. (2020) developed a model to choose the appropriate locations for the construction of the medical emergency centers.

We conducted a literature survey to classify the studies treating hospital location selection problems with respect to the multi-attribute decision-making methods utilized:

- AHP (Sinuany-Stern et al., 1995; Wu et al., 2007a; Chatterjee and Mukherjee, 2013; Chiu and Tsai, 2013; Şahin et al., 2019)
- ANP (Onut et al., 2008)
- Fuzzy AHP (Wu et al., 2007b; Lin et al., 2008; Aydın and Arslan, 2010)
- Fuzzy ANP (Wu et al., 2009)
- AHP and Pareto technique (Zeferino et al., 2021)
- AHP and Geographic Information System integration (Eldemir and Onden, 2016; Mirzahossein et al., 2020)
- AHP and Gray Relational Analysis integration (Şen and Demiral, 2016)
- ANP and TOPSIS integration (Lin and Tsai, 2009)
- Fuzzy AHP and Geographic Information System integration (Vahidnia et al., 2009)
- Fuzzy ANP and Geographic Information System integration (Soltani and Marandi, 2011)
- Hesitant Fuzzy Sets and TOPSIS integration (Şenvar et al., 2016)
- Additive Ratio Assessment method with Gray values (Sen, 2017)
- Data Envelopment Analysis (Lin et al., 2010)

As can be seen, Analytic Hierarchy Process (AHP) and its variants are widely used. Some of these studies offer an appropriate location for a hospital across the country. Sinuany-Stern et al. (1995) developed mathematical models for hospital location selection, which were planned to be established to ensure smooth distribution of the population in the Negev region. They have identified six alternative cities. They utilized AHP to evaluate the alternatives. Wu et al. (2007a) determined which one of the three alternatives would be appropriate to provide a competitive advantage for a hospital to be built in Taiwan by utilizing AHP and made a sensitivity analysis for the importance of the attributes. Chatterjee and Muherjee (2013) used AHP to evaluate three potential rural hospital sites in India based on three attributes and 11 sub-attributes. To build the planned regional teaching hospital in Taiwan's Yunlin County, Chiu and Tsai (2013) have determined which one of the two major regions of the county would be appropriate using AHP.

On the other hand, some studies offer an appropriate location in a city or region. Şahin et al. (2019) have developed a hospital location selection model for the province Muğla of Turkey by utilizing AHP. This model determined counties as alternatives, and evaluation is made based on six attributes and 19 sub-attributes.

There are also studies utilizing the ANP method. For a medium-sized hospital planned to be established in Istanbul, Onut et al. (2008) have decided which one of the three alternative counties would be appropriate by utilizing ANP. Wu et al. (2009) have determined with Fuzzy ANP which of the three alternative cities would be convenient for the planned establishment of the hospital in Taiwan.

Like Wu et al. (2009), some researchers used fuzzy sets and algorithms to treat the problem. Wu et al. (2007b) made the location selection for the planned establishment of a regional hospital in Taiwan by Fuzzy AHP. Lin et al. (2008) evaluated three alternatives according to six attributes and 18 sub-attributes using Fuzzy AHP to establish a competitive regional hospital in Taiwan. Aydın and Arslan (2010) decided which of the five candidate regions would be appropriate to establish a new hospital in Ankara by Fuzzy AHP. According to six attributes with the ARAS-G method, Sen (2017) evaluated three alternative locations for a new public hospital.

Another group of studies treated the location selection problem with the integration of more than one MADM method. Zeferino et al. (2021) proposed integrating the AHP and Pareto technique to assess the factors that can be used to select a suitable quarantine facility for COVID-19 patients based on facility location experts' opinions. Şen and Demiral (2016) presented a model using six attributes to select a new public hospital location. They determined the importance of these attributes utilizing AHP, and they utilized the Gray Relational Analysis to choose one of three alternative locations. To make a hospital investment in the most appropriate location for the investors in the cities of three major particular economic regions of China, Lin and Tsai (2009) have used the TOPSIS method to rank the alternatives after determining attributes importance by ANP method. Şenvar et al. (2016) set a TOPSIS model which uses Hesitant Fuzzy sets for the hospital location selection model.

In some studies, we confronted an integration of a MADM method and an additional method such as Geographic Information System (GIS). Mirzahossein et al. (2020) developed a model to select the appropriate locations for the construction of the medical emergency centers to elevate the new silk road's safety measures utilizing AHP in the ArcGIS platform. Eldemir and Onden (2016) determined attributes' importance by using AHP and then evaluated alternatives based on 13 attributes by using GIS. Vahidnia et al. (2009) have set 5 lands as alternatives with GIS and have evaluated these alternatives with Fuzzy AHP. For the planned establishment of a new hospital in the fifth Region of Iran's Şiran town, Soltani and Marandi (2011) have set three alternatives by Fuzzy AHP and GIS and then determined the area to establish the hospital by fuzzy ANP. In the proposed model, we evaluated alternative locations according to three attributes and ten sub-attributes, and they applied this model to a sample in the city of Dallas.

When studies on hospital location selection problems were examined, we realized that the studies were not using actual data, assessing the subjective judgments of the experts instead. Most of the studies in the literature utilize AHP or ANP to derive performance values of alternatives with respect to attributes from pairwise comparisons (Sinuany-Stern et al., 1995; Lin et al., 2008; Önüt et al., 2008; Lin and Tsai, 2009; Vahidnia et al., 2009; Wu et al., 2009; Wu et al., 2007a; Wu et al., 2007b;

Aydın and Arslan, 2010; Chatterjee and Muherjee, 2013; Chiu and Tsai, 2013; Şen and Demiral, 2016, Şen, 2017; Şahin et al., 2019). In the remaining few studies using subjective data, the inadequacy of the used attributes is noticeable (Soltani and Marandi 2011; Kim et al., 2015). The list of attributes used in these previous studies is presented in Appendix. Based on the literature review results, the most widely used hospital location selection attribute can be stated as land cost, followed by population density.

3 Methods

Multi-attribute value function methods aggregate the performances of alternatives with respect to attributes on a normative basis. These methods are too rich to be reliable in some cases (Brans & Vincke, 1985). On the other hand, outranking methods aggregate decision makers' preferences on alternatives on a descriptive basis. These methods enrich the dominance relation, which is too poor to be useful at multi-attribute decision problems. Therefore, in this study, PROMETHEE, a widely used outranking method, was used to reveal decision makers' preferences on possible hospital locations with respect to conflicting attributes.

Compared to other MADM methods, the difference of the PROMETHEE method is using an evaluation system employing the preference function. Through this preference function, characteristics and preferences of different attributes are transmitted better to the decision model. The threshold values have significant meaning in terms of alternatives. PROMETHEE is a user-friendly method that is easy to use and to interpret the parameters (Al-Shemmeri et al., 1997). In terms of the effect of minor deviations in the values of threshold parameters on final recommendations, the PROMETHEE method can be considered more stable than the ELECTRE method, another widely-used outranking method (Brans et al., 1986). Furthermore, Brans and Mareschal (1994) have developed a decision support system called GAIA to visualize the PROMETHEE results.

PROMETHEE, like most of the MADM methods, cannot reveal the importance of attributes. That is why an additional method can be used to prioritize the attributes (Macharis et al., 2004). As will be discussed in the next section, due to relationships among evaluation attributes, ANP, which takes dependencies and feedback into account, was used.

In the literature, there are various studies that use ANP to assess the importance of attributes and PROMETHEE to evaluate the preferences on alternatives. Kabak and Dağdeviren (2014) proposed a MADM approach that combines ANP and PROMETHEE to present to aid students at their selection among universities. Peng and Xiao (2013) utilized ANP and PROMETHEE methods in their hybrid decision model for material selection. Kilic et al. (2015) used the aggregation of ANP and PROMETHEE methods to select an ERP system for a small-medium enterprise. Govindan et al. (2015) presented an integrated MADM model where Dematel, ANP, and PROMETHEE methods were utilized together to evaluate green production practices.

Similarly, the proposed MADM model in this study is an integrated MADM model utilizing ANP and PROMETHEE. The details of these two methods are given in the following subsections.

3.1 Analytic Network Process

Saaty (1996) has introduced the ANP method for complicated and unstructured problems. ANP uses a network structure to model a complex decision problem with interconnections (dependencies and feedback). A network model with dependence and feedback improves the priorities derived from judgments and makes a prediction, especially prioritizes the attributes much more accurately. This method allows groups or individuals to deal with the interconnections between factors of complex structure in the decision-making process.

ANP assesses decision makers' judgments through pairwise comparisons of the attributes in the network. Possible replies may be the values in Saaty's 1–9 scale (1 = equally important, 3 = moderately important, 5 = strongly important, 7 = very strongly important, 9 = absolutely important, and reciprocals for inverse comparisons).

The ANP consists of six steps:

1. *Identifying elements and clusters*: At this stage, the aim of the problem must be stated clearly. Relate attributes and alternatives are determined. These elements, i.e., attributes and alternatives, are grouped under clusters.
2. *Assessing relations*: The relations among elements are identified, i.e., which element affects which one. If these exist relations, a network structure is created consisting of inner dependencies, outer dependencies, and feedback.
3. *Comparing elements*: For each element, decision makers are asked to compare the influence degrees of the affecting elements in pairs. A pairwise comparison matrix of the elements is created for each affected element by using the responses.
4. *Computing eigenvectors*: Then, the eigenvector of each matrix is calculated:

$$\mathbf{Aw} = \boldsymbol{\lambda}_{\mathbf{max}}\mathbf{w} \tag{1}$$

where $\mathbf{A}$ represents the pairwise comparison matrix, $\mathbf{w}$ represents the eigenvector of the matrix, and $\boldsymbol{\lambda}_{\mathbf{max}}$ represents the largest eigenvalues of matrix $\mathbf{A}$.

5. *Constructing special matrices*: Eigenvectors are placed as entries into a particular matrix called a supermatrix. Supermatrix represents each element at one row and one respective column. There, an eigenvector is read at a column of an affected element and rows of affecting elements

Then, column totals in supermatrix are normalized to have a column stochastic matrix called a weighted supermatrix.

By raising that matrix to significantly high degrees, the limit matrix with stable and converged values is obtained.

6. *Revealing priorities*: Any column of limit matrix presents the global priority values of the elements. These priorities would be the importance of attributes or preference for the alternatives based on each cluster's type of elements

3.2 PROMETHEE

PROMETHEE, which was brought into literature first in a conference held in Canada in 1982 by Jean Pierre Brans, is a multi-attribute decision aid method that allows building outranking relations among alternatives. It is an acronym representing the phrase Preference Ranking Organization METHod for Enrichment Evaluations (Brans and Vincke, 1985).

PROMETHEE, as an outranking method, includes two phases (Brans and Vincke, 1985), i.e., the construction of an outranking relation and the exploitation of this relation to aid decision maker.

1. *Constructing the outranking relation*

First of all, the preference of decision maker for an alternative a_i compared to another alternative a_j for each attribute k, a preference function $P_k(a_i, a_j)$ is defined:

$$P_k\left(a_i, a_j\right) = \begin{cases} 0 & if \;\; f_k\left(a_i\right) \le f_k\left(a_j\right) \\ p\left[f_k\left(a_i\right), f\left(a_j\right)\right] & if \;\; f_k\left(a_i\right) > f_k\left(a_j\right) \end{cases} \tag{2}$$

where $f_k(a_i)$ represents the performance value of alternative a_i with respect to attribute k.

Decision maker chooses among six generalized criteria which are used to find the values of preference function. Figure 1 shows the functions used for the generalized criteria, while Fig. 2 shows the representation of these functions.

Here, the difference between the performance values of alternatives a_i and a_j with respect to attribute k is calculated to find the value of d; on the other hand, the decision maker identifies necessary parameters such as indifference threshold (q) and preference threshold (p).

Then, the weighted average of the preference functions is computed to reveal a multi-attribute preference index:

$$\pi\left(a_i, a_j\right) = \sum_k w_k P_k\left(a_i, a_j\right) \tag{3}$$

2. *Exploiting the outranking relation*

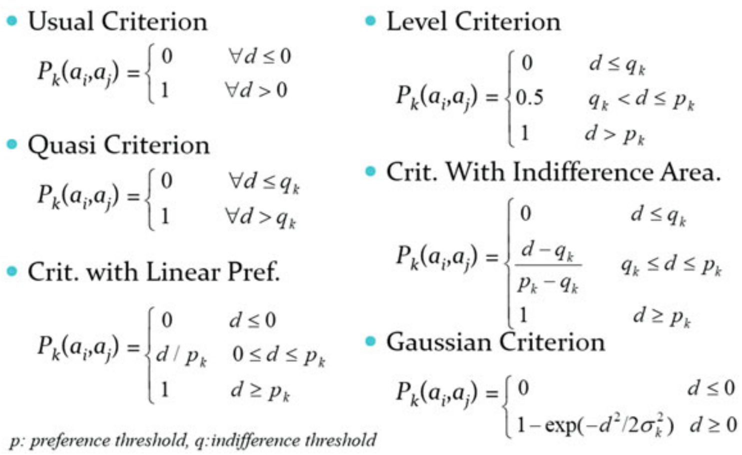


Fig. 1 Generalized criteria functions (Source: Topcu et al. (2020), p. 6)

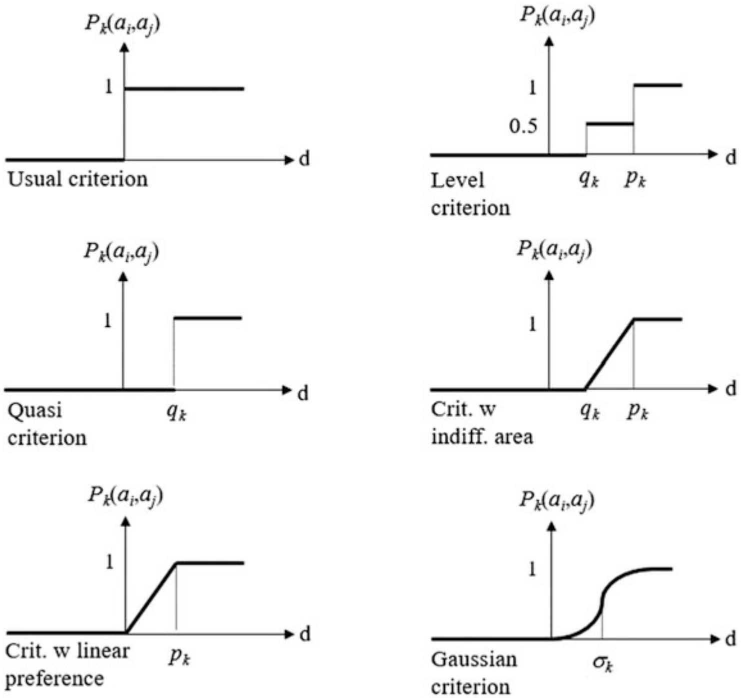


Fig. 2 Representation of generalized criteria functions (Source: Topcu et al. (2020), p. 6)

Based on the preference indices, three types of flows are calculated for an alternative a_i, i.e., the leaving flow ($\Phi^+(a_i)$), the entering flow ($\Phi^-(a_i)$), and the net flow ($\Phi(a_i)$):

$$\Phi^+ (a_i) = \sum_{a_j \in A} \pi \left(a_i, a_j\right) \tag{4}$$

$$\Phi^- (a_i) = \sum_{a_j \in A} \pi \left(a_j, a_i\right) \tag{5}$$

$$\Phi (a_i) = \Phi^+ (a_i) - \Phi^- (a_i) \tag{6}$$

Based on these flows, three types of ranking are revealed:

- Ranking of alternatives based on the decreasing order of leaving flows.
- Ranking of alternatives based on the increasing order of entering flows.
- Ranking of alternatives based on the decreasing order of net flows.

PROMETHEE I uses the intersection of the first two rankings and yields a partial pre-order of alternatives where preference, indifference, and incomparability among them are allowed.

PROMETHEE II, on the other hand, uses the third ranking and yields a complete pre-order of alternatives where preference and indifference among them are allowed.

4 Application of the Proposed Approach

4.1 Structuring the Problem

Trabzon province, located in the Eastern Black Sea Region of Turkey at the seaside, is a bridge between the Caucasus, Central Asia, and the West on the "Trans Caucasian Corridor" passing through the Caucasus into the Middle East. The population of the city is 811,901 as of 2019 (Turkish Statistical Institute, 2021). The location of Turkey in the world (URL1, n.d.) and Trabzon in Turkey (URL2, n.d.) as well as the layout of the counties of Trabzon (URL3, n.d.) are shown in Fig. 3.

As aforementioned, Trabzon is expected to become a leading international medical tourism center. Hospitals in Trabzon aim to provide treatment and rehabilitation services to international patients, especially from Azerbaijan, Georgia, Iran, and Arab countries. In Trabzon, currently, there are 22 hospitals, including 12 public hospitals, four branch hospitals, one university hospital, four private hospitals, and one training research hospital affiliated to the Ministry of Health. Although three new hospital buildings have been built in Trabzon in the last decade, there is still a hospital shortage for both local people and those who come for health tourism in

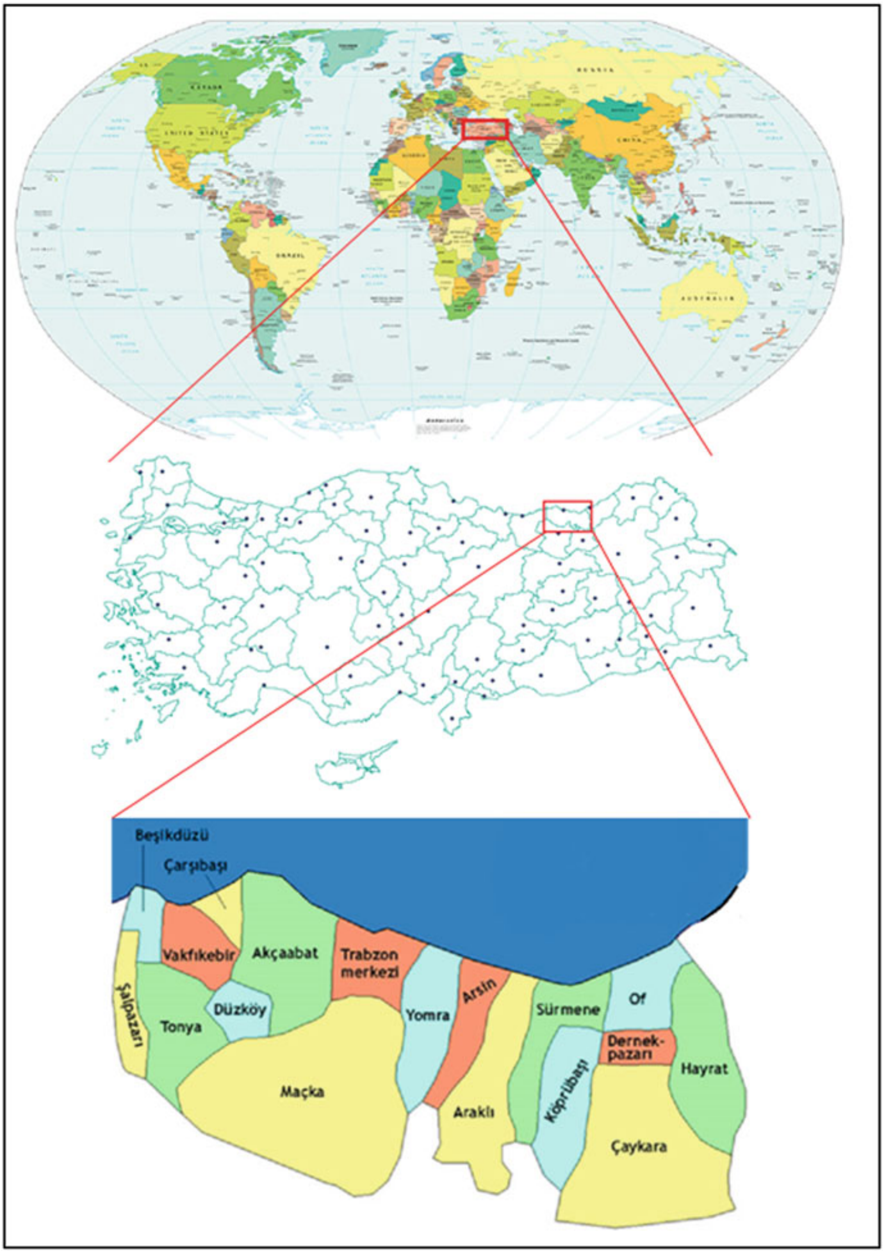


Fig. 3 Location of Trabzon Province and its counties (Sources: URL1, URL2, URL3)

conjunction with increasing demand. For this reason, in this study, we aimed to find an appropriate location for a new hospital building in Trabzon.

It has seemed more appropriate to construct a new building for the hospital instead of converting existing buildings used for other purposes because it is necessary for a hospital building to meet many physical, environmental, infrastructural, and ergonomic attributes. Therefore, the alternatives were not selected as existing buildings. Since the hospital to be established as a non-profit public hospital, we assumed that every land can be accessible and can be determined as alternatives.

Some quantitative information such as population and density is accessible only based on counties in the region. In addition, attributes such as the cost of land and traffic density differ dramatically among counties. Hence, the counties of Trabzon were considered as alternatives to this multi-attribute decision problem. There are 18 counties in this province, namely, Ortahisar (city center), Akçaabat, Araklı, Arsin, Beşikdüzü, Çarşıbaşı, Çaykara, Dernekpazarı, Düzköy, Hayrat, Köprübaşı, Maçka, Of, Sürmene, Şalpazarı, Tonya, Vakfıkebir, and Yomra.

To evaluate these counties on an MADM basis, first, as aforementioned in section "Literature Review," a detailed literature review was conducted. Then, the list of attributes used in the previous studies was revealed as given in Appendix. After that, we interacted with 11 experts to finalize the list. Among the experts, five were medical doctors, another five were authorities working at Trabzon Provincial Health Directorate. The last one was an authority working at Trabzon Regional Directorate of Transport and Infrastructure.

Based on their advice, some attributes used in the previous studies conducted for location selection of private hospitals were disabled because, in this study, the location would be selected for a public hospital. Accordingly, level of income, management objective, the rank of competitors, health sector and activities of competitors, and distance of competing hospitals have been excluded. Some attributes evaluating lands or buildings were discarded as counties were determined as the alternatives. These attributes were lease cost, building rearrangement cost, landscape cost, water and electricity supply, enlargement opportunity, city plan compliance, and parking area. Besides, the attribute of the pharmaceutical sector was not taken into consideration as many pharmacies would be opened around the new hospital in a short period. On the other hand, attributes such as construction cost and labor cost have been eliminated, as they would not vary between counties of a medium-sized city. As a result, after discussing the remaining attributes in the literature with experts, we made a final list of 14 attributes classified under four clusters as can be seen in Table 1.

The population of a county (a1) and all of the attributes of Cluster B (b1, b2, . . . b7) are self-explanatory and need objective evaluation. The sources are given in Table 1. Other objective attributes are population density (a2), centrality (a3), and distance to the main road (c1). Population density is the ratio of the population of a county to its area. Centrality is the sum of ratios of the population of each county to

Table 1 The evaluation attributes

Attribute groups	Attributes	Data source
A. Demographic characteristics	a1. Population	Turkish Statistical Institute
	a2. Population density	By calculation
	a3. Centrality	By calculation
B. The health sector and medical applications	b1. The number of family health centers in the county	Trabzon Provincial Public Health Office
	b2. The number of physicians in the family health centers	Trabzon Provincial Public Health Office
	b3. The number of public hospitals in the county	Turkey Public Hospitals Authority, Trabzon Public Hospitals Association, General Secretary, and Public Health Agency of Turkey
	b4. Total number of beds in the county's public hospitals	Same as the source of b3
	b5. The number of branch hospitals in the county	Same as the source of b3
	b6. Total number of beds in the county's branch hospitals	Same as the source of b3
	b7. The number of private hospitals in the county	General Secretary and Public Health Agency of Turkey
C. Environmental effects	c1. Distance to the main road	Turkey General Directorate of Highways
	c2. Traffic congestion	Experts' opinion
	c3. Noise center	Experts' opinion
D. Cost	d1. Land cost	Experts' opinion

its distance to the county whose centrality is computed. Finally, the distance to the main road is the distance of a county center to the state highway that runs along the coastline.

The last three attributes, namely traffic congestion (c2), noise center (c3), and land cost (d1), need subjective evaluation. Traffic congestion in the county and average land cost of possible areas in that county are self-explanatory. The distance of an alternative land to the nearest noise center is taken into account in the literature. However, in this study, the noise center attribute was used as the average noise level from the noise centers such as factories and garbage collection areas in a given county as counties were considered alternatives.

4.2 Constructing the Decision Model

4.2.1 Determination of the Relations among Attributes

The experts were then requested to identify the existence of the effects among attributes. We constructed an adjacency matrix based on experts' judgments as presented in Table 2. In the matrix, if an entry is filled with "X," the attribute in the row of that entry affects the attribute in the column of the same entry.

As can be seen in Table 2, population attribute directly affects 11 attributes. Centrality attribute directly affects nine attributes, and distance to the main road affects six attributes. On the other hand, 11 attributes affect population attribute. Six attributes affect traffic congestion and land cost attributes.

Table 2 The adjacency matrix

	a1	a2	a3	b1	b2	b3	b4	b5	b6	b7	c1	c2	c3	d1
a1 Population		X	X	X	X	X	X			X	X	X	X	X
a2 Population density											X	X		X
a3 Centrality						X	X	X	X	X	X	X	X	X
b1 The number of family health centers (FHC) in the county	X				X									
b2 The number of physicians in the county's FHC	X													
b3 The number of public hospitals in the county	X						X					X		X
b4 Total number of beds in the county's public hospitals	X													
b5 The number of branch hospitals in the county	X								X			X		X
b6 Total number of beds in the county's branch hospitals	X													
b7 The number of private hospitals in the county	X													
c1 Distance to the main road	X		X							X		X	X	X
c2 Traffic congestion	X										X			
c3 Noise center	X													
d1 Land cost	X									X	X			

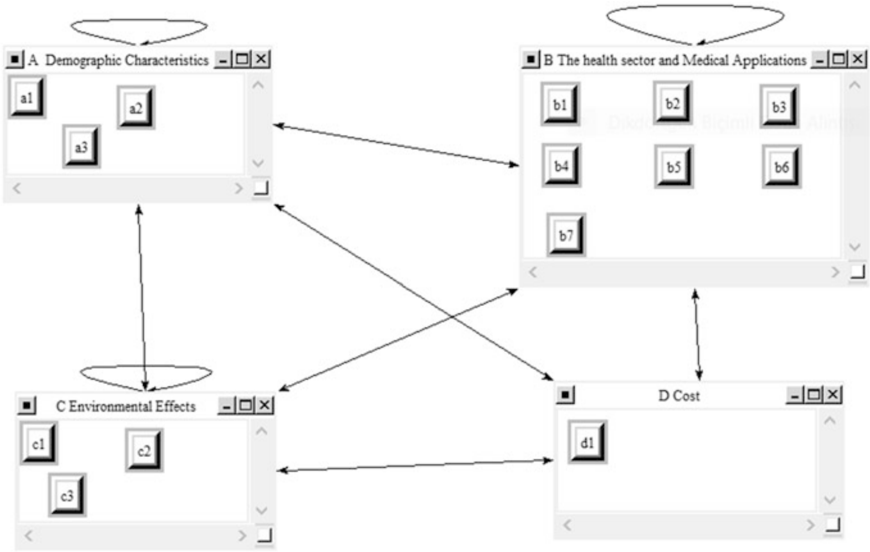


Fig. 4 The Decision Network

After determining the groups of attributes and attributes affecting each other, creating clusters and dependencies, a network structure was formed at Super Decisions software which works based on two multi-attribute decision-making methods, AHP and ANP. The decision network representing the dependencies among clusters is presented in Fig. 4.

4.2.2 Assessing the Importance of the Attributes

In this study, the importance of attributes was calculated using ANP as dependencies exist among the attributes. In accordance with ANP, the experts were asked to reply to a set of pairwise comparison questions as a further step. We computed the geometric means of the replies of the experts for each pairwise comparison question assessed on a 1-9 scale to aggregate their judgments. These judgments were then entered into Super Decisions software which computed the eigenvectors of the pairwise comparison matrices, the supermatrix, the weighted supermatrix, and the limit matrix. The entries of any column of the limit matrix revealed the importance of the attributes. We also aggregated the judgments of the medical doctors and the civil authorities and computed the importance of the attributes accordingly. The overall importance and importance for the two groups of participants are given in Table 3.

Table 3 The importance of attributes

	Attribute	Priorities according to MD	Priorities according to civil authorities	Overall priorities
A	a1. Population	0.2909	0.2898	0.2891
	a3. Centrality	0.1856	0.1700	0.1780
	a2. Population density	0.0208	0.0409	0.0318
B	b3. Number of public hospitals	0.1160	0.1246	0.1198
	b5. Number of branch hospitals	0.0469	0.0505	0.0496
	b4. Total number of beds in the public hospitals	0.0235	0.0346	0.0305
	b1. Number of FHC	0.0316	0.0143	0.0203
	b7. Number of private hospitals	0.0179	0.0172	0.0180
	b6. Number of beds in the branch hospitals	0.0128	0.0137	0.0135
	b2. Number of physicians in FHC	0.0103	0.0085	0.0095
C	c1. Distance to main road	0.1121	0.1101	0.1114
	c2. Traffic congestion	0.05937	0.05985	0.0598
	c3. Noise center	0.01481	0.00904	0.0112
D	d1. Land cost	0.05744	0.05699	0.0572

As shown in Table 3, there is no significant difference between the two groups, and the ranking is the same. According to the overall results as well as the groups' results, the population attribute (a1) is the most important attribute, followed by centrality (a2), the number of public hospitals (b3), and distance to the main road (c1). On the other hand, the number of physicians in the county's family health centers attribute (b2) becomes the least important attribute.

4.2.3 Decision Matrix

As we aimed to evaluate the counties of Trabzon according to the identified attributes in a multi-attribute decision-making basis, a decision matrix representing alternatives (i.e., the counties) at the rows, attributes at the columns, and the performance values of alternatives with respect to the attributes at the entries were generated. Most of the performance values were gathered from databases such as the Turkish Statistical Institute (TUIK), Turkey Public Hospitals Authority, Trabzon's provincial public health office, Trabzon Public Hospitals Association General Secretary, and Public Health Agency of Turkey, official websites of the private hospitals in Trabzon, and the Republic of Turkey General Directorate of Highways.

As aforementioned in section "Structuring the problem," population density (a2) value was computed by dividing the population of a county by its area while the centrality (a3) value of a county was computed as the sum of the ratios of the population of other counties to the distances of them to that county. However, there is no specific data for traffic congestion (c2), noise center (c3), and land cost (d1). For this reason, the values of the counties with respect to these attributes were obtained from experts on a scale of 0-100. The arithmetic averages of these subjective evaluations were computed. At this step, five authorities working at Trabzon Regional Directorate of Highways determined the average traffic congestion of the counties. Five academicians from the City and Regional Planning Department of Karadeniz Technical University determined the values of the noise center attribute. Finally, three land surveyors working in Trabzon determined the average land cost of possible areas in that county. The resulting decision matrix is shown in Table 4.

4.3 Analyzing the Problem

To select the most appropriate county for building a new hospital, we constructed outranking relations among counties utilizing PROMETHEE. The alternatives (i.e., the counties) and the attributes were introduced into Visual PROMETHEE software, and then the performance values of each alternative with respect to each attribute were entered.

In accordance with PROMETHEE, the preference functions were defined to represent the preferences of experts among alternatives with respect to each attribute. The experts were also consulted to determine the corresponding indifference and preference threshold values (Table 5). Each expert specified threshold values individually. Then, their arithmetic means were computed.

Based on this information, necessary computations were made on the decision matrix given in Table 4 utilizing Visual PROMETHEE software, and we came up with the partial and complete pre-orders of the counties.

PROMETHEE I provides the partial pre-order of the counties based on the decreasing order of their leaving flows and the increasing order of their entering flows as given in Fig. 5. As a result, a subset of counties appropriate for a new hospital building location in Trabzon was revealed. Akçaabat, Beşikdüzü, Of, and Araklı counties outranked most of the counties while Ortahisar county was found incomparable with the others.

We need a complete pre-order to propose the most appropriate county for a new hospital location in Trabzon, instead of the subset of the countries found above. PROMETHEE II provides the ranking of the counties based on the decreasing order

Table 4 The decision matrix

[p]

	a1	a2	a3	b1	b2	b3	b4	b5	b6	b7	c1	c2	c3	d1
Criteria importance	0.2891	0.03186	0.1780	0.0203	0.0095	0.1198	0.0305	0.0496	0.0135	0.0180	0.1114	0.0598	0.0112	0.0572
Unit	(person)	(centrality)	(person/m^2)	(FHC)	(physician)	(public hospital)	(bed)	(branch hospital)	(bed)	(private hospital)	(km)	(0.100)	(0.100)	(TL)
Akçaabat	115,939	309.1707	141993.8774	13	33	1	221	0	0	2	0	87	0.1333	700,000
Araklı	47,960	103.3621	71817.9022	1	14	1	106	0	0	0	0	44	0.8836	900,000
Arsin	28,208	179.6688	64934.4800	2	10	1	5	0	0	0	0	42	1.4013	550,000
Beşikdüzü	21,870	260.3571	44015.9544	1	6	0	0	0	0	0	0	40	0.7143	800,000
Çarşıbaşı	15,596	236.3030	39632.4214	2	5	0	0	0	0	0	0	41	0.0000	350,000
Çaykara	13,854	24.1359	25351.2921	3	4	1	15	1	88	0	26	28	0.0000	400,000
Dernekpazarı	3803	42.7303	18458.3615	1	1	0	0	0	0	0	18	19	0.0000	500,000
Düzköy	14,527	116.2160	29889.7561	3	5	1	10	0	0	0	28	19	0.0000	500,000
Hayrat	7631	31.2746	21138.9758	2	3	0	0	0	0	0	11	19	0.0000	250,000
Köprübaşı	4940	26.1376	20451.7782	1	2	1	10	0	0	0	14	22	0.0000	1,000,000
Maçka	24,232	26.1968	43320.2633	2	6	0	0	1	110	0	26	38	0.1622	700,000
Of	42,405	164.3605	58101.9159	2	11	1	79	0	0	0	0	47	0.1938	750,000
Ortahisar	314,246	67.9010	332203.9963	66	98	2	999	2	430	2	0	98	0.0983	7,000,000
Sürmene	10,903	65.6807	38769.7518	2	8	1	61	0	0	0	0	42	0.2711	1,000,000
Şalpazarı	26,421	164.1056	37276.0245	1	5	1	10	0	0	0	16	22	0.0000	650,000
Tonya	15,217	86.4602	28489.9815	2	5	1	5	0	0	0	20	23	0.5114	600,000
Vakfıkebir	26,636	188.9078	49459.6786	3	7	1	132	0	0	0	0	46	0.0000	600,000
Yomra	32.394	161.970	77112.6864	4	10	1	616	0	0	0	0	71	0.3750	1,250,000

Table 5 Information on preference functions

Attribute	Preference function type	Indifference threshold value	Preference threshold value
Population	Linear preference	.	140707.03
Population density	Linear preference	.	172.51
Centrality	Indifference area	85107.30	143940.17
The number of FHC in the county	Linear preference	.	28.47
The number of officials as physicians in the county's FHC	Indifference area	5.00	43.26
The number of public hospitals in the county	Level	0.56	1.30
Total number of beds in the county's public hospitals	Indifference area	50.00	250.00
The number of branch hospitals in the county	Level	0.66	1.07
Total number of beds in the county's branch hospitals	Linear preference	.	250.00
The number of private hospitals in the county	Level	0.81	1.230
Distance to the main road	Linear preference	.	14.00
Traffic congestion	Indifference area	21.35	45.93
Noise center	Indifference area	0.25	0.578
Land cost	Indifference area	500,000	1,500,000

of their net flows. As shown in Fig. 6, Akçaabat county, which had a net flow value of 0.1892, outranked all other counties and may be recommended as the most appropriate hospital location. On the other hand, Ortahisar county, which had a net flow value of 0.1503, followed Akçaabat county with a slight difference. Despite Ortahisar, which had the highest leaving flow, was at the best position (0.5082), it took second place in the final ranking because of its very high entering flow where less is better. According to the entering flow, Beşikdüzü, which had the lowest entering flow, was at the best position. However, it took fifth place in the final ranking due to its low leaving flow value.

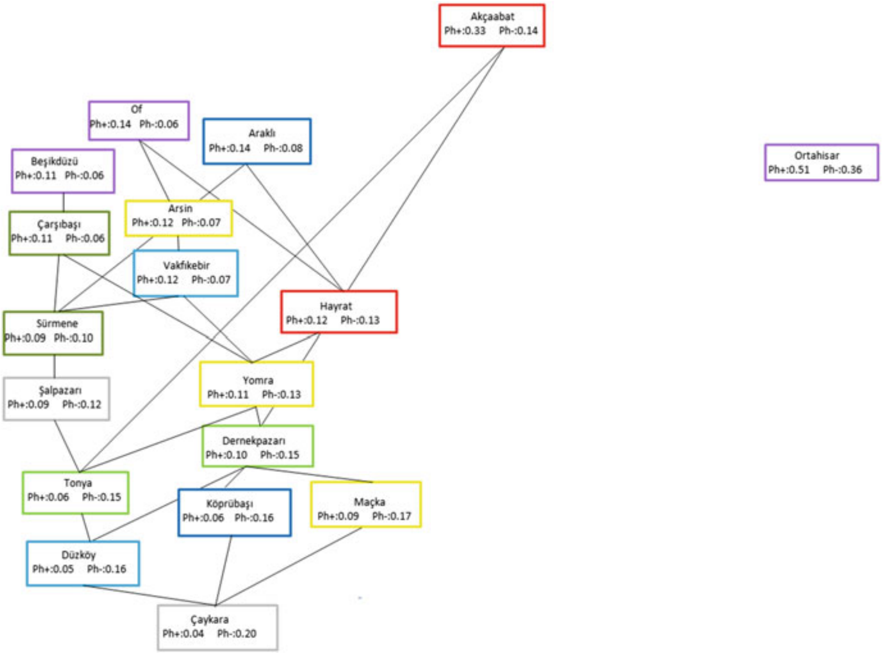


Fig. 5 The partial pre-order of the counties

4.4 Sensitivity Analysis Findings

As a final step, we conducted a sensitivity analysis to assess the robustness of the results of the proposed approach. For this purpose, we observed the ranking changes when importance of attributes changes.

PROMETHEE II results were found to be insensitive to the change in the importance of most of the attributes. When the importance of the attributes such as population density (a2), number of branch hospitals in the county (b5), number of beds in the county's branch hospitals (b6), distance to the main road (c1), and noise center (c3) change, Akçaabat county maintains its first place. Akçaabat also holds its first place even if there is a moderate change in the importance of the remaining attributes except for land cost (d1). Akçaabat loses its first place only when there is a slight change (0.02 points) in the importance of the land cost attribute. Readers may find more detailed information on the conducted sensitivity analysis in İmamoğlu's (2015) study.

PROMETHEE Flow Table

Rank	action	Phi	Phi+	Phi-
1	Akçaabat	0,1892	0,3302	0,1410
2	Ortahisar	0,1503	0,5082	0,3579
3	Of	0,0725	0,1360	0,0636
4	Araklı	0,0636	0,1393	0,0757
5	Beşikdüzü	0,0577	0,1137	0,0560
6	Çarşıbaşı	0,0518	0,1098	0,0580
7	Arsin	0,0473	0,1180	0,0707
8	Vakfıkebir	0,0404	0,1152	0,0748
9	Hayrat	-0,0066	0,1228	0,1294
10	Sürmene	-0,0083	0,0886	0,0969
11	Yomra	-0,0250	0,1100	0,1350
12	Şalpazarı	-0,0353	0,0861	0,1214
13	Dernekpazarı	-0,0510	0,1012	0,1523
14	Maçka	-0,0791	0,0893	0,1683
15	Tonya	-0,0949	0,0564	0,1514
16	Köprübaşı	-0,1018	0,0643	0,1661
17	Düzköy	-0,1117	0,0532	0,1648
18	Çaykara	-0,1588	0,0401	0,1989

Fig. 6 The complete pre-order of the counties

5 Conclusions and Further Suggestions

The health sector is in constant development due to advanced technology and increasing healthcare needs. Within this development, new hospital requirements and the problem of selecting appropriate locations for these hospitals arise. Additionally, the COVID-19 outbreak makes this decision very important and urgent.

In this study, an integrated MADM model was presented for finding an appropriate hospital location in Trabzon province of Turkey at the level of counties. In the proposed model, ANP was utilized to determine the importance of the attributes, and PROMETHEE was used to rank the counties.

Based on an extensive literature review and the point of view of experts, we determined 14 attributes under four clusters to evaluate counties considered as alternatives to hospital locations. Then, we identified the relations among the attributes, and we found that a network model was appropriate for the analysis. Based on the number of relations among attributes, the population attribute can be considered the most central attribute, i.e., had the highest number of attributes affected by itself and it affected the highest number of attributes. We utilized the ANP method to assess the importance of the attributes. The population, the most central attribute, was also found as the most important attribute. As a further step, we used PROMETHEE to evaluate counties with respect to objective and subjective attributes. Akçaabat county was the most appropriate location for building a new hospital, according to PROMETHEE II. Besides, we found that the first place of Akçaabat was not sensitive to the changes in the importance of the attributes. If policymakers want to build more than one hospital, the counties such as Akçaabat, Ortahisar, Beşikdüzü, Of, and Araklı may be recommended according to PROMETHEE I results.

The findings are justified as Akçaabat is a county with one of the highest populations and a central one in Trabzon. Besides, the current state-owned hospital in the county was not meeting the demand even before the COVID-19 outbreak.

The generated MADM model is a generic model that can be applied in other cities and other countries by changing experts as well as updating pairwise comparison questions and attributes importance.

Furthermore, different aspects such as the point of view of the potential patients can be included in the model as decision makers. Finally, to improve this study, a decision model can be developed to select a suitable location for a new hospital in Akçaabat county.

Appendix: The List of Attributes Used in Previous Studies

	Construction/ Lease Cost	Building rearrange-ment cost	Landscape cost	Labour cost /Oppor-tunity	Land cost/ opportu-nity	Water and Elec-tricity supply	Enlargement opportunity	Population	Population density	Population age structure	Level of income	Management objective	Rank of competi-tors	Policymaker's attitude	Puplic support	Hospital man-age-ment sector	Medical practice and phar-maceu-tical sector
Wu et al.(2007a)	X			X	X			X	X	X		X	X	X		X	X
Wu et al.(2007b)	X			X	X			X	X	X		X	X	X		X	X
Lin et al.(2008)				X	X			X	X	X		X	X	X		X	X
Önüt et al.(2008)	X			X	X			X	X		X	X	X	X		X	X
Lin and Tsai (2009)	X			X	X						X			X		X	X
Vahidnia et al. (2009)																	
Wu et al. (2009)	X			X	X			X	X	X		X	X	X		X	X
Aydın and Arslan (2010)	X	X	X						X		X						
Sinuany-Stern et al. (1995)					X				X								

(continued)

	Construction/ Lease Cost	Building rearrange-ment cost	Landscape cost	Labour cost /Oppor-tunity	Land cost/ opportu-nity	Water and Elec-tricity supply	Enlargement opportunity	Population	Population density	Population age structure	Level of income	Management objective	Rank of competi-tors	Policymaker's attitude	Puplic support	Hospital man-age-ment sector	Medical practice and pharma-ceutical sector
Soltani and Marandi (2011)					X				X								
Chatterjee and Muherjee (2013)			X		X		X		X		X						
Chiu and Tsai (2013)	X				X			X	X				X			X	X
Kim et al.(2015)						X		X		X	X			X	X		X
Sen and Demi-ral(2016)	X		X		X		X										
Şen (2017)	X		X		X		X										
Şahin et al.(2019)				X		X		X		X	X						X

	Health sector and activities of competitors	Distance of competing hospitals	Regulation and standards of establishment of hospitals	Efforts to promote medical network	Centrality	Distance from main road	Transport time to the hospital / traffic density	Proximity to settlements	City plan compliance	Proximity to noise centers	Public Transportation	Parking area	Significant change in market demand	Significant fluctuations in production costs	Significant changes in the financial market	Reduction local unemployment rate	Providing population distribution
Wu et al.(2007a)	X		X	X										X	X		
Wu et al.(2007b)	X		X	X										X	X		
Lin et al.(2008)	X		X	X									X	X	X		
Önüt et al.(2008)	X												X		X		
Lin and Tsai (2009)	X		X														
Vahidnia et al. (2009)						X	X			X							
Wu et al. (2009)			X	X										X	X		
Aydın and Arslan (2010)	X	X			X			X	X	X		X					
Sinuany-Stern et al. (1995)																X	X
Soltani and Marandi (2011)		X				**X**	X										

(continued)

	Health sector and activities of competitors	Distance of competing hospitals	Regulation and standards of establishment of hospitals	Efforts to promote medical network	Centrality	Distance from main road	Transport time to the hospital / traffic density	Proximity to settlements	City plan compliance	Proximity to noise centers	Public Transportation	Parking area	Significant change in market demand	Significant fluctuations in production costs	Significant changes in the financial market	Reduction local unemployment rate	Providing population distribution
Chatterjee and Muherjee (2013)						X	X	X			X						
Chiu and Tsai (2013)	X					X			X		X	X					
Kim et al.(2015)	X					X					X						
Sen and Demiral(2016)										X	X	X					
Şen (2017)										X	X	X					
Şahin et al.(2019)	X	X				**X**				X	X		X				

References

Abdel-Basset, M., Gamal, A., Chakrabortty, R. K., & Ryan, M. (2021). A new hybrid multi-criteria decision-making approach for location selection of sustainable offshore wind energy stations: A case study. *Journal of Cleaner Production, 280*(2), 124462. https://doi.org/10.1016/j.jclepro.2020.124462

Agrebi, M., & Abed, M. (2021). Decision-making from multiple uncertain experts: Case of distribution center location selection. *Soft Computing, 25*, 4525–4544. https://doi.org/10.1007/s00500-020-05461-y

Al-Shemmeri, T., Al-Kloub, B., & Pearman, A. (1997). Model choice in multi-criteria decision aid. *European Journal of Operational Research, 97*(3), 550–560. https://doi.org/10.1016/S0377-2217(96)00277-9

Atılgan, E. (2016). The technical efficiency of hospital inpatient care services: An application for Turkish public hospitals. *Business and Economics Research Journal, 7*(2), 203–214. https://doi.org/10.20409/berj.2016217537

Aydin, O., & Arslan, G. (2010). Optimal hospital location with fuzzy AHP. *The Business Review, Cambridge, 15*, 262–268.

Brans, J. P., & Mareschal, B. (1994). The PROMCALC and GAIA decision support system for multi-criteria decision aid. *Decision Support Systems, 12*(4-5), 297–310. https://doi.org/10.1016/0167-9236(94)90048-5

Brans, J. P., & Vincke, P. (1985). Note—A preference ranking organization method: (the PROMETHEE method for multiple criteria decision-making). *Management Science, 31*(6), 647–656. https://doi.org/10.1287/mnsc.31.6.647

Chatterjee, D., & Mukherjee, B. (2013). Potential hospital location selection using AHP: A study in rural India. *International Journal of Computer Applications, 71*(17).

Cheng, E. W., Li, H., & Yu, L. (2007). A GIS approach to shopping mall location selection. *Building and Environment, 42*(2), 884–892. https://doi.org/10.1016/j.buildenv.2005.10.010

Chiu, J. E., & Tsai, H. H. (2013). Applying analytic hierarchy process to select optimal expansion of hospital location: The case of a regional teaching hospital in Yunlin. In *Proceedings from ICSSSM 2013: In Service Systems and Service Management, 10th International Conference* (pp. 603–606). IEEE. https://doi.org/10.1109/ICSSSM.2013.6602588

Chou, T. Y., Hsu, C. L., & Chen, M. C. (2008). A fuzzy multi-criteria decision model for international tourist hotels location selection. *International Journal of Hospitality Management, 27*(2), 293–301. https://doi.org/10.1016/j.ijhm.2007.07.029

Demirel, T., Demirel, N. Ç., & Kahraman, C. (2010). Multi-criteria warehouse location selection using Choquet integral. *Expert Systems with Applications, 37*(5), 3943–3952. https://doi.org/10.1016/j.eswa.2009.11.022

Dey, B., Bairagi, B., Sarkar, B., & Sanyal, S. K. (2013). A hybrid fuzzy technique for the selection of warehouse location in a supply chain under a utopian environment. *International Journal of Management Science and Engineering Management, 8*(4), 250–261. https://doi.org/10.1080/17509653.2013.825075

Dock, J. P., Song, W., & Lu, J. (2015). Evaluation of dine-in restaurant location and competitiveness: Applications of gravity modeling in Jefferson County, Kentucky. *Applied Geography, 60*, 204–209. https://doi.org/10.1016/j.apgeog.2014.11.008

Ehsanifar, M., Wood, D. A., & Babaie, A. (2020). UTASTAR method and its application in multi-criteria warehouse location selection. *Operations Management Research.*https://doi.org/10.1007/s12063-020-00169-6

Eldemir, F., & Onden, I. (2016). Geographical information systems and multi-criteria decisions integration approach for hospital location selection. *International Journal of Information Technology and Decision Making, 15*(05), 975–997. https://doi.org/10.1142/S0219622016500218

Fragkiadakis, G., Doumpos, M., Zopounidis, C., & Germain, C. (2016). Operational and economic efficiency analysis of public hospitals in Greece. *Annals of Operations Research, 247*(2), 787–806. https://doi.org/10.1007/s10479-014-1710-7

Furtado, P. A. X., & Sola, A. V. H. (2020). Fuzzy complex proportional assessment applied in location selection for installation of photovoltaic plants. *Energies, 13*, 6260. https://doi.org/10.3390/en13236260

Gang, J., Tu, Y., Lev, B., Xu, J., Shen, W., & Yao, L. (2015). A multi-objective bi-level location planning problem for stone industrial parks. *Computers and Operations Research, 56*, 8–21. https://doi.org/10.1016/j.cor.2014.10.005

Govindan, K., Kannan, D., & Shankar, M. (2015). Evaluation of green manufacturing practices using a hybrid MCDM model combining DANP with PROMETHEE. *International Journal of Production Research, 53*(21), 6344–6371. https://doi.org/10.1080/00207543.2014.898865

Guneri, A. F., Cengiz, M., & Seker, S. (2009). A fuzzy ANP approach to shipyard location selection. *Expert Systems with Applications, 36*(4), 7992–7999. https://doi.org/10.1016/j.eswa.2008.10.059

İmamoğlu, G. (2015). *Analitik ağ süreci ve PROMETHEE teknikleri ile hastane yer seçimi: Trabzon örneği [Unpublished master's thesis]*. Istanbul Technical University.

Ishizaka, A., Nemery, P., & Lidouh, K. (2013). Location selection for the construction of a casino in the greater London region: A triple multi-criteria approach. *Tourism Management, 34*, 211–220. https://doi.org/10.1016/j.tourman.2012.05.003

Kabak, M., & Dağdeviren, M. (2014). A hybrid MCDM approach to assess the sustainability of students' preferences for university selection. *Technological and Economic Development of Economy, 20*(3), 391–418. https://doi.org/10.3846/20294913.2014.883340

Kabir, G., & Sumi, R. S. (2014). Power substation location selection using fuzzy analytic hierarchy process and PROMETHEE: A case study from Bangladesh. *Energy, 72*, 717–730. https://doi.org/10.1016/j.energy.2014.05.098

Kampf, R., Průša, P., & Savage, C. (2011). Systematic location of the public logistic centres in Czech Republic. *Transport, 26*(4), 425–432. https://doi.org/10.3846/16484142.2011.635424

Kannan, D., Moazzeni, S., Darmian, S. M., & Afrasiabi, A. (2021). A hybrid approach based on MCDM methods and Monte Carlo simulation for sustainable evaluation of potential solar sites in east of Iran. *Journal of Cleaner Production, 279*, 122368. https://doi.org/10.1016/j.jclepro.2020.122368

Kılcı, F., Kara, B. Y., & Bozkaya, B. (2015). Locating temporary shelter areas after an earthquake: A case for Turkey. *European Journal of Operational Research, 243*(1), 323–332. https://doi.org/10.1016/j.ejor.2014.11.035

Kilic, H. S., Zaim, S., & Delen, D. (2015). Selecting "the best" ERP system for SMEs using a combination of ANP and PROMETHEE methods. *Expert Systems with Applications, 42*(5), 2343–2352. https://doi.org/10.1016/j.eswa.2014.10.034

Kim, J. I., Senaratna, D. M., Ruza, J., Kam, C., & Ng, S. (2015). Feasibility study on an evidence-based decision-support system for hospital site selection for an aging population. *Sustainability, 7*(3), 2730–2744. https://doi.org/10.3390/su7032730

Kumar, B. V., & Srikanth, N. V. (2015). Optimal location and sizing of unified power flow controller (UPFC) to improve dynamic stability: A hybrid technique. *International Journal of Electrical Power and Energy Systems, 64*, 429–438. https://doi.org/10.1016/j.ijepes.2014.07.015

Lee, Y.-H., Keum, H.-J., Han, K.-Y., & Hong, W.-H. (2020). A hierarchical flood shelter location model for walking evacuation planning. *Environmental Hazards*. https://doi.org/10.1080/17477891.2020.1840327

Liang, F., Verhoeven, K., Brunelli, M., & Rezaei, J. (2021). Inland terminal location selection using the multi-stakeholder best-worst method. *International Journal of Logistics Research and Applications, ahead-of-print*.https://doi.org/10.1080/13675567.2021.1885634

Li, C., Xu, C., & Li, X. (2020). A multi-criteria decision-making framework for site selection of distributed PV power stations along high-speed railway. *Journal of Cleaner Production, 277*, 124086. https://doi.org/10.1016/j.jclepro.2020.124086

Lin, C. T., & Tsai, M. C. (2009). Development of an expert selection system to choose ideal cities for medical service ventures. *Expert Systems with Applications, 36*(2), 2266–2274. https://doi.org/10.1016/j.eswa.2007.12.056

Lin, C. T., Lee, C., & Chen, Z. J. (2010). An expert system approach to medical region selection for a new hospital using data envelopment analysis. *iBusiness, 2*(02), 128–138. https://doi.org/10.4236/ib.2010.22016

Lin, C. T., Wu, C. R., & Chen, H. C. (2008). The study of construct key success factors for the Taiwanese hospitals of location selection by using the fuzzy AHP and sensitivity analysis. *International Journal of Information and Management Sciences, 19*(1), 175–200.

Liu, P., & Li, Y. (2020). Multiattribute decision method for comprehensive logistics distribution center location selection based on 2-dimensional linguistic information. *Information Sciences, 538*, 209–244. https://doi.org/10.1016/j.ins.2020.05.131

Lu, X. X., Yang, S. W., & Zheng, N. (2011). Location-selection of wireless network based on restricted Steiner tree algorithm. *Procedia Environmental Sciences, 10*, 368–373. https://doi.org/10.1016/j.proenv.2011.09.060

Macharis, C., Springael, J., De Brucker, K., & Verbeke, A. (2004). PROMETHEE and AHP: The design of operational synergies in multi-criteria analysis.: Strengthening PROMETHEE with ideas of AHP. *European Journal of Operational Research, 153*(2), 307–317. https://doi.org/10.1016/S0377-2217(03)00153-X

Magaji, N., & Mustafa, M. W. (2011). Optimal location and signal selection of UPFC device for damping oscillation. *International Journal of Electrical Power and Energy Systems, 33*(4), 1031–1042. https://doi.org/10.1016/j.ijepes.2011.01.020

Mahmoudi, M., Aydi, A., & Ibrahim, H. (2021). Site selection for artificial recharge with treated wastewater with the integration of multi-criteria evaluation and ELECTRE III. *Environmental Science and Pollution Research. ahead-of-print.*https://doi.org/10.1007/s11356-021-12354-6

Meng, Q. B. L. J. T., & Lu, X. X. (2011). An approximate algorithm to solve the location-selection of recycle Water Station problem. *Procedia Environmental Sciences, 10*, 363–367. https://doi.org/10.1016/j.proenv.2011.09.059

Milovanovic, Z., Milovanovic, S., Milovanovic, V. J., Dumonjic-Milovanovic, S., & Brankovic, D. (2021). Modeling of the optimization procedure for selecting the location of new thermal power plants (TPP). *International Journal of Mathematical, Engineering and Management Sciences, 6*(1), 118–165. https://doi.org/10.33889/IJMEMS.2021.6.1.009

Mirzahossein, H., Sedghi, M., Habibi, H. M., & Jalali, F. (2020). Site selection methodology for emergency centers in silk road based on compatibility with Asian highway network using the AHP and ArcGIS. *Innovative Infrastructre Solutions, 5*, 113. https://doi.org/10.1007/s41062-020-00362-3

Mokhtarian, M. N., & Hadi-Vencheh, A. (2012). A new fuzzy TOPSIS method based on left and right scores: An application for determining an industrial zone for dairy products factory. *Applied Soft Computing, 12*(8), 2496–2505. https://doi.org/10.1016/j.asoc.2012.03.042

Onut, S., Tuzkaya, U. R., & Kemer, B. (2008). An analytical network process approach to the choice of hospital location (Hastane yeri secimine bir analitik ag sureci yaklasimi). *Muhendislik ve Fen Bilimleri Dergisi, 25*(4), 367–379.

Peng, A. H., & Xiao, X. M. (2013). Material selection using PROMETHEE combined with analytic network process under hybrid environment. *Materials and Design, 47*, 643–652. https://doi.org/10.1016/j.matdes.2012.12.058

Rahmaniani, R., & Ghaderi, A. (2015). An algorithm with different exploration mechanisms: Experimental results to capacitated facility location/network design problem. *Expert Systems with Applications, 42*(7), 3790–3800. https://doi.org/10.1016/j.eswa.2014.12.051

Saaty, T., & L. (1996). *Decision making with dependence and feedback: The analytic network process* (Vol. 4922). RWS Publications.

Şahin, T., Ocak, S., & Top, M. (2019). Analytic hierarchy process for hospital site selection. *Health Policy and Technology, 8*(1), 42–50. https://doi.org/10.1016/j.hlpt.2019.02.005

Sen, H. (2017). Hospital location selection with ARAS-G. *The Eurasia Proceedings of Science Technology Engineering and Mathematics, 1*, 359–365.

Şen, H., & Demiral, M. F. (2016). Hospital location selection with grey system theory. *European Journal of Economics and Business Studies, 2*(2), 66–79.

Şenvar, O., Otay, I., & Bolturk, E. (2016). Hospital site selection via hesitant fuzzy TOPSIS. *IFAC-PapersOnLine, 49*(12), 1140–1145. https://doi.org/10.1016/j.ifacol.2016.07.656

Shen, C. Y., & Yu, K. T. (2009). A generalized fuzzy approach for strategic problems: The empirical study on facility location selection of authors' management consultation client as an example. *Expert Systems with Applications, 36*(3), 4709–4716. https://doi.org/10.1016/j.eswa.2008.06.035

Sinuany-Stern, Z., Mehrez, A., Tal, A. G., & Shemuel, B. (1995). The location of a hospital in a rural region: The case of the Negev. *Location Science, 3*(4), 255–266. https://doi.org/10.1016/0966-8349(96)00002-2

Soltani, A., & Marandi, E. Z. (2011). Hospital site selection using two-stage fuzzy multi-criteria decision making process. *Journal of Urban and Environmental Engineering, 5*(1), 32–43. https://doi.org/10.4090/juee.2011.v5n1.032043

Tercan, E., Tapkın, S., Latinopoulos, D., Dereli, M. A., Tsiropoulos, A., & Ak, M. F. (2020). A GIS-based multi-criteria model for offshore wind energy power plants site selection in both sides of the Aegean Sea. *Environmental Monitoring and Assessment, 192*, 652. https://doi.org/10.1007/s10661-020-08603-9

Turkish Statistical Institute. (2021). Population of province/district Centers, towns/villages by provinces and districts and annual growth rate of population [data file]. https://data.tuik.gov.tr/Bulten/Index?p=The-Results-of-Address-Based-Population-Registration-System-2020-37210.

Topcu, Y. İ., Ulengin, F., Kabak, Ö., Önsel, E. Ş., & Ünver, B. (2020). A decision support methodology for increasing the efficiency of the largest border crossing between Europe and Turkey, *research in transportation economics*, 80/may. *Art, 100825*. https://doi.org/10.1016/j.retrec.2020.100825

URL1. https://tr.wikipedia.org/wiki/Dosya:World_map_pol_2005_v02.svg

URL2. https://www.harita.gov.tr/urun/turkiye-mulki-idare-sinirlari/232

URL3. https://trabzon.csb.gov.tr/ilcelerimiz-i-1279

Vahidnia, M. H., Alesheikh, A. A., & Alimohammadi, A. (2009). Hospital site selection using fuzzy AHP and its derivatives. *Journal of Environmental Management, 90*(10), 3048–3056. https://doi.org/10.1016/j.jenvman.2009.04.010

Wey, W. M. (2015). Smart growth and transit-oriented development planning in site selection for a new metro transit station in Taipei. *Taiwan. Habitat International, 47*, 158–168. https://doi.org/10.1016/j.habitatint.2015.01.020

Wu, C. R., Lin, C. T., & Chen, H. C. (2007a). Optimal selection of location for Taiwanese hospitals to ensure a competitive advantage by using the analytic hierarchy process and sensitivity analysis. *Building and Environment, 42*(3), 1431–1444. https://doi.org/10.1016/j.buildenv.2005.12.016

Wu, C. R., Lin, C. T., & Chen, H. C. (2007b). Evaluating competitive advantage of the location for Taiwanese hospitals. *Journal of Information and Optimization Sciences, 28*(5), 841–868. https://doi.org/10.1080/02522667.2007.10699777

Wu, C. R., Lin, C. T., & Chen, H. C. (2009). Integrated environmental assessment of the location selection with fuzzy analytical network process. *Quality and Quantity, 43*(3), 351–380. https://doi.org/10.1007/s11135-007-9125-z

Xie, W. J., Huang, S. G., Zhao, Y. L., Xin, L. I. U., & Gu, W. Y. (2013). Selection of PCEs' location in multi-domain optical networks. *The Journal of China Universities of Posts and Telecommunications, 20*(6), 62–68. https://doi.org/10.1016/S1005-8885(13)60110-7

Yang, L., Ji, X., Gao, Z., & Li, K. (2007). Logistics distribution centers location problem and algorithm under fuzzy environment. *Journal of Computational and Applied Mathematics, 208*(2), 303–315. https://doi.org/10.1016/j.cam.2006.09.015

Yucenur, G. N., & Ipekci, A. (2021). SWARA/WASPAS methods for a marine current energy plant location selection problem. *Renewable Energy, 163*, 1287–1298. https://doi.org/10.1016/j.renene.2020.08.131

Yunna, W., & Geng, S. (2014). Multi-criteria decision making on selection of solar–wind hybrid power station location: A case of China. *Energy Conversion and Management, 81*, 527–533. https://doi.org/10.1016/j.enconman.2014.02.056

Zeferino, E. F. S., Makinde, O. A., Mpofu, K., Ramatsetse, B. I., & Daniyan, I. A. (2021). Prioritizing factors influencing the selection of a suitable quarantine facility for COVID-19 patients using Pareto-enhanced analytical hierarchy process. *Facilities, ahead-of-print.*https://doi.org/10.1108/F-04-2020-0043

Determining the Priority of University Technology Transfer Office Activities for Each Group of the Stakeholder Using SWARA Method

Melis Almula Karadayi and Yeliz Ekinci

Abstract This chapter is an attempt to summarize the University Technology Transfer Office (UTTO) activities firstly. The main objective is to determine the priority of these activities for each group of the stakeholder; namely, students, academicians, and industry. UTTOs promote the conversion of knowledge created at university to marketable goods; which makes them crucial for the countries. Moreover, UTTOs serve to the economic development of countries by bringing companies and researchers together and motivating them for technology transfer. In this study, we firstly examine the UTTOs to understand the dynamics of these offices. Then, we list the critical activities of the UTTOs and conduct three surveys specific to each group of the stakeholder where the participants evaluate the relative importance of the activities determined. Importance/priority of UTTO activities is determined using Stepwise Weight Assessment Ratio Analysis (SWARA) method, for each group of the stakeholder. Finally, the results are presented, which will shed light on the decisions in managing and organizing UTTO activities.

Keywords Technology transfer · University Technology Transfer Offices · Multiple-criteria decision-making · Stepwise weight assessment ratio analysis (SWARA)

M. A. Karadayi (✉)
Department of Industrial Engineering, Istanbul Medipol University, Istanbul, Turkey
e-mail: makaradayi@medipol.edu.tr

Y. Ekinci
Department of Management Information Systems, Istanbul Bilgi University, Istanbul, Turkey
e-mail: yeliz.ekinci@bilgi.edu.tr

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_17

1 Introduction

Recently, technology transfer has gained a significant role because of the rapid changes in the convergence of high tech industry and globalization in the world (Lee et al., 2012). With the growing perception of the importance of seamless interactions between science and industry, and thus technology transfer, industrial science links have become a central concern of many government policies in recent years (Macho-Stadler et al., 2007).

There are many definitions of technology transfer. However, one of the popular definitions of it is; "*the transfer of new knowledge, products or processes from one organization to another for business benefit*" (Wittamore et al., 1998). Technology transfer terminology brings universities into the scene, since universities are seen as the centers of research, development, and technology creation. Universities aim at creating positive impacts through science commercialization. Therefore, it is observed that they have become more entrepreneurial, and have increased technology transfer activities by increasing their communication with the industry (Baglieri et al., 2018).

As Ekinci and Uray (2016) state; "businesses have increasingly found 'pure knowledge spillover' channels such as scientific publications and patents to be a less effective means of dissemination and instead they promote more tailored knowledge transfer channels, including collaborative research and information contacts." These changes in the field lead to the formation of University Technology Transfer Offices (UTTO).

The expectation of the universities from the UTTOs has been "resolving the different institutional logics of academic and commercial research in seeking technology commercialization" (Lee & Jung, 2021). The expectation of the students has been supporting and motivation for entrepreneurship activities, mostly. As Baglieri et al. (2018) state "technology transfer processes enable universities to increase their positive impact on society by following their entrepreneurial mission in various ways." In fact, entrepreneurship activities for new technology-based startups are supported by the governments since these startups are viewed as a critical source of new job creation (Chapple et al., 2005).

To the best of our knowledge, this is the first study that analyzes the priority/importance of the critical UTTO activities through surveys conducted with each group of the stakeholder; student, academician, and industry. In this study, we firstly perform a literature review to understand the dynamics of these offices. Technology transfers, UTTOs and their roles, goals are summarized, as well. This review helps us to determine the list of main activities conducted by the UTTOs. The existing literature in this research area shows that we must divide the stakeholders in the process into three bodies such as students, academicians, and industry. UTTOs have a key role in providing link among these parties (Tunca & Kanat, 2019). Therefore, we conduct three types of surveys specific to each body and these bodies evaluate the relative importance of the UTTO activities/ services determined. Stepwise weight assessment ratio analysis (SWARA) approach is used to analyze the evaluations of these bodies. Finally, the results are presented to provide important pathways to decision makers.

This chapter is organized into six more sections. Section "University Technology Transfer Offices" explains the technology transfer and TTO terminologies and also summarizes the existing literature on UTTOs together with their activities of them. Section "The Proposed Methodology" gives the details on the proposed methodology. Section "Findings" presents the application on the calculations and results of determination of importance. Section "Discussion" discusses the findings and section "Conclusions" finalizes the chapter with conclusions and future research directions.

2 University Technology Transfer Offices

Before analyzing technology transfer, it is better to give the definition of technology. Boer (1999) defines technology as adaptation of knowledge for practical purposes. Having defined the technology, the definition of technology transfer can be understood better now. "Technology transfer is the movement of data, designs, inventions, materials, software, know-how or trade secrets from one organization to another or from one purpose to another" (Twi-global, 2021). Another definition focuses on transforming scientific knowledge and innovation into marketable goods. Patent, license, and royalty agreements are forms of formal technology transfer.

Based on the definition of technology transfer, Technology Transfer Offices (TTO) can be defined as the places where the scientific knowledge is transformed into marketable goods. Holgersson and Aaboen (2019) state that the main role of TTOs is to support the commercialization of research output. TTOs tend to specialize in providing services at different stages of the value chain of technology transfer activities (Olaya-Escobar et al., 2020). TTO operation capability can be characterized by a TTO's ability to leverage its resources in generating patent registrations and applications. TTOs evaluate the commercial potential of scientific research and decide whether to apply for patent applications and register the inventions (Lee & Jung, 2021). TTOs also need to develop systems to take care of their employees and to support their knowledge relevant to technology commercialization (Lee & Jung, 2021).

Undoubtedly, universities are responsible for a large amount of knowledge creation. This makes university technology transfer a principal activity in releasing the economic potential of innovative technologies (Battaglia et al., 2017). Hence, this fundamental activity should be managed and organized by an organizational structure, This structure has been realized to be the technology transfer offices (TTO) within a university. These offices develop relations with industry, and are specialized in support services, most notably, partner search, management of intellectual property, and business development (Macho-Stadler et al., 2007). University TTOs are enigmatic actors in the academic entrepreneurship arena (O'kane et al., 2015).

Simply, the knowledge created at the universities should be converted to a marketable good in order to be used for public welfare and this is accomplished by University Technology Transfer Offices (UTTO). UTTOs work under university's administrative structure and they report to the Vice President (Rectorate). They are strongly committed to the university in terms of implementation policies. Although universities have always been a source of science and technology, they have made significant progress in strengthening technology transfer, especially in the last three decades, by improving cooperation between industry and academics (Sart, 2014). In this progress, UTTOs have an important place, since they are established to ensure the commercialization efforts of research activities; hence the commercial value the academicians generate stays in the university.

UTTOs basically protect university's intellectual property and support research and discoveries while working with industry (Markman et al., 2005). They provide their university the opportunity of gaining income from patents or licensing and they use their income to carry out researches by providing fund to the researchers, in attempt to produce more marketable goods and patents. The establishment of an UTTO can help to increase the awareness of academic researchers about obtaining academic patents under the name of a university, and provide them with additional benefits (van Burg et al., 2021). Another important role of UTTOs is that they serve the economic development of the countries by creating and supporting startup firms. As Gül and Ekinci (2015) state "UTTOs play a key role in economic development, which puts them in a position to not only monitor, facilitate and regulate transactions between university and industry, but also create an environment for both parties to foster technology transfer activities."

The main stakeholders of an UTTO are industry, students, and academicians. These bodies perceive some services of the UTTO more important than others. For instance, students are expected to be more interested in entrepreneurship and startup activities. Academicians are expected to benefit from research-based services more while industry is expected to be more interested in startup activities and research-based collaborations with the university. Although we have these inferences, there is a need to understand the true perceptions of the stakeholders of an UTTO through an empirical study. In this way, instead of guessing the perceptions and expectations of the students, academicians, and industry about an UTTO, it will be better to determine these by a scientific method. The opinions of the stakeholders should be taken by a survey and should be evaluated by a Multi-Criteria Decision-Making (MCDM) method. Such a study will be beneficial for both these bodies and policy makers. Policy makers will use the outputs of this study while planning the services of UTTOs.

To the best of our knowledge, there has been no study in the literature that analyze the perceived importance of the main UTTO activities by different stakeholders. Hence, we decided to study this topic in order to fill this gap in the existing literature. In this chapter, we focus on the rating of the UTTO activities from the point of view of academicians, students, and industry.

It is important to list the main activities of the UTTOs, firstly. Therefore, we conducted a review in the existing literature in order to make this list and then we

Table 1 Studies on UTTO activities

UTTO activity	Studies mentioning the activity
C1: Support for writing R&D project proposals to get funds	Baglieri et al., 2018; Munari et al., 2018; Holgersson & Aaboen, 2019; Link et al., 2007
C2: Commercialization of R&D results (patenting, licensing, intellectual property)	Antonelli, 2008; Bekkers & Freitas, 2008; Link et al., 2007; Macho-Stadler et al., 2007; Baglieri et al., 2018; Olaya-Escobar et al., 2020; Lee & Jung, 2021; Holgersson & Aaboen, 2019; Etzkowitz & Zhou, 2021; van Burg et al., 2021
C3: Trainings and seminars on technology transfer	Antonelli, 2008; Bekkers & Freitas, 2008; Hewitt-Dundas, 2013; Link et al., 2007; Macho-Stadler et al., 2007; Baglieri et al., 2018; Munari et al., 2018; Olaya-Escobar et al., 2020; Etzkowitz & Zhou, 2021; van Burg et al., 2021
C4: Support for bringing together researchers and industry, partner search	Bekkers & Freitas, 2008; Hewitt-Dundas, 2013; Baglieri et al., 2018;Munari et al., 2018; O'Kane, 2018; Lee & Jung, 2021; Tunca & Kanat, 2019; Etzkowitz & Zhou, 2021
C5:Providing information on financial support and funding opportunities	Baglieri et al., 2018; Munari et al., 2018; Tunca & Kanat, 2019; Holgersson & Aaboen, 2019; Etzkowitz & Zhou, 2021
C6:Startup opportunities	Battaglia et al., 2017 Markman et al., 2005; Sart, 2014; Baglieri et al., 2018; Olaya-Escobar et al., 2020; Holgersson & Aaboen, 2019; Etzkowitz & Zhou, 2021
C7: Talks, seminars, trainings, and events on entrepreneurship	Markman et al., 2005; Sart, 2014; Munari et al., 2018; Olaya-Escobar et al., 2020; Holgersson & Aaboen, 2019; Etzkowitz & Zhou, 2021

verified them with three UTTO professionals. Finally, we came up with the activities in Table 1.

3 The Proposed Methodology

The TTO activities that are listed in Table 1 include services that can be benefited by all three bodies; students, academicians, and industry. Therefore, these can be used in a survey conducted with all three bodies. In this study, the priority of these activities will be analyzed by an MCDM method and final priorities will be calculated. Multiple Criteria Decision-Making (MCDM) is the collective term for those formal decision support methods, that take multiple conflicting criteria into consideration explicitly. MCDM has already been used in various fields such as business, engineering, healthcare, energy, and environment. MCDM methods are generally used to find the important weights of several criteria. In this study, instead of criteria we have the services given by the UTTOs. As an MCDM method, we will use stepwise weight assessment ratio analysis (SWARA) method to analyze the pairwise comparison evaluations given by the participants from each group of the stakeholder and to calculate the priorities.

"MCDM is a set of methods that constitute a sub-branch of Decision Science and incorporate different approaches. MCDM is based on the process of modeling the decision process according to the criteria and analyzing it in a way that maximizes the benefit that the decision-maker will obtain at the end of the process." There are various methods for determining the weights of decision attributes in the MCDM literature like analytic hierarchy process (AHP) (Saaty, 1980), analytic network process (ANP) (Saaty & Vargas, 2001), entropy (Keršulienė & Turskis, 2011; Shannon, 1948; Sušinskas et al., 2011), and factor relationship (FARE) (Ginevicius, 2011). Furthermore, in some methods, calculations are very complicated, and the accuracy of the methods is not very high (Aghdaie et al., 2014).

The fundamental feature of SWARA (Stepwise Weight Assessment Ratio Analysis) method, which is an MCDM technique, is that experts' opinions, their own implicit knowledge, information, and experiences are applied throughout the evaluation process. The most important and influencing criterion gets the first rank, and the least important criterion gets the last rank (Ghorshi Nezhad et al., 2015). To sum up, the ability and mastery of the experts are the most vital and influential points in determining the importance of each criterion in the SWARA method (Keršulienė et al., 2010). Moreover, it is not considered to be complicated and time consuming (Hashemkhani Zolfani et al., 2013; Mardani et al., 2017). Moving from this point on, the SWARA method is selected in order to analyze the UTTO activities' priorities/importances. The activities are listed after an extensive literature review in section "University Technology Transfer Office."

The SWARA method was first introduced in 2010 by Keršuliene, Zavadskas, and Turskis (Keršulienė et al., 2010). Over the last decade, it has been used in many fields such as health, environment, tourism, management, and economy.

Keršulienė and Turskis (2011) employed SWARA for architect selection. Hashemkhani Zolfani, Farrokhzad, and Turskis (2013) used SWARA in order to determine the best alternative of mechanical longitudinal ventilation of tunnel pollutants during automobile accidents. Hashemkhani Zolfani et al. (2013) investigated important factors of online games based on explorer for game producer companies and designers of online games. Hashemkhani Zolfani et al. (2013) applied SWARA for evaluating different R&D project alternatives. Popović (2018) prioritizes tourism development strategies utilizing SWARA. Taghavi et al. (2019) applied SWARA to evaluate the preparedness of a university in Iran by determining important factors for successful ERP establishment.

Furthermore, in due course of time many researchers have integrated SWARA method with other MCDM techniques. Karabašević et al. (2015) suggested a framework for the recruitment and selection process of companies using the SWARA and ARAS methods. Işık and Adalı (2016) developed a model for hotel selection by integrating SWARA and OCRA methods. Juodagalviene et al. (2017) proposed a model for evaluating the layout of a single-family residential house using SWARA and EDAS. Eghbali-Zarch et al. (2018) evaluated treatment alternatives for type 2 diabetes using SWARA and modified MULTIMOORA methods. Vesković et al. (2018) combined Delphi, SWARA, and MABAC methods to evaluate railway management models in Bosnia and Herzegovina. Dersen and

Yontar (2020) integrated SWARA and TOPSIS methodologies to determine the most suitable renewable energy source. Ansari et al. (2020) proposed a hybrid model combining SWARA and COPRAS methodologies to evaluate and rank solutions to mitigate the sustainable remanufacturing supply chain risks. Ayyildiz et al. (2021) utilized SWARA and DEA approaches for assessing the performance of wastewater treatment plants in Turkey.

To sum up, it can be concluded that SWARA methodology helps decision makers to take into consideration their own priorities in many decision problems. The role of experts determined as decision makers is much more important in this method compared to other MCDM techniques used for weighting procedure (Hashemkhani & Saparauskas, 2013). Moreover, unlike the methods such as AHP and ANP, the evaluation process can be performed without evaluating or ranking the criteria. In addition, the number of comparisons made for criteria weighting in the SWARA method is less than the AHP method.

In this method, experts list all the criteria from the most important one to the least important one based on their own knowledge and experience. The SWARA method has a six-step implementation process. These stages can be summarized as follows (Keršulienė et al., 2010):

Step 1: In the first stage, the criteria to be included in the decision problem and decision committee consisting of decision makers are determined. It is assumed that there are n criteria (C_n, $n = 1,2, ..., n$) in the problem and there are m decision makers on the decision committee.

Step 2: At this stage, decision makers evaluate the criteria in order of importance from best to worst, based on their knowledge and experience.

Step 3: At this stage, the relative importance levels of the criteria are determined. For this, criterion j is compared with the previous criterion (j-1). Keršulienė et al. (2010) called this ratio "the comparative importance of the mean value" and indicated it with the symbol s_j.

Step 4: The coefficient k_j is determined as follows:

$$\boldsymbol{k_j = \begin{cases} 1 & j = 1 \\ s_j + 1 & j > 1 \end{cases}} \tag{1}$$

Step 5: The recalculated weight q_j is computed as follows:

$$\boldsymbol{q_j = \begin{cases} 1 & j = 1 \\ \frac{q_{j-1}}{k_j} & j > 1 \end{cases}} \tag{2}$$

Step 6: Determine the relative weight w_j of the evaluation criteria using the following equation:

$$\boldsymbol{w_j = \frac{q_j}{\sum_{k=1}^{n} q_k}} \tag{3}$$

where w_j represents the relative weight of the j-th criterion, and n shows the number of the criteria.

As mentioned before, the stakeholders of the UTTOs are students, academicians, and industry. Therefore the subjective evaluations of these bodies should be taken in order to determine the priorities of the UTTO activities. For each group of the stakeholder, three decision makers made the evaluations; who are three students and three academicians from different disciplines, as well as participants from three different industrial sectors. Although increasing the number of participants will definitely represent the whole population; since three DMs are seen as sufficient in MCDM literature (Karabašević et al., 2015; Prajapati et al., 2019; Yücenur & Ipekçi, 2021), we conducted the surveys with three participants from each group of the stakeholder.

4 Findings

In order to apply the SWARA method, each expert first determines the importance of each criterion. All the criteria are sorted according to each expert's opinion from first to last. Hence, UTTO activities are prioritized by each expert from most important one to least important one. Experts considered multiples of 5% as a scale of evaluation to determine the priority of each criterion. Using Eqs. (1-3), the comparative importance s_j, coefficient k_j, recalculated weight q_j, and relative weight value $w_j$are calculated by gathering information from each participant. In order to calculate sj, information gained privately from each expert. Experts considered multiples of 5% as a scale of evaluation. Results are presented in Table 2.

After obtaining Table 2, the final weight is calculated by taking the arithmetic average of experts' evaluations. Finally, priorities are calculated (over 1) as in Table 3. The results show that, the most important two activities have priority values over 0.2. Hence, the first two should be analyzed in detail. The importance of the other five activities is distributed more evenly by the industry with values changing between 0.097 and 0.149. The range for students is 0.080-0.163 and it is 0.090-0.147 for academicians.

5 Discussion

Looking at the results in section "Findings," we can discuss that academicians perceive *providing information on financial support and funding opportunities* as the most important activity, while both students and industry perceive *startup opportunities* as the most important activity of UTTO. Second order is given to *providing information on financial support and funding opportunities* by the industry; which was in the first order by the academicians. S*upport for writing R&D project proposals to get funds* is ranked second by the academicians and

Table 2 The resulting weight of UTTO activities

The resulting weight obtained from "Students"					The resulting weight obtained from "Academicians"					The resulting weight obtained from "Industry"				
DM1	s_j	k_j	q_j	w_j	DM1	s_j	k_j	q_j	w_j	DM1	s_j	k_j	q_j	w_j
C6		1	1	0.313	C5		1	1	0.238	C6		1	1	0.216
C1	0.5	1.5	0.667	0.209	C1	0.2	1.2	0.833	0.198	C4	0.3	1.3	0.769	0.166
C7	0.4	1.4	0.476	0.149	C2	0.3	1.3	0.641	0.152	C5	0.1	1.1	0.699	0.151
C5	0.3	1.3	0.366	0.115	C4	0.15	1.15	0.557	0.133	C1	0.1	1.1	0.636	0.137
C4	0.4	1.4	0.262	0.082	C3	0.3	1.3	0.429	0.102	C7	0.15	1.15	0.553	0.119
C3	0.2	1.2	0.218	0.068	C6	0.1	1.1	0.390	0.093	C3	0.1	1.1	0.503	0.108
C2	0.05	1.05	0.208	0.065	C7	0.1	1.1	0.354	0.084	C2	0.05	1.05	0.479	0.103
DM2	s_j	$\boldsymbol{k_j}$	$\boldsymbol{q_j}$	$\boldsymbol{w_j}$	**DM2**	s_j	$\boldsymbol{k_j}$	$\boldsymbol{q_j}$	$\boldsymbol{w_j}$	**DM2**	s_j	$\boldsymbol{k_j}$	$\boldsymbol{q_j}$	$\boldsymbol{w_j}$
C7		1	1	0.282	C5		1	1	0.205	C5		1	1	0.249
C6	0.3	1.3	0.769	0.217	C1	0.05	1.05	0.952	0.195	C4	0.4	1.4	0.714	0.178
C1	0.5	1.5	0.513	0.145	C4	0.2	1.2	0.794	0.163	C6	0.1	1.1	0.649	0.162
C2	0.3	1.3	0.394	0.111	C2	0.2	1.2	0.661	0.136	C1	0.3	1.3	0.500	0.124
C3	0.3	1.3	0.303	0.086	C3	0.25	1.25	0.529	0.109	C7	0.2	1.2	0.416	0.104
C5	0.05	1.05	0.289	0.082	C7	0.1	1.1	0.481	0.099	C3	0.1	1.1	0.378	0.094
C4	0.05	1.05	0.275	0.078	C6	0.05	1.05	0.458	0.094	C2	0.05	1.05	0.360	0.090
DM3	s_j	$\boldsymbol{k_j}$	$\boldsymbol{q_j}$	$\boldsymbol{w_j}$	**DM3**	s_j	$\boldsymbol{k_j}$	$\boldsymbol{q_j}$	$\boldsymbol{w_j}$	**DM3**	s_j	$\boldsymbol{k_j}$	$\boldsymbol{q_j}$	$\boldsymbol{w_j}$
C6		1	1	0.278	C1		1	1	0.227	C6		1	1	0.285
C7	0.45	1.45	0.69	0.192	C5	0.1	1.1	0.909	0.206	C5	0.5	1.5	0.667	0.190
C1	0.4	1.4	0.493	0.137	C2	0.35	1.35	0.673	0.153	C1	0.4	1.4	0.476	0.136
C4	0.15	1.15	0.428	0.119	C4	0.2	1.2	0.561	0.127	C4	0.3	1.3	0.366	0.104
C5	0.2	1.2	0.357	0.099	C3	0.25	1.25	0.449	0.102	C2	0.05	1.05	0.349	0.099
C2	0.1	1.1	0.325	0.090	C6	0.05	1.05	0.428	0.097	C7	0.05	1.05	0.332	0.095
C3	0.05	1.05	0.309	0.086	C7	0.1	1.1	0.389	0.088	C3	0.05	1.05	0.316	0.090

Table 3 Final weights of the TTO activities with respect to each group of the stakeholder

UTTO activity	Final weight with respect to "Students"	Final weight with respect to "Academicians"	Final weight with respect to "Industry"
Startup opportunities	0.269	0.216	0.221
Talks, seminars, trainings and events on entrepreneurship	0.208	0.207	0.197
Support for writing R&D project proposals to get funds	0.163	0.147	0.149
Providing information on financial support and funding opportunities	0.098	0.141	0.132
Support for bringing together researchers and industry, partner search	0.093	0.104	0.106
Commercialization of R&D results (patenting, licensing, intellectual property)	0.089	0.095	0.098
Trainings and seminars on technology transfer	0.080	0.090	0.097

talks, seminars, trainings, *and events on entrepreneurship* by students. Looking at the most important two activities, one can conclude that students are interested in entrepreneurship activities of UTTO, while academicians mainly focus on research projects for funding and industry is in between.

S*upport for writing R&D project proposals to get funds,* which was ranked second by the academicians, is ranked third by students, and fourth by industry. Third order is given to *commercialization of R&D results (patenting, licensing, intellectual property)* by the academicians, and to *support for bringing together researchers and industry, partner search* by the industry; which is also ranked fourth by academicians. *Providing information on financial support and funding opportunities,* which was perceived as most important by academicians and second most important by the industry; is given fourth order by the students.

Academicians rank *trainings and seminars on technology transfer* and *startup opportunities* as the fifth and sixth order. While students give the fifth and sixth order to support for *bringing together researchers and industry, partner search* and *commercialization of R&D results (patenting, licensing, intellectual property).* Industry perceives *talks, seminars, trainings, and events on entrepreneurship* and *trainings and seminars on technology transfer* as the fifth and sixth important activities of UTTO. The least important activity is seen as *talks, seminars, trainings, and events on entrepreneurship* by the academicians, *trainings, and seminars on technology transfer* for students; and *commercialization of R&D results (patenting, licensing, intellectual property)* for industry. These results show that academicians are not very interested in entrepreneurship and startup activities. On the other hand, students and industry are not very interested in technology transfer and intellectual property activities.

6 Conclusions

University Technology Transfer Offices promote the conversion of knowledge created at university to the marketable goods; which makes them important for the countries. UTTOs create a link between the companies and universities, and they are stated as strategic bodies and motivators for the university–industry research cooperation.

In this study, we try to shed a light on technology transfer literature and the technology transfer offices of universities. The main idea in this chapter is to understand the services provided by the UTTOs and the importance given to these services by the stakeholders of UTTOs. The stakeholders of UTTOs are industry, academicians, and students as known. Therefore, their opinions about the importance of the UTTO services should be taken by surveys and analyzed by a scientific method. Multi-Criteria Decision-Making methods are suitable for these kinds of problems, since they help us to determine the subjective priorities of the criteria (the activities performed by the UTTOs).

In this study, Stepwise Weight Assessment Ratio Analysis (SWARA) method is used as an MCDM method. Stakeholders' (students, academicians, and industry in this study) viewpoints are the major determinants of this method. To summarize the method, participants from each group of the stakeholder rank the services of UTTO from the first to last, then makes a pairwise comparison of these services. After taking the subjective evaluations of each decision maker, the overall ranks of the decision model are determined by using the mediocre value of ranks (Kersuliene & Turskis, 2011).

To the best of our knowledge, this study is the first study that analyzes the priority/importance of the critical UTTO activities through surveys conducted with participants from each group of the stakeholder; student, academician, and industry. Each group of the stakeholder of the UTTOs perceives some services of the UTTO more important than others. Although we have these inferences, there is a need to understand the true perceptions of the stakeholders of an UTTO through an empirical study. The opinions of the stakeholders should be taken by a survey and should be evaluated by a Multiple Criteria Decision-Making (MCDM) method. Such a study will be beneficial for both these bodies and policy makers. Policy makers will use the outputs of this study while planning the services of UTTOs.

In this study, we firstly performed a literature review and conducted interviews with UTTO professionals to determine the list of main activities/services offered by the UTTOs and they are listed below:

- Support for writing R&D project proposals to get funds
- Commercialization of R&D results (patenting, licensing, intellectual property)

- Trainings and seminars on technology transfer
- Support for bringing together researchers and industry, partner search
- Providing information on financial support and funding opportunities
- Startup opportunities
- Talks, seminars, trainings, and events on entrepreneurship

The results of the surveys that include the rankings of these activities (services) based on pairwise comparisons are analyzed by the SWARA method. After necessary calculations, the final priorities of each activity for each group of the stakeholder are determined. Results show that students are mostly interested in entrepreneurship activities of UTTO, while academicians focus on research projects for funding and industry is in between. These results are in fact parallel with the literature. Looking at the least important activities, it is seen that, academicians are not very interested in entrepreneurship and startup activities. On the other hand, students and industry are not very interested in technology transfer and intellectual property activities.

The results of this study will provide a pathway to decision makers and policy makers working in the area of university technology transfer. The UTTO administration, government, and universities should consider the findings of this study while planning their services. For instance, if they want to focus on academicians, they should increase services on *providing information on financial support and funding opportunities,* if they want to target students or industry, they should increase *startup opportunities.* Similarly, companies and students will prefer to work with UTTOs which are good at startup opportunities and academicians will prefer to work with UTTOs, which focus on providing information on financial support and funding opportunities.

Based on the numerical application presented in this study, it can also be concluded that SWARA method is simple, easy to use, adaptable, and applicable in this area and can be applied by UTTOs effectively. Hence, it can be safely used in similar applications to rank criteria and obtain weight values based on evaluations of decision makers.

As a direction of future research, other MCDM methods can also be used to prioritize UTTO activities and obtained findings can be compared with the data in this study. Moreover, the scope of the survey can be extended and university administrators (president, rector, dean, etc.) can participate in the study since their judgements may be different than academicians.

References

Aghdaie, M. H., Hashemkhani Zolfani, S., & Zavadskas, E. K. (2014). Sales branches performance evaluation: a multiple attribute decision making approach, 8th International Scientific Conference on Business and Management. *Vilnius.*https://doi.org/10.3846/bm.2014.001

Ansari, Z. N., Kant, R., & Shankar, R. (2020). Evaluation and ranking of solutions to mitigate sustainable remanufacturing supply chain risks: a hybrid fuzzy SWARA-fuzzy COPRAS framework approach. *International Journal of Sustainable Engineering, 13*(6), 473–494. https://doi.org/10.1080/19397038.2020.1758973

Antonelli, C. (2008). The new economics of the university: A knowledge governance approach. *The Journal of Technology Transfer, 33*(1), 1–22. https://doi.org/10.1007/s10961-007-9064-9

Baglieri, D., Baldi, F., & Tucci, C. L. (2018). University technology transfer office business models: One size does not fit all. *Technovation, 76*, 51–63. https://doi.org/10.1016/j.technovation.2018.05.003

Battaglia, D., Landoni, P., & Rizzitelli, F. (2017). Organizational structures for external growth of University Technology Transfer Offices: An explorative analysis. *Technological Forecasting and Social Change, 123*, 45–56. https://doi.org/10.1016/j.techfore.2017.06.017

Bekkers, R., & Freitas, I. M. B. (2008). Analysing knowledge transfer channels between universities and industry: To what degree do sectors also matter? *Research Policy, 37*, 1837–1853. https://doi.org/10.1016/j.respol.2008.07.007

Boer, F. P. (1999). *The valuation of technology business and financial issues in R&D* (pp. 4–6). Wiley.

Chapple, W., Lockett, A., Siegel, D., & Wright, M. (2005). Assessing the relative performance of UK university technology transfer offices: parametric and non-parametric evidence. *Research Policy, 34*(3), 369–384. https://doi.org/10.1016/j.respol.2005.01.007

Dersen, O., & Yontar, E. (2020). SWARA-TOPSIS Yöntemi ile En Uygun Yenilenebilir Enerji Kaynağinin Belirlenmesi. *Endüstri Mühendisliği, 31*(3), 389–419. https://doi.org/10.46465/endustrimuhendisligi.798063

Eghbali-Zarch, M., Tavakkoli-Moghaddam, R., Esfahanian, F., Sepehri, M. M., Azaron, A. (2018). Pharmacological therapy selection of type 2 diabetes based on the SWARA and modified MULTIMOORA methods under a fuzzy environment. *Artif Intell Med, 87*, 20–33. https://doi.org/10.1016/j.artmed.2018.03.003

Ekinci, Y., & Uray, N. (2016). *The Role of Technology Transfer Offices In Innovation Success of Organizations: Perceptions of The Stakeholders, 16th Production Research Symposium*, 12-14 October, İstanbul, Turkey, 462-467 (ISBN: 978-605-320-530-2).

Etzkowitz, H., & Zhou, C. (2021). Licensing life: The evolution of Stanford university's technology transfer practice. *Technological Forecasting and Social Change, 168*, 120764.

Ghorshi Nezhad, M. R., Hashemkhani Zolfani, S., Moztarzadeh, F., Zavadskas, E. K., & Bahrami, M. (2015). Planning the priority of high tech industries based on SWARA-WASPAS methodology: The case of the nanotechnology industry in Iran. *Economic research-Ekonomska istraživanja, 28*(1), 1111–1137. https://doi.org/10.1080/1331677X.2015.1102404

Ginevicius, R. (2011). A new determining method for the criteria weights in multi-criteria evaluation. *International Journal of Information Technology & Decision Making, 10*, 1067–1095. https://doi.org/10.1142/S0219622011004713

Gül, A., & Ekinci, Y. (2015). *Awareness about the Services of University Technology Transfer Offices, Conference Proceeding ICHSS 2015*, 25-26 September, Rome, Italy, 84-88 (ISBN: 978-88-909163-6-6).

Hashemkhani, Z. S., & Saparauskas, J. (2013). New application of SWARA method in prioritizing sustainability assessment indicators of energy system. *Inzinerine Ekon Eng Econ, 24*(5), 408–414. https://doi.org/10.5755/j01.ee.24.5.4526

Hashemkhani Zolfani, S., Esfahani, M.H., Bitarafan,M., Zavadskas E.K & Arefi, S.L. (2013). Developing a new hybrid MCDM method for selection of the optimal alternative of mechanical longitudinal ventilation of tunnel pollutants during automobile accidents, *Transport,* 28, 89–96. https://doi.org/10.3846/16484142.2013.782567

Hashemkhani Zolfani, S., Farrokhzad, M., & Turskis, Z. (2013). Investigating on successful factors of online games based on explorer. *E & M Ekonomie a Management, 16*, 161–169.

Hewitt-Dundas, N. (2013). The role of proximity in university-business cooperation for innovation. *The Journal of Technology Transfer, 38*(2), 93–115. https://doi.org/10.1007/s10961-011-9229-4

Holgersson, M., & Aaboen, L. (2019). A literature review of intellectual property management in technology transfer offices: From appropriation to utilization. *Technology in Society, 59*, 101132.

Işık, A. T., & Adalı, E. A. (2016). A new integrated decision making approach based on SWARA and OCRA methods for the hotel selection problem. *International Journal of Advanced Operations Management, 8*(2), 140–151.

Juodagalviene, B., Turskis, Z., Saparauskas, J., & Endriukaityte, A. (2017). Integrated Multi-Criteria Evaluation of House's Plan Shape Based on the EDAS and SWARA Methods. *Engineering Structures and Technologies, 9*(3), 117.

Karabašević, D., Stanujkić, D., & Urošević, S. (2015). The MCDM Model for Personnel Selection Based on SWARA and ARAS Methods. *Management*, (1820-0222), 20(77). https://doi.org/10.7595/management.fon.2015.0029

Keršulienė, V., & Turskis, Z. (2011). Integrated fuzzy multiple criteria decision making model for architect selection. *Technol. Econ. Dev. Econ., 17*, 645–666. https://doi.org/10.3846/20294913.2011.635718

Keršulienė, V., Zavadskas, E. K., & Turskis, Z. (2010). Selection of rational dispute resolution method by applying new step-wise weight assessment ratio analysis (SWARA). *J Bus Econ Manag, 11*(2), 243–258. https://doi.org/10.3846/jbem.2010.12

Lee, K., & Jung, H. J. (2021). Does TTO capability matter in commercializing university technology? Evidence from longitudinal data in South Korea. *Research Policy, 50*(1), 104133. https://doi.org/10.1016/j.respol.2020.104133

Lee, S., Kim, W., Kim, Y. M., & Oh, K. J. (2012). Using AHP to determine intangible priority factors for technology transfer adoption. *Expert Systems with Applications, 39*(7), 6388–6395. https://doi.org/10.1016/j.eswa.2011.12.030

Link, A. N., Siegel, D. S., & Bozeman, B. (2007). An empirical analysis of the propensity of academics to engage in informal university technology transfer. *Industrial and Corporate Change, 16*(4), 641–655. https://doi.org/10.1093/icc/dtm020

Macho-Stadler, I., Pérez-Castrillo, D., & Veugelers, R. (2007). Licensing of university inventions: The role of a technology transfer office. *International Journal of Industrial Organization, 25*(3), 483–510. https://doi.org/10.1016/j.ijindorg.2006.06.001

Mardani, A., Nilashi, M., Zakuan, N., Loganathan, N., Soheilirad, S., Saman, M. Z. M., & Ibrahim, O. (2017). A systematic review and meta-Analysis of SWARA and WASPAS methods: Theory and applications with recent fuzzy developments. *Applied Soft Computing, 57*, 265–292. https://doi.org/10.1016/j.asoc.2017.03.045

Markman, G. D., Phan, P. H., Balkin, D. B., & Gianiodis, P. T. (2005). Entrepreneurship and university-based technology transfer. *Journal of Business Venturing, 20*(2), 241–263. https://doi.org/10.1016/j.jbusvent.2003.12.003

Munari, F., Sobrero, M., & Toschi, L. (2018). The university as a venture capitalist? Gap funding instruments for technology transfer. *Technological Forecasting and Social Change, 127*, 70–84.

O'kane, C., Mangematin, V., Geoghegan, W., & Fitzgerald, C. (2015). University technology transfer offices: The search for identity to build legitimacy. *Research Policy, 44*(2), 421–437. https://doi.org/10.1016/j.respol.2014.08.003

O'Kane, C. (2018). Technology transfer executives' backwards integration: An examination of interactions between university technology transfer executives and principal investigators. *Technovation, 76*, 64–77.

Olaya-Escobar, E. S., Berbegal-Mirabent, J., & Alegre, I. (2020). Exploring the relationship between service quality of technology transfer offices and researchers' patenting activity. *Technological Forecasting and Social Change, 157*, 120097.

Popović, G. (2018). Prioritization of Strategies for Tourism Development by Applying a SWOT-SWARA Analysis: The Case of Sokobanja Spa. *Teme-Časopis za Društvene Nauke, 42*(3), 999–1016. https://doi.org/10.22190/TEME1803999P

Prajapati, H., Kant, R., & Shankar, R. (2019). Prioritizing the solutions of reverse logistics implementation to mitigate its barriers: A hybrid modified SWARA and WASPAS approach. *Journal of Cleaner Production, 240*, 118219.

Saaty, L. T. (1980). *The analytic hierarchy process*. McGraw Hill Company.
Saaty, L. T., & Vargas, L. G. (2001). *Models, methods, concepts & applications of the analytical hierarchy process*. Kluwer Academic Publishers.
Sart, G. (2014). Effects of Technology Transfer Offices on Capacity Building in Creativity. *Innovation and Entrepreneurship*.https://doi.org/10.13140/2.1.2269.8244
Shannon, C. E. (1948). The mathematical theory of communication. *Bell System Technical Journal, 27*, 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x
Sušinskas, S., Zavadskas, E. K., & Turskis, Z. (2011). Multiple criteria assessment of pilecolumns alternatives. *The Baltic Journal of Road and Bridge Engineering, 6*, 77–83.
Taghavi, F., Antucheviciene, J., & Yaghobian, S. A. (2019). Assessment of Universities' and Higher Education Centers' Preparedness for Successful Establishment of Enterprise Resource Planning Based on SWARA Method. *Engineering Economics, 30*(4), 496–506.
Tunca, F., & Kanat, Ö. N. (2019). Harmonization and Simplification Roles of Technology Transfer Offices for Effective University–Industry Collaboration Models. *Procedia Computer Science, 158*, 361–365.
van Burg, E., Du, J., & Kers, J. G. (2021). When do academics patent outside their university? An in-depth case study. *Technovation, 107*, 102287.
Vesković, S., Stević, Željko, Stojić, G., Vasiljević, M., & Milinković, S. (2018). Evaluation of the railway management model by using a new integrated model DELPHI-SWARA-MABAC. Decision Making: *Applications in Management and Engineering, 1*(2), 34–50. https://doi.org/10.31181/dmame1802034v
Yücenur, G. N., & Ipekçi, A. (2021). SWARA/WASPAS methods for a marine current energy plant location selection problem. *Renewable Energy, 163*, 1287–1298.
Wittamore, K., Bahns, R., Brown, A., Carter, P., Clements, G., & Young, C. (1998). International technology transfer – A developing empirical model, management of technology, sustainable development and ecoefficiency. In *The seventh international conference on management of technology* (pp. 16–20). February.
Twi-Global (2021). What is Technology Transfer? https://twi-global.com

Picture Fuzzy Extension of DEMATEL and its Usage in Educational Quality Evaluation

Sait Gül

Abstract Multiple attribute decision-making (MADM) models are strength tools for assessing alternatives when more than one attribute should be considered. The experts consulted need scales to state their judgments and preferences. The contemporary versions of fuzzy sets provide various kinds of scales to let the experts specify different characteristics of their ideas. The concept of picture fuzzy sets (PFS) includes four elements. The first three of them can be independently assignable: the positive, the neutral, and the negative membership degrees. The only constraint is that their sum must not exceed 1. The difference between 1 and the sum of the assigned degrees is called refusal degree and it represents the expert's choice of refusing the sharing of the preference. PFS has greater representation strength than other fuzzy extensions have since it exposes an additional fourth component, refusal degree. In this study, picture fuzzy version of DEMATEL (DEcision-MAking Trial and Evaluation Laboratory), which is developed for revealing the potential causal dependencies among the multiple attributes, is introduced for considering the hesitancy and refusal degrees of experts. In this new version, no constant linguistic evaluation scale is proposed. Instead, the evaluations of experts are grouped into four categories: YES, NO, ABSTAIN, and REFUSE. The strength of the influence between two attributes is determined according to the aggregation of these statements. The proposition is applied to a real case regarding an evaluation of the educational quality and a comparative analysis including different applications is conducted to validate the method.

Keywords Picture fuzzy sets · DEMATEL · Linguistic assessment · Network relation map

---

S. Gül (✉)
Management Engineering Department, Bahçeşehir University, Faculty of Engineering and Natural Sciences, İstanbul, Turkey
e-mail: sait.gul@eng.bau.edu.tr

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_18

1 Introduction

As one of the most applied multiple attribute decision-making (MADM) tools, DEcision-MAking Trial and Evaluation Laboratory (DEMATEL) which was introduced by the Science and Human Affairs Program of the Battelle Memorial Institute of the Geneva Research Centre aims to figure out the associated contents of the decision problems (Fontela & Gabus, 1972). The method processes the expert assessments in revealing the influences or causal dependencies between the attributes. DEMATEL recognizes the influences between attributes and appoints them to either cause or effect group. While cause group covers attributes largely impacting the other attributes, effect group attributes are accepted as being impacted by the others (Wang et al., 2012). The implication of this finding is the prioritization of attributes in terms of enhancement effort scheduled for the groups which means that any enhancement in cause-type attributes can potentially generate an enhancement in effect-type attributes (Yazdi et al., 2020). DEMATEL approach also calculates the subjective weights of attributes by considering their central roles played in the decision process (Cebi, 2013). These weights represent the importance of the attributes for the decision problem at hand and can be used in hybrid MADM applications. For instance, after collecting the performance score of the different courses of actions (alternatives) with respect to the attributes, these scores and the attribute weights obtained by DEMATEL can be processed together to find the most appropriate or best course of action. This finding is declared as the decision proposition of the MADM method utilized.

Most MADM problems often lack a proper and easily quantifiable data set consisting of cardinal measures. The experts who are consulted favor linguistic terms in stating their evaluations. Various fuzzy concepts provide linguistic evaluation scales possessing appropriate fuzzy number correspondences. Zadeh (1965) is the pioneer of the field of fuzzy sets in illustrating human judgments. In his classical definition of ordinary fuzzy sets, the evaluations are shown by a sole membership degree (μ) ranging between 0 and 1. The membership degree has a positive perspective and represents the agreement level of the expert to a linguistic term. To enrich the representation of the uncertainty and vagueness in human judgments, Atanassov (1986) emerged the concept of intuitionistic fuzzy sets (IFS) via adding a new element to the fuzzy set definition: a negative membership (or non-membership) degree (v). Since the experts now are capable of indicating their disagreement opinions, IFSs deliver flexibility to the uncertainty depiction of human evaluations. Therefore, nonmembership degree has a negative perspective. In IFS, there is a novel element called indeterminacy degree defining the experts' neutral opinion: $\pi = 1 - \mu - v$. Consequently, IFS was the first three-dimensional fuzzy set concept that can exhibit human judgments more thoroughly. However, the limitation of IFS is that the indeterminacy degree is not independently appointable by the experts.

During the decades after the first appearance of fuzzy sets, novel fuzzy set definitions have been developed in the literature. Pythagorean fuzzy sets (Yager, 2013), q-Rung orthopair fuzzy sets (Yager, 2017), and Fermatean fuzzy sets (Senapati

& Yager, 2020) extended the representation domain of the human judgments but they just take the separately assignable positive and negative membership degrees into account. An independent hesitancy degree is considered in neutrosophic sets (Smarandache, 1999) and spherical fuzzy sets (Kutlu Gündoğdu & Kahraman, 2019). Considering the mentioned and more fuzzy concepts such as hesitant fuzzy sets (Torra, 2010) or dense fuzzy locksets (De, 2017), none have handled the refusal preference of the expert until the picture fuzzy sets (PFS) concept emerged.

PFS concept which recognizes the three independent membership degrees were introduced by Cuong and Kreinovich (2013) as a generalization of fuzzy sets and IFSs. Along with the positive (μ), neutral (η), and negative (v) membership degrees, PFS adds a new item called refusal degree which can be obtained by $\pi = 1 - \mu - \eta - v$ into the fuzzy set definition. Hence, PFS imposes the sum of these four elements in the set definition to be equal to 1. The novelty provided by PFS is the consideration and representation of the expert's choice of refusing the idea sharing. Voting environments are good examples for illustrating PFSs. Cuong (2014) propounded that each voter is a nominee for being included by one of the four groups identifying the voting behavior: one can vote for the candidate (positive membership), one can abstain to vote (hesitancy/neutral membership), one can vote against the candidate (negative membership), and one can refuse to vote for any reason (refusal degree). Thus, a PFS is a fuzzy set definition possessing an extensive representation power as it is the only fuzzy set definition that can handle the refusal opinions among all the fuzzy set definitions.

As explained in the literature review section, DEMATEL has been extended into various fuzzy environments. In terms of three-dimensional fuzzy sets, there are many neutrosophic set versions and few spherical fuzzy set versions of it. These current versions are not able to handle both the hesitancy and the refusal opinions of the experts. So, this study aims to develop a picture fuzzy numbers (PFN) oriented DEMATEL approach (PF-DEMATEL) to deal with this issue. In this manner, DEMATEL has been become strengthened for comprehending many aspects of the expert. Consequently, we aim to make DEMATEL a more intelligent structural attribute modeling approach. For this purpose, a new data gathering process is proposed with the aim of easing the data collection and providing more freedom to the experts in expressing their judgments.

Educational quality is one of the most important issues for the futures of our societies, countries, and even the world. During the emergency occurrences such as today's Pandemic conditions, educational activities need some modifications in order to be prevented from the potential negative effects of these bad conditions. For this purpose, we formed a committee consisting of four educational experts for initiating a discussion about the educational quality evaluations of higher educational institutions. The data collection process was designed in accordance with the rules of the voting environment and the data collected is aggregated as PFNs. Then, the applicability of the proposed PF-DEMATEL is shown in revealing the influences among the attributes affecting the educational quality and their weights. The results are compared with the results of traditional DEMATEL and

IF-DEMATEL (IFS-based DEMATEL version) and it is shown that the differences between results emerge from the fuzzy set concepts on which the methods are based.

The chapter is organized as follows. A literature review focusing on the three-dimensional fuzzy set extensions of DEMATEL is summarized in Sect. 2. While Sect. 3 explains the definitions and operations on PFS, the proposed PF-DEMATEL proposition is detailed in Sect. 4. The method's applicability and validity are shown in an application of evaluating the educational quality of higher educational institutions in Sect. 5 along with the comparison of the results with DEMATEL and IF-DEMATEL. Section 5 also includes the discussion on the findings of PF-DEMATEL. Section 6 concluded the study with limitations, remarks, contributions, and future works.

2 Literature Review on 3D Extensions of DEMATEL

The Battelle Memorial Institute of the Geneva Research Centre established DEcision-Making Trial and Evaluation Laboratory (DEMATEL) tool intending to figure out solutions for complex problems involving interrelating attributes. As a MADM method, DEMATEL considers the preferences and expertise of decision makers as the data required to report the causal influences among the attributes of the decision problem at hand (Gabus & Fontela, 1972, 1973).

DEMATEL is a recognized MADM approach that can be benefitted in determining the cause and effect relationships among attributes. The DEMATEL literature is fruitful with applications in real-life problems. Kabak et al. (2016) evaluated 27 critical success factors of the iron and steel industry in Turkey through strengthening fuzzy DEMATEL with Delphi and a web-based survey that is conducted for gathering the preferences of 36 experts. Govindan et al. (2020) integrated fuzzy ANP (analytic network process), fuzzy DEMATEL, and multi-objective mixed-integer linear programming model for circular supplier selection problem and order allocation in a multiproduct circular closed-loop supply chain. Gül et al. (2020) used DEMATEL to reveal the interrelations among two well-known occupational health and safety procedures of the shipping industry with the aim of building safer and reliable working conditions in vessels. Kim and Nguyen (2021) first conducted a two-round Delphi to understand the attributes affecting the international development project decisions and then DEMATEL reveals the complexities between these attributes. Kouhizadeh et al. (2021) utilized DEMATEL to explore adoption barriers between blockchain technology and sustainable supply chains by aggregating the ideas of 12 practitioners and 35 academics. The interested readers can see Si et al. (2018) presenting the state-of-the-art of DEMATEL applications and Gül (2020) introducing the literature on contemporary fuzzy extensions of the method, such as type-2, intuitionistic, hesitant, and Pythagorean fuzzy sets, and neutrosophic sets.

In the literature review, we will focus on DEMATEL modifications which are extensions under three-dimensional (3D) fuzzy environments including three independently assignable membership degrees: positive, negative, and hesitancy degrees. Intuitionistic, q-Rung orthopair, Pythagorean, and Fermatean fuzzy sets are not included here since their hesitancy degree consideration is based on the positive and negative membership degree determination.

Table 1 shows the results of the literature review on 3D fuzzy extensions of DEMATEL. There are 15 papers in the SCOPUS platform on this specific field of research. These studies were published after 2017 and 8 of them (53%) were published in 2020 and 2021. It is obvious that this research theme is under consideration by the present soft computing researchers today. While 2 of the studies modified DEMATEL with SFSs, 11 papers (73%) developed single-valued NS (SVNS) extensions and 2 studies approached DEMATEL from the view of interval-valued NS (IVNS). So, it is seen that the literature has studies coping with the hesitant or indeterminant ideas of the experts as well as their positive and negative opinions. Eight papers integrated their DEMATEL extensions with different MADM methods (4 applications of TOPSIS, 2 applications of ANP, 1 application of AHP, etc.). All the modified DEMATEL versions are implemented in real cases such as evaluation of e-commerce websites, supplier selection, project selection, coastal erosion factors, and comparison of municipalities.

The 3D DEMATEL versions lack conveniently fuzzified steps that are dedicated to keeping the computations fuzzy until the very end of the process: 10 articles (67%) utilized defuzzification or deneutrosophication of the decision tables in early steps. It is an understandable and empathized selection since the matrix operations (especially, taking the inverse of the direct influence matrix of DEMATEL) are not handled in fuzzy environments easily.

In the study, our main contribution is the proposition of a picture fuzzy version of DEMATEL (PF-DEMATEL) for the first time in the literature because no study handles the problems that can be dealt with by PFS, namely the refusal preference of the expert. Current DEMATEL versions such as NS and SFS consider the positive, negative, and hesitant characteristics of the expert opinions but they are not capable of dealing with the refusal preferences which are the situations the experts do not give feedback or share their ideas or cast a vote in a voting environment. This refusal degree consideration is the main peculiarity of the study. After the definitions of and operations defined on PFSs are given in the next chapter, the proposed method of PF-DEMATEL is explained in detail.

3 Preliminaries: Picture Fuzzy Sets

Cuong and Kreinovich (2013) presented the PFSs as a generalization of IFS theory to model complicated and ambiguous evaluations of decision makers in MADM problems (Meksavang et al., 2019). There are three separately identifiable explanatory elements in a PFS: positive membership degree (μ), neutral membership degree (η), and negative membership degree (ν). Since the definition of PFS allows these three elements' sum to be smaller than 1, it renames the difference between 1 and the aforementioned sum as refusal degree (π), which exhibits the novel aspect of the

Table 1 3D fuzzy extensions of DEMATEL

Paper	Type of FS	Integrated methods	Application	Aspect
Liang et al. (2017)	SVNS	–	Evaluation of e-commerce websites	Early deneutro-sophication
Abdel-Basset et al. (2018)	SVNS	–	Developing supplier selection criteria	Early deneutro-sophication
Abdel-Basset et al. (2018)	SVNS	AHP	Deriving a security estimation framework of smart supply chain management	Early deneutro-sophication
Awang et al. (2018)	SVNS	–	Investigating factors of coastal erosion	Early deneutro-sophication
Yang and Pang (2018)	IVNS	TOPSIS	Truck selection for the logistics industry	Operations for each element in the NS
Abdel-Basset et al. (2019)	SVNS	TOPSIS	Project selection	Early deneutro-sophication
Awang et al. (2019)	SVNS	ANP	Investigating factors of coastal erosion	Operations for each element in the NS
Abdel-Basset et al. (2020)	SVNS	CRITIC	Evaluation of innovation value proposition for smart systems	Early deneutro-sophication
Al-Quran et al. (2020)	IVNS	–	Hospital service quality evaluation	Early deneutro-sophication
Awang et al. (2020)	SVNS	–	Investigating factors of coastal erosion	The multiplicative inverse of the matrix consideration
Tan and Zhang (2020)	SVNS	Linear assign-ment model	Typhoon disaster assessment	Early deneutro-sophication
Kilic and Yalcin (2020)	SVNS	TOPSIS	Comparison of municipalities in terms of environmental sustainability	Operations for each element in the NS
Nabeeh et al. (2021)	SVNS	GRA, ANP, TOPSIS	Evaluation of green credit rating for the manufacturing industry	Early deneutro-sophication
Gül (2020)	SFS	–	Building contractor selection	Operations for each element in the SFS
Xie et al. (2020)	SFS	–	Determining the ways to improve the public acceptance for nuclear energy investment in energy importing countries	Early defuzzification

Source: Author

PFS definition. These four elements implicate the potential vote types such as yes, no, abstain, and refuse. As detailed in the next step, this voting based understanding of PFS is the backbone of the data gathering approach included in the proposition.

Let X be a universal set. Then a PFS P on X is defined as follows:

$$P = \left\{\left(x, \mu_P(x), \eta_P(x), \nu_P(x)\right) \mid x \in X\right\} \tag{1}$$

where $\mu_P(x)$ is called positive, $\eta_P(x)$ is neutral, and $\nu_A(x)$ is negative membership degrees of $x \in A$. μ_P, η_P, ν_P must satisfy the conditions: $0 \le \mu_P(x), \eta_P(x), \nu_P(x) \le 1$ and $0 \le \mu_P(x) + \eta_P(x) + \nu_P(x) \le 1$. $\pi_P(x) = 1 - \mu_P(x) - \eta_P(x) - \nu_P(x)$ is called refusal membership degree of x in A (Cuong, 2014).

The subsethood, equality, union, intersection, and complement for every two PFSs P and S are defined by Cuong (2014) as follows:

1. $P \subseteq S$ if $\forall x \in X, \mu_P(x) \le \mu_S(x),\ \eta_P(x) \le \eta_S(x), \nu_P(x) \ge \nu_S(x)$;
2. $P = S$ iff $P \subseteq S$ and $S \subseteq P$;
3. $P \cup S = \{\langle x, \max(\mu_P(x), \mu_S(x)), \min(\eta_P(x), \eta_S(x)), \min(\nu_P(x), \nu_S(x))\rangle \mid x \in X\}$
4. $P \cap S = \{\langle x, \min(\mu_P(x), \mu_S(x)), \min(\eta_P(x), \eta_S(x)), \max(\nu_P(x), \nu_S(x))\rangle \mid x \in X\}$
5. $P^c = \{\langle x, \nu_P(x), \eta_P(x), \mu_P(x)\rangle \mid x \in X\}$.

Let $P = (\mu_P, \eta_P, \nu_P)$ and $S = (\mu_S, \eta_S, \nu_S)$ are two PFNs. Their operations are as follows (Wang et al., 2017):

$$P \oplus S = \left(1 - (1 - \mu_P)\left(1 - \mu_S\right), \eta_P \eta_S, \left(\eta_P + \nu_P\right)\left(\eta_S + \nu_S\right) - \eta_P \eta_S\right) \tag{2}$$

$$P \otimes S = \left(\left(\mu_P + \eta_P\right)\left(\mu_S + \eta_S\right) - \eta_P \eta_S\right), \eta_P \eta_S, 1 - (1 - \nu_P)(1 - \nu_S)\right) \tag{3}$$

$$\tau P = \left(1 - (1 - \mu_P)^\tau, \left(\eta_P\right)^\tau, \left(\eta_P + \nu_P\right)^\tau - \left(\eta_P\right)^\tau\right) \tag{4}$$

$$P^\tau = \left(\left(\mu_P + \eta_P\right)^\tau - \left(\eta_P\right)^\tau, \left(\eta_P\right)^\tau, 1 - (1 - \nu_P)^\tau\right) \tag{5}$$

For comparing PFSs, the algorithm given below can be utilized where score function is $sc(P) = \mu_P - \nu_P$ and accuracy function is $ac(P) = \mu_P + \eta_P + \nu_P$ (Wang et al., 2017):

(i) If $sc(P) > sc(S)$, then $P > S$ means P is superior to S.
(ii) If $sc(P) = sc(S)$, then

   1. $ac(P) > ac(S)$ implies that $P > S$ which means P is superior to S.
   2. $ac(P) = ac(S)$ implies that $P = S$ which means P is equivalent to S.

The score function definition of Wang et al. (2017) does not consider the hesitancy/indeterminacy degree of the expert. A more convenient and comprehensive score function definition is given by Xu et al. (2019). For obtaining the crisp value of a PFS P, the two-step defuzzification procedure can be used as follows:

Step 1. Distribute the hesitancy degree to the positive and negative membership degrees: $\mu'_P = \mu_P + \frac{\eta_P}{2}$ and $\nu'_P = \nu_P + \frac{\eta_P}{2}$.

Step 2. Calculate the defuzzified value $sc(P)$ by $sc(P) = \mu'_P + \frac{\pi_P}{2}\left(1 + \mu'_P - \nu'_P\right)$.

After a reorganization, the two-step defuzzification approach can be reduced to Eq. (6).

$$sc(P) = \mu_P + \frac{\eta_P}{2} + \frac{1 - \mu_P - \eta_P - \nu_P}{2}(1 + \mu_P - \nu_P) \tag{6}$$

Non-compensatory MADM methods such as TOPSIS, VIKOR, EDAS, and CODAS require distance measure definitions because their common decision point is the distances between alternatives and the ideal and/or nadir solutions. For this purpose, Euclidean ($d_E(P, S)$) and Hamming ($d_H(P, S)$) distances are developed as follows (Cuong & Kreinovich, 2013):

$$d_E(P, S) = \left[\frac{1}{n}\sum_{i=1}^{n}\left(\left(\mu_P - \mu_S\right)^2 + \left(\eta_P - \eta_S\right)^2 + \left(\nu_P - \nu_S\right)^2\right)\right]^{\frac{1}{2}} \tag{7}$$

$$d_H(P, S) = \frac{1}{n}\sum_{i=1}^{n}\left(\left|\mu_P - \mu_S\right| + \left|\eta_P - \eta_S\right| + \left|\nu_P - \nu_S\right|\right) \tag{8}$$

4 Proposed PF-DEMATEL

DEMATEL is a MADM tool that uses the assessments of expert opinions to reveal the hidden causal dependencies among the attributes. In the study, we aim to develop a picture fuzzy version of DEMATEL to allow the decision analyst to handle decision situations having various types of characteristics of the expert judgments, such as positive and negative thoughts as well as hesitant and refusal preferences. The steps are as follows:

Step 1. Construction of the decision model.

Suppose m experts who are capable of sharing their expertise to solve the decision problem and n attributes that can potentially affect the solution are considered ($e = 1, \ldots, m$; $i, j = 1, \ldots, n$; $i \neq j$). Since the main aim of DEMATEL is to reveal the influences among the attributes, each expert is asked to give his/her idea regarding the potential influence between each pair of attributes. For this purpose, the experts need a scale showing the linguistic terms that are used in indicating their judgments. These linguistic terms in PF environment are YES, NO, ABSTAIN, and REFUSE because these are the four potential characteristics of an expert judgment that should be modeled. Hence, a voting procedure is used here. As a novelty, we propose gradation in specifying the positive/agreement idea that is represented by YES. Examples of votes are given as follows:

- If an expert is sure that attribute i has a complete influence on attribute j, he/she can state this idea by giving an agreement vote of YES with a membership degree of 1.00, i.e., YES (1.00).
- Assume an expert determines a partial influence, e.g., 0.6, of attribute i over attribute j. The agreement level of this opinion is specified by 0.6 so that the evaluation is represented by YES (0.6).
- Another expert can see no influence between the attributes. Thus, his/her answer will be NO which shows the disagreement level regarding the comparison.
- If the expert is not sure about stating an idea, he/she can stay hesitant and give a vote of ABSTAIN which shows the hesitant or indeterminant characteristic of the idea about the comparison.
- Last, an expert can say that he/she refuses to vote without explaining the reason behind this decision. The corresponding evaluation will be REFUSE which represents the refusal characteristic of the expert about the comparison.

Step 2. Construction of the direct PF influence matrix.

The influence of attribute i over attribute j is obtained from the experts and finally, there are m different evaluation matrices that have a size of $n*n$. These matrices should be aggregated to initiate the DEMATEL process. Joshi (2020a, 2020b) proposed a statistical methodology for aggregation of the expert evaluations which are determined from a voting perspective, but this process cannot deal with the newly proposed gradation in the agreement levels of the experts. Arya and Kumar (2020) and Jovcic et al. (2020) used the same method and they do not make this consideration too. For aggregation purposes, we improved the statistical method developed by Joshi (2020a, 2020b).

An evaluation matrix shows the individual's pairwise comparisons of the attributes in terms of the influences between them. The aggregated matrix is called the direct PF influence matrix ($\tilde{X}$) and shown as given in Eq. (9).

$$\tilde{X} = \begin{bmatrix} - & \langle \mu_{12}, \eta_{12}, v_{12} \rangle & \langle \mu_{1(n-1)}, \eta_{1(n-1)}, v_{1(n-1)} \rangle & \langle \mu_{1n}, \eta_{1n}, v_{1n} \rangle \\ \langle \mu_{21}, \eta_{21}, v_{21} \rangle & - & \langle \mu_{2(n-1)}, \eta_{2(n-1)}, v_{2(n-1)} \rangle & \langle \mu_{2n}, \eta_{2n}, v_{2n} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mu_{n1}, \eta_{n1}, v_{n1} \rangle & \langle \mu_{n2}, \eta_{n2}, v_{n2} \rangle & \langle \mu_{n(n-1)}, \eta_{n(n-1)}, v_{n(n-1)} \rangle & - \end{bmatrix} \tag{9}$$

where $\langle \mu_{ij}, \eta_{ij}, v_{ij} \rangle$ shows the positive, neutral, and negative membership degrees showing the influence of attribute i over attribute j ($i \neq j$). Here μ, η, and v are the corresponding membership degrees of the linguistic terms YES, ABSTAIN, and NO. Thus, the refusal degree regarding the concerned influence (REFUSE) is

$\pi_{ij} = 1 - \mu_{ij} - \eta_{ij} - v_{ij}$. The elements of PFNs stated in Eq. (9) are calculated as follows:

$$\mu_{ij} = \frac{\sum_{e=1}^{m} \text{YES}_{ij}\left(\mu_{ij}^{e}\right)}{m} \tag{10}$$

$$v_{ij} = \frac{\delta_{\text{NO}_{ij}}}{m} \tag{11}$$

$$\eta_{ij} = \frac{\delta_{\text{ABSTAIN}_{ij}}}{m} + \frac{\frac{\delta_{\text{YES}_{ij}}}{m} - \mu_{ij}}{2} \tag{12}$$

$$\pi_{ij} = \frac{\delta_{\text{REFUSE}_{ij}}}{m} + \frac{\frac{\delta_{\text{YES}_{ij}}}{m} - \mu_{ij}}{2} \tag{13}$$

where δ_* shows the number of votes of YES (YES_{ij}), NO (NO_{ij}), ABSTAIN (ABSTAIN_{ij}), and REFUSE (REFUSE_{ij}); m is the number of experts; $\text{YES}_{ij}\left(\mu_{ij}^{e}\right)$ shows the gradation value in the positive idea (YES) of the expert e about the comparison of attributes i and j.

In Eqs. (12 and 13), there are two parts. The first part represents the average score of the votes saying ABSTAIN or REFUSE. The second part covers half the unassigned part which is remained from the answers of YES. After computing the average positive opinions of the experts, unassigned gradation is left behind. The most rational manner of keeping this important but unassigned information in computations is the distribution of it equally to the ABSTAIN and REFUSE degrees. There can be different distribution manners if the decision analyst has any evidence to validate it. For example, the unassigned part can be added either to the REFUSE or ABSTAIN levels. Examples of the distribution are given in Sect. 5.

Step 3. Construction of the direct influence matrix.

DEMATEL has some matrix operations such as taking the inverse of a matrix or matrix multiplications. When these matrix operations are separately performed, the fuzziness of the data is frequently disrupted. To avoid this data distortion, most of the researchers, e.g., Xie et al. (2020), Abdel-Basset et al. (2018, 2018, 2020), Al-Quran et al. (2020), Tan and Zhang (2020), prefer an early defuzzification as explained in the literature survey section. Thus, it seems the most appropriate way of protecting the method's fundamentals is the defuzzification of the direct influence PF evaluation matrix in this step.

The crisp influence of attribute i over attribute j is denoted by x_{ij} and the direct influence matrix is shown in Eq. (14) where the defuzzification is performed for each entry of $\tilde{X}$ via Eq. (15).

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{bmatrix} \quad (14)$$

$$x_{ij} = \mu_{ij} + \frac{\eta_{ij}}{2} + \frac{1 - \mu_{ij} - \eta_{ij} - \nu_{ij}}{2}\left(1 + \mu_{ij} - \nu_{ij}\right) \quad (15)$$

The remaining steps are identical to the traditional DEMATEL (Cebi, 2013; Hwang et al., 2016; Si et al., 2018).

Step 4. Establishment of the initial direct influence matrix.

X is here normalized and recalled as the initial direct influence matrix (Z). $Z = [z_{ij}]_{n*n}$ is formed in Eq. (16), in which all principal diagonal entries are zero. k is called the normalization index. Thus, $z_{ij} = \frac{x_{ij}}{k}$.

$$Z = \frac{X}{k} \text{where } k = \max\left[\max_i \sum_{j=1}^{n} x_{ij}, \max_j \sum_{i=1}^{n} x_{ij}\right] \quad (16)$$

Step 5. Computation of the total influence matrix.

The total influence can be computed by summing the direct influences (Z) and all the indirect influences (Z^2, Z^3, ..., Z^∞). The total indirect influence is the aggregation of the powers of the matrix Z. A convergent solution is found by the matrix inversion, similar to an absorbing Markov chain matrix.

$$T = Z + Z^2 + Z^3 + Z^4 + \cdots + Z^\infty = Z(I - Z)^{-1} \quad (17)$$

where $T = [t_{ij}]_{n*n}$ exhibits the total influences between each attribute pair.

Step 6. Computation of Prominence and Relation values.

DEMATEL calculates the vector of row sum (R) and the vector of column sum (C):

$$R = [r_i]_{n*1} = \left[\sum_{j=1}^{n} t_{ij}\right] \quad (18)$$

$$C = [c_j]_{1*n} = \left[\sum_{i=1}^{n} t_{ij}\right]' \quad (19)$$

While the vector of row sums (R) shows the total influences of attribute i over other attributes and indicates the attribute's general strength, the vector of column sum (C) represents the situations in which attribute j has been influenced by other attributes and indicates the general weakness of the attribute.

When $i = j$ and $i,j = 1, \ldots, n$, "$R + C$" which is called Prominence, represents the total influences given and received by attribute i and "$R – C$" which is called Relation, states the net influence that attribute i contributes to the system. According to the sign of Relation values, the attributes are assigned to either group of "cause" or "effect."

(a) If $R – C > 0$, attribute i generally influences other attributes so that it is called a "cause" attribute.
(b) If $R – C < 0$, attribute i is generally influenced by others so that it is called an "effect" attribute.

The attributes are interpreted in terms of their groups such that any improvement achieved in the cause attributes potentially generates an indirect enhancement in the effect attributes. Thus, when resources are limited, the investment or any enhancement priority should be directed to the cause attributes.

Step 7. Drawing Influential Relation Map.

The influential results of DEMATEL are depicted in Influential Relation Map (IRM). Prominence values are shown in the horizontal axis and Relation values are shown in the vertical axis. The cause attributes are clustered on the upper side while the effect attributes are located on the lower side of IRM. The overall relations can be represented on IRM. However, the higher complexity the higher number of exhibited relations. In order to show only the significant influences on IRM for reducing the complexity, T matrix can be filtered by applying a threshold value of α. Some potential determination ways of α are summarized by Si et al. (2018) such as determined by an expert through discussions, the general average of the entries in T, or the maximum value of the diagonal entries of T.

Step 8. Weighting of attributes.

The prominence value shows the degree of the central role that attribute i plays in the problem. Therefore, these values can be used in calculating the importance of the attributes as given in Eq. (20).

$$w_i = \frac{r_i + c_i}{\sum_{i=1}^{n} (r_i + c_i)} \tag{20}$$

The proposed PF-DEMATEL approach is implemented in a real decision problem regarding the assessment of the educational quality of a higher educational institution.

5 Application: Evaluation of Educational Quality

The service quality has a hard-to-maintain structure because of its specific features of intangibility, the inseparability of production and consumption, and heterogeneity (Parasuraman et al., 1985). The service quality assessments should consider its unique properties, which are unrepeatability and impalpability, and their use at the right time they are supplied as well as changeability, which is a significant factor in terms of education. These aspects of service quality are impacted by steadily increasing requirements (Nedeliakova et al., 2014). As one of the most important services in the national economies, education satisfies a crucial need of countries since it has a direct influence on their future. Therefore, the education quality must be examined in detail, especially in times of challenging conditions like the present COVID-19 Pandemic.

The educational quality literature consists of many important examples. For example, while Noaman et al. (2017) developed an extensive higher education quality assessment model for improvement of university services using AHP which involves 8 main criteria (curriculum, staff, infrastructure, services, etc.) and 53 sub-criteria, Snijders et al. (2020) investigate hypotheses on the associations between students' perceptions of the quality of their relationship with their educational faculty and staff and students' involvement via partial least squares structural equation modeling. Koza Ciftci and Karadag (2016) developed an education quality scale to appraise students' perceptions of the quality of Mathematics courses through factor analysis. Guangli (2016) examined the efficacy of the higher education quality assessment systems and claimed that establishing a social accountability system is mandatory in national size.

Cheng et al. (2016) integrated DEMATEL with importance-performance and gap analysis and quality function deployment to develop a quality improvement model for identifying critical education deficiencies. First, they considered 28 attributes of educational service quality and then prioritized 12 attributes that should be focused on first. To test their educational quality measurement model as well as our newly proposed PF-DEMATEL approach, we formed a committee in a private university of Turkey involving four experts who are academicians having education expertise of more than 5 years. The steps of the application are given below:

Step 1. For emphasizing the educational challenges that are faced by the universities during the COVID-19 Pandemic, we selected 7 potentially related attributes ($i,j = A, \ldots, G;\ i \neq j$), namely teaching labs and support equipment (A), campus Internet and Wi-Fi access (B), internship programs (C), personal taking initiative in solving students' problems (D), fair and unbiased evaluations of lecturers (E), well-thought courses (F), and fair and unbiased treatment of each student (G). There are four experts in the committee ($e = 1, 2, 3, 4$). Each is asked to fulfill the comparison matrix of attributes via linguistic terms of YES, NO, ABSTAIN, and REFUSE. Also, they are informed that if they have a gradation in their opinion of YES, they can state it with a number between 0.01 and 1.00. 0 level of agreement naturally means NO so that 0 cannot be used for grading YES. Their answers are shown in Table A1 in Appendix.

Step 2. The individuals' evaluation matrices are aggregated for building the direct PF influence matrix. Equations (10, 11 and 12) are performed for this purpose. Noticed that there is no need of performing Eq. (13) because it can be generated from the other membership degrees such that $\pi_{ij} = 1 - \mu_{ij} - \eta_{ij} - \nu_{ij}$. The direct

PF influence matrix ($\tilde{X}$) is given in Table 2. For illustration purposes, some examples are given below to clarify the aggregation process:

(i) Let us think about the aggregation of the evaluations regarding the influences between attributes G and C. The answers are ABSTAIN, NO, YES (0.6), and REFUSE, respectively. The corresponding PFN is found as (0.15, 0.30, 0.25).

$$\mu_{GC} = \frac{\sum_{e=1}^{4} \text{YES}_{GC}\left(\mu_{GC}^{e}\right)}{4} = \frac{0.6}{4} = 0.15;\ v_{GC} = \frac{\delta_{\text{NO}_{GC}}}{4} = \frac{1}{4} = 0.25$$

$$\eta_{GC} = \frac{\delta_{\text{ABSTAIN}_{GC}}}{4} + \frac{\frac{\delta_{\text{YES}_{GC}}}{4} - \mu_{GC}}{2} = \frac{1}{4} + \frac{\frac{1}{4} - 0.15}{2} = 0.25 + 0.05 = 0.30$$

$$\pi_{GC} = 1 - 0.15 - 0.30 - 0.25 = 0.30 \text{ OR}$$

$$\pi_{GC} = \frac{\delta_{\text{REFUSE}_{GC}}}{4} + \frac{\frac{\delta_{\text{YES}_{GC}}}{4} - \mu_{GC}}{2} = \frac{1}{4} + \frac{\frac{1}{4} - 0.15}{2} = 0.25 + 0.05 = 0.30$$

(ii) Now let us calculate the aggregated evaluation regarding the influences between attributes A and D. The answers are YES (0.2), YES (1.0), NO, and NO, respectively. The corresponding evaluation is a PFN of (0.3, 0.1, 0.5).

$$\mu_{AD} = \frac{\sum_{e=1}^{4} \text{YES}_{AD}\left(\mu_{AD}^{e}\right)}{4} = \frac{0.2+1}{4} = 0.30;\ v_{AD} = \frac{\delta_{\text{NO}_{AD}}}{4} = \frac{2}{4} = 0.50$$

$$\eta_{AD} = \frac{\delta_{\text{ABSTAIN}_{AD}}}{4} + \frac{\frac{\delta_{\text{YES}_{AD}}}{4} - \mu_{AD}}{2} = \frac{0}{4} + \frac{\frac{2}{4} - 0.30}{2} = 0.10$$

$$\pi_{AD} = 1 - 0.30 - 0.10 - 0.50 = 0.10$$

(iii) The aggregated evaluation regarding the influences between attributes G and F is found now. The respective answers are ABSTAIN, YES (0.8), YES (0.6), and YES (0.5). The corresponding PFN evaluation is (0.475, 0.388, 0).

$$\mu_{GF} = \frac{\sum_{e=1}^{4} \text{YES}_{GF}\left(\mu_{GF}^{e}\right)}{4} = \frac{0.8+0.6+0.5}{4} = 0.475;\ v_{GF} = \frac{\delta_{\text{NO}_{GF}}}{4} = \frac{0}{4} = 0$$

$$\eta_{GF} = \frac{\delta_{\text{ABSTAIN}_{GF}}}{4} + \frac{\frac{\delta_{\text{YES}_{GF}}}{4} - \mu_{GF}}{2} = \frac{1}{4} + \frac{\frac{3}{4} - 0.475}{2} = 0.25 + 0.138 = 0.388$$

$$\pi_{AD} = 1 - 0.475 - 0.388 - 0 = 0.137$$

Step 3. The defuzzification of matrix $\tilde{X}$ gives the direct influence matrix of X. Eq. (15) is performed for each entry of $\tilde{X}$. For illustration purposes, the defuzzification of (0.15, 0.30, 0.25) is $x_{GC} = \mu_{GC} + \frac{\eta_{GC}}{2} + \frac{\pi_{GC}}{2}\left(1 + \mu_{GC} - \nu_{GC}\right) = 0.15 + \frac{0.30}{2} + \frac{0.30}{2}(1 + 0.15 - 0.25) = 0.435$. Table 3 exhibits the matrix of X.

Step 4. This step covers the normalization of the direct influence matrix and the resulting matrix is called the initial direct influence matrix (**Z**). Table 3 shows

Table 2 The direct PF influence matrix

	A	*B*	*C*	*D*	*E*	*F*	*G*
A	(0, 0, 0)	(0.000, 0.250, 0.750)	(0.400, 0.175,0.250)	(0.300, 0.100, 0.500)	(0.250, 0.125, 0.250)	(0.725, 0.013, 0)	(0.150, 0.050, 0.750)
B	(0.675, 0.163, 0)	(0, 0, 0)	(0.100, 0.075, 0.500)	(0.175, 0.038, 0.750)	(0, 0, 1)	(0.350, 0.200, 0.250)	(0.075, 0.088, 0.750)
C	(0.275, 0.238, 0.250)	(0, 0, 1)	(0, 0, 0)	(0.025, 0.113, 0.750)	(0, 0, 1)	(0.300, 0.100, 0.500)	(0.050, 0.100, 0.750)
D	(0.400, 0.175, 0.250)	(0, 0, 1)	(0.200, 0.150, 0.500)	(0, 0, 0)	(0.375, 0.063, 0.500)	(0.400, 0.175, 0.250)	(0.650, 0.050, 0.250)
E	(0.075, 0.088, 0.750)	(0, 0, 1)	(0.175, 0.038, 0.500)	(0.300, 0.100, 0.500)	(0, 0, 0)	(0.425, 0.288, 0)	(0.650, 0.050, 0.250)
F	(0.700, 0.100, 0)	(0, 0, 1)	(0.600, 0.075, 0.250)	(0.300, 0.100, 0.500)	(0.725, 0.138, 0)	(0, 0, 0)	(0.725, 0.138, 0)
G	(0.175, 0.288, 0.500)	(0, 0.250, 0.750)	(0.150, 0.300, 0.250)	(0.425, 0.288, 0.250)	(0.625, 0.313, 0)	(0.475, 0.388, 0)	(0, 0, 0)

Source: Author

Table 3 The direct influence matrix

	A	*B*	*C*	*D*	*E*	*F*	*G*	$\sum_{j=1}^{n} x_{ij}$
A	0.000	0.125	0.588	0.390	0.500	0.958	0.185	2.746
B	0.892	0.000	0.235	0.202	0.000	0.560	0.133	2.022
C	0.515	0.000	0.000	0.097	0.000	0.390	0.115	1.117
D	0.588	0.000	0.328	0.000	0.434	0.588	0.710	2.647
E	0.133	0.000	0.291	0.390	0.000	0.774	0.710	2.297
F	0.903	0.000	0.688	0.390	0.912	0.000	0.912	3.805
G	0.331	0.125	0.435	0.591	0.832	0.770	0.000	3.084
$\sum_{i=1}^{n} x_{ij}$	3.363	0.250	2.565	2.059	2.678	4.040	2.765	$k = 4.040$

Source: Author

Table 4 The initial direct influence matrix

	A	*B*	*C*	*D*	*E*	*F*	*G*
A	0.000	0.031	0.146	0.097	0.124	0.237	0.046
B	0.221	0.000	0.058	0.050	0.000	0.139	0.033
C	0.128	0.000	0.000	0.024	0.000	0.097	0.028
D	0.146	0.000	0.081	0.000	0.107	0.146	0.176
E	0.033	0.000	0.072	0.097	0.000	0.192	0.176
F	0.223	0.000	0.170	0.097	0.226	0.000	0.226
G	0.082	0.031	0.108	0.146	0.206	0.191	0.000

Source: Author

Table 5 The total influence matrix

	A	*B*	*C*	*D*	*E*	*F*	*G*
A	0.950	0.190	1.383	1.020	1.411	1.980	1.149
B	1.535	0.073	0.903	0.703	0.753	1.422	0.809
C	0.894	0.044	0.407	0.403	0.464	0.894	0.523
D	1.415	0.092	1.157	0.698	1.394	1.715	1.566
E	0.966	0.077	1.039	0.989	0.926	1.723	1.512
F	1.921	0.121	1.713	1.268	2.059	1.598	1.999
G	1.325	0.200	1.330	1.287	1.808	1.983	1.097

Source: Author

the row and column sums of $\boldsymbol{X}$ as well as the k (normalization index which is the maximum of maximums of the row and column sums). After performing Eq. (16), the matrix of $\boldsymbol{Z}$ is obtained and shown in Table 4.

Step 5. The initial direct influence matrix is raised to the total influence matrix (T) which comprises all the direct and indirect influences. Equation (17) produced the matrix of T (Table 5).

Step 6–7. R and C values represent the row and column sums of Table 5 and their addition ($R + C$) is called Prominence while subtraction (R–C) is called Relation. According to the sign of Relation values, the groups of attributes are obtained. All the mentioned values are given in Table 6.

Cause group covers campus Internet and Wi-Fi access (B), personal taking initiative in solving students' problems (D), and fair and unbiased treatment of each

Table 6 The results of PF-DEMATEL

	R	C	$R+C$	w	R–C	Group
A	8.082	9.006	17.089	0.162	−0.924	Effect
B	6.197	0.796	6.994	0.066	5.401	Cause
C	3.629	7.932	11.561	0.109	−4.303	Effect
D	8.037	6.368	14.404	0.136	1.669	Cause
E	7.233	8.814	16.047	0.152	−1.582	Effect
F	10.678	11.316	21.994	0.208	−0.638	Effect
G	9.031	8.654	17.685	0.167	0.377	Cause

Source: Author

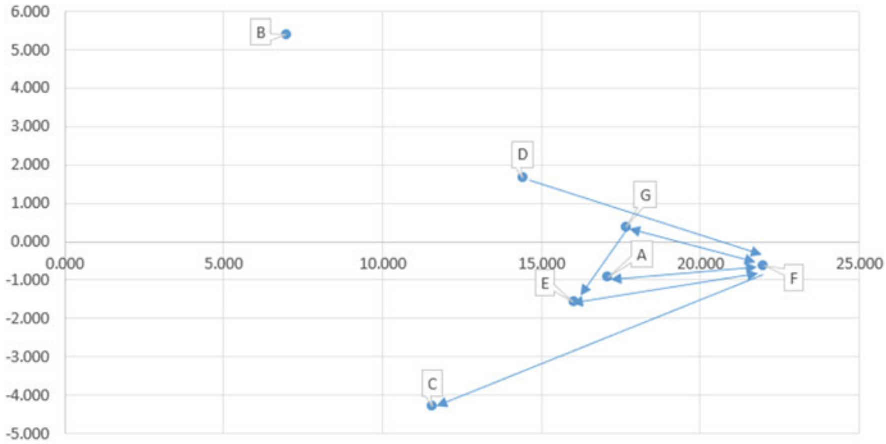


Fig. 1 IRM of the educational quality evaluation model. Source: Author

student (G). Teaching labs and support equipment (A), internship programs (C), fair and unbiased evaluations of lecturers (E), and well-thought courses (F) are found as the effect group attributes of the study. IRM is drawn according to the Prominence and Relation values as depicted in Fig. 1. For filtering the total influences between attributes, α is set to the maximum value of the diagonal entries of T ($\alpha = 1.598$). On the IRM, the significant influences are represented by directional arrows.

Step 8. The normalization of Prominence values gives the weights of the attributes (the fifth column of Table 6). The importance ranking of attributes is found as $F \succ G \succ A \succ E \succ D \succ C \succ B$. It is seen that the most important attribute is well-thought courses (F) while the least important one is campus Internet and Wi-Fi access (B).

DEMATEL finds the importance of attributes via respecting the influences between the attributes. DEMATEL and ANP are integrated in the literature and this hybrid approach is called D-ANP. To keep the flow of the text, we avoid giving further information regarding D-ANP, but the interested readers can benefit from Gölcük and Baykasoğlu (2016).

5.1 Comparison

For clarifying the importance of the usage of PFNs in DEMATEL, we conducted two different applications. The same problem is solved first by the traditional DEMATEL (Si et al., 2018) and then the intuitionistic fuzzy version of DEMATEL—IF-DEMATEL (Zhou et al., 2018; Hinduja & Pandey, 2019). The influence data gathered from the experts are used in these applications too after being processed in different manners.

In traditional DEMATEL application, the four evaluation matrices are converted to the conventional representation style in DEMATEL by applying some rules. No Influence (NI) represents the linguistic terms of NO, ABSTAIN, and REFUSAL because traditional DEMATEL cannot handle these kinds of information. The gradation levels in the linguistic term of YES are converted to the appropriate levels of influence. We determined intervals for respective evaluations in DEMATEL: Weak Influence (WI) covers the gradation levels between 0.001 and 0.349; Moderate Influence (M) covers the degrees between 0.350 and 0.699; and finally, the degrees ranging from 0.700 to 1.000 are represented by Strong Influence (S). The last step is the assignment of crisp values: 0 for NI, 1 for W, 2 for M, and 3 for S. All the other steps are operated as given in the classical definition of DEMATEL.

In IF-DEMATEL, the evaluation matrices are transformed into intuitionistic fuzzy numbers. For this purpose, the equations given below are derived from Eqs. (10 and 11). The answers of NO, ABSTAIN, and REFUSE can be represented by the nonmembership degree of an IFS because they have negative meanings rather than positive meanings.

$$\mu_{ij} = \frac{\delta \mathrm{YES}_{ij}}{m} \tag{21}$$

$$v_{ij} = \frac{\delta \mathrm{NO}_{ij} + \delta \mathrm{ABSTAIN}_{ij} + \delta \mathrm{REFUSE}_{ij}}{m} \tag{22}$$

After converting the linguistic evaluations of the experts to IFNs, the two-step defuzzification process is operated as defined by Ocampo and Yamagishi (2020).

1. The IFNs are converted to the standard fuzzy subsets using $\mu_{ij}^F = \left(1 + \mu_{ij} - v_{ij}\right) / 2$.
2. The corresponding μ_{ij}^F membership degrees are assigned to a triangular fuzzy number—TFN (0, 3, 3), where *0* represents NI since it is the minimum level of influence evaluations and *3* represents S since it is the maximum level of influence evaluations in DEMATEL. By a transformation formula, all the membership degrees will be converted to crisp values. For this purpose, using the transformation formula for the left side will be sufficient: $\mu_{ij}^F = (x - l) / (m - l) = (x - 0) / (3 - 0)$. After the reorganization of the equation, the corresponding defuzzification formula is obtained: $x = 3 * \mu_{ij}^F$.

All the other steps are identical to the classical definition of DEMATEL.

After converting and aggregating the experts' data sets, Tables A2 and A3 in Appendix are obtained. The results of the applications are summarized in Table 7.

In terms of attribute ranking, it is seen that the first (*F*) and the last three (*D-C-B*) attributes are kept their positions. A and E takes the third and fourth order alternately in PF-DEMATEL and DEMATEL while attributes *A*, *E*, and *G* change their ranks in the comparison of PF-DEMATEL and IF-DEMATEL. Similarly, the groups of the attributes are slightly different. From *A* to *E*, the attributes are kept in their group assignments in various applications of DEMATEL. At the same time, *F* and *G* changed their groups in DEMATEL and PF-DEMATEL implementations while just *G* changed its group from "effect" to "cause" when we compare the results of IF-DEMATEL and PF-DEMATEL. Thus, in these limited-content applications of DEMATEL versions, it is seen that even though the differences are not huge, it is obvious that the consideration of hesitancy and refusal opinions of the experts by PF-DEMATEL produces different findings.

5.2 Discussion

The findings depicted in Table 6 and Fig. 1 are discussed in this section. For the educational quality evaluation MADM model, some important points are highlighted as follows:

(i) *Weight ranking of attributes:* The most important attribute is found as well-taught courses (*F*). Since the *raison d'étre* of educational institutions is to provide quality education to the students, this finding is an expected result, and it is independent of the states of nature. During the pandemic periods, educational activity should also be kept in an acceptably good condition. The second rank is taken by fair and unbiased treatment of each student. In the pandemic era, the students are not at school so that they need to be treated most equitably. Transparency should be ensured for having the students feel comfortable in the processes. The third important attribute is the lab and supporting equipment. It is reasonable because distance education has challenging conditions in terms of many aspects of educational actions such as online learning platforms, tablets, or laptops for reaching the learning platforms, the software required for teaching the coding in the software engineering department, or simulating the experiments in chemical or electronics engineering departments, etc. All these applications and experiment-based aspects of education should also be maintained in distance education too. The least important attribute is campus Internet and Wi-Fi access since the students as well as the lecturers are mostly participating in the online courses at home so that the level of Wi-Fi usage loses its importance and significance in the pandemic era. Similar remarks can be given for the remaining attributes. For administrative purposes, the weights obtained by the proposed PF-DEMATEL can be used in further evaluations of some strategies. For instance, the university administrators can evaluate their faculties or departments to reveal their current educational quality performance.

Table 7 Results of DEMATEL, IF-DEMATEL, and proposed PF-DEMATEL

	DEMATEL					IF-DEMATEL					PF-DEMATEL				
	$R+C$	w	RANK	$R–C$	Group	$R+C$	w	RANK	$R–C$	Group	$R+C$	w	RANK	$R–C$	Group
A	38.752	0.153	*4*	−2.658	Effect	55.041	0.159	*2*	−7.621	Effect	17.089	0.162	*3*	−0.924	Effect
B	16.195	0.064	7	16.195	Cause	23.792	0.069	7	23.792	Cause	6.994	0.066	7	5.401	Cause
C	26.654	0.105	6	−9.586	Effect	41.758	0.121	6	−10.135	Effect	11.561	0.109	6	−4.303	Effect
D	36.982	0.146	5	3.339	Cause	50.464	0.146	5	4.976	Cause	14.404	0.136	5	1.669	Cause
E	40.186	0.158	*3*	−4.448	Effect	52.145	0.151	*3*	−2.298	Effect	16.047	0.152	*4*	−1.582	Effect
F	53.090	0.209	1	0.651	*Cause*	70.639	0.204	1	−2.028	Effect	21.994	0.208	1	−0.638	*Effect*
G	42.016	0.165	2	−3.493	*Effect*	51.981	0.150	*4*	−6.687	*Effect*	17.685	0.167	*2*	0.377	*Cause*

Source: Author

These performance scores and the weights can be processed in a PFS-based MADM method such as PF-ARAS (Jovcic et al., 2020), PF-EDAS (Zhang et al., 2019), PF-TOPSIS, PF-VIKOR (Si et al., 2020), and PF-TODIM (Tian & Peng, 2020).

(ii) *Groups and significant influences in IRM:* When we look at the group assignments, it is seen that the cause group includes campus Internet and Wi-Fi access, personal taking initiatives in solving problems, and fair and unbiased treatment of each student while the others are listed in effect group. Thus, the priority for solving issues in these three attributes is required for building a more quality education system during a Pandemic occurrence since any enhancement achieved in them can indirectly create a positive effect on the others. The most significant influences between attributes are shown in Fig. 1. Accordingly, it is seen that there are three bidirectional influences between well-taught courses and fair evaluations of lecturers, fair treatment of each student, and teaching labs and support equipment, separately. Therefore, any enhancement in well-taught courses generates developments in these three attributes and vice versa. Also, there are 3 directional influences: attribute of well-taught courses affects the internship programs and is affected by personal taking initiatives in solving problems, also the fair treatment of students potentially affects the fair evaluation of their success in the courses. The priorities should be distributed in accordance with these findings and the objectives of the higher educational institution.

6 Conclusions

PFS concept is one of the latest developments in the fuzzy logic field and it is capable of handling expert opinions more comprehensively. PFS takes the positive and negative expert opinions as positive membership and negative membership degrees, respectively. These two elements are coming from IFS concept but PFS is developed as a generalization of IFS. IFS calculates the hesitancy of the expert from the pre-recognized membership degrees and it does not allow the expert to determine the hesitancy degree independently. PFS considers hesitancy degree as a third individual element. According to the definition of PFS, the sum of the mentioned three elements should not exceed 1 while each ranges between 0 and 1. In addition, PFS defines a fourth element which is called refusal degree being computed from the pre-defined three elements. Refusal degree is the basic peculiarity of PFS because PFS is the only fuzzy set concept that can deal with the refusal preference of the expert.

This study aims to extend DEMATEL approach which is one of the most powerful and well-known MADM tools into picture fuzzy environment. To the best of our knowledge, there is no proposition of DEMATEL that can cope with both the hesitancy degree and refusal degree of the experts. In the proposed PF-DEMATEL approach, a voting-based data collection process is developed. In this novel system, the experts can represent their opinions about the influences among the attributes by linguistic terms of YES, NO, ABSTAIN, and REFUSE. They are respectively accepted as the positive, negative, hesitancy, and refusal membership degrees and they are combined as PFNs. The first contribution of the study is the novel data collection process and the representation of the data as PFNS.

The second contribution is the development of a picture fuzzy DEMATEL version that is capable of handling the data collected from the experts. At this point, a limitation of the study emerges. After collecting and aggregating the expert data and constructing the initial PF influence matrix, we had to defuzzy it for further operations. It is a common way of handling fuzzy numbers in DEMATEL applications, but this early defuzzification action can potentially cause an information loss. In future studies, the whole process should be kept fuzzy until the very end of the application. It is not applied in the study since required matrix operations in PFS have not been developed yet in the literature. After the definitions of matrix operations such as matrix inversion or matrix multiplication in picture fuzzy mathematics appear, this early defuzzification issue will be solved.

The proposed method is applied in a group decision-making environment to demonstrate its applicability in real-life problems. To understand the educational qualification requirements in higher educational institutions, we formed a committee consisting of four experts. These experts determined 7 attributes that are supposed to affect the educational quality, especially during unexpected occurrences such as Pandemics. According to the results, the first three important attributes are prioritized as well-thought courses, fair and unbiased treatment of students, and teaching laboratories and supporting equipment. The least important attribute is found as campus Internet and Wi-Fi access as expected because there is no educational activity in the buildings. The influences among the attributes are revealed at the end of the application and they are discussed in the previous section for showing their managerial implications. The results are compared with the application of traditional DEMATEL and IF-DEMATEL, and the findings are discussed. In short, it is seen that the consideration of separated hesitancy and refusal degrees of the expert opinions make a significant difference.

The proposed method is applied here in the educational quality evaluation problem, but the method can be used for any future MADM problem that requires the consideration of the influences among the decision attributes. The proposed PF-DEMATEL version provides a group decision-making approach because it extensively handles many parts of the different expert opinions and aggregates them to reach a group decision. In near future studies, the applicability of PF-DEMATEL method and its enhanced versions will be tested in solving various MADM problems such as face mask selection, public transportation evaluation, and Pandemic relaxation protocols during the Pandemic era.

Appendix

Table A.1 Expert evaluations

	A	*B*	*C*	*D*	*E*	*F*	*G*		*A*	*B*	*C*	*D*	*E*	*F*	*G*
A	–	NO	YES (0.5)	YES (0.2)	REFUSE	REFUSE	NO	*A*	–	NO	YES (0.8)	YES (1.0)	YES (0.8)	YES (1.0)	YES (0.6)
B	YES (0.7)	–	NO	NO	NO	YES (0.1)	NO	*B*	YES (0.6)	–	YES (0.2)	NO	NO	NO	NO
C	YES (0.2)	NO	–	NO	NO	YES (0.2)	YES (0.2)	*C*	NO	NO	–	YES (0.1)	NO	NO	NO
D	YES (0.5)	NO	NO	–	YES (0.7)	YES (0.2)	YES (0.7)	*D*	YES (1.0)	NO	NO	–	NO	YES (1.0)	NO
E	NO	NO	NO	NO	–	YES (0.1)	NO	*E*	YES (0.3)	NO	NO	YES (0.4)	–	YES (1.0)	YES (1.0)
F	YES (0.4)	NO	NO	NO	YES (0.6)	–	YES (0.6)	*F*	YES (0.8)	NO	YES (0.6)	YES (0.6)	YES (0.9)	–	YES (0.9)
G	ABSTAIN	ABSTAIN	ABSTAIN	ABSTAIN	ABSTAIN	ABSTAIN	–	*G*	YES (0.7)	NO	NO	NO	YES (1.0)	YES (0.8)	–
	A	*B*	*C*	*D*	*E*	*F*	*G*		*A*	*B*	*C*	*D*	*E*	*F*	*G*
A	–	NO	YES (0.3)	NO	YES (0.2)	YES (0.9)	NO	*A*	–	ABSTAIN	NO	NO	NO	YES (1.0)	NO
B	YES (0.6)	–	YES (0.2)	YES (0.7)	NO	YES (0.5)	NO	*B*	YES (0.8)	–	NO	NO	NO	YES (0.8)	YES (0.3)
C	YES (0.7)	NO	–	NO	NO	NO	NO	*C*	YES (0.2)	NO	–	NO	NO	YES (1.0)	NO
D	NO	NO	YES (0.3)	–	NO	NO	YES (0.9)	*D*	YES (0.1)	NO	YES (0.5)	–	YES (0.8)	YES (0.4)	YES (1.0)
E	NO	NO	YES (0.7)	NO	–	YES (0.4)	YES (0.6)	*E*	NO	NO	REFUSE	YES (0.8)	–	YES (0.2)	YES (1.0)
F	YES (0.6)	NO	YES (0.8)	NO	YES (0.4)	–	YES (0.3)	*F*	YES (1.0)	NO	YES (1.0)	YES (0.6)	YES (1.0)	–	YES (0.8)
G	NO	NO	YES (0.6)	YES (0.7)	YES (0.5)	YES (0.6)	–	*G*	NO	NO	REFUSE	YES (1.0)	YES (1.0)	YES (0.5)	–

Source: Author

Table A.2 Initial direct influence matrix of DEMATEL

	A	*B*	*C*	*D*	*E*	*F*	*G*
A	0.000	0.000	1.500	1.000	1.000	2.250	0.500
B	2.500	0.000	0.500	0.750	0.000	1.500	0.250
C	1.250	0.000	0.000	0.250	0.000	1.000	0.250
D	1.000	0.000	0.750	0.000	1.500	1.500	2.250
E	0.250	0.000	0.750	1.250	0.000	1.750	2.000
F	2.500	0.000	2.000	1.000	2.500	0.000	2.250
G	0.750	0.000	0.500	1.500	2.000	1.750	0.000

Source: Author

Table A.3 Initial direct influence matrix of IF-DEMATEL

	A	*B*	*C*	*D*	*E*	*F*	*G*
A	0.000	0.000	2.250	1.500	1.500	2.250	0.750
B	3.000	0.000	1.500	0.750	0.000	2.250	0.750
C	2.250	0.000	0.000	0.750	0.000	1.500	0.750
D	2.250	0.000	1.500	0.000	1.500	2.250	2.250
E	0.750	0.000	0.750	1.500	0.000	3.000	2.250
F	3.000	0.000	2.250	1.500	3.000	0.000	3.000
G	0.750	0.000	0.750	1.500	2.250	2.250	0.000

Source: Author

References

Abdel-Basset, M., Atef, A., & Smarandache, F. (2019). A hybrid neutrosophic multiple criteria group decision making approach for project selection. *Cognitive Systems Research, 57*, 216–227. https://doi.org/10.1016/j.cogsys.2018.10.023

Abdel-Basset, M., Manogaran, G., Gamal, A., & Smarandache, F. (2018). A hybrid approach of neutrosophic sets and DEMATEL method for developing supplier selection criteria. *Design Automation for Embedded Systems, 22*, 257–278. https://doi.org/10.1007/s10617-018-9203-6

Abdel-Basset, M., Manogaran, G., & Mohamed, M. (2018). Internet of things (IoT) and its impact on supply chain: A framework for building smart, secure and efficient systems. *Future Generation Computer Systems, 86*, 614–628. https://doi.org/10.1016/j.future.2018.04.051

Abdel-Basset, M., Mohamed, R., & Elhoseny, M. (2020). A novel framework to evaluate innovation value proposition for smart product–service systems. *Environmental Technology & Innovation, 20*, 101036. https://doi.org/10.1016/j.eti.2020.101036

Al-Quran, A., Hashim, H., & Abdullah, L. (2020). A hybrid approach of interval Neutrosophic vague sets and DEMATEL with new linguistic variable. *Symmetry, 12*, 275. https://doi.org/10.3390/sym12020275

Arya, V., & Kumar, S. (2020). A new picture fuzzy information measure based on shannon entropy with applications in opinion polls using extended VIKOR-TODIM approach. *Computational and Applied Mathematics, 39*, 197. https://doi.org/10.1007/s40314-020-01228-1

Atanassov, K. T. (1986). Intuitionistic fuzzy sets. *Fuzzy Sets and Systems, 20*, 87–96. https://doi.org/10.1016/S0165-0114(86)80034-3

Awang, A., Abdullah, L., Ghani, A. T. A., Aizam, N. A. H., & Ahmad, M. F. (2020). A fusion of decision-making method and neutrosophic linguistic considering multiplicative inverse matrix for coastal erosion problem. *Soft Computing, 24*, 9595–9609. https://doi.org/10.1007/s00500-019-04467-5

Awang, A., Aizam, N. A. H., & Abdullah, L. (2019). An integrated decision-making method based on Neutrosophic numbers for investigating factors of coastal erosion. *Symmetry, 11*, 328. https://doi.org/10.3390/sym11030328

Awang, A., Ghani, A. T. A., Abdullah, L., & Ahmad, M. F. (2018). A DEMATEL method with single valued neutrosophic set (SVNS) in identifying the key contribution factors of Setiu Wetland's coastal erosion. *AIP Conference Proceedings, 1974*, 020011. https://doi.org/10.1063/1.5041542

Cebi, S. (2013). Determining importance degrees of website design parameters based on interactions and types of websites. *Decision Support Systems, 54*, 1030–1043. https://doi.org/10.1016/j.dss.2012.10.036

Cheng, C. C., Tsai, M. C., & Lin, C. L. (2016). Quality education service: Put your feet in their shoes. *Current Issues in Tourism, 19*(11), 1120–1135. https://doi.org/10.1080/13683500.2013.839633

Cuong, B. C. (2014). Picture fuzzy sets. *Journal of Computer Science and Cybernetics, 30*(4), 409–420. https://doi.org/10.15625/1813-9663/30/4/5032

Cuong, B. C., & Kreinovich, V. (2013). Picture fuzzy sets – A new concept for computational intelligence problems. In *Proceeding of the 3rd World Congress on Information and Communication Technologies*, December 15–18, Hanoi, Vietnam, pp. 1–6. https://doi.org/10.1109/WICT.2013.7113099.

De, S. K. (2017). Triangular dense fuzzy lock sets. *Soft Computing, 22*, 7243–7254. https://doi.org/10.1007/s00500-017-2726-0

Fontela, E., & Gabus, A. (1972). *The DEMATEL observer*. Battelle Geneva Research Centre.

Gabus, A., & Fontela, E. (1972). *World problems, an invitation to further thought within the framework of DEMATEL*. Battelle Geneva Research Centre.

Gabus, A., & Fontela, E. (1973). *Perceptions of the world problematique: Communication procedure, communicating with those bearing collective responsibility (DEMATEL report No. 1)*. Battelle Geneva Research Centre, Geneva, Switzerland.

Gölcük, İ., & Baykasoğlu, A. (2016). Analysis of DEMATEL approaches for criteria interaction handling within ANP. *Expert Systems with Applications, 46*, 346–366. https://doi.org/10.1016/j.eswa.2015.10.041

Govindan, K., Mina, H., Esmaeili, A., & Gholami-Zanjani, S. M. (2020). An integrated hybrid approach for circular supplier selection and closed loop supply chain network design under uncertainty. *Journal of Cleaner Production, 242*, 118317. https://doi.org/10.1016/j.jclepro.2019.118317

Guangli, Z. (2016). The effectiveness of the higher education quality assessment system: Problems and countermeasures in China. *Chinese Education & Society, 49*(1–2), 39–48. https://doi.org/10.1080/10611932.2016.1192385

Gül, S. (2020). Spherical fuzzy extension of DEMATEL (SF-DEMATEL). *International Journal of Intelligent Systems, 35*(9), 1329–1353. https://doi.org/10.1002/int.22255

Gül, S., Kabak, Ö., & Topçu, Y. I. (2020). Coordinating the ISM code and OHSAS procedures to improve the occupational health and safety at sea. *Journal of Multi-Criteria Decision Analysis, 27*(5–6), 286–303. https://doi.org/10.1002/mcda.1710

Hinduja, A., & Pandey, M. (2019). Assessment of healthcare waste treatment alternatives using an integrated decision support framework. *International Journal of Computational Intelligence Systems, 12*(1), 318–333. https://doi.org/10.2991/ijcis.2018.125905685

Hwang, W., Hsiao, B., Chen, H. G., & Chern, C. C. (2016). Multiphase assessment of project risk interdependencies: Evidence from a university ISD project in Taiwan. *Project Management Journal, 47*(1), 59–75. https://doi.org/10.1002/pmj.21563

Joshi, R. (2020a). A new picture fuzzy information measure based on Tsallis–Havrda–Charvat concept with applications in presaging poll outcome. *Computational and Applied Mathematics, 39*, 71. https://doi.org/10.1007/s40314-020-1106-z

Joshi, R. (2020b). A novel decision-making method using *R*-norm concept and VIKOR approach under picture fuzzy environment. *Expert Systems with Applications, 147*, 113228. https://doi.org/10.1016/j.eswa.2020.113228

Jovcic, S., Simic, V., Prusa, P., & Dobrodolac, M. (2020). Picture fuzzy ARAS method for freight distribution concept selection. *Symmetry, 12*, 1062. https://doi.org/10.3390/sym12071062

Kabak, Ö., Ülengin, F., Çekyay, B., Önsel, Ş., & Özaydın, Ö. (2016). Critical success factors for the iron and steel industry in Turkey: A fuzzy DEMATEL approach. *International Journal of Fuzzy Systems, 18*(3), 523–536. https://doi.org/10.1007/s40815-015-0067-7

Kilic, H. S., & Yalcin, A. S. (2020). Comparison of municipalities considering environmental sustainability via neutrosophic DEMATEL based TOPSIS. *Socio-Economic Planning Sciences*, In Press. https://doi.org/10.1016/j.seps.2020.100827

Kim, S. Y., & Nguyen, M. V. (2021). Mapping the complexity of international development projects using DEMATEL technique. *Journal of Management in Engineering, 37*(2). https://doi.org/10.1061/(ASCE)ME.1943-5479.0000882

Kouhizadeh, M., Saberi, S., & Sarkis, J. (2021). Blockchain technology and the sustainable supply chain: Theoretically exploring adoption barriers. *International Journal of Production Economics, 231*, 107831. https://doi.org/10.1016/j.ijpe.2020.107831

Koza Ciftci, S., & Karadag, E. (2016). Developing a mathematics education quality scale. *African Education Review, 13*(3–4), 211–234. https://doi.org/10.1080/18146627.2016.1224590

Kutlu Gündoğdu, F., & Kahraman, C. (2019). Spherical fuzzy sets and spherical fuzzy TOPSIS method. *Journal of Intelligent & Fuzzy Systems, 36*(1), 337–352. https://doi.org/10.3233/JIFS-181401

Liang, R., Wang, J., & Zhang, H. (2017). Evaluation of e-commerce websites: An integrated approach under a single-valued trapezoidal neutrosophic environment. *Knowledge-Based Systems, 135*, 44–59. https://doi.org/10.1016/j.knosys.2017.08.002

Meksavang, P., Shi, H., Lin, S. M., & Liu, H. C. (2019). An extended picture fuzzy VIKOR approach for sustainable supplier management and its application in the beef industry. *Symmetry, 11*, 468. https://doi.org/10.3390/sym11040468

Nabeeh, N. A., Abdel-Basset, M., & Soliman, G. (2021). A model for evaluating green credit rating and its impact on sustainability performance. *Journal of Cleaner Production, 280*, 124299. https://doi.org/10.1016/j.jclepro.2020.124299

Nedeliakova, E., Sekulova, J., Nedeliak, I., & Loch, M. (2014). Methodics of identification level of service quality in railway transport. *Procedia - Social and Behavioral Sciences, 110*, 320–329. https://doi.org/10.1016/j.sbspro.2013.12.876

Noaman, A. Y., Ragab, A. H. M., Madbouly, A. I., Khedra, A. M., & Fayoumi, A. G. (2017). Higher education quality assessment model: Towards achieving educational quality standard. *Studies in Higher Education, 42*(1), 23–46. https://doi.org/10.1080/03075079.2015.1034262

Ocampo, L., & Yamagishi, K. (2020). Modeling the lockdown relaxation protocols of the Philippine government in response to the COVID-19 pandemic: An intuitionistic fuzzy DEMATEL analysis. *Socio-Economic Planning Sciences, 72*, 100911. https://doi.org/10.1016/j.seps.2020.100911

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing, 49*(1), 41–50. https://doi.org/10.1177/002224298504900403

Senapati, T., & Yager, R. R. (2020). Fermatean fuzzy sets. *Journal of Ambient Intelligence and Humanized Computing, 11*, 663–674. https://doi.org/10.1007/s12652-019-01377-0

Si, A., Das, S., & Kar, S. (2020). Extension of TOPSIS and VIKOR method for decision-making problems with picture fuzzy number. In J. K. Mandal & S. Mukhopadhyay (Eds.), *Proceedings of the global AI congress 2019, advances in intelligent systems and computing 1112* (pp. 563–577). Springer Nature Singapore. https://doi.org/10.1007/978-981-15-2188-1_44

Si, S. L., You, X. Y., Liu, H. C., & Zhang, P. (2018). DEMATEL technique: A systematic review of the state-of-the-art literature on methodologies and applications. Mathematical Problems in Engineering, 2018 33. https://doi.org/10.1155/2018/3696457. 3696457

Smarandache, F. (1999). *A unifying field in logics neutrosophy: Neutrosophic probability, set and logic*. American Research Press.

Snijders, I., Wijnia, L., Rikers, R. M. J. P., & Loyens, S. M. M. (2020). Building bridges in higher education: Student-faculty relationship quality, student engagement, and student loyalty. *International Journal of Educational Research, 100*, 101538. https://doi.org/10.1016/j.ijer.2020.101538

Tan, R., & Zhang, W. (2020). Multiple attribute decision making method based on DEMATEL and fuzzy distance of trapezoidal fuzzy neutrosophic numbers and its application in typhoon disaster evaluation. *Journal of Intelligent & Fuzzy Systems, 39*(3), 3413–3439. https://doi.org/10.3233/JIFS-191758

Tian, C., & Peng, J. J. (2020). An integrated picture fuzzy ANP-TODIM multi-criteria decision-making approach for TOURISM attraction recommendation. *Technological and Economic Development of Economy, 26*(2), 331–354. https://doi.org/10.3846/tede.2019.11412

Torra, V. (2010). Hesitant fuzzy sets. *International Journal of Intelligent Systems, 25*(6), 529–539. https://doi.org/10.1002/int.20418

Wang, W. C., Lin, Y. H., Lin, C. L., Chung, C. H., & Lee, M. T. (2012). DEMATEL-based model to improve the performance in a matrix organization. *Expert Systems with Applications, 39*, 4978–4986. https://doi.org/10.1016/j.eswa.2011.10.016

Wang, C., Zhou, X., Tu, H., & Tao, S. (2017). Some geometric aggregation operators based on picture fuzzy sets and their application in multiple attribute decision making. *Italian Journal of Pure and Applied Mathematics, 37*, 477–492.

Xie, Y., Peng, Y., Yüksel, S., Dinçer, H., Uluer, G. S., Çağlayan, Ç., & Li, Y. (2020). Consensus-based public acceptance and mapping of nuclear energy investments using spherical and Pythagorean fuzzy group decision making approaches. *IEEE Access, 8*, 206248–206263. https://doi.org/10.1109/ACCESS.2020.3037344

Xu, X. G., Shi, H., Xu, D. H., & Liu, H. C. (2019). Picture fuzzy petri nets for knowledge representation and Acquisition in Considering Conflicting Opinions. *Applied Sciences, 9*, 983. https://doi.org/10.3390/app9050983

Yager, R. R. (2013). Pythagorean fuzzy subsets. In *Proceedings of joint IFSA World Congress and NAFIPS Annual Meeting*, June 24–28, Edmonton, Canada, pp. 57–61. https://doi.org/10.1109/IFSA-NAFIPS.2013.6608375.

Yager, R. R. (2017). Generalized Orthopair fuzzy sets. *IEEE Transactions on Fuzzy Systems, 25*(5), 1222–1230. https://doi.org/10.1109/TFUZZ.2016.2604005

Yang, W., & Pang, Y. (2018). New multiple attribute decision making method based on DEMATEL and TOPSIS for multi-valued interval Neutrosophic sets. *Symmetry, 10*, 115. https://doi.org/10.3390/sym10040115

Yazdi, M., Khan, F., Abbasi, R., & Rusli, R. (2020). Improved DEMATEL methodology for effective safety management decision-making. *Safety Science, 127*, 104705. https://doi.org/10.1016/j.ssci.2020.104705

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*, 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X

Zhang, S., Wei, G., Gao, H., Wei, C., & Wei, Y. (2019). Edas method for multiple criteria group decision making with picture fuzzy information and its application to green suppliers selections. *Technological and Economic Development of Economy, 25*(6), 1123–1138. https://doi.org/10.3846/tede.2019.10714

Zhou, L., Dai, G., Qin, R., Tang, M., & Qiu, J. (2018). Risk analysis of gob coal spontaneous combustion in MethaneRich, combustion-prone coal seam based on intuitionistic fuzzy DEMATEL. *Journal of Failure Analysis and Prevention, 18*, 975–987. https://doi.org/10.1007/s11668-018-0492-7

A Contingency Approach to Multi-Criteria Decision-Making: A Search for Validity Through Rigor and Relevance

Enrique Mu and Orrin Cooper

Abstract Addressing the contingent dimensions (content and context) in multi-criteria decision-making is very important to ensure the validity of a study. While this approach is widely accepted in the strategic decision-making community, it is argued here that this practice is not properly addressed and/or reported in many cases and that it must be applied in all MCDM decisions to ensure the rigor and relevance of the outcome. To explore the extent to which contingent factors are addressed in the literature, a sample of 46 MCDM group decision-making papers from a single year of publication was examined with regard to a well-known contingent dimension: group decision-making. More specifically, the following four critical variables were examined: group membership, group process, aggregation of perspectives, and group engagement. The study found that the percentage of papers that addressed these variables in a reasonable way was 23.9%, 17.4%, 26.1%, and 19.6%, respectively. These results suggest that MCDM analysts are not, for the most part, properly addressing (or reporting) group decision-making and similar contingent dimensions. For this reason, this research is a call to authors and journal editors to include and properly address all MCDM applicable contingent dimensions to improve MCDM rigor and relevance; that is, the overall validity of MCDM studies.

Keywords Contingency dimensions · Rigor · relevance · validity · Group decision-making · MCDM

E. Mu (✉)
Carlow University, Pittsburgh, PA, USA
e-mail: emu@carlow.edu

O. Cooper
University of Memphis, Memphis, TN, USA
e-mail: olcooper@memphis.edu

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_19

1 Introduction

Every year, a large number of multi-criteria decision-making/analysis (MCDM/MCDA) papers are submitted for publication in prestigious journals. In general, editors are concerned about academic rigor and relevance. Typically, relevance is considered in terms of the discipline and the specific journal, but it is rarely discussed in terms of the nature of the specific decision and its contextual factors. This is a weakness because MCDM may have originated from the need to structure decision-making within the framework of specific decisions (Franklin, 1772). MCDM can be defined as "making choices, supporting those choices, and understanding them, in the presence of multiple, conflicting criteria" (Koksalan, Wallenius et al. 2011). Because of this, MCDM applications are expected to be relevant, at the very least in relation to the decision for which the criteria and alternatives were developed.

In other words, relevance and academic rigor are linked like Siamese twins. The answer to the famous question, *"Are managers from Mars and academicians from Venus?"* proposed by Baldridge, Floyd et al. (2004) is a resounding "no." Rigor and relevance should go hand in hand in MCDM. This is even more important now that MCDM is commonly used for high-level managerial decision-making (Kasanen, Wallenius, et al., 2000; Wallenius, Dyer, et al., 2008). For this reason, it is necessary to move beyond our narrow focus on the rigor of the mathematical decision-making analysis in MCDM studies and also explore their validity and relevance in terms of specific decisions. MCDM model validity has been defined as "support that the decision-making model within its range of applicability possesses, a satisfactory range of accuracy consistent with the intended application of the model" (Mu, Cooper, et al., 2020). How does one reach MCDM model validity? It is proposed here that it is necessary to use a contingency approach and an overall concern for quality as part of best practices in multi-criteria decision-making analysis. Group decision-making is a well-known and established form of decision-making in the field and therefore can serve as a representative contingent dimension in MCDM for this study. Exploring the extent to which the most common and salient variables of group decision-making are addressed and reported in published studies can provide insight into how well contingency factors are addressed overall in MCDM models.

2 Literature Review/Theoretical Framework

Contingency theory, which has been used for many years, states that managerial activity must take into account situational factors (Hatch, 1997). In the MCDM discipline, the managerial activity of interest is decision-making, which according to the contingency approach, must take into account situational factors that may affect the decision-making process and/or outcome by providing boundary conditions. The following two kinds of contingencies are usually considered for MCDM purposes:

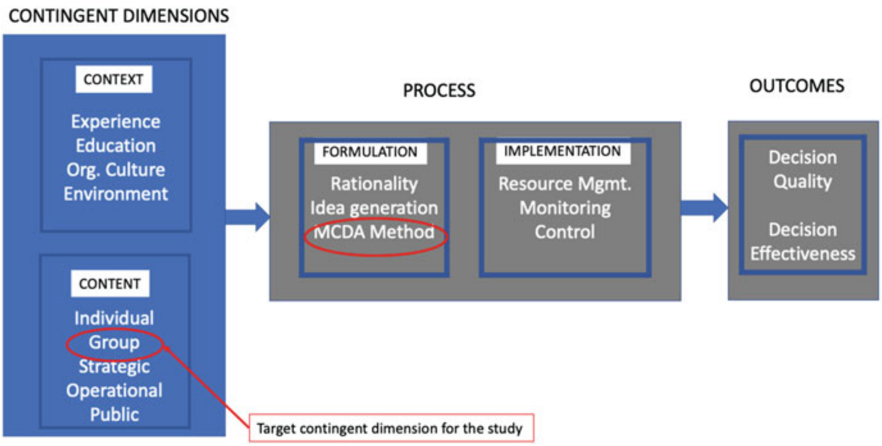


Fig. 1 Conceptual framework (Adapted from Papadakis, et al., 2010)

content and context. Both of these contingencies influence the choices, expected benefits and processes to be applied (Bell, Bromley, et al., 1998). While most of this literature originated within the context of strategic decision-making, it is argued here that the consideration of contingent dimensions (also referred to as factors in this study) is also important for the validity and relevance of any MCDM analysis.

Content refers to the type of decision (Fig. 1). There are many possible types of decisions such as individual, group, strategic, operational, conflict resolution, and others. Strategic decision-making (SDM) has become an important area for MCDM applications and a field in its own right. These decisions tend to deal with unstructured topics in uncertain environments and are performed at the team or organizational level and sometimes at the national level.

Other decision characteristics that will affect the decision-making process are uncertainty and importance. In particular, the importance or magnitude of the impact of the decision has been found to affect the rationale and comprehensiveness of the decision-making process (Papadakis, Lioukas, et al., 1998). Unfortunately, calling a decision "strategic" has become a cheap way to suggest that the decision is important; however, it is true that these decisions are particularly important and have a large impact on all those involved in making the decision or who are affected by the decision outcome. The MCDM analyst should indicate who (and how) will be affected by the decision and the extent of impact related to the outcome (importance). Similarly, simply because some sort of a team was somehow involved in the decision does not ensure a certain level of quality and importance of the decision. In statistics larger samples are generally a good thing; but that is not always the case in group decisions.

Applying all the new developments of MCDM in the analysis of strategic decisions makes perfect sense and is currently being done (Baldridge, Floyd, et al., 2004). However, it is also important to consider the contingent factors of decisions. Strategic decisions have a set of characteristics, in addition to their importance, such as long-term consequences, the extent of resources involved, infrequent choices, and high-risk level. In general, all decisions whether strategic or not, have specific contingent factors associated with them. These factors must be considered in the MCDM analysis because they will influence both the decision process and outcome. Strategic decisions usually require consideration of factors such as organizational strengths and weaknesses, benefits and costs, opportunities and risks, and the discussion of possible scenarios. The proposed decision analysis must be in accordance with accepted best practices for the specific decision (e.g., benefit/cost/risk analysis for financial investment decisions). The need to consider diverse scenarios is associated with the two following aspects of MCDM: assumptions and sensitivity. In general, decisions are made under certain, usually unspoken, assumptions. The MCDM analyst is expected to make these assumptions explicit and consider different possible scenarios in their studies to provide a boundary for the decision outcome.

Many times, decision-making is carried out at the team level, more specifically by the organization's top management team. In this case, there are additional considerations to take into account. Given that a group or team is formed by several individual decision makers, reaching consensus in a group decision involves people presenting, arguing, and finally synthesizing their opinions into a single "group decision" either by agreement, voting, or any other predetermined mechanisms (Mu & Stern, 2014). Group decision-making usually follows a two-stage approach which starts with divergence (generation of ideas to propose alternatives and criteria) and moves to convergence (evaluation and selection of alternatives based on the proposed criteria). There is extensive literature on high-impact teams and agreement that the actual process is key in group decision-making (Eden & Ackermann, 2010; Lingham & Richley, 2018). The use of a procedural justice process makes it more likely that the team members will be engaged in the decision-making process and will accept and "own" the decision to ensure its successful implementation. A procedural justice approach has been recommended for this purpose. Procedural justice involves a clear explanation of the decision-making process, setting realistic expectations, and ensuring engagement (Kim & Mauborgne, 1995). This means that the MCDM analyst should report how the group was formed to incorporate the different relevant perspectives particularly of those who will be affected (stakeholders) in addition to the results. Several MCDM methods allow aggregation of perspectives through specific techniques and the analysis of the extent of agreement or disagreement (dispersion) of the team members with respect to specific judgments or the final outcome. Furthermore, it is also possible to identify clear outliers or "extremists" to analyze what the decision would be without the participation of the outliers or even to give different weights to different perspectives or team members. Different decisions will involve the need for different MCDM group analysis techniques. Any MCDM study that involves group decision-making must explain the group (perspective) aggregation process, expectations (objectives), and team membership criteria, team engagement and dispersion of agreement on the decision among others.

Context refers to the environment in which a decision is made (Fig. 1). Internal and external factors influence the available choices and even the way the MCDM analysis is conducted. Contextual external factors play an important role in strategic and public decision-making. Some internal factors to consider at the individual level are risk aversion, experience, and education. Size, top management team, performance and collective mindfulness, among others should be considered at the organizational level. In the public sector, the broader context requires consultation or even participatory decision-making with the stakeholders involved. A stakeholder is defined as any group or individual who can affect the achievement or is affected by the achievement of the organization's objectives (Freeman, 1984; Freeman, 2010). In many countries, one key distinction between public and private sector management is that the former requires a great deal of transparency and interaction with a large number of stakeholders, many of whom are quite vocal and organized, and whose actions will play a decisive role in the adoption and public acceptance of proposed actions (Bryson, 2004). However, independent of how vocal or influential stakeholders are, there is an ethical consideration to inform and obtain feedback from the people and organizations that will be affected by the public decision. Furthermore, it is expected that with stakeholder participation the decision-making process benefits from the inclusion of various perspectives and wider acceptance of the public decision (OECD, 2001) Therefore, identifying stakeholders and incorporating them into the decision-making process either through direct consultation or with the inclusion of representatives that can communicate their perspectives is important. For this reason, it is recommended that stakeholder identification occurs at the beginning of the strategic planning cycle in the public context of a decision in order to form a team that represents and provides the different stakeholders' perspectives (Bryson 2004). This is not an easy task because in many cases, major decisions in public organizations involve more complexity, dynamism, intervention, and interruption than in private companies (Rainey, Ronquillo, et al., 2010). Another complication arises when the public decision involves an innovation because many of the critical stakeholders cannot be identified in advance. Mu and Stern (2012) proposed a self-identification approach based on the public participation literature to deal with this problem. Finally, ethical considerations, while present in most if not all decisions, become even more salient in public policy decision-making. The MCDM analyst should recognize this distinctive context and report, beyond the MCDA, the procedures and protocols that were followed concerning the public nature of the decision. Stakeholders should be identified and prioritized to justify either their inclusion or exclusion from the decision-making team. Environmental scanning, SWOT (Strengths, Weaknesses, Opportunities, and Threats), and pertinent public decision-making techniques should also be addressed. The MCDA context should be recognized to understand the decision constraints.

A classic tragic example of the importance of contingency factors is the decision to launch the Challenger space shuttle on the cold morning of January 28, 1986, despite the initial objections of the Morton-Thiokol engineers. This flawed decision led to the loss of seven US astronauts. A subsequent investigation determined that the rushed way in which the decision was made was highly influenced by the culture of risk that NASA had gradually adopted, more specifically, the normalization of deviance. Also, the broader context of political pressure to meet the launch schedule and the added public and media expectations surrounding the launch of the first civilian passenger, teacher Krista McAuliffe who lost her life along with the rest of the Challenger crew, influenced the flawed decision to launch (Vaughan, 1996).

In summary, a contingency approach that considers both the content and the context of the decision is important not only to provide rigor and relevance to the MCDM study, but more importantly, to provide validity to the overall decision-making process and outcomes. Recently, some efforts have been made to address the validity of MCDM analyses in terms of its decisional content and context (Mu, Cooper et al., 2020), as well as addressing MCDM applications while taking into account their contingency dimensions, but these efforts are still incipient (Nutt, 2008; Mu & Stern, 2018). For this reason, this chapter argues that it is important to consider the specific contingent factors in decision-making analysis and intends to explore to what extent MCDM studies are currently taking into account contingency factors in the context of group decision-making (target dimension) as shown in Fig. 1. In this figure, the list of elements in each category is not intended to be exhaustive. However, it is crucial to emphasize that contingent considerations are independent from the specific methodology used for the decision analysis as is also shown in Fig. 1 and emphasized by the fact that the MCDA method is encircled.

3 Research Design

A representative contingent dimension listed in Fig. 1, group decision-making, was chosen for this study because group decision-making is a widely used and well-studied contingency. In effect, group decision-making has received considerable attention in the literature with regard to theory and application and therefore addressing the main variables should be expected. Yu and Lai (2011) organized multi-criteria group decision-making into three stages. The first stage is where a subset of decision makers formulates a group decision-making framework. Second, a MCDM process is used to obtain the decision makers preferences. Third, the individual inputs are aggregated. Aggregation can be done concurrently with the group present where the group members agree on a common preference or fuzzy range. Likewise, many aggregation procedures can be used to combine individual preferences (Mohd & Abdullah 2017; Mardani, Nilashi, et al., 2018). One primary focus of the theoretical side of group decision-making has been measuring and improving group consensus (Fedrizzi, 1990; Herrera & Herrera-Viedma, 1996; Wu, Dai, et al., 2018). Group consensus refers to the general level of agreement in a group. It is also important to differentiate between consensus and conformity (Allen & Levine, 1969). Unfortunately, it is possible to achieve conformity and mistakenly consider it consensus which can lead to groupthink and result in flawed decisions as in the case of the Challenger explosion. Decision support systems have been developed to deal with these issues and improve group decision-making (DeSanctis

& Gallupe, 1987; Carneiro, Alves, et al., 2021). Other issues include trust (Chen, Yu, et al., 2020), conflict detection (Liu, Zhou, et al., 2019), the minimum cost of consensus (Cheng, Zhou, et al., 2018), and noncooperative behaviors (Palomares, Martinez, et al., 2013; Chao, Kou, et al., 2021). In summary, MCDM group decision-making has been largely studied and many of its contingent factors have been clearly identified.

The extent of the proper use of and reporting of contingent variables when performing MCDM group decisions was explored using a convenient sample that included papers published in a single year that integrated the ANP with other MCDM methods (e.g., TOPSIS, DEMATEL, PROMETHEE, VIKOR, and DEA). At least two concurrent methodologies were used in each study to avoid the results being attributed to a single methodology. Forty-six papers that used two or more concurrent methodologies were identified and reviewed[1] (Table 1). The median impact factor of all the journals with papers reviewed was 1.719 (Thompson, 2015). All of the papers reviewed contained at least two MCDM methods (with ANP being one of them). Also, at least two or more other MCDM methods were combined with ANP in 54.3% of the papers that were reviewed. The ANP was combined with 3 or more MCDM methods in 17.4% of papers. DEMATEL was the most common method that was combined with ANP with 56.5% of the papers using this method. Fuzzy was the next most common method that was applied with the ANP in 39.1% of the papers. The Delphi technique was combined with the ANP in 17.4% of the papers. The TOPSIS method was applied with the ANP in 13.0% of the paper. The VIKOR and balanced scorecard were both combined with the ANP in 10.9% of the papers.

The purpose of this study was to assess to what extent MCDM analysts are considering contingent variables in their analyses. Group decision-making was selected as the decision content type because the contingent variables related to group decision-making (e.g., team membership, group processes, aggregation of perspectives, group engagement, and group consensus) have been clearly identified and used for many years. Therefore, it would be expected that MCDM analysts would pay particular attention to these variables and customarily address them in their group decision-making analyses. The number of group decision-making variables to address is dependent on the context and content of a specific decision. Initially, six contingent variables were identified for this study. However, depending on the desired level of detail, it was also possible to consider a much larger set of variables including elements such as cooperative or noncooperative behaviors. Furthermore, expecting researchers to encounter and report on every single variable that can be reported in a group study would not be fair. Nor would every study have addressed items like noncooperative behavior. Therefore, this research focuses on the group decision-making dimension and four of its well-known and essential

[1] One of the methods was the Analytic Network Process (ANP) because the sample used for this study was a sample of convenience, which constituted a subset of an original larger ANP dataset corresponding to 2015, the year of the data collection.

Table 1 List of papers with integrated methods reviewed

Paper #	Citation	Paper title	Nature of the decision	MCDM method(s) in addition to ANP
1	Ramkumar and Jenamani (2015)	Sustainability in supply chain through e-procurement—An assessment framework based on DANP and Liberatore score	Identify 26 driving factors to evaluate the criteria that introduce sustainability into the supply chain through e-procurement system adoption	DEMATEL
2	Tavana, Khalili-Damghani, et al. (2015)	A hybrid fuzzy MCDM method for measuring the performance of publicly held pharmaceutical companies	A complex model to measure and evaluate the performance of publicly held pharmaceutical companies	Balanced Scorecard, Data Envelopment Analysis, DEMATEL, Fuzzy
3	Lin (2015)	A novel hybrid decision-making model for determining product position under consideration of dependence and feedback	A model to prioritize vehicle telematics systems product/services solutions	DEMATEL, Principal Component Analysis, VIKOR
4	Nilashi, Zakaria, et al. (2015)	MCPCM: A DEMATEL-ANP-based multi-criteria decision-making approach to evaluate the critical success factors in construction projects	Determine the critical success factors that impact construction	DEMATEL, Gray relational analysis
5	Wudhikarn, Chakpitak, et al. (2015)	Use of an Analytic Network Process and Monte Carlo analysis in new product formula selection decisions	A framework to capture uncertainties when selecting new roof tile product formulas	Monte Carlo simulations
6	Ortíz, Felizzola, et al. (2015)	A contrast between DEMATEL-ANP and ANP methods for six sigma project selection: a case study in healthcare industry	Selecting six sigma projects in the healthcare industry	DEMATEL

7	Uygun, Kaçamak, et al. (2015)	Readiness assessment model for institutionalization of SMEs using fuzzy hybrid MCDM techniques	Measure multidimensional institutionalization readiness of SMEs	DEMATEL, Fuzzy, TOPSIS
8	Uygun, Kahveci et al. (2015)	An integrated DEMATEL and fuzzy ANP techniques for evaluation and selection of outsourcing provider for a telecommunication company	Selection of an outsourcing provider for a telecommunications company	DEMATEL, Fuzzy
9	Zhang & Li (2015)	A hybrid performance evaluation model of TPL providers in agricultural products based on fuzzy ANP-TOPSIS	An evaluation index for third-party logistics providers for agricultural products	Fuzzy, TOPSIS
10	Chou (2015)	Multiple technique approach for improving performance measurement and management system: Action research in a mining company	A model to select and weight key performance indicators for a performance management system in the mining industry	Delphi
11	Chen & Chen (2015)	The assessment of intellectual capital for the information and communication technology industry in Taiwan applying a hybrid MCDM model	Identify the most valuable forms of intellectual capital in the information and communication technology industry	DEMATEL
12	Pan & Nguyen (2015)	Achieving customer satisfaction through product service systems	Identify the key performance metrics for manufacturing firms that implement product service systems	Balanced Scorecard, Delphi, DEMATEL
13	Lai (2015)	Using fuzzy Analytic Network Process to explore the factors of performance housing refurbishment in Taiwan	Prioritize the most important indicators when selecting which housing refurbishment projects to subsidize	Fuzzy

(continued)

Table 1 (continued)

Paper #	Citation	Paper title	Nature of the decision	MCDM method(s) in addition to ANP
14	Gupta & Narain (2015)	A fuzzy-based ANP approach in the selection of the best e-business strategy and to assess the impact of e-procurement on organizational performance	Select the best e-business strategy to increase organizational performance	Fuzzy
15	Kuleli Pak, Albayrak, et al. (2015)	Renewable energy perspective for Turkey using sustainability indicators	A framework to include sustainable and renewable perspectives in energy planning	TOPSIS
16	Abdollahi, Arvan, et al. (2015)	An integrated approach for supplier portfolio selection: Lean or agile?	A framework for supplier evaluation and selection	Data Envelopment Analysis, DEMATEL, Fuzzy
17	Kilic, Zaim, et al. (2015)	Selecting "The Best" ERP system for SMEs using a combination of ANP and PROMETHEE methods	Selecting an enterprise software with an emphasis on small- and medium-sized businesses	PROMETHEE
18	Moghaddam (2015)	Fuzzy multiobjective model for supplier selection and order allocation in reverse logistics systems under supply and demand uncertainty	Ranking suppliers and calculating the respective optimal order quantities	Fuzzy, Monte Carlo Simulation, Multi-objective Optimization
19	Ahmadi, Nilashi, et al. (2015)	Organizational decision to adopt hospital information system: An empirical investigation in the case of Malaysian public hospitals	Develop a model to identify the most important criteria when adopting a hospital information system	DEMATEL
20	He, Luo, et al. (2015)	Measuring the complexity of mega construction projects in China-A fuzzy Analytic Network Process analysis	Identify and measure the complexities in mega construction projects	Delphi, Fuzzy

21	Jeng & Huang (2015)	Strategic project portfolio selection for national research institutes	A model to systematically select research and development projects for national research institutes to fund	Delphi, DEMATEL
22	Zhao & Li (2015)	Evaluating the performance of thermal power enterprises using sustainability balanced scorecard, fuzzy Delphic and hybrid multi-criteria decision-making approaches for sustainability	Evaluate the sustainable performance of thermal power enterprises	Balanced Scorecard, Delphi, Fuzzy, TOPSIS
23	Kumru & Kumru (2015)	A fuzzy ANP model for the selection of 3D coordinate-measuring machine	A model to choose a 3D coordinate measuring machine for a manufacturing company	Fuzzy
24	Hu, Wen, et al. (2015)	Measuring the performance of knowledge resources using a value perspective: Integrating BSC and ANP	A framework to measure knowledge resources that can be developed through an organization's research and development	Balanced Scorecard
25	Chang, Liao, et al. (2015)	A hybrid MCDM approach for Taiwanese tour guides selection	A personnel selection model to select the optimal tour guides	Delphi, Fuzzy, TOPSIS
26	Lin, Wang, et al. (2015)	A multi-criteria decision-making for innovation services attributes: An empirical study of mobile banking system	Identify which services are most important to mobile banking consumers	DEMATEL, Zero-One Goal Programming
27	Hung Chen, Kang, et al. (2015)	Strategies, decisions and operations for keeping exploitative and exploratory activities balanced	A model to prioritize funding that integrates and balances exploitative and exploratory factors applied in the context of flat panel display manufacturing	Delphi, Fuzzy, Goal Programming

(continued)

Table 1 (continued)

Paper #	Citation	Paper title	Nature of the decision	MCDM method(s) in addition to ANP
28	Lu, Hu, et al. (2015)	Evaluating the implementation of business-to-business m-commerce by SMEs based on a new hybrid MADM model	Identify the most important factors to address when implementing business-to-business m-commerce	DEMATEL, VIKOR
29	Chuang, Lin, et al. (2015)	Exploring the triple reciprocity nature of organizational value cocreation behavior using multi-criteria decision-making analysis	A model to prioritize the factors that influence the value of cocreation behavior	DEMATEL
30	Sangari, Razmi, et al. (2015)	Developing a practical evaluation framework for identifying critical factors to achieve supply chain agility	Comprehensive model to identify the critical factors for achieving supply chain agility	DEMATEL, Fuzzy
31	Ju, Wang, et al. (2015)	Emergency alternative evaluation and selection based on ANP, DEMATEL, and TI-TOPSIS	A framework to select emergency response plans	DEMATEL, TOPSIS
32	Najafinasab, Karbassi et al. (2015)	Fuzzy analytic network process approach to evaluate land and sea criteria for land use planning in coastal areas	Integrate land and marine environments criteria in land use planning	DEMATEL, FUZZY
33	Masoumik, Abdul-Rashid, et al. (2015)	The development of a strategic prioritization method for green supply chain initiatives	A practical tool to prioritize green strategies	Structural Equation Modelling
34	Tsai, Xue et al. (2015)	Establishing a criteria system for green production	Determine the most important criteria to use when assessing how to accommodate profit generation and green production in high-tech manufacturing	DEMATEL

35	Kiakojuri, Shamshirband, et al. (2015)	Analysis of the social capital indicators by using Dematel approach: The case of Islamic Azad University	A model to identify the most important factors of social capital in a university that managers should focus on building	DEMATEL
36	Akyuz (2015)	A hybrid accident analysis method to assess potential navigational contingencies: The case of ship grounding	To identify the most crucial causal factors contributing to grounding accidents	Accident Analyse Mapping
37	Lu, Tzeng, et al. (2015)	Exploring mobile banking services for user behavior in intention adoption: Using new hybrid MADM model	Determine the best method to understand how to improve evaluations of mobile banking services for user behavior	DEMATEL, VIKOR
38	Shen & Tzeng (2015)	A decision rule-based soft computing model for supporting financial performance improvement of the banking industry	Provide a model to diagnose the financial performance improvement of commercial banks	DEMATEL, Dominance Based Rough Set, VIKOR
39	Hosseini, Banaitis, et al. (2015)	Combination of fuzzy-AHP and DEMATEL-ANP with GIS in a new hybrid MCDM model used for the selection of the best space for leisure in a blighted urban site	Integrate stakeholder input and GIS data to determine which site to develop into a leisure space	DEMATEL, FUZZY
40	Lee (2015)	Navigating SWOT-FANP with GSM method to prioritize the strategic location	A model for foreign companies to choose a location in second-tier cities in China	Fuzzy, Grand Strategy Method
41	Hsu, Shen, et al. (2015)	Toward successful commercialization of university technology: Performance drivers of university technology transfer in Taiwan	Identify critical performance drivers in university technology transfers	Delphi, Fuzzy

(continued)

Table 1 (continued)

Paper #	Citation	Paper title	Nature of the decision	MCDM method(s) in addition to ANP
42	Lam (2015)	Designing a sustainable maritime supply chain: A hybrid QFD-ANP approach	A framework to prioritize resources to ensure sustainable development in the maritime industry supply chain	Quality Function Deployment
43	Chang, Chen, et al. (2015)	Evaluation framework for alternative fuel vehicles: Sustainable development perspective	Evaluating alternative fuel vehicles according to sustainable development	DEMATEL
44	Kuo, Hsu, et al. (2015)	Developing a green supplier selection model by using the DANP with VIKOR	Evaluate green suppliers in an electronics company	DEMATEL, VIKOR
45	Staš, Lenort, et al. (2015)	Green transport Balanced Scorecard Model with Analytic Network Process support	The ability to evaluate the performance of green transportation and support the effective implementation of green transportation strategies	Balanced Scorecard
46	Yang, Yuan, et al. (2015)	Key determinant derivations for information technology disaster recovery site selection by the multi-criteria decision-making method	A framework to select an information technology disaster recovery site	DEMATEL

Source: Authors

minimum number of variables that should be reported in a group decision-making study (team membership, group processes, aggregation of perspectives, and group engagement) (Table 2). This choice of variables is consistent with the classic literature about group work that establishes that group composition (Hunt, 1992) and group process are critical for the group success (Elwyn, Greenhalgh et al., 2016, p. 24). Also, recent literature on group decision-making dynamics has shown that group engagement is extremely important for effective team performance (Lingham and Ritchley, 2018; Chao, Peng et al., 2021). Addressing these MCDM group decision-making variables is important because it affects the overall validity of the decision. The sample dataset that was examined will be discussed next.

4 Findings

Each of the 46 papers was reviewed using the four contingent variables listed in Table 2 corresponding to the group decision dimension. Proper reporting of these variables was rated on a scale from 0 to 2. A score of 0 meant the item was not addressed. For example, in some cases with respect to G1-Team Membership, the only information given regarding the team/group membership was a simple statement such as "evaluated by a group." In these cases, there was no reference to the number of group members, how they were selected, the group size, or any rationale for the team membership. A score of 1 was assigned if any information, even if potentially underreported, was provided regarding the metric. For example, if the paper included statements such as "a group of 5 experts" or a "group of 6 experts with an average of 4 years of experience." While it could be argued that such short statements convey no actual information about the group membership and practically these papers should not score above a 0, a very conservative approach was used for this study and these papers were given a score of 1. A score of 2 was assigned when a reasonable or very detailed report about group membership was provided. For example, in many cases, an article contained a table with multiple pieces of demographic data describing group members including their background, stakeholder representation, and the rationale for their selection, or statements were made about the group such as it included "two members from the executive team, one member from the production department, and two members from the supplier's management, each had a minimum of 4 years of experience." A summary of this contingent variable evaluation scale is provided in Table 3.

Table 3 shows that the expected situation if the MCDM analysts are addressing these well-known group decision considerations, would be to assign a score of 2 to all the contingent variables (reasonable or a very detailed report). The worst situation is certainly a score of 0, while a score of 1 is only slightly better. A score of 1 means that this variable was recognized, but was barely addressed; however, to evaluate conservatively these papers were assigned a value above 0.

As can be seen in Table 4, most studies provided some information, even if minimal or underreported, about the team membership (65.2%) and group process

Table 2 Group decision contingent variables checklist. Source: Authors

Item	Contingent Variable	Variable Definition	Group decisions contingent variables checklist	√
G1	Team membership	(Mu, Wormer,et al., 2012; Mu, Cooper, et al., 2020)	What was the rationale for team membership?	____
			How were the team members selected?	____
			What were the team member's qualifications, credentials, experience?	____
G2	Group process	(Saaty & Peniwati, 2013; Carneiro, Alves, et al., 2021)	Is there a description of the group process, was the decision reached via consensus, voting, etc.?	____
			What information were group members given about the decision-making methodology to be used?	____
G3	Aggregation of perspectives	(Mohd & Abdullah, 2017; Mardani, Nilashi, et al., 2018)	Were team member perspectives aggregated? At what level?	____
			How was the aggregation done: Geometric mean, arithmetic mean, weighted geometric mean, etc.?	____
G4	Engagement	(Fedrizzi, 1990; Palomares, Martinez, et al., 2013; Cheng, Zhou, et al., 2018; Chao, Kou, et al., 2021)	Is there any description of the protocol used to ensure team member engagement?	____
			Were details about the extent of dispersion (or consensus) of group members included?	____
			Were the decision makers asked to revise judgments or given any feedback? What method(s) was used in the revision process?	____
			Was there any mechanism in place to improve consensus or to deal with non-cooperative behaviors?	____

Table 3 Group decision contingent variable evaluation scale

Code	Description
0	Variable is not addressed at all
1	Any information, even if underreported, is provided
2	Reasonable or very detailed report is provided

Source: Authors

Table 4 Summary of findings

Code	Description	Team membership	Group process	Aggregation of perspectives	Engagement
0	Variable is not addressed at all	10.9%	13.0%	47.8%	73.9%
1	Any information, even if minimal or underreported, is provided	65.2%	69.6%	26.1%	19.6%
2	Reasonable or very detailed report is provided	23.9%	17.4%	26.1%	6.5%

Source: Authors

(69.6%). By definition, coding 1 for any reporting is extremely forgiving. Reporting on the aggregation of perspectives was even more scarce as shown in Table 4. Only 26.1% reported aggregation reasonably while a similar amount provided minimal information on the subject. Finally, the great majority of studies (73.9%) did not discuss any engagement considerations such as the extent of group consensus. These findings, as well as their implications for MCDM analysis will be discussed in the next section.

5 Discussion

The ability to obtain valid results and identify how generalizable they are depends on the knowledge that at a minimum the contingent variables in Figs. 2, 3, 4, and 5 have been thoughtfully considered and addressed during the decision-making process and then reported in the literature. A lack of reporting raises potential concerns about the validity of the results.

In group decision-making, it is important to know about the group composition to be able to make inferences about the validity and generalizability of the team's judgments (Mu, Wormer, et al., 2012). This is very important because the team members typically provide the criteria, alternatives, and judgments to evaluate and prioritize the alternatives. *However, only 23.9% of the published studies provided reasonable details about the group membership so that readers could make inferences about the generalizability of the results*. While 65.2% of the published papers provided minimal but very insufficient information about the

Fig. 2 G1 team membership (Note – 0 is none, 1 is some but insufficient and 2 is reasonable)

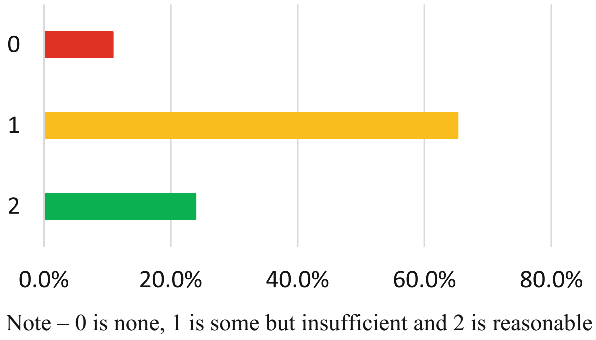


Fig. 3 G2 Group Process

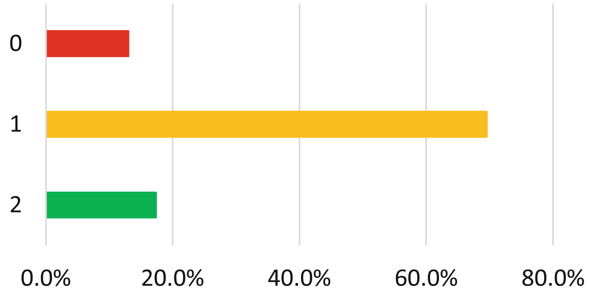


Fig. 4 G3 Aggregation of perspectives

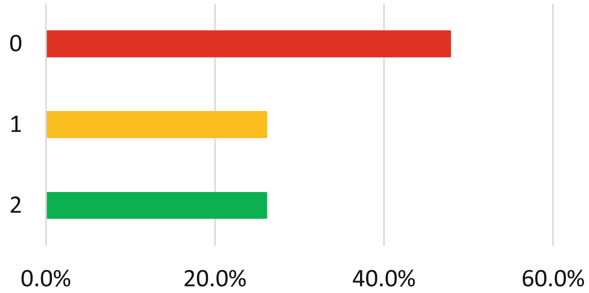
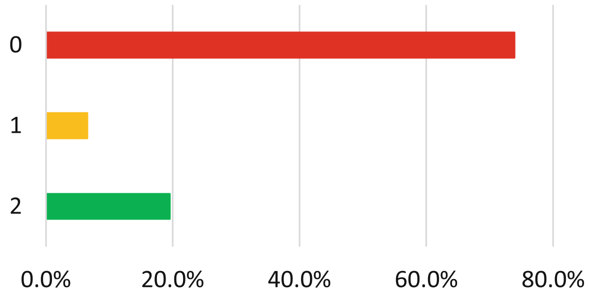


Fig. 5 G4 engagement



group members, 10.9% of the papers only included information about the group members such as "a group of experts." There was no information about the group size, selection criteria, group member qualifications, stakeholder representation, etc., or any rationale or justification for the decisions. Of the 65.2% that provided some information about the team membership, the information was still limited and insufficient such as "a group of 5 experts." Were the group members all academicians such as a faculty member and one graduate student? Were any of the group members from the industry? How much experience did they have? Were the members a convenient sample, probability sampling, or obtained through nonprobability sampling techniques like snowball sampling (Goodman 1961) or other sampling methods?[2] In terms of stakeholders, it is important to provide the details about which stakeholder groups were identified and included, what the individual stakeholder's qualifications are, and whether they represent the stakeholder group at large. This information should be relatively simple to assemble, would not require a significant amount of space in the article and is essential for the credibility and validity of the decision.

The importance of group process in the decision outcome validity has been widely discussed and accepted in the MCDM literature (Hunt, 1992; Saaty & Peniwati, 2013; Elwyn, Greenhalgh, et al., 2016). This process is important to the likelihood that team members accept the overall decision-making outcome as fair and take ownership of it (Eden & Ackermann, 2010). However, *only 17.4% of selected papers provided sufficient transparency about the group process to demonstrate a thoughtful consideration of the group process and the lack of information casts significant doubts on the validity of the findings.* As shown in Fig. 3, 13.0% of the studies that were reviewed were largely silent about the group process. There was no information about how the alternative priorities were obtained. The discussion was as limited as "a group of experts provided their input and the following priorities were obtained." What was the process for obtaining the experts' opinions? Another 69.6% of the studies provided some information such as "the experts provided their input by answering a survey." In these cases, no additional information was provided. Were the questions administered via an interview or self-administered? Were the decision makers familiar with the MCDM method in use? Were they given any instruction or training about the MCDM method and how to use it? Did the decision makers have any way to obtain feedback and review their judgments? Did the group meet collectively and agree on single judgments for each comparison?

Group decision-making requires aggregation of the members' perspectives which can be done in different ways (Mohd & Abdullah, 2017). Therefore, specifying how this aggregation occurs is very critical to understanding how the decision was reached and for the overall validity of the decision outcome. Dalio (2017) explains (p. 64) "I believe that for a group decision-making system to be effective, the people

[2] Snowball sampling is a nonprobability sampling technique that is used when participants are hard to obtain. Participants are recruited from prior participants social and professional networks.

using it have to believe that it's fair." This has important implications for the third and fourth contingent variables. It is important to explain to the group members at the outset what measures are in place to make sure the process is fair. This helps motivate them to participate and feel valued and enables them to feel confident their voices will be heard. Most studies were largely silent about what, if anything, the group did to aggregate perspectives, measure or address the group feedback, or about a revision process.

As shown in Fig. 4, only 26.1% of the papers reported this critical variable in a reasonable and proper way. Another 26.1% of the papers contained some, but insufficient information about how the group members' priorities were aggregated, while *a staggering 47.8% did not provide any information about this critical aggregation variable at the core of group MCDM.* For example, if the group members did not meet and collectively decide on every single judgment throughout the process, how were the individual priorities aggregated at the end? Were the priorities combined via the geometric mean, arithmetic mean, etc. (Mardani, Nilashi et al., 2018)? Subsequently, there needs to be not only communication but specific procedures in place to help the group receive feedback, improve, and reach a desirable level of consensus.

Group engagement is key to highly effective teams (Lingham & Richley, 2018). Lack of engagement will lead to only nominal participation of the team members without actual involvement in the decision process. While engagement may be difficult to measure, a proxy variable used in this study was the extent of agreement or dispersion reached by the group. Reporting details about the dispersion of the group members' priorities is also invaluable. This is particularly important because the analyst has many different methods to use to accomplish this (Herrera & Herrera-Viedma, 1996; Saaty & Vargas, 2007; Dong & Saaty, 2014; Herrera-Viedma, Cabrerizo, et al., 2014; Wu, Liu, et al., 2019). If and how decision makers were asked to revise their judgments is another crucial component of the decision-making process and should be reported. Measuring and reporting on the consensus is essential for the generalizability of the results. This can also provide external decision makers with relevant information for appropriate representation in future studies.

As shown in Fig. 5, *the majority of published studies in the sample, 73.9%, were silent about the level of group consensus and any procedures in place to measure the level of group consensus.* Of the remaining 26.1% of papers, those that included any discussion at all about group consensus, feedback, or a revision process were almost three times as likely to have addressed it well. This seems to imply that this is an area where individuals who recognize the importance of this issue are also well versed in how to address it.

This information is essential context to provide along with the results of the decision. For example, were group members' priorities widely dispersed so that there were no real groupings/clusters of priorities? While this could mean there is a diversity of strong opinions and experience, it could also mean there is no clarity about the drivers in the decision. Are there clusters of group members and do they share common characteristics like being suppliers or senior management? Or are the clusters essentially the different groups of stakeholders? This information can influence the mechanisms used to revise judgments and improve consensus. Is there a tight grouping of the decision makers with only a few outliers? In this type of situation, it is helpful to explore if there is an information bias. What is the experience and level of trust among the outliers? Are those members exhibiting noncooperative behaviors, and if so, how should they be dealt with? As group decision-making continues to evolve and mature, it is crucial to include these additional details to improve the quality of the publications.

6 Conclusions

This study has argued for the need to produce MCDM studies that have rigor and relevance. While rigor is mainly associated with the mathematical analysis, relevance is more related to the adequacy of the decision and its generalization. Both aspects are important to the validity of the MCDM study. MCDM validity can be defined as the support that the decision-making model within its range of applicability possesses, a satisfactory range of accuracy consistent with the intended application of the model (Mu, Cooper, et al., 2020). The analysis of contingent variables related to the nature of the decision itself (content) and its situational aspects (context) are important for the validity of the study. While these considerations are widely accepted (although not always followed) for strategic decision-making, it is argued here that all decisions (whether strategic or not) have contingent factors that should be considered. For example, while group decision-making may be used in either strategic or nonstrategic decisions, the important contingent variables for group decision-making are well-known, have been studied, and should be addressed in the studies.

To explore the extent to which MCDM analysts address contingent variables for group decision-making, a convenience sample of 46 integrated method papers (i.e., using two or more MCDM methods covering ANP and TOPSIS, DEMATEL, PROMETHEE, VIKOR, and DEA) corresponding to a single year of publication in Web of Science journals was examined with regard to the following four well-known group decision-making contingent variables: team membership, group process, aggregation of perspectives, and group engagement. The results, summarized in Fig. 6, showed that only 23.9%, 17.4%, 26.1%, and 19.6% of the studies satisfactorily addressed the proposed contingent variables. The other studies either ignored the contingent variables or mentioned them without really addressing them.

It is argued here that the conclusions identified in this study, incomplete or improper treatment of group decision-making contingent variables, are most likely generalizable to other contingent factors of MCDM decisions. Group decision-making was chosen because it is a widely used and studied decision-making approach and the factors that need to be addressed (e.g., team membership and aggregation of perspectives) in an MCDM study are well-known. This suggests that addressing other less established contingent dimensions would probably produce

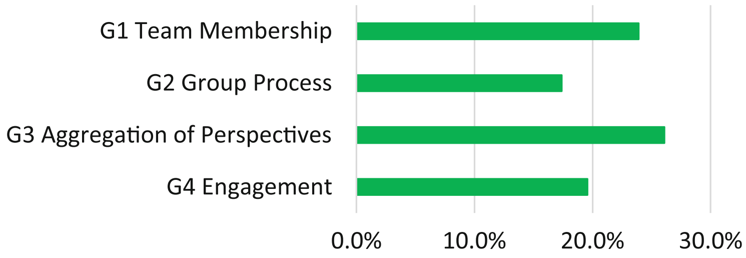


Fig. 6 Percent of studies satisfactorily addressing group decision-making contingencies

worse results than what is found in this study. Also, while the study was performed in a sample of convenience that combined ANP and other methodologies such as DEMATEL, TOPSIS, DEA, and others, the contingent dimensions are independent of the MCDM methodology used for the analysis (Fig. 1). Therefore, it is argued that the conclusions of this study are generalizable for other contingent factors as well as any other MCDM methodology.

The poor results could also be explained by a more general observation. While MCDM analysts may be rigorous with the mathematical underpinning of their analysis, they tend to be less rigorous with the overall context of the study, including complementary methodologies. For example, the majority of the group decisions in this sample used surveys to obtain participants' judgments. However, even though survey research has long and well-established best practices for implementation and reporting (e.g., number of surveys sent, received, valid, and response rate), most of the papers in our sample lacked the rigor and analysis that should accompany survey research, such as pretesting, reporting how the sample was obtained, response rate, sample questions (or including the survey as an appendix). It is an unfortunate situation for the MCDM community, which is not only the fault of the analysts but also of the journals that allow these studies to be published. The expectation of this study is to contribute to the improvement of MCDM studies and literature by calling MCDM authors and journals to be more rigorous in addressing the contingent factors in decision-making in order to make the studies more relevant.

References

Abdollahi, M., Arvan, M., & Razmi, J. (2015). An integrated approach for supplier portfolio selection: Lean or agile? *Expert Systems with Applications, 42*(1), 679–690.

Ahmadi, H., Nilashi, M., & Ibrahim, O. (2015). Organizational decision to adopt hospital information system: An empirical investigation in the case of Malaysian public hospitals. *International Journal of Medical Informatics, 84*(3), 166–188.

Akyuz, E. (2015). A hybrid accident analysis method to assess potential navigational contingencies: The case of ship grounding. *Safety Science, 79*, 268–276.

Allen, V. L., & Levine, J. M. (1969). Consensus and conformity. *Journal of Experimental Social Psychology, 5*(4), 389–399.

Baldridge, D., Floyd, S., & Markoczy, L. (2004). Are managers from Mars and academicians from Venus? *Strategic Management Journal, 25*, 1063–1074.

Bell, G., Bromley, P., & Bryson, J. (1998). Spinning a complex web between strategic decision making context, content, process and outcome. In V. Papadakis & P. Barwise (Eds.), *Strategic decisions*. Kluwer.

Bryson, J. M. (2004). What to do when stakeholders matter: Stakeholder identification and analysis techniques. *Public Management Review, 6*(1), 21–53.

Carneiro, J., Alves, P., Marreiros, G., & Novais, P. (2021). Group decision support systems for current times: Overcoming the challenges of dispersed group decision-making. *Neurocomputing, 423*, 735–746.

Chang, D. S., Chen, S. H., Hsu, C. W., Hu, A. H., & Tzeng, G. H. (2015). Evaluation framework for alternative fuel vehicles: Sustainable development perspective. *Sustainability, 7*(9), 11570.

Chang, K. L., Liao, S. K., & Chen, Y. C. (2015). A hybrid MCDM approach for Taiwanese tour guides selection. *Journal of Multiple-Valued Logic & Soft Computing, 25*(6), 605–621.

Chao, X., Kou, G., Peng, Y., & Viedma, E. H. (2021). Large-scale group decision-making with non-cooperative behaviors and heterogeneous preferences: An application in financial inclusion. *European Journal of Operational Research, 288*(1), 271–293.

Chen, J. K., & Chen, I. S. (2015). The assessment of intellectual capital for the information and communication technology industry in Taiwan applying a hybrid MCDM model. *European Journal of International Management, 9*(1), 88–107.

Chen, X.-G., Yu, G. F., Wu, J., & Yang, Y. (2020). A minimum trust discount coefficient model for incomplete information in group decision making with intuitionistic fuzzy soft set. *International Journal of Fuzzy Systems, 22*(6), 2025–2040.

Cheng, D., Zhou, Z., Cheng, F., Zhou, Y., & Xie, Y. (2018). Modeling the minimum cost consensus problem in an asymmetric costs context. *European Journal of Operational Research, 270*(3), 1122–1137.

Chou, H. H. (2015). Multiple-technique approach for improving a performance measurement and management system: Action research in a mining company. *Engineering Management Journal, 27*(4), 203–217.

Chuang, H. M., Lin, C. Y., & Chen, Y. S. (2015). Exploring the triple reciprocity nature of organizational value cocreation behavior using multicriteria decision making analysis. *Mathematical Problems in Engineering, 2015*, 15.

Dalio, R. (2017). *Principles: Life and work*. Simon and Schuster.

DeSanctis, G., & Gallupe, R. B. (1987). A foundation for the study of group decision support systems. *Management Science, 33*(5), 589–609.

Dong, Q., & Saaty, T. L. (2014). An analytic hierarchy process model of group consensus. *Journal of Systems Science and Systems Engineering, 23*(3), 362–374.

Eden, C., & Ackermann, F. (2010). Decision making in groups: Theory and practice. In D. Wilson (Ed.), *Nutt, P* (pp. 231–272). Handbook of Decision Making.

Elwyn, G., Greenhalgh, T., Macfarlane, F., & Koppel, S. (2016). *Groups: A guide to small group work in healthcare, management, education and research*. CRC Press.

Fedrizzi, M. (1990). On a consensus measure in a group MCDM problem. In J. Kacprzyk & M. Fedrizzi (Eds.), *Multiperson decision making models using fuzzy sets and possibility theory* (pp. 231–241). Springer.

Franklin, B. (1772). *Letter to Joseph priestly*. Fawcett.

Freeman, R. E. (1984). *Stakeholder management: A strategic approach*. Pitman.

Freeman, R. E. (2010). *Strategic management: A stakeholder approach*. Cambridge University Press.

Goodman, L. A. (1961). Snowball sampling. *The Annals of Mathematical Statistics, 32*, 148–170.

Gupta, M., & Narain, R. (2015). A fuzzy ANP based approach in the selection of the best e-business strategy and to assess the impact of e-procurement on organizational performance. *Information Technology and Management, 16*(4), 339–349.

Hatch, M. J. (1997). *Organization theory*. Oxford University Press.

He, Q., Luo, L., Hu, Y., & Chan, A. P. (2015). Measuring the complexity of mega construction projects in China—A fuzzy analytic network process analysis. *International Journal of Project Management, 33*(3), 549–563.

Herrera, F., & Herrera-Viedma, E. (1996). A model of consensus in group decision making under linguistic assessments. *Fuzzy Sets and Systems, 78*(1), 73–87.

Herrera-Viedma, E., Cabrerizo, F. J., Kacprzyk, J., & Pedrycz, W. (2014). A review of soft consensus models in a fuzzy environment. *Information Fusion, 17*, 4–13.

Hosseini, A., Banaitis, A., Nasiri, H., Banaitienė, N., & Tzeng, G.-H. (2015). Combination of fuzzy-AHP and DEMATEL-ANP with GIS in a new hybrid MCDM model used for the selection of the best space for leisure in a blighted urban site AU - Pourahmad, Ahmad. *Technological and Economic Development of Economy, 21*(5), 773–796.

Hsu, D., Shen, Y.-C., Yuan, B. J. C., & Chou, C. J. (2015). Toward successful commercialization of university technology: Performance drivers of university technology transfer in Taiwan. *Technological Forecasting and Social Change, 92*, 25–39.

Hu, Y., & Wen, J.& Yan, Y. (2015). Measuring the performance of knowledge resources using a value perspective: Integrating BSC and ANP. *Journal of Knowledge Management, 19*(6), 1250–1272.

Hung Chen, H., Kang, H.-Y., Lee, A. H. I., & Chen, S. (2015). Strategies, decisions and operations for keeping exploitative and exploratory activities balanced. *Knowledge Management Research & Practice, 13*(2), 198–213.

Hunt, J. (1992). *Groups in organizations. Managing people at work*. McGraw-Hill.

Jeng, D. J.-F., & Huang, K.-H. (2015). Strategic project portfolio selection for national research institutes. *Journal of Business Research, 68*(11), 2305–2311.

Ju, Y., Wang, A., & You, T. (2015). Emergency alternative evaluation and selection based on ANP, DEMATEL, and TL-TOPSIS. *Natural Hazards, 75*(2), 347–379.

Kasanen, E., Wallenius, H., & Wallenius, J.& Zionts, S. (2000). A study of high-level managerial processes, with implications for MCDM research. *European Journal of Operational Research, 120*, 496–510.

Kiakojuri, D., Shamshirband, S., Anuar, N. B., & Abdullah, J. (2015). Analysis of the social capital indicators by using DEMATEL approach: The case of Islamic Azad University. *Quality & Quantity, 49*(5), 1985–1995.

Kilic, H. S., Zaim, S., & Delen, D. (2015). Selecting "the best" ERP system for SMEs using a combination of ANP and PROMETHEE methods. *Expert Systems with Applications, 42*(5), 2343–2352.

Kim, W. C., & Mauborgne, R. A. (1995). A procedural justice model of strategic decision making: Strategy content implications in the multinational. *Organization Science, 6*(1), 44–61.

Koksalan, M., Wallenius, J., & Zionts, S. (2011). *Multiple criteria decision making: From early history to the 21st century*. World Scientific.

Kuleli Pak, B., Albayrak, Y. E., & Erensal, Y. C. (2015). Renewable energy perspective for Turkey using sustainability indicators. *International Journal of Computational Intelligence System's, 8*(1), 187–197.

Kumru, M., & Kumru, P. Y. (2015). A fuzzy ANP model for the selection of 3D coordinate-measuring machine. *Journal of Intelligent Manufacturing, 26*(5), 999–1010.

Kuo, T. C., Hsu, C.-W., & Li, J.-Y. (2015). Developing a green supplier selection model by using the DANP with VIKOR. *Sustainability, 7*(2), 1661.

Lai, P. P. Y. (2015). Using fuzzy analytic network process to explore the factors of performance housing refurbishment in Taiwan. *Acta Oeconomica, 65*(s2), 223–237.

Lam, J. S. L. (2015). Designing a sustainable maritime supply chain: A hybrid QFD–ANP approach. *Transportation Research Part E: Logistics and Transportation Review, 78*, 70–81.

Lee, Y. H. (2015). Navigating SWOT-FANP with GSM method to prioritize the strategic location. *Technological and Economic Development of Economy, 21*(1), 140–163.

Lin, C. L. (2015). A novel hybrid decision-making model for determining product position under consideration of dependence and feedback. *Applied Mathematical Modelling, 39*(8), 2194–2216.

Lin, W. R., Wang, Y. H., & Shih, K. H. (2015). A multi-criteria decision making for innovation services attributes: An empirical study of mobile banking system. *Journal of Testing and Evaluation, 43*(4), 20130220.

Lingham, T., & Richley, B. A. (2018). *High-impact engagement: A two-phase approach to individual and development ream*. iUniverse.

Liu, B., Zhou, Q., Ding, R. X., Palomares, I., & Herrera, F. (2019). Large-scale group decision making model based on social network analysis: Trust relationship-based conflict detection and elimination. *European Journal of Operational Research, 275*(2), 737–754.

Lu, M.-T., Hu, S.-K., Huang, L.-H., & Tzeng, G.-H. (2015). Evaluating the implementation of business-to-business m-commerce by SMEs based on a new hybrid MADM model. *Management Decision, 53*(2), 290–317.

Lu, M.-T., Tzeng, G.-H., Cheng, H., & Hsu, C.-C. (2015). Exploring mobile banking services for user behavior in intention adoption: Using new hybrid MADM model. *Service Business, 9*(3), 541–565.

Mardani, A., Nilashi, M., Zavadskas, E. K., Awang, S. R., Zare, H., & Jamal, N. M. (2018). Decision making methods based on fuzzy aggregation operators: Three decades review from 1986 to 2017. *International Journal of Information Technology & Decision Making, 17*(02), 391–466.

Masoumik, S. M., Abdul-Rashid, S. H., & Olugu, E. U. (2015). The development of a strategic prioritisation method for green supply chain initiatives. *PLoS One, 10*(11), e0143115.

Moghaddam, K. S. (2015). Fuzzy multi-objective model for supplier selection and order allocation in reverse logistics systems under supply and demand uncertainty. *Expert Systems with Applications, 42*(15–16), 6237–6254.

Mohd, W. R. W., & Abdullah, L. (2017). Aggregation methods in group decision making: A decade survey. *Informatica, 41*(1), 71–86.

Mu, E., Cooper, O., & Peasley, M. (2020). Best practices in analytic network process studies. *Expert Systems with Applications, 159*(30), 113536.

Mu, E., & Stern, H. (2012). Structured stakeholder self-identification approach for the deployment of public information systems: The case of surveillance technology in the City of Pittsburgh. *Journal of Information Technology Management XXIII, 4*, 50–66.

Mu, E., & Stern, H. (2014). The City of Pittsburgh goes to the cloud: A case study of cloud strategic selection and deployment. *Journal of Information Technology Teaching Cases, 4*, 70–85.

Mu, E., & Stern, H. (2018). A contingent/assimilation framework for public IS inter-organizational decisions. *International Journal of Information Technology and Decision Making, 17*(6), 1611–1658.

Mu, E., Wormer, S., Barkon, B., Foizey, R., & Vehec, M. (2012). Group modelling and integration of multiple perspectives in the functional selection of a new technology: The case of a next-generation electronic portfolio system. *Journal of Multi-Criteria Decision Analysis, 19*(1–2), 15–31.

Najafinasab, F., Karbassi, A. R., & Ghoddousi, J. (2015). Fuzzy analytic network process approach to evaluate land and sea criteria for land use planning in coastal areas. *Ocean & Coastal Management, 116*, 368–381.

Nilashi, M., Zakaria, R., Ibrahim, O., Majid, M. Z. A., Zin, R. M., & Farahmand, M. (2015). MCPCM: A DEMATEL-ANP-based multi-criteria decision-making approach to evaluate the critical success factors in construction projects. *Arabian Journal for Science and Engineering, 40*(2), 343–361.

Nutt, P. C. (2008). Investigating decision making processes. *Journal of Management Studies, 45*(2), 425–455.

OECD. (2001). *Citizens as partners: Information, consultation and public participation in policy-making*. OECD Publishing.

Ortíz, M. A., Felizzola, H. A., & Isaza, S. N. (2015). A contrast between DEMATEL-ANP and ANP methods for six sigma project selection: A case study in healthcare industry. *BMC Medical Informatics and Decision Making, 15*(3), S3.

Palomares, I., Martinez, L., & Herrera, F. (2013). A consensus model to detect and manage noncooperative behaviors in large-scale group decision making. *IEEE Transactions on Fuzzy Systems, 22*(3), 516–530.

Pan, J.-N., & Nguyen, H. T. N. (2015). Achieving customer satisfaction through product–service systems. *European Journal of Operational Research, 247*(1), 179–190.

Papadakis, V., Thanos, I., & Barwise, P. (2010). Research on strategic decisions: Taking stock and looking ahead. In P. Nutt & D. Wilson (Eds.), *Hanbook of decision making* (pp. 31–69). John Wiley and Sons.

Papadakis, V. M., Lioukas, S., & Chambers, D. (1998). Strategic decision-making processes: The role of management and context. *Strategic Management Journal, 19*(2), 115–147.

Rainey, H. G., Ronquillo, J. C., & Avellaneda, C. N. (2010). Decision making in public organizations. In P. Nutt & D. Wilson (Eds.), *Handbook of decision making* (pp. 349–378). John Wiley and Sons.

Ramkumar, M., & Jenamani, M. (2015). Sustainability in supply chain through e-procurement—An assessment framework based on DANP and liberatore score. *IEEE Systems Journal, 9*(4), 1554–1564.

Saaty, T. L., & Peniwati, K. (2013). *Group decision making: Drawing out and reconciling differences*. RWS Publications.

Saaty, T. L., & Vargas, L. G. (2007). Dispersion of group judgments. *Mathematical and Computer Modelling, 46*(7–8), 918–925.

Sangari, M. S., Razmi, J., & Zolfaghari, S. (2015). Developing a practical evaluation framework for identifying critical factors to achieve supply chain agility. *Measurement, 62*, 205–214.

Shen, K.-Y., & Tzeng, G.-H. (2015). A decision rule-based soft computing model for supporting financial performance improvement of the banking industry. *Soft Computing, 19*(4), 859–874.

Staš, D., Lenort, R., Wicher, P., & Holman, D. (2015). Green transport balanced scorecard model with analytic network process support. *Sustainability, 7*(11), 15243.

Tavana, M., Khalili-Damghani, K., & Rahmatian, R. (2015). A hybrid fuzzy MCDM method for measuring the performance of publicly held pharmaceutical companies. *Annals of Operations Research, 226*(1), 589–621.

Thompson, R. (2015). InCites Journal Citation Reports.

Tsai, S.-B., Xue, Y.-Z., Huang, P.-Y., Zhou, J., Li, G. D., Guo, W.-F., Lau, H., & Shang, Z.-W. (2015). Establishing a criteria system for green production. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 229*(8), 1395–1406.

Uygun, Ö., Kaçamak, H., & Kahraman, Ü. A. (2015). An integrated DEMATEL and fuzzy ANP techniques for evaluation and selection of outsourcing provider for a telecommunication company. *Computers & Industrial Engineering, 86*, 137–146.

Uygun, Ö., Kahveci, T. C., Taşkın, H., & Piriştine, B. (2015). Readiness assessment model for institutionalization of SMEs using fuzzy hybrid MCDM techniques. *Computers & Industrial Engineering, 88*, 217–228.

Vaughan, D. (1996). *The challenger launch decision: Risky technology, culture and deviance at NASA*. University of Chicago Press.

Wallenius, J., Dyer, J. S., Fishburn, P. C., Steuer, R. E., Zionts, S., & Deb, K. (2008). Multiple criteria decision making, multiattribute utility theory: Recent accomplishments and what lies ahead. *Management Science, 54*(7), 1336–1349.

Wu, J., Dai, L., Chiclana, F., Fujita, H., & Herrera-Viedma, E. (2018). A minimum adjustment cost feedback mechanism based consensus model for group decision making under social network with distributed linguistic trust. *Information Fusion, 41*, 232–242.

Wu, T., Liu, X., Qin, J., & Herrera, F. (2019). Consensus evolution networks: A consensus reaching tool for managing consensus thresholds in group decision making. *Information Fusion, 52*, 375–388.

Wudhikarn, R., Chakpitak, N., & Neubert, G. (2015). Use of an analytic network process and Monte Carlo analysis in new product formula selection decisions. *Asia-Pacific Journal of Operational Research, 32*(02), 1550007.

Yang, C.-L., Yuan, B. J. C., & Huang, C.-Y. (2015). Key determinant derivations for information technology disaster recovery site selection by the multi-criterion decision making method. *Sustainability, 7*(5), 6149.

Yu, L., & Lai, K. K. (2011). A distance-based group decision-making methodology for multi-person multi-criteria emergency decision support. *Decision Support Systems, 51*(2), 307–315.

Zhang, S., & Li, X. (2015). A hybrid performance evaluation model of TPL providers in agricultural products based on fuzzy ANP-TOPSIS. *Custos e Agronegocio online, 11*(3), 144–165.

Zhao, H., & Li, N. (2015). Evaluating the performance of thermal power enterprises using sustainability balanced scorecard, fuzzy Delphic and hybrid multi-criteria decision making approaches for sustainability. *Journal of Cleaner Production, 108*, 569–582.

Analyzing the Interaction of Renewable Energy Penetration with the Wealth of Nations Using Bayesian Nets

Mine Isik, Özay Özaydın, Şule Önsel Ekici, and Y. Ilker Topcu

Abstract Recently, countries are trying to improve their economies while increasing the number of positive steps they are taking against climate change and minimizing greenhouse carbon emission. However, this effort is futile unless these countries turn their attention to renewable energy sources. The shift from conventional energy to renewable energy will contribute to economic growth, employment opportunities, and human welfare while meeting climate goals in the long-term. In this study, using the data provided mainly by The World Bank (WB) and The International Renewable Energy Agency (IRENA), the authors aim to construct a Bayesian Network to analyze the interaction of renewable energy penetration with the wealth of nations. In this context, initially, the factors related to Renewable Energy will be determined, and then a Bayesian Network is going to be developed. Using multiple what-if analyses, the resulting model will act as a diagnostic tool for policymakers in their attempts to understand and manage the renewable energy system. The what-if analyses conducted from the resulting model show that if renewable energy consumption increases and fossil fuel energy consumption decreases, CO_2 intensity as well as health expenditures will be expected to decrease. Similarly several other scenarios are constructed and reflected in the study.

Keywords Renewables · Energy · Bayesian network

M. Isik (✉)
Independent Researcher, London, UK

Ö. Özaydın
Department of Industrial Engineering, Dogus University, Istanbul, Turkey
e-mail: oozaydin@dogus.edu.tr

Ş. Ö. Ekici
Department of Industrial Engineering, Istanbul University-Cerrahpasa, Istanbul, Turkey
e-mail: sule.ekici@iuc.edu.tr

Y. I. Topcu
Department of Industrial Engineering, Istanbul Technical University, Istanbul, Turkey
e-mail: ilker.topcu@itu.edu.tr

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_20

1 Introduction

A cleaner future is a common goal for many people, organizations, governments, and societies. Reaching this goal requires a collaborative effort involving all stakeholders including scientists, urban planners, businesspeople, investors, government officials, and policy makers. The required multi-dimensional approach includes perspectives of energy, transportation, production, and food. Among these perspectives, energy plays a pivotal role for reaching environmental goals. It is indisputable that there should be a global switch to 100% clean power, which in turn can be achieved through renewable energy.

Global energy-related CO_2 emissions have been exacerbated by the fuel consumption and reached its historic high level (33.1 Gt CO_2) (*Global Energy & CO2 Status Report*, 2019). However, the power sector alone accounted for the highest share, with two-thirds of the emissions growth. Climate change is one of the most significant global threats. Unless actions at a scale are realized, the emission level continues to expand. Hence, the concerns regarding global warming have given rise to the need for low carbon transition and sustainable development of the global energy systems. In that respect, almost every country bears a certain degree of responsibility to abate their emissions. Accordingly, worldwide efforts have been directed toward minimizing the reliance on fossil-based energy sources by increasing the penetration of renewable options. As a response, governments have defined key targets for their energy sector to enhance sustainability and limit global CO_2 emissions all around the World. Although the energy sector related emissions have started changing in a desired direction, statistics confirm the insufficiency of the existing efforts. As response, policy makers started to declare even more aggressive emission targets including the European Green Deal and achieving net-zero. Renewable energy sources constitute an integral part of those targets, to achieve the rapid emissions reductions needed. Hence, the role of renewable energy is becoming even more significant.

According to the Intergovernmental Panel on Climate Change (IPCC)'s report limiting average global temperature increases to 1.5 °C is compulsory to avert the disruption of the climate system (Masson-Dellmotte et al., 2018). It necessitates an immediate transformation of the power sector leading to a low carbon and even zero-carbon generation. Although renewable energy use constitutes a significant portion of the electricity generation globally, this transformation requires increased use of renewables higher than the existing share.

In 2015, the United Nations declared "the deployment of renewable energy" as one of the sustainable development goals (Singh et al., 2019). Besides playing an essential role in emission mitigation, the transition of renewable energy technologies regarded as a promising opportunity for economic development. It can bring numerous social and economic benefits to the wealth of nations. Renewable energy serves economic independencies in terms of imports reduction.

Investment for renewable energy technologies holds great potential for job creation (Cai et al., 2011). Many of the declared energy transformation scenarios project a considerable job growth especially due to solar photovoltaic, bioenergy, and wind (IRENA, 2020). According to the statistics more than 11.5 million people work in the renewable energy industry worldwide in 2019 (*Number of Renewable Energy Jobs Worldwide from 2012 to 2019*, 2020). IRENA report (2020) indicates that under the transforming energy scenario, renewable energy sector is expected to stimulate the even more intense job growth. Accordingly, the employment in renewable energy sector can reach up to 42 million jobs by 2050 among every level of the energy sector varying from white collar engineers to technicians, factor and construction workers.

Renewable energy deployment can also create additional income for rural landowners and farmers which boost sustainable rural development (Benedek et al., 2018). IRENA report also indicates that the benefits of renewable energy deployment will not be limited to job creation and GDP but also social welfare due to decrease in health expenditure and increase in the spending on education (IRENA, 2020). Penetration of renewable energy sources can directly transform the electricity generation (fossil share in the electricity generation, related air and greenhouse gas emissions, reliability factor, price of electricity, etc.) and the energy sector as a whole. Besides, renewables have significant impact on the electricity demanded by other sectors, the transformation in the energy sector bears great potential also to boost material production and consumption and the investment level in related sectors.

While scientific articles, policy papers and reports prepared by international organizations envision a positive impact of renewable energy sources on economic development, factor based quantitative impact of the renewable energy transition is not clear. Energy industry by nature affects numerous other factors such as social, economic, and environmental. Any changes in energy-related parameters have a direct and indirect effect on the whole system.

To analyze this complex system, in this study, we aim to shed light on the impact of renewable energy deployment on the wealth of nations by utilizing Bayesian Networks (BNs). The World Bank includes four categories into the definition of wealth, namely; produced capital, natural capital, human capital, and net foreign assets (Lange et al., 2018). BNs are useful tools to study and analyze complex systems by establishing cause and effect relationships between the factors included in the model relying on the information available. BNs can question the effect of not only one variable on another variable but also observe the effect of many variables changing at once on either one variable or the whole system. Additionally, BNs are visual which makes analyses easier to comprehend.

The organization of this paper is as follows. The next section provides an overview of renewable energy options and their relationship with the wealth of nations. Section 3 gives insight into the comprehensive data compilation and review methodology adopted in this study and the proposed model. The results are presented and discussed in Sect. 4. The final section concludes the study with a discussion of major findings together with further suggestions.

2 Renewable Energy and Wealth of Nations

The concerns regarding global warming give rise to the declaration of emission mitigation action plans all around the world during recent years. Due to its environmental benefits, renewable energy options are placed in the center of those action plans globally. Although its positive impacts on greenhouse gas emissions are beyond dispute, the benefits of renewable energy sources are not limited to this. Renewable energy options can contribute to the Earth's future in many different ways.

The impact of renewable energy on GDP has been examined abundantly in the literature. Majority of these studies adopted multi country approach by analyzing G7 countries (Ike et al., 2020; Zafar et al., 2019), BRICS countries (Bloch et al., 2012; Liu et al., 2020), OECD countries (Cheng et al., 2019; Dogru et al., 2020), or randomly selected heterogeneous countries (Mohsin et al., 2021; Saidi & Omri, 2020).

Speaking of the studies evaluating renewable energy consumption from a broader perspective, Ali et al. (2020) investigated the long-run causal relationship between total reserves, energy consumption in general, renewable energy use in particular and GDP (Ali et al., 2020). They collected the country specific data for 100 different countries covering the period from 1995 to 2017. They revealed that the hypothesis, claiming these variables have a long-run causal impact on GDP, is only valid in politically free countries. Bhattacharya et al. (2017) investigated the financial advantages resulted from renewable energy consumption in 85 countries and reported the positive impact of renewables on economic growth (Bhattacharya et al., 2017).

In the literature, there exist several studies that want to explore the impact of renewable energy on national economic development. Magazzino et al. (2021) focus solely on solar and wind production rather than covering every available renewable energy option. They tried to find the causal relation among electricity generation from solar and wind, coal consumption, economic growth, and CO_2 emissions by considering single-country data (Magazzino et al., 2021). Shahzad et al. (2020) employed the autoregressive distributed lag model for Pakistan and demonstrated that renewable energy consumption could significantly reduce health expenditures (Shahzad et al., 2020).

Besides geographical coverage, studies are diversified with respect to the variables included. Although the impact of renewables on GDP has been widely investigated, the literature's grounding body also emphasizes the potential of renewable energy on job creation (Cuesta et al., 2020). Akella et al. (2009) mentioned that renewable energy system has a potential to create comparatively more jobs per each unit of currency invested than conventional generation options (Akella et al., 2009). Bulavskaya and Reynès (2018) also develop a neo-Keynesian Computable General Equilibrium Model (CGAM) for the Netherlands (Bulavskaya & Reynès, 2018). They investigated the key macroeconomic impacts of shifting the electricity generation mix toward renewable sources and they project more than 50,000 jobs by 2030. IRENA report (2020) also declares that every region globally is expected to experience the relative employment gain varying from 10% to 115% (which also outweigh the job loss in the fossil-fuel based sectors). Zakeri et al. (2018) analyze the renewable energy integration of the Nordic-UK power markets through a power transmission line and they concluded that regional overall social welfare will be expected to improve by approximately 230 million euro annually (Zakeri et al., 2018). Konstantelos et al. (2017) also indicate that especially integrated offshore wind energy is expected to increase social welfare as a result of asset sharing and related investment cost reduction. There also exist some controversial studies that identify the potential negative impacts of a high level of renewable energy penetration (Konstantelos et al., 2017). While considering those effects, Ahn et al. (2021) focus on social welfare. They find that the increase of uncertainty level due to the energy policies leads to a higher risk in return on investments and proportionally has a net negative impact on social welfare. While they just focus on the U.S. energy sector, in the long run, they show that even 10% increase in the share of renewable energy within the overall mix can create 0.8% additional social welfare. They claim that the negative effect of economic uncertainty outweighs the gains of reduced damage on the climate (Ahn et al., 2021).

Caglar (2020) examines the relationship between renewable energy consumption and foreign direct investment besides non-renewable energy consumption, economic growth, and carbon emissions by deploying a bootstrap autoregressive distributed lag approach. Although he employs the algorithm for nine different countries with the highest Climate Change Performance index, the study identifies a significant relationship in some countries (Caglar, 2020). Khan et al. (2021) also investigate the same relationship for 69 countries of the Belt and Road Initiative between 2000 and 2014 by deploying robust standard error regression and dynamic generalized method of moments estimators (Khan et al., 2021). Their study shows that foreign direct investments have shown to be a negative determinant of renewable energy. Apart from economic and social benefits, Nazir et al. (2020) demonstrate other potential impacts of the deployment of clean-energy solutions such as reduced dependence on imported petrochemical fuels, positive market impact and technological impact including high system performance (Nazir et al., 2020). They also mention that renewable energy sites especially wind energy can be classified as tourist attractions. The same study also points out the adverse impacts of renewable energy deployment. Building renewable energy sites and related activities such as road construction and digging can harm landscape conservation. Speaking of wind farms, they also mention the adverse visualization impacts, soil erosion, and deforestation problems. Sarpong et al. (2020) investigate the impact of renewable energy consumption on quality of life by implementing panel data analysis in a time frame between 1995 and 2017 (Sarpong et al., 2020). They show that quality of life is negatively associated with renewable energy consumption for 8 Southern African countries.

From a technical, environmental, sociological, and economic point of view, a plethora of the existing literature focuses on renewable energy deployment benefits. However, there exist other studies that argue the potential risks associated with the exacerbated deployment of clean-energy resources. Statistics reveal the urgent need for energy transition. Hence, all the extreme decarbonization goals rely on renewable energy sources. Renewable energy can affect economic patterns, social institutions and the environment at a global level. This research aims to investigate the combined effects of renewable energy penetration and national wealth on a number of factors. These impacts are explored through conditional dependencies using Bayesian Nets.

3 The Methodology

In this paper, we used Bayesian Nets to analyze the complex structure of renewable energy system. BNs act as multivariate models and they can be used effectively under uncertainty. They are used in several studies ranging from health and medicine to social sciences (Ekici & Ekici, 2016). As a type of graphical models, they use joint probability distribution for the variables of the system. They do not only help the decision-makers to analyze both direct and indirect relations between various variables of a system, but also help to conduct what-if analyses by using conditional probabilities.

The advantage of BN compared to regression models is that the correlation between variables causes multicollinearity in regression, whereas correlation is used to define the conditional probability within systems of interconnected variables in BNs (Sebastiani & Perls, 2008). This property of BNs helps us to use them as effective tools to make both predictive and diagnostic reasoning within the system (Anderson & Vastag, 2004; Lauría & Duchessi, 2007). Moreover, BNs also deal with nonlinearity. What-if analysis can be conducted in a very efficient way.

A BN is a directed acyclic graph (Borunda et al., 2016). There are three main components in a BN: (1) nodes that represent variables, (2) directed arrows that represent the dependency relation between nodes in the acyclic graphical network, and (3) tables that quantifies the conditional probabilities between nodes. Assume that we have two variables: X_1 and X_2. If there is an arrow from X_1 to X_2, that means that X_1 “influences” X_2. We name X_1 as the parent of X_2 and X_2 as the child of X_1. If a node has no parent node, it is defined through a prior probability distribution (Korb & Nicholson, 2010). Nodes with parents are defined through conditional probability distributions. So, if we deal with a BN having n variables, a specific value in the joint distribution is characterized by $P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$. The chain rule of probability theory allows us to factorize joint probabilities as it is given in the formula (1).

$$\begin{aligned} P\left(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\right) &= P\left(x_1, x_2, \ldots, x_n\right) \\ &= P\left(x_1\right) \cdot P\left(x_2|x_1\right) \cdot \cdots \cdot P\left(x_n|x_1, \ldots, x_{n-1}\right) = \prod_i P\left(x_i/x_1, \ldots, x_{i-1}\right) \end{aligned} \tag{1}$$

The structure of a BN indicates that the value of a specific node is conditional only on the values of its parent nodes, that is:

$$P\left(x_1, x_2, \ldots, x_n\right) = \prod_i P\left(x_i / Parents\left(x_i\right)\right) \tag{2}$$

BN methodology consists of three stages: (1) determination of variables, (2) determination of network structure, and (3) determination of parameters to quantify the relations between variables.

We used the same procedure in our study. Initially, the variables of the system were determined by a literature survey. Next, a BN was constructed through structural learning using GeNIe software (*GeNIe*, 2020). Then, parameter learning was conducted and in the last step, several sensitivity analyses as well as what-if analyses were conducted to investigate the effects of the changes that occur in the system in detail.

3.1 Determination of the Variables

As it is stated when constructing the BN, the first step is to determine the variables. We have analyzed the current literature about renewable energy sources, mainly focusing on the International Renewable Energy Agency (IRENA) publications. IRENA works for international co-operation and a driver of action for the transformation of the global energy system. According to the "Renewable Energy Benefits: Measuring The Economics" report, published by IRENA, Ferroukhi et al. (2016) focus on the link between macroeconomic variables and renewable energy sources and indicate that besides its importance for climate change, renewable energy usage also has great socio-economic benefits that must be considered (Ferroukhi et al., 2016). The increase in the share of renewable energy sources is expected to create an increase not only in GDP and human welfare but also in direct and indirect employment. That is why, in this study the variables listed in Table 1 were taken into consideration to analyze the renewable energy effect on socio-economic status of countries.

Most of the data explained above were obtained from the World Bank databases, the codes for accessing these are given in the Table 1 for easier access. Initially, the time period was selected as 2007–2020 and number of countries included was 151. As the analyses need a complete dataset, the resulting time period was 2007–2012 including 91 countries.

We initially discretized our data into five homogeneous states so that each state (very low, low, medium, high, and very high) has approximately equal amounts of

Table 1 List of variables

Variable name	Abbreviation	Worldbank Data Code
Economic aspects		
GDP per capita	GDP	NY.GDP.PCAP.KD
Households and NPISHs final consumption expenditure per capita	HCE	NE.CON.PRVT.PC.KD
Gross capital formation	GCF	NE.GDI.TOTL.KD
Net trade in goods and services	NTIGAS	BN.GSR.GNFS.CD
Fuel exports	FE	TX.VAL.FUEL.ZS.UN
Fuel imports	FI	TM.VAL.FUEL.ZS.UN
Employment to population ratio	ETPR	SL.EMP.TOTL.SP.ZS
Energy imports	EIN	EG.IMP.CONS.ZS
Environmental aspects		
Total greenhouse gas emissions	TGHGE	EN.ATM.GHGT.KT.CE
Fossil fuel energy consumption	FFEC	EG.USE.COMM.FO.ZS
Alternative and nuclear energy	AANE	EG.USE.COMM.CL.ZS
Renewable energy consumption	REC	EG.FEC.RNEW.ZS
Electric power consumption	EPC	EG.USE.ELEC.KH.PC
CO_2 emissions from electricity and heat production	CO2EFEAHP	EN.CO2.ETOT.ZS
Electricity production from renewable sources excl hydroelectric	EPFRSEHE	EG.ELC.RNWX.ZS
Electricity production from oil gas and coal sources	EPFOACS	EG.ELC.FOSL.ZS
Electricity production from nuclear sources	EPFNS	EG.ELC.NUCL.ZS
Electricity production from natural gas sources	EPFNGS	EG.ELC.NGAS.ZS
Electricity production from hydroelectric sources	EPFHES	EG.ELC.HYRO.ZS
Electricity production from coal sources	EPFCS	EG.ELC.COAL.ZS
CO_2 intensity	CO2I	EN.ATM.CO2E.EG.ZS
Social aspects		
Current health expenditure	CHE	SH.XPD.CHEX.GD.ZS
Government expenditure on education, total (% of GDP)	GEOE	SE.XPD.TOTL.GD.ZS

data samples. Netica software does this discretization process automatically. Table 2 shows the variables and their related state ranges.

3.2 The Bayesian Net

There are mainly two ways to construct a BN: (1) by using expert judgement to determine the conditional relations between variables and (2) by using data and with the help of software packages to conduct specific learning algorithms.

Table 2 Variables and their related state ranges

Parameter	Very Low	Low	Medium	High	Very High
GDP (x1,000)	0–2.8	2.8–6	6–14	14–43	43–103
HCE (x1,000)	0–1.8	1.8–4	4–8	8–23	23–41
GCF (x1,000,000,000)	0–6	6–20	20–50	50–120	120–145
ETPR	0–51	51–56	56–59	59–64	64–86
CHE	0–4	4–6.2	6.2–7.7	7.7–9.1	9.1–11.4
GEOE	0–3.4	3.4–4.4	4.4–5	5–5.9	5.9–9.7
TGHGE (x1,000)	0–29	29–60	60–110	110–370	370–301
NTIGAS	0 to −7	−7 to −2.1	−2.1–0.6	0.6–16	16–1450
FE	0–1.8	1.8–5	5–9	9–20	20–99
FI	0–10	10–13	13–17	17–22	22–41
EIN	0 to −16	−16–23	23–47	47–70	70–100
FFEC	0–48	48–69	69–77	77–89	89–100
AANE	0–0.9	0.9–3.6	3.6–8	8–18	18–47
REC	0–8	8–15	15–28	28–46	46–92
EPC (x1,000)	0–0.9	0.9–2.1	2.1–4	4–7	7–54
CO2EFEAHP	0–26	26–35	35–42	42–51	51–81
EPFRSEHE	0–0.09	0.09–1.4	1.4–3.4	3.4–10	10–49
EPFOACS	0–24	24–45	45–66	66–89	89–100
EPFNS	0–0	0–6	6–23	23–39	39–75
EPFNGS	0–1.1	1.1–10	10–24	24–50	50–100
EPFHES	0–2.6	2.6–9	9–29	29–59	59–100
EPFCS	0–0	0–6	6–18	18–37	37–100
CO2I	0–1.5	1.5–2	2–2.35	2.35–2.6	2.6–3.5

In this study, we used data to determine the structure of the net since we wanted to derive the net totally depending on hard data. To do this, we provided input to the GeNIe software on the 23 variables of 91 countries listed in Table 1 and used the "greedy thick thinning" algorithm to determine the conditional dependency structure (*GeNIe*, 2020). "Greedy Thick Thinning" algorithm initially starts with no relations between the variables and then step by step add one arc at a time to generate an increase in the marginal likelihood value of the net. The same method is used for arc deletion that again aims to result an increase in the marginal likelihood. The resulting net was transferred to Netica to quantify the relations between variables (*Netica*, 2018). The resulting net is given in Fig. 1.

Netica software validates the constructed BN according to 3 measurement scores well known in the literature: Logarithmic Loss, Quadratic Loss, and Spherical Payoff (Kontkanen et al., 1999; Winkler et al., 1996). Since the most central variable is GDP in our BN, we have tested the performance of this variable specifically

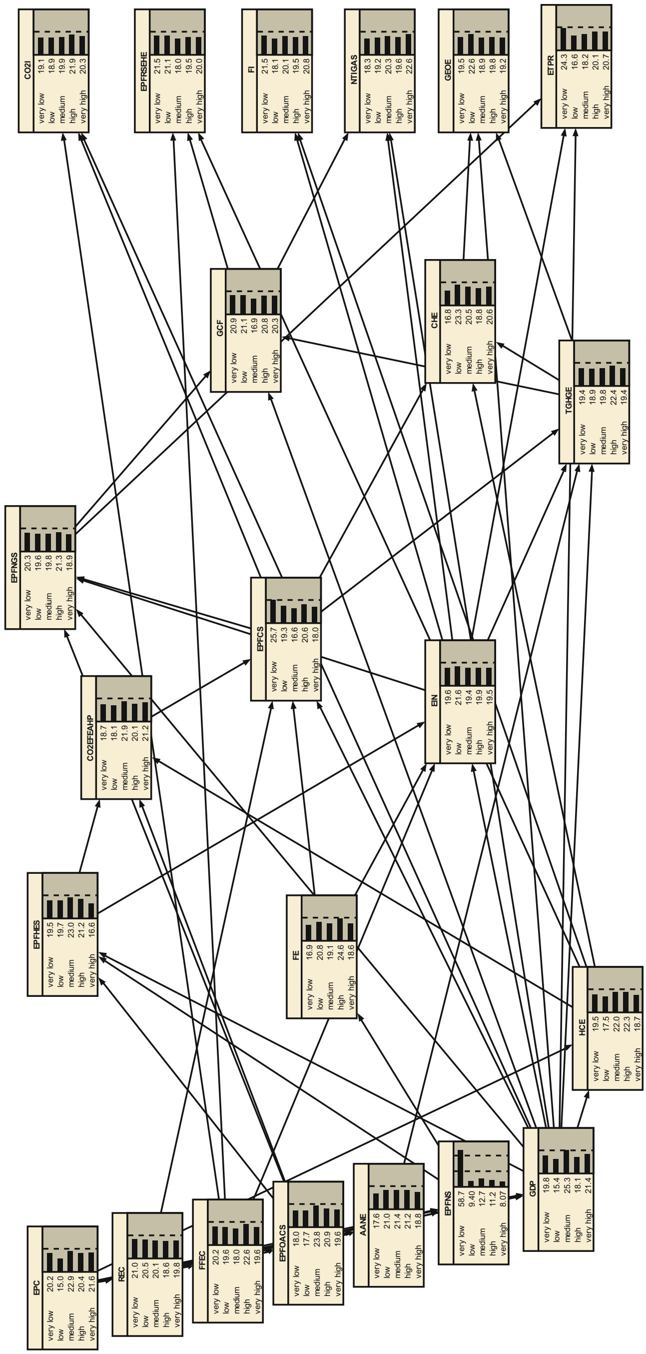


Fig. 1 Developed Bayesian network

and the scores calculated are found as 0.020 (must be in between 0 and infinity; 0 is the best score); 0.013 (must be in between 0 and 2; 0 is the best score); and 0.993 (must be in between 0 and; 1 is the best score), respectively, from which we can conclude that the BN's performance on capturing the conditional probability relations between variables is extremely satisfactory.

The net is organized according to the graph theory indices (Özesmi & Özesmi, 2003). The total number of outgoing arrows from a variable gives the out-degree (od) of that variable and the total number of incoming arrows to a variable denotes the in-degree (id) of that variable. Calculating od and id values for each variable makes it possible for us to create a hierarchical network just like given in Fig. 1. As it can be seen from the Figure, at the top of the hierarchy (the right-hand side of the net), the variables with 0-od values are located while at the bottom (the left-hand side of the net), there is just one variable (EPC) since it is the only variable with a 0-id value. The rest of the net is organized by the ratio of "od" to "id" values for each variable. Moreover, the central variable whose summation of its "id" and "od" values is the highest can be quantified as the variable with the highest contribution to the system. In our net, "GDP," with a centrality value of 13 (11-od and 2-id values), is the most central variable. Energy Imports have the second highest centrality value (9) and Electricity Production from Coal Sources has the third highest centrality value (8). In our BN, totally, there are 59 arcs between 23 nodes and 19,455 conditional probabilities in total.

When we investigate Fig. 1, we can have an idea about the related state probabilities for each variable. For example, (with no specified posterior probabilities), Electricity Production from Renewable Sources Excluding Hydroelectric (EPFRSEHE) is in the medium state with a 18% probability, in the low state with a 21.1% probability and, in the very low state with a 21.5% probability. Similarly, we see that for the countries that are given as input to the net, most of them (%25.3) has a medium GDP per capita value ranging from \$6000 to \$14,000.

3.3 Analyzing the System with Bayesian Net

One of the advantages of using BNs is to be able to conduct investigations that include both diagnostic reasoning (i.e., "upwards" reasoning from effects to cause) and predictive reasoning (i.e., "downwards" reasoning from cause to effect) (Lauría & Duchessi, 2007). So, after determining the conditional dependency structure and learning the parameters, we can give evidence to one or more variable(s) and observe the changes in the whole system, including both direct and indirect relations. Besides, we also can conduct sensitivity analysis which means that we can identify the variables with the most explanatory power on a variable.

Netica allows us to conduct sensitivity analysis in an easy way so that we can identify factors with the highest explanatory power on an observed variable. Sensitivity analysis is conducted by calculating the expected variance reduction in the real value of query variable (Q) due to a finding at varying variable (F). Giving an evidence to the F variable as input causes a change in the values of probable values of the variable Q. That difference gives us information about the explanatory power for variables. The variable with the greatest variance reduction rate is expected to be the one that changes the beliefs of the observed variable most.

For CO_2 intensity variable, for example, the sensitivity analysis result is given in Fig. 2. As can be seen, fossil fuel energy consumption (FFEC), renewable energy consumption (REC), and GDP per capita (GDP) have the highest explanatory power on CO_2 intensity (CO2I). More specifically, changes in CO2I are explained by FFEC about 20%, by REC about 13%, and by GDP about 7%. For instance, if an evidence about the FFEC of a country is given to the system, the variance in CO2I will be reduced by 19.6%.

According to the information given in Fig. 2, Fig. 3 shows the change in the probabilities of the states of CO2I variable when an information about the state of a specific variable is fed to the network. The probabilities, without the information about FFEC and CO2I, are given in the left-hand side of Fig. 3. As can be seen, whenever the information about the probability of being in “high” state is given for FFEC, we observe that, for CO2I, the possibility of being in “high” state also increases to 45.7% and being in “very high” state to 26.9%, as well. Contrarily, whenever the information about the probability of being in “high” state is given for REC, we observe that, the possibility of being in “high” state decreases to 15.5% and being in “very high” state to 9.1% for CO2I. On the other hand, if the information about the probability of being in “high” state is given for GDP, the possibility of being in “high” state slightly increases to 27.8% and being in “very high” state to 29.8% for CO2I.

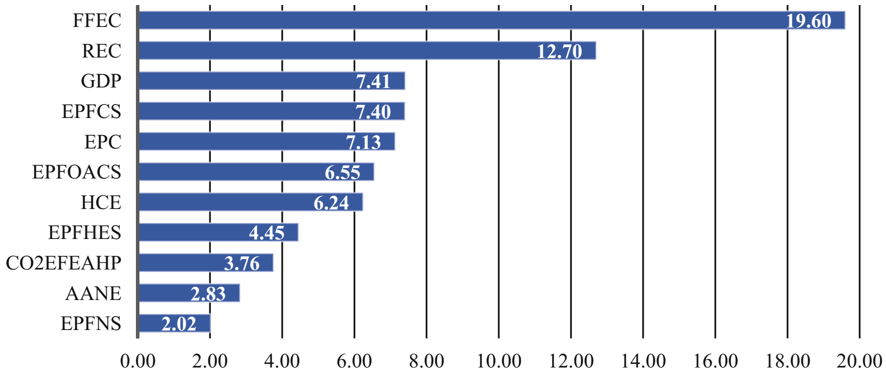


Fig. 2 Sensitivity analysis for CO_2 intensity

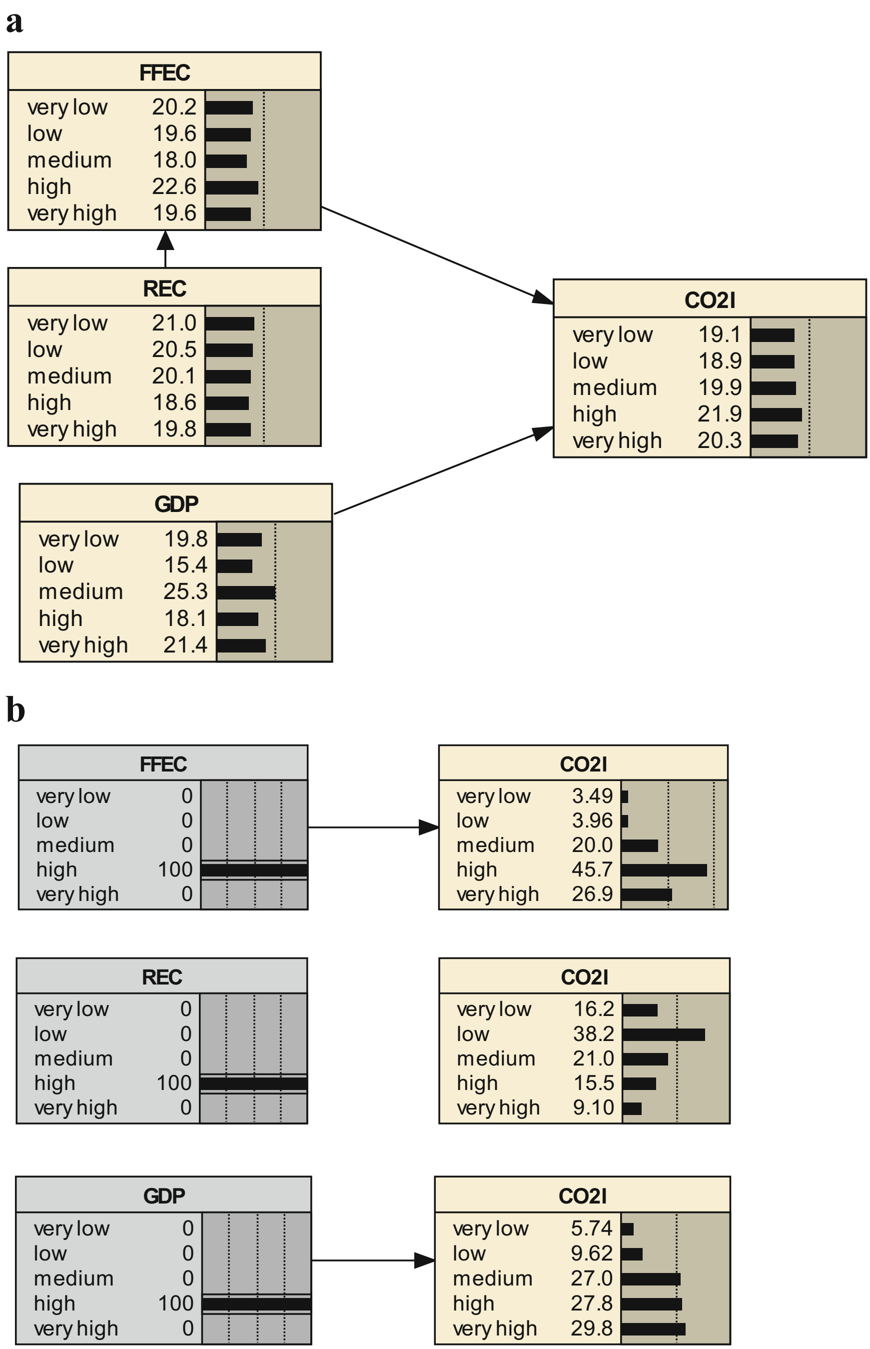


Fig. 3 Detailed sensitivity analyses for CO_2 Intensity (**a**) the marginal probabilities (**b**) the change in the probabilities of CO_2 intensity when an evidence is given to FFEC, REC, and GDP

As can be seen in Fig. 4, whenever the information about the probability of being in "very high" state is given for Net trade in goods and services (NTIGAS), we observe that, for GDP, the possibility of being in "very high" state also increases to 56.1% while for gross capital formation (GCF), the possibility of being in "very high" state also increases to 43.9% and being in "high" state increases to 32.2%. On the other hand, whenever the same information is given for NTIGAS, the possibility of being in "very high" state slightly increases to 28.7% and being in "high" state

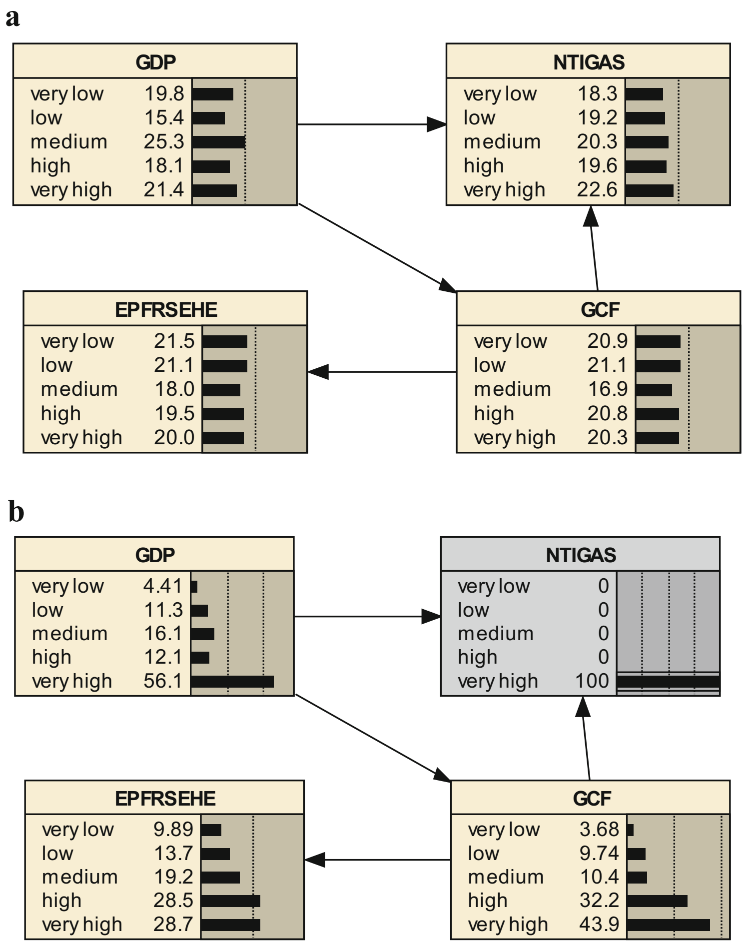


Fig. 4 Detailed sensitivity analysis for Net Trade in Goods and Services (**a**) the marginal probabilities (**b**) the change in the probabilities of GDP, GCF, EPFRSEHE when an evidence is given to NTIGAS

to 28.5% for electricity production from renewable sources excluding hydroelectric (EPFRSEHE). As can be observed from the figure, there is no direct link between NTIGAS and EPFRSEHE. However, from this what-if analysis, we can conclude that an economy in which "Net trade in goods and services" is known to be "very high," renewable sources excluding hydroelectric is also expected to be high (with a total probability of 57.2%).

Figure 5 represents the results of sensitivity analyses for Current Health Expenditure (CHE). If the information about the probability of being in "very low" state is given for fossil fuel energy consumption (FFEC), the possibility of being in "low" state also increases to 34.3% and being in "very low" state increases to 21.1% for CHE. On the other hand, if the information about the probability of being in "very

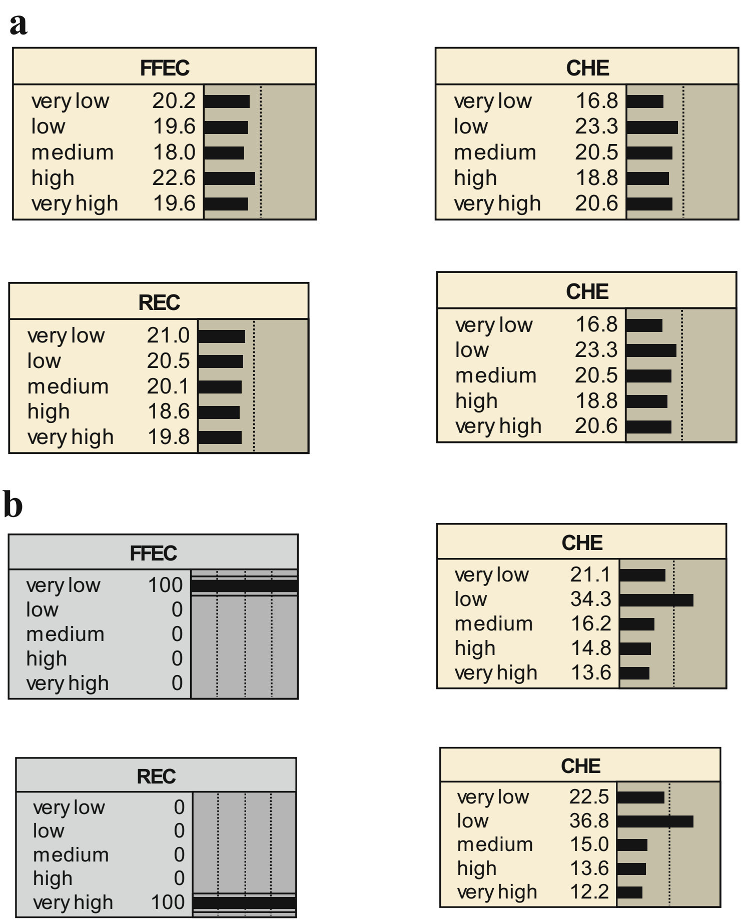


Fig. 5 Detailed sensitivity analyses for Current Health Expenditure (**a**) the marginal probabilities (**b**) the change in the probability of CHE when an evidence is given to FFEC and REC

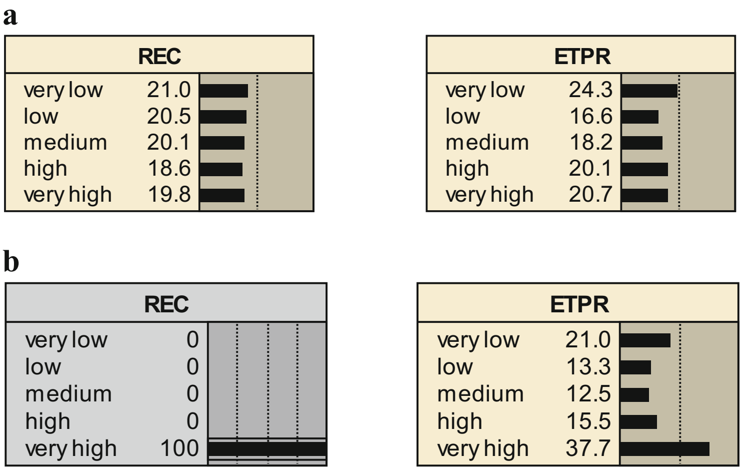


Fig. 6 Detailed sensitivity analysis for Employment to population ratio (**a**) the marginal probabilities (**b**) the change in the probability of ETPR when an evidence is given to REC

high" state is given for renewable energy consumption (REC), the possibility of being in "low" state also increases to 36.8% and being in "very low" state increases to 22.5% for CHE.

As can be seen in Fig. 6, whenever the information about the probability of being in "very high" state is given for REC, we observe that, for employment to population ratio (ETPR), the possibility of being in "very high" state also increases to 37.7.

As a final sensitivity analysis, we examined the changes in the probabilities of electric power consumption (EPC) and electricity production for renewable sources excluding hydroelectric (EPFRSEHE) when the probable state of Government Expenditure on Education (GEOE) as well as that of gross capital formation (GCF) is known (Fig. 7). Whenever the information about the probability of being in "very high" state is given for both GEOE and GCF, we observe that, the possibility of being in "very high" state also increases to 51.4% for EPC and the possibility of being in "high" state increases to 37.7% for EPFRSEHE.

The analyses conducted are summarized in the following Table 3:

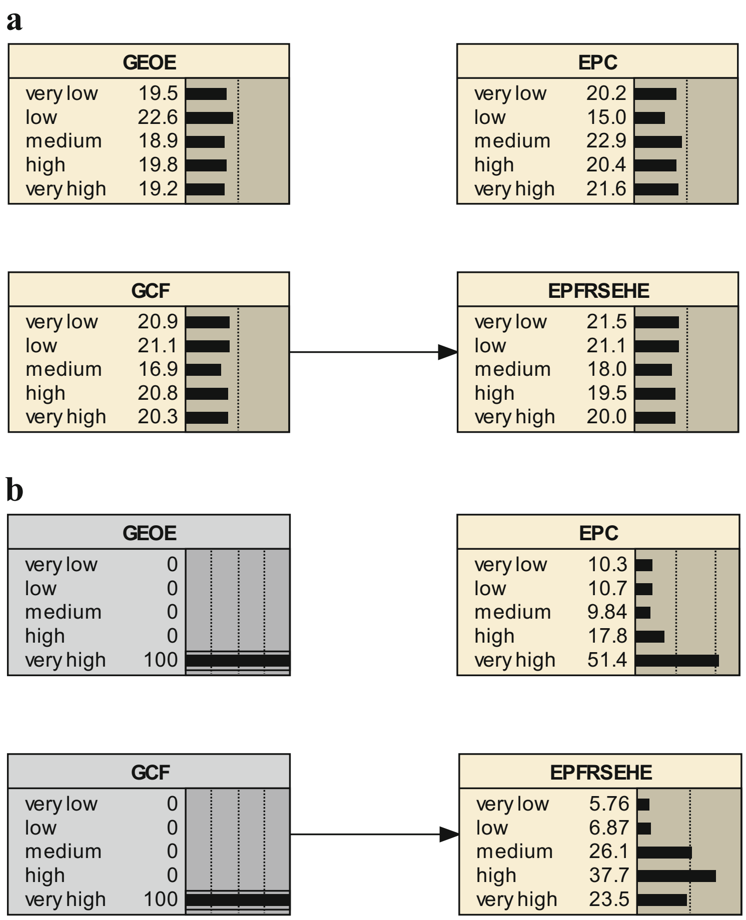


Fig. 7 Detailed sensitivity analysis for EPC and EPFRSEHE (**a**) the marginal probabilities (**b**) the change in the probabilities of EPC and EPFRSEHE when an evidence is given to GEOE and GCF

Table 3 Presentation of all analysis results

Analysis	Effected variable	Changed variable	State	Marginal Prob.	Prob.	Change
1	CO2I	FFEC-h	vl	19.1	3.49	−15.61
1	CO2I	FFEC-h	l	18.9	3.96	−14.94
1	CO2I	FFEC-h	m	19.9	20	0.1
1	CO2I	FFEC-h	h	21.9	45.7	23.8
1	CO2I	FFEC-h	vh	20.3	26.9	6.6
1	CO2I	GDP-h	vl	19.1	5.74	−13.36
1	CO2I	GDP-h	l	18.9	9.62	−9.28
1	CO2I	GDP-h	m	19.9	27	7.1
1	CO2I	GDP-h	h	21.9	27.8	5.9
1	CO2I	GDP-h	vh	20.3	29.8	9.5
1	CO2I	REC-h	vl	19.1	16.2	−2.9
1	CO2I	REC-h	l	18.9	38.2	19.3
1	CO2I	REC-h	m	19.9	21	1.1
1	CO2I	REC-h	h	21.9	15.5	−6.4
1	CO2I	REC-h	vh	20.3	9.1	−11.2
2	GDP	NTIGAS-vh	vl	19.8	4.41	−15.39
2	GDP	NTIGAS-vh	l	15.4	11.3	−4.1
2	GDP	NTIGAS-vh	m	25.3	16.1	−9.2
2	GDP	NTIGAS-vh	h	18.1	12.1	−6
2	GDP	NTIGAS-vh	vh	21.4	56.1	34.7
2	GCF	NTIGAS-vh	vl	20.9	3.68	−17.22
2	GCF	NTIGAS-vh	l	21.1	9.74	−11.36
2	GCF	NTIGAS-vh	m	16.9	10.4	−6.5
2	GCF	NTIGAS-vh	h	20.8	32.2	11.4
2	GCF	NTIGAS-vh	vh	20.3	43.9	23.6
2	EPFRSEHE	NTIGAS-vh	vl	21.5	9.89	−11.61
2	EPFRSEHE	NTIGAS-vh	l	21.1	13.7	−7.4
2	EPFRSEHE	NTIGAS-vh	m	18	19.2	1.2
2	EPFRSEHE	NTIGAS-vh	h	19.5	28.5	9
2	EPFRSEHE	NTIGAS-vh	vh	20	28.7	8.7
3	CHE	FFEC-vl	vl	16.8	21.1	4.3
3	CHE	FFEC-vl	l	23.3	34.3	11
3	CHE	FFEC-vl	m	20.5	16.2	−4.3
3	CHE	FFEC-vl	h	18.8	14.8	−4
3	CHE	FFEC-vl	vh	20.6	13.6	−7
3	CHE	CHE-vh	vl	16.8	22.5	5.7
3	CHE	CHE-vh	l	23.3	36.8	13.5
3	CHE	CHE-vh	m	20.5	15	−5.5
3	CHE	CHE-vh	h	18.8	13.6	−5.2
3	CHE	CHE-vh	vh	20.6	12.2	−8.4

(continued)

Table 3 (continued)

Analysis	Effected variable	Changed variable	State	Marginal Prob.	Prob.	Change
4	ETPR	REC-vh	vl	24.3	21	−3.3
4	ETPR	REC-vh	l	16.6	13.3	−3.3
4	ETPR	REC-vh	m	18.2	12.5	−5.7
4	ETPR	REC-vh	h	20.1	15.5	−4.6
4	ETPR	REC-vh	vh	20.7	37.7	17
5	EPC	GEOE-vh	vl	20.2	10.3	−9.9
5	EPC	GEOE-vh	l	15	10.7	−4.3
5	EPC	GEOE-vh	m	22.9	9.84	−13.06
5	EPC	GEOE-vh	h	20.4	17.8	−2.6
5	EPC	GEOE-vh	vh	21.6	51.4	29.8
5	EPC	REC-vh	vl	21.5	5.76	−15.74
5	EPC	REC-vh	l	21.1	6.87	−14.23
5	EPC	REC-vh	m	18	26.1	8.1
5	EPC	REC-vh	h	19.5	37.7	18.2
5	EPC	REC-vh	vh	20	23.5	3.5

4 Discussion

Mitigating the adverse effects of climate change is of the utmost importance. To address this challenge, institutional investors and governmental organizations direct their money to the renewable energy sector. Although the primary purpose is to ensure environmental sustainability, renewable energy is considered to be beneficial also for economy and society. Numerous factors have been identified in the literature to evaluate the interaction between renewable energy on a country's wellbeing. GDP is considered the primary indicator of countries' welfare. Accordingly, the majority of the studies focus on the relationship between renewable energy investments and GDP. Since GDP lacks measuring the wealth of a country, such as health, education, and clean air, solely focusing on GDP to observe the impact of renewable energy investments is not sufficient. To cover the interaction between renewable energy and national economic development, additional factors are identified by benefiting the governmental and international reports as well as scientific articles. After the selection of factors, the BN framework employed in this study, as a consequence of its structure, not only direct impacts of the parameters but also indirect relations are observed. The analysis is conducted in a way to observe direct impact of one factor on another. The what-if scenarios were constructed in this manner.

According to the BN results, although renewable energy consumption has significant explanatory power on CO_2 intensity variable, the impact of fossil fuel consumption is higher. Ending the domination of fossil fuels is expected to be more beneficial to decarbonize the system rather than just focusing on renewable energy investments. Renewable energy projects on hard-to-decarbonize industries are appeared to be more impactful.

Regarding the impact of the parameters that define the wealth on the renewable energy sector, we observe that the net trade in goods and services does not have a significant amount of strength to boost renewable-based electricity generation. Hence, decision-makers put a particular focus on the possibility of increased electricity demand due to the trade volume. This demand may exacerbate the need for fossil-based energy generation and put an additional burden on emission mitigation actions. This study's conclusion also justifies the European Union's proposal on "carbon border adjustment mechanism," which enforces an increase in renewable energy-based production.

Governmental reports reveal the quantitative impact of renewables. Renewable energy can reduce CO_2 as well as sulfur dioxide (SO_2), nitrogen oxides (NOx) which are major sources of air pollution that can harm human health (*The Multiple Benefits of Energy Efficiency and Renewable Energy, Part One*, 2016). As pointed out earlier, our findings also depict the direct impact of renewable energy consumptions on the reduction of health expenditure in line with the literature (Pizarro-Loaiza et al., 2021; Shahzad et al., 2020). The BN results indicate that renewable energy consumption has a significant impact on overall health expenditure (not just respiratory and cardiovascular illnesses that are directly related to air pollutants), which ensures the importance of renewables.

In summary, when all the overall results of the sensitivity analysis are considered, this study shows the strength of each component on a system level.

5 Conclusion and Further Suggestions

BNs are especially efficient for the analyses of real-world where there are complex relations among variables. They can also deal with uncertainty in a graphical structure with the help of probability and help the decision-makers to gain a better understanding on the investigated system.

In this study by the help of BNs, we aimed to focus on the interaction of renewable energy consumption with the wealth of nations by investigating the socio-economic benefits of increasing the share of renewable energy sources. The policy makers may utilize the proposed model as a diagnostic tool to manage renewable energy systems by conducting sensitivity analyses.

To the best of our knowledge, there is one major study that focuses on a similar subject, that is "Renewable Energy Benefits: Measuring the Economics" report published by International Renewable Energy Agency. However, this study uses a macro-econometric approach dealing with 59 countries in the context of 13 inputs and 6 outputs. Our study, on the other hand, not only expands the number of the countries taken into consideration (91 countries) but also analyzes the interaction in a Bayesian network structure where both direct and indirect relations can be evaluated easily. Besides, we try to clarify the factor based quantitative impact of the renewable energy transition.

In the literature, the benefits of renewable energy technologies are discussed: the increased use of renewables creates new jobs, decreases health expenditures, increases in the spending on education, and improves social welfare as well as national economic development. The quantitative results of our study also justify the renewable energy transition to be a promising opportunity for economic development. We conducted several what-if analyses by using conditional probabilities. According to our results, this transition is expected to positively impact the factors related to environment, economy, health, employment, and education. As aforementioned in Sect. 3, if renewable energy consumption increases and fossil fuel energy consumption decreases, CO_2 intensity as well as health expenditures will be expected to decrease. On the other hand, renewable energy consumption has a positive relation with the ratio of employment to population. Similarly, electricity production from renewable sources excluding hydroelectric has a positive relation with net trade in goods and services, gross capital formation, and government expenditure on education.

As a further suggestion, the countries can be clustered initially, and a cluster variable can be added to the network which in turn will help to repeat similar analyses according to the cluster a country belongs to. The variable set can also be expanded to take into account a more detailed view of the system.

References

Ahn, K., Chu, Z., & Lee, D. (2021). Effects of renewable energy use in the energy mix on social welfare. *Energy Economics, 96*, 105174. https://doi.org/10.1016/j.eneco.2021.105174

Akella, A. K., Saini, R. P., & Sharma, M. P. (2009). Social, economical and environmental impacts of renewable energy systems. *Renewable Energy, 34*(2), 390–396. https://doi.org/10.1016/j.renene.2008.05.002

Ali, Q., Raza, A., Narjis, S., Saeed, S., & Khan, M. T. I. (2020). Potential of renewable energy, agriculture, and financial sector for the economic growth: Evidence from politically free, partly free and not free countries. *Renewable Energy, 162*, 934–947. https://doi.org/10.1016/j.renene.2020.08.055

Anderson, R. D., & Vastag, G. (2004). Causal modeling alternatives in operations research: Overview and application. *European Journal of Operational Research, 156*(1), 92–109. https://doi.org/10.1016/S0377-2217(02)00904-9

Benedek, J., Sebestyén, T.-T., & Bartók, B. (2018). Evaluation of renewable energy sources in peripheral areas and renewable energy-based rural development. *Renewable and Sustainable Energy Reviews, 90*, 516–535. https://doi.org/10.1016/j.rser.2018.03.020

Bhattacharya, M., Awaworyi Churchill, S., & Paramati, S. R. (2017). The dynamic impact of renewable energy and institutions on economic output and CO_2 emissions across regions. *Renewable Energy, 111*, 157–167. https://doi.org/10.1016/j.renene.2017.03.102

Bloch, H., Rafiq, S., & Salim, R. (2012). Coal consumption, CO2 emission and economic growth in China: Empirical evidence and policy responses. *Energy Economics, 34*(2), 518–528. https://doi.org/10.1016/j.eneco.2011.07.014

Borunda, M., Jaramillo, O. A., Reyes, A., & Ibargüengoytia, P. H. (2016). Bayesian networks in renewable energy systems: A bibliographical survey. *Renewable and Sustainable Energy Reviews, 62*, 32–45. https://doi.org/10.1016/j.rser.2016.04.030

Bulavskaya, T., & Reynès, F. (2018). Job creation and economic impact of renewable energy in the Netherlands. *Renewable Energy, 119*, 528–538. https://doi.org/10.1016/j.renene.2017.09.039

Caglar, A. E. (2020). The importance of renewable energy consumption and FDI inflows in reducing environmental degradation: Bootstrap ARDL bound test in selected 9 countries. *Journal of Cleaner Production, 264*, 121663. https://doi.org/10.1016/j.jclepro.2020.121663

Cai, W., Wang, C., Chen, J., & Wang, S. (2011). Green economy and green jobs: Myth or reality? The case of China's power generation sector. *Energy, 36*(10), 5994–6003. https://doi.org/10.1016/j.energy.2011.08.016

Cheng, C., Ren, X., & Wang, Z. (2019). The impact of renewable energy and innovation on carbon emission: An empirical analysis for OECD countries. *Energy Procedia, 158*, 3506–3512. https://doi.org/10.1016/j.egypro.2019.01.919

Cuesta, M. A., Castillo-Calzadilla, T., & Borges, C. E. (2020). A critical analysis on hybrid renewable energy modeling tools: An emerging opportunity to include social indicators to optimise systems in small communities. *Renewable and Sustainable Energy Reviews, 122*, 109691. https://doi.org/10.1016/j.rser.2019.109691

Dogru, T., Bulut, U., Kocak, E., Isik, C., Suess, C., & Sirakaya-Turk, E. (2020). The nexus between tourism, economic growth, renewable energy consumption, and carbon dioxide emissions: Contemporary evidence from OECD countries. *Environmental Science and Pollution Research, 27*(32), 40930–40948. https://doi.org/10.1007/s11356-020-10110-w

Ekici, A., & Ekici, S. O. (2016). A Bayesian network analysis of ethical behavior. *Journal of Macromarketing, 36*(1), 96–115. https://doi.org/10.1177/0276146715607620

Ferroukhi, R., Lopez-Peña, A., Kieffer, G., Nagpal, D., Hawila, D., Khalid, A., El-Katiri, L., Vinci, S., & Fernandez, A. (2016). *Renewable energy benefits: Measuring the economics*. IRENA. https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2016/IRENA_Measuring-the-Economics_2016.pdf

GeNIe (3.0.5905). (2020). *C++; Windows 32-bit*. BayesFusion. https://download.bayesfusion.com/files.html?category=Academia.

Global Energy & CO2 Status Report (IEA Publications). (2019). International Energy Agency. https://www.iea.org/reports/global-energy-co2-status-report-2019/emissions

Ike, G. N., Usman, O., Alola, A. A., & Sarkodie, S. A. (2020). Environmental quality effects of income, energy prices and trade: The role of renewable energy consumption in G-7 countries. *Science of the Total Environment, 721*, 137813. https://doi.org/10.1016/j.scitotenv.2020.137813

IRENA. (2020). Global renewables outlook: Energy transformation 2050. International Renewable Energy Agency. https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2020/Apr/IRENA_Global_Renewables_Outlook_2020.pdf

Khan, A., Chenggang, Y., Hussain, J., & Kui, Z. (2021). Impact of technological innovation, financial development and foreign direct investment on renewable energy, non-renewable energy and the environment in Belt & Road Initiative countries. *Renewable Energy, 171*, 479–491. S0960148121002500. https://doi.org/10.1016/j.renene.2021.02.075

Konstantelos, I., Pudjianto, D., Strbac, G., De Decker, J., Joseph, P., Flament, A., Kreutzkamp, P., Genoese, F., Rehfeldt, L., Wallasch, A.-K., Gerdes, G., Jafar, M., Yang, Y., Tidemand, N., Jansen, J., Nieuwenhout, F., van der Welle, A., & Veum, K. (2017). Integrated North Sea grids: The costs, the benefits and their distribution between countries. *Energy Policy, 101*, 28–41. https://doi.org/10.1016/j.enpol.2016.11.024

Kontkanen, P., Myllymäki, P., Silander, T., & Tirri, H. (1999). On Supervised Selection of Bayesian Networks. In *Proceedings of the 15th International Conference on Uncertainty in Artificial Intelligence (UAI'99)* (p. 342).

Korb, K. B., & Nicholson, A. E. (2010). *Bayesian artificial intelligence*. CRC Press. https://doi.org/10.1201/b10391

Lange, G.-M., Wodon, Q., & Carey, K. (2018). *The changing wealth of nations 2018: Building a sustainable future*. World Bank. https://doi.org/10.1596/978-1-4648-1046-6

Lauría, E. J. M., & Duchessi, P. J. (2007). A methodology for developing Bayesian networks: An application to information technology (IT) implementation. *European Journal of Operational Research, 179*(1), 234–252. https://doi.org/10.1016/j.ejor.2006.01.016

Liu, J.-L., Ma, C.-Q., Ren, Y.-S., & Zhao, X.-W. (2020). Do real output and renewable energy consumption affect CO2 emissions? Evidence for selected BRICS countries. *Energies, 13*(4), 960. https://doi.org/10.3390/en13040960

Magazzino, C., Mele, M., & Schneider, N. (2021). A machine learning approach on the relationship among solar and wind energy production, coal consumption, GDP, and CO2 emissions. *Renewable Energy, 167*, 99–115. https://doi.org/10.1016/j.renene.2020.11.050

Masson-Dellmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D. C., Skea, J., Shukla, P. R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., Matthews, J. B. R., Chen, Y., Zhou, X., Gomis, M. I., Lonnoy, E., Maycock, T., Tignor, M., & Waterfield, T. (Eds.). (2018). *Global warming of 1.5°C. an IPCC special report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty* [Special Report]. The Intergovernmental Panel on Climate Change (IPCC).

Mohsin, M., Kamran, H. W., Atif Nawaz, M., Sajjad Hussain, M., & Dahri, A. S. (2021). Assessing the impact of transition from nonrenewable to renewable energy consumption on economic growth-environmental nexus from developing Asian economies. *Journal of Environmental Management, 284*, 111999. https://doi.org/10.1016/j.jenvman.2021.111999

Nazir, M. S., Bilal, M., Sohail, H. M., Liu, B., Chen, W., & Iqbal, H. M. N. (2020). Impacts of renewable energy atlas: Reaping the benefits of renewables and biodiversity threats. *International Journal of Hydrogen Energy, 45*(41), 22113–22124. https://doi.org/10.1016/j.ijhydene.2020.05.195

Netica (6.06). (2018). *Windows*. Norsys Software Corp.

Number of renewable energy jobs worldwide from 2012 to 2019. (2020). Statista. https://www.statista.com/statistics/859908/employment-in-renewable-energy-sector-globally/

Özesmi, U., & Özesmi, S. (2003). A participatory approach to ecosystem conservation: Fuzzy cognitive maps and stakeholder group analysis in Uluabat Lake, Turkey. *Environmental Management, 31*(4), 518–531. https://doi.org/10.1007/s00267-002-2841-1

Pizarro-Loaiza, C. A., Antón, A., Torrellas, M., Torres-Lozada, P., Palatsi, J., & Bonmatí, A. (2021). Environmental, social and health benefits of alternative renewable energy sources. Case study for household biogas digesters in rural areas. *Journal of Cleaner Production, 297*, 126722. https://doi.org/10.1016/j.jclepro.2021.126722

Saidi, K., & Omri, A. (2020). The impact of renewable energy on carbon emissions and economic growth in 15 major renewable energy-consuming countries. *Environmental Research, 186*, 109567. https://doi.org/10.1016/j.envres.2020.109567

Sarpong, S. Y., Bein, M. A., Gyamfi, B. A., & Sarkodie, S. A. (2020). The impact of tourism arrivals, tourism receipts and renewable energy consumption on quality of life: A panel study of southern African region. *Heliyon, 6*(11), e05351. https://doi.org/10.1016/j.heliyon.2020.e05351

Sebastiani, P., & Perls, T. T. (2008). Complex genetic models. In O. Pourret, P. Naim, & B. Marcot (Eds.), *Statistics in practice* (pp. 53–72). Wiley. https://doi.org/10.1002/9780470994559.ch4

Shahzad, K., Jianqiu, Z., Hashim, M., Nazam, M., & Wang, L. (2020). Impact of using information and communication technology and renewable energy on health expenditure: A case study from Pakistan. *Energy, 204*, 117956. https://doi.org/10.1016/j.energy.2020.117956

Singh, N., Nyuur, R., & Richmond, B. (2019). Renewable energy development as a driver of economic growth: Evidence from multivariate panel data analysis. *Sustainability, 11*(8), 2418. https://doi.org/10.3390/su11082418

The Multiple Benefits of Energy Efficiency and Renewable Energy, Part One (p. 17). (2016). U.S. Environmental Protection Agency.

Winkler, R. L., Muñoz, J., Cervera, J. L., Bernardo, J. M., Blattenberger, G., Kadane, J. B., Lindley, D. V., Murphy, A. H., Oliver, R. M., & Ríos-Insua, D. (1996). Scoring rules and the evaluation of probabilities. *TEST, 5*(1), 1–60. https://doi.org/10.1007/BF02562681

Zafar, M. W., Zaidi, S. A. H., Sinha, A., Gedikli, A., & Hou, F. (2019). The role of stock market and banking sector development, and renewable energy consumption in carbon emissions: Insights from G-7 and N-11 countries. *Resources Policy, 62*, 427–436. https://doi.org/10.1016/j.resourpol.2019.05.003

Zakeri, B., Price, J., Zeyringer, M., Keppo, I., Mathiesen, B. V., & Syri, S. (2018). The direct interconnection of the UK and Nordic power market – Impact on social welfare and renewable energy integration. *Energy, 162*, 1193–1204. https://doi.org/10.1016/j.energy.2018.08.019

The Role of Logistics Management in Food Supply Chains

João Gilberto Mendes dos Reis and Sivanilza Teixeira Machado

Abstract Food supply chains are essential for human survival. However, the path from 'farm to fork' is a challenge to society due to the distance between markets and producers, traditional production processes, acquisition costs, household income, and so on. Globalisation intends to reduce these issues through the specialisation of production areas, use of machinery and chemical compounds, low-labour cost, and efficient logistics network. But this means more transportation, gas emissions, and high dependence on foreign countries. The objective of this chapter is to discuss the impact of logistics management in food supply chains using two case studies: Soybean and Beef production. To do so, we collected the volume of production in 2019 from the Food and Agriculture Organisation from the United Nations, and the volume of exports by origin state, port, and country destination from the Brazilian Ministry of Economy. Brazil was chosen for analysis due to its important role in world agriculture/livestock production, its importance as a producer and exporter of goods, and the situation of being a developing country in Latin America with enormous geographic size. Moreover, we used Social Network Analysis to understand the relationship among these players. Our findings indicated a great number of importers in the network, but China has a central role in soybean and beef exports. This influence causes an imbalance flow of these commodities in the network and causes a dependence of agriculture producers for single marketing.

J. G. M. dos Reis (✉)
Research Group in Supply Networks, Postgraduate Program in Production Engineering and Postgraduate Program in Business Administration, Universidade Paulista – UNIP, São Paulo, Brazil

Centro de Ciências Socias Aplicadas, Universidade Presbiteriana Mackenzie-UPM, São Paulo, Brazil.
e-mail: joao.reis@docente.unip.br

S. T. Machado
Napole Research Group, Instituto Federal de Educação Tecnológica de São Paulo, Campus Suzano, Suzano, Brazil
e-mail: sivanilzamachado@ifsp.edu.br

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_21

Keywords Agriculture commodities · Food logistics · Food supply chains · Social Network Analysis

1 Introduction

The world population growth has transformed Food Supply Chains (FSCs) into a central problem for governments worldwide. According to the Food and Agriculture Organisation from the United Nations (FAO, 2017), more people live now in cities than in rural areas creating a transition in dietary patterns with great impacts on food systems. An issue that can be aggravated by the prediction of 9.73 billion inhabitants in 2050, the pandemic crisis of COVID-19, and other possible challenges and diseases.

FSCs are vital to providing daily meals to the community; however, food losses and waste, poor logistics management and infrastructure, poverty, and unemployment compromise access to food. In this sense, we need to understand that food access is different from the availability of food. The former is social, economic, and political revolving around employment status, household income, and eligibility to receive social services or financial assistance while the latter is a technical matter dependent on agricultural productivity, infrastructure efficiency, and practices related to the transportation and food processing (Loo, 2019).

Thapa Karki et al. (2020) address the paradox of tonnes of food wasted while people go hungry and recognise that donation and distribution of surplus food to people in need is a potential solution. Dani & Deep (2010) point out how FSCs are fragile and suggest that quick action and efficient logistics can make a crisis manageable, citing Walmart's actions during Hurricane Katrina in the United States in 2005. Coluccia et al. (2021), studying an exhaustive compilation of relevant macroeconomic data regarding the status of the agri-food sector after the shock of COVID-19, identified that fresh and perishable products, which production or harvest took place during the first wave of the pandemic, have suffered price level effects, while storable products have not registered significant impacts. The uncertainty around supply and demand has a straight impact on economies pressuring consumers to prioritise food acquisition instead of other goods.

Food is a fragile and complex supply chain due to its characteristics and necessity to manage perishable items. Dani (2015) reinforces the complexity of FSCs showing that a fish caught on the western coast of the United States is frozen and shipped to China where it is defrosted, filleted, packed, and refrozen before reaching retail shelves in Europe or the own United States. Indeed, FSCs are intensely connected with the production areas, valuable markets, and massive logistics networks. Brazil, for instance, is definitely an agriculture/livestock producer for more than 300 crops and exports 350 types of products to 180 countries (Abraham et al., 2020). The low price of arable and fertile land, tropical climate conditions, the Brazilian Real exchange currency rate against the US Dollar and Euro, and competitive maritime freight rates, make the country an important supplier of many agriculture commodities.

In this chapter, our objective is to study the role of logistics management in FSCs. To do so, we conducted two case studies in Brazil using data from the Food and Agriculture Organisation from the United Nations (FAO) and the Brazilian Ministry of Economy (ME).

Adopting a Social Network Analysis approach, we analysed Soybean and Beef production based on the Brazilian export network. Our results seek to demonstrate the advantages and bottlenecks of these supply chains considering the logistics and marketing point of view.

The chapter is organised as follows: the next section presents the literature review regarding food supply chains, logistics, and local or global food sourcing. Then, it moves for methods adopted, followed by our findings. Finally, the discussion section is provided, and the chapter finishes with a conclusion section drawing the main ideas and suggesting future research.

2 Literature Review

2.1 Food Supply Chains

There is nothing more essential than food for human survival. In this regard, agriculture and livestock production have a fundamental role in creating conditions to sustain this human life. Initially, they were organised locally, but step by step, local production was substituted by global agriculture connected by an extensive maritime network.

In particular, this change was provoked by the Age of Discovery (fifteenth–eighteenth century), and the developments of the Industrial Revolution that allowed countries commercialise their food surplus overseas. During the twentieth century, maritime transportation consolidated with the main system for agriculture commodities exports arising a complete variety of vessels for different purposes and turning agriculture more global than ever.

This scenario is a one-way round, global food trade increased 37% in ten years (2010–2019) (FAO, 2021). Globalisation leads to more movements of products so that consumers can find food items from other parts of the world, and fruits and vegetables regardless of the season (Behnke & Janssen, 2020).

Global agriculture raised the role of FSCs that are complex and interlinked with a large number of producers, processors, retailers, and consumers, each interacting with many other actors in the chain, becoming more and more global and dependent on this increasing number of actors (Behnke & Janssen, 2020; de Gorter et al., 2020).

FSCs have several characteristics that differentiate them from other product supply chains such as a continuous change in the quality of food products throughout the entire supply chain involving freshness and safety (Nagurney et al., 2013). These aspects have been raising concerns about losses and waste reduction, technology adoption, and food safety and security.

The FSCs technically span the entire 'farm-to-fork' process (Dani & Deep, 2010) and its study from an operational perspective is essential because they not only influence everyday life but also affect businesses and livelihoods (Dani, 2015). Therefore, the literature presents many studies focusing on FSCs, some of which are given in Table 1.

As it can be seen in Table 1, the actual literature is studying the FSCs at the level of food waste and losses, environmental issues, pandemic crisis, and technology implementation. In this chapter, our focus is to study FSCs in the context of logistics management.

2.2 Logistics Management

Broadly speaking, logistics is a science that studies the flow of goods, services, and information through the interaction among suppliers and clients connected by a network. Thus, logistics can be defined as the process of strategically managing the procurement, movement and storage of materials, parts and finished inventory and, the related information flows through the organisation and its marketing channels in such a way that current and future profitability are maximised through the cost-effective fulfilment of orders (Christopher, 2011).

In agriculture and FSCs, logistics involve purchasing, transportation, warehousing, loading and unloading, handling, packaging, processing, distribution, and information processing (Kukovic et al., 2014; Paciarotti & Torregiani, 2021). Another important aspect of logistics in FSCs is the necessity to deal with perishability and freshness. Food logistics consists of the movement of food through the supply chain until it reaches the consumer's plate (Dani, 2015).

The performance of logistics is vital to ensuring trade and delivery of contracted items. Therefore, World Bank developed a Logistics Performance Index (LPI) based on a survey of logistics characteristics of the countries that affect trade (Arvis et al., 2018). Mendes dos Reis et al. (2020) identified that a one-point level increase in exporter LPI could represent an increase of 45% in the soybean trade. Nevertheless, performance is not only important for large international supply chains but logistics best practices also have the potential to improve the efficiency and effectiveness of regional FSCs (Mittal et al., 2018).

Wajszczuk (2016) demonstrated that agricultural products pressure the logistics operations, due to: (1) considerable costs of domestic transport resulting from multiple journeys on account of the spatial character in regards to agricultural production, (2) seasonality that occupies a large warehouse area, (3) harvest may cause a considerable loss in the yield interrupting transportation of products, (4) diversity of loads transported that determines a wide range of means of transportation, (5) one-way haulage that consists in the non-existence of backhaul cargo, (6) short-distance haulage on poor-quality roads, and (7) low transport and storage capacity of numerous agricultural products, resulting from their short shelf life and high perishability.

Table 1 Food supply chain literature

Authors/Year	Purpose	Approach	Results
Behnke & Janssen, (2020)	To identify boundary conditions for sharing assurance information to improve traceability using Blockchain.	Four case studies in the food supply chains using a template analysis of 16 interviews.	Supply chain systems have first to be modified and organisational measures need to be taken to fulfil the boundary conditions before blockchain can be used successfully.
Chaboud & Moustier (2020)	To assess the volume of food losses and waste along a food supply chain and analyse the roles that supermarket and non-supermarket channels play in dealing with these losses and wastes.	A case study of tomato chain in Cali, Colombia.	They found relatively low food losses and waste due to the diversity of consumer preferences, the reduced harvest-to-sale time, the marketing strategies adopted to sell downgraded and damaged products, and the overlap and complementarity between supermarket and non-supermarket channels.
Coluccia et al. (2021)	To assess the resilience level of the agri-food sector to the coronavirus pandemic, analysing its effect on commodity prices and focusing on the supply and value chain.	Time-series statistical methods.	Fresh and perishable products, whose production or harvest took place during the first wave of COVID-19, have suffered price level effects, while storable products have not registered significant impacts.
de Gorter et al. (2020)	To contribute to the existing literature by outlining a general analytical framework to analyse food waste along an entire food supply chain.	Theoretical economic model.	The amount of food waste always declines at the stage of the supply chain where the rate of waste declines. The reduction in the waste rate has a direct quantity impact on how much of the purchased food can be consumed.

(continued)

Table 1 (continued)

Authors/Year	Purpose	Approach	Results
Despoudi (2020)	To explore the challenges that the EU producers are facing in their efforts to reduce food losses using Contingency Theory.	Semi-structured interviews with Greek producers.	The impact of climatic change as well as collaboration are the main challenges for reducing food losses.
Gholami-Zanjani et al. (2021)	To propose a comprehensive two-stage scenario-based mathematical model to design a resilient food supply chain under demand uncertainty and epidemic disruptions.	A Monte Carlo procedure is developed to generate plausible scenarios to avoid disruptions.	The established capacity and lost-sale cost are found to be the most influencing factors for the yielded food supply chain network desirability in terms of profit, fill rate, and product freshness.
Glover & Touboulic (2020)	To explore how actors across the supply chain have been able to pass risk and responsibility for environmental impacts down the supply chain, in addition to associated economic costs.	They use vignettes to relay farmers' experiences and voices, which remain seldom heard in both practice and research.	The hidden practice and consequences of what authors refer to as 'passing the environmental buck'. Bypassing these environmental risks, powerful players in supply chains are also able to pass financial risk and costs down the supply chain.
Li & Hanafi (2013)	To focus on multimodal transportation planning and optimal strategies in a UK food supply chain case under carbon emissions control. They investigated and identified the impact of the policies on logistics performance.	Optimisation Model.	A network design approach for fresh produce logistics services under carbon emission control was proposed.

Messner et al. (2021)	To offer a systems-based understanding of food waste, which allows for an account of the interconnected processes that underpin waste creation along the whole supply chain.	A qualitative inquiry on practices and processes of surplus and waste creation in the Australian horticulture industry was conducted.	The results offered recommendations for industry, policy, and research: that approaches addressing systemic processes of waste creation are essential to unlocking food waste prevention, that food waste prevention should target the identified system processes contributing to food chain lock-ins, and that transparent monitoring and disclosure of food surplus is a prerequisite for systemic food waste prevention across the whole supply chain.
Stranieri et al. (2021)	To explore the impact of the blockchain technology on food supply chains.	An integrated conceptual framework including dimensions discussed in the literature: efficiency, flexibility, responsiveness, food quality, and transparency of supply chains.	Blockchain technology impacts positively on the profit and/or return on investment of supply chains and increase extrinsic food quality attributes.
Thapa Karki et al. (2020)	To explore actors and organisations in the value chain of surplus food distribution at the city level.	Semi-structured interviews and participant observation.	The need for a coordinated effort between actors is an essential arrangement to capture the value of surplus food.

Source: Authors

Considering all those aspects, logistics management is a complex task for any kind of FSCs both large and small. However, larger international supply chains have economic advantages of scale that bring gains of productivity and production costs, reducing consumer final prices, and leading to the globalisation approach.

2.3 Local or Global Food Sourcing (Smallholders or Large-Scale farm producers)

A topic of much public debate is local food production and Short Food Supply Chains (SFSCs) can be seen as an alternative for these massive international supply chains (Jones et al., 2004; Loiseau et al., 2020; Mittal et al., 2018; Thomé et al., 2021). However, it is necessary to recognise that logistics favour the largest chains against the local ones (dos Reis et al., 2016; Mendes dos Reis et al., 2020; Mittal et al., 2018).

Nowadays, local food production and SFSCs are at the core of the discussion about food production. The definition of local food revolves around the idea of items produced and consumed within a narrowly defined geographical area but is different from locality food, which has an identifiable geographical origin, but which is often marketed and sold nationally and internationally (Jones et al., 2004).

SFSCs emerged from changes in consumer habits regarding industrialisation and the globalisation of food (Thomé et al., 2021). In these supply chains, there are few or no intermediaries, and production, sorting and packaging, storage, transportation, sale, and the final consumer trip are connected to the producer (Loiseau et al., 2020).

The point is to strengthen smallholders and generate value for the local communities. Despite making some sense, SFSCs reduce variety, present inefficient logistics, have higher prices, and are highly dependent on local climate conditions reducing the resilience of the supply. On the other hand, large international supply chains present several intermediaries increasing the risk of contamination and shortage of supply.

One big issue for local food production is the workforce. For many countries, due to the movement of people to the cities, enough workers are not available. 'Hard work under the solid sun' to obtain a low income is not a dream for the youth.

The fact is that local food is more related to environmental protection and empowerment of local communities, on the other hand, better logistical solutions in regional food systems require a substantial private and public financial investment for implementation, which prevents most local producers from adopting them (Mittal et al., 2018).

In our study, we show two case studies where perhaps the adoption of local food and short supply is almost impossible due to the great use of natural resources and area of production.

3 Methods

Illustrating the role of logistics management in FSCs, this chapter focuses on two supply chains: (1) Soybean and (2) Beef. They were chosen due to their importance to human nutrition and the connection between them. Animal protein production is highly dependent on the soybean chain to animal feeding and consequently, the decisions made in one of them affect another and vice versa. Moreover, both chains allow us to demonstrate the unbalanced of global FSCs, the dependence of foreign markets with arable land, the low cost of land, and the importance of logistics to connect the global markets.

3.1 Data Collection

In this study, data were obtained from two sources: (1) the Food and Agriculture Organisation from the United Nations (FAO, 2021) and (2) the Ministry of Economy from Brazil (ME, 2021).

The data are volume of soybean and beef produced in 2019 (last year available in FAO dataset), and the volume of soybean and beef exported in 2020 (last year available in ME dataset)[1] considering state producer, origin port, and importing country. The data were arranged in four matrices using Microsoft Excel© 2016:

- World soybean production by exporting country and importing country.
- World beef production by exporting country and importing country.
- Brazilian soybean exports by state, origin port, and importing country.
- Brazilian beef exports by state, origin port, and importing country.

3.2 Data Analysis

The data were analysed using Microsoft Excel© 2016, UCINET© 6.698, and NetDraw© 2.138. In Microsoft Excel 2016, we plotted a world map with the soybean and beef production based on the volume of production while UCINET and NetDraw were used to establish a Social Network Analysis (SNA) approach.

SNA is a technique adopted in studies of economics, management, and sociology to establish relationship patterns among actors to identify the most influential ones (Huang et al., 2019; Palomo-Campesino et al., 2018; Yustiawan et al., 2015). The SNA uses network models where nodes are linked by arcs that determine a

[1] In this chapter, we adopt the last data available instead to approximate for the same year providing, therefore, the most recent data.

relationship: a connection, a trade, a communication, monetary values, and so on (Cheah & Shimul, 2018).

The SNA can be analysed through different measures as centrality, betweenness, and closeness (Everett & Borgatti, 2012; Krackhardt, 1990; Marsden, 2002). In our study, we used two metrics: (1) Freeman's degree centrality and (2) k-core.

Degree centrality refers to the number of actors in which a specific actor is connected without an intermediary (Agneessens et al., 2017). A node is considered central when bonded to many actors (Kim et al., 2011). Eq. 1 presents the mathematical configuration of degree centrality (Wasserman & Faust, 1994; Wichmann & Kaufmann, 2016). Degree centrality in a supply chain is measured by the number of direct ties to a node (Wasserman & Faust, 1994; Wichmann & Kaufmann, 2016).

$$C_d\left(n_i\right) = \sum_j x_{ij} = \sum_j x_{ji} \tag{1}$$

where $C_d(n_i)$ corresponds to the centrality degree for an actor i and x_{ij} or x_{ji} is the binary variable equal to 1 (or another data) when occurring a link between n_i and n_j; or 0 otherwise (Freeman, 1978; Nieminen, 1973).

The approach of degree centrality establishes two axes, outdegree when one flow is initiated or indegree when one flow is received. Indegree and outdegree centrality present the size of the adjacent upstream and downstream tiers (Powell et al., 1996). The outdegree centrality equation can be seen in Eq. 2 and indegree centrality in Eq. 3 (Wasserman & Faust, 1994).

$$C'_d out\left(n_i\right) = \sum_{j=1}^{n} X_{ij} \tag{2}$$

$$C'_d in\left(n_i\right) = \sum_{j=1}^{n} X_{ji} \tag{3}$$

A k-core is a subgraph in which every actor has a degree k or more with the other actors in the subgraph with the purpose to find cohesive subgroups (Borgatti et al., 2013). Moreover, the k-core allows identifying coherent subgroups. In this work, we used the k-core to identify main relationships, subnetworks in soybean and beef supply chains considering the Brazilian export network (Hanneman & Riddell, 2011).

3.3 Computational Procedures

The maps in Microsoft Excel 2016 were created using the latitude and longitude of each country. The strength of colours considers the volume of soybean or beef produced by each country. The coordinates are related to the volumes using the choropleth map function of Excel.

Regarding the networks, the excel matrix is converted to UCINET files and opened in NetDraw environment. There, we included relations' strength, Freeman's degree centrality (indegree and outdegree), and the *k*-core calculations. One model of the network includes the volumes of soybean/beef import and export for each country, the other include the volumes of Brazilian soybean/beef moved from Brazilian states producers to the Brazilian Ports, and from Brazilian Ports to the other countries.

Afterward, the indegree and outdegree metrics were computed in UCINET performing Freeman's degree centrality (Freeman, 1978; Hanneman & Riddle, 2014). First using the volumes among actors and after dichotomising the values to binary data where 1 represents relationship existence and 0 otherwise. The results were organised in tables according to dichotomised values.

4 Findings

4.1 Soybean Supply Chain

Soybean is the fourth most important agricultural world commodity behind corn, wheat, and rice (International Grains Council [IGC], 2021). The main uses of soybean are as food grain, as oil, as animal feedstuff, and in a small quantity as biofuel. A soybean grain contains 17% of oil and 63% of meal where 50% of this meal is protein (Britannica, 2021; United States Department of Agriculture [USDA], 2019).

Despite the fact of being a grain, due to its characteristics, the soybean is classified as an oilseed commodity. About 85% of the world's soybeans are crushed where approximately 98% of the meal is further processed into animal feed; with the balance used to make soybean flour and proteins while the oil fraction, 95% is consumed as edible oil; and the rest is used for industrial products such as fatty acids, soaps, and biodiesel (HighQuest Group, 2020).

To understand how the world production is divided, we plotted Fig. 1 considering the 2019 year using data from FAO.

In 2019, Brazil, the United States, and Argentina were responsible for 77% of world production (33%, 28%, and 16%, respectively). Usually, Brazil and the United States have similar production competing for the first place; however, due to climate issues in 2019, the Brazilian production was slightly higher than the United States. Other producing countries are China (5%), India (4%), Paraguay (2%), Canada (2%), Russia (1%), Ukraine (1%), Bolivia (1%), and Uruguay (1%). Note that none of them have significant production. Indeed, soybean production is highly concentrated in the American continent (Mendes dos Reis et al., 2020).

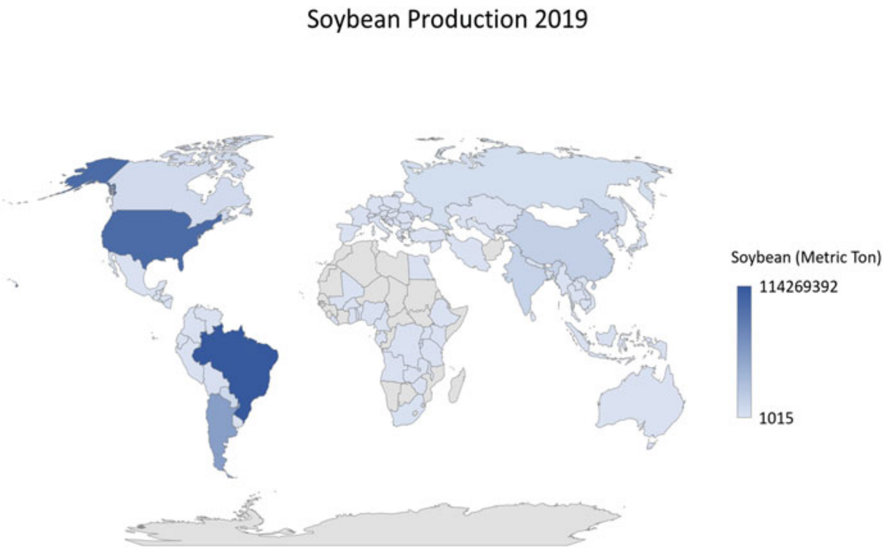


Fig. 1 Producers of soybean (Source: Adapted from FAO (2021) using database 2019 and Microsoft Excel© 16)

Soybean is an East Asian seed being introduced in North America in 1765 and South America at the end of the nineteenth century. The increase of this production occurred during World War II due to the war efforts when Canada and the United States produced oils and lubricants using soybean. In South America, Argentina and Brazil started producing soybean on a commercial scale in the 1940s, mainly motivated by the breeding of soybean cultivars with a long juvenile period, which allowed for the production at very low latitudes (Chang et al., 2015). However, the great expansion raised after the 1990s motivated by the impact of China on the economy. With one-fifth of the world's population and growing industrialisation, the country observed an increase in GDP per capita changing diet and food consumption. Considering that China has between 8 and 10% of arable land, the demand for import soybean raised with the purpose of animal feed (mostly pork production) and edible oil (Sly, 2017). Consequently, it was a matter of time to China country move the direction to the American continent that was producing soybean with available arable land to obtain the amount necessary of oilseeds to attend its internal demand. Besides that, the operation was facilitated by the low cost of sea freight rates (dos Reis et al., 2016).

Based on the perspective of a small number of soybean producing countries, we expected a small number of exporting countries and a higher number of importing countries. Given that, rather than plot a map for importers and exporters, we decided to cross the data using SNA as presented in the methodology section. Also, using data from FAO (2021), UCINET 6.698, and NetDraw 2.138 were plotted Fig. 2.

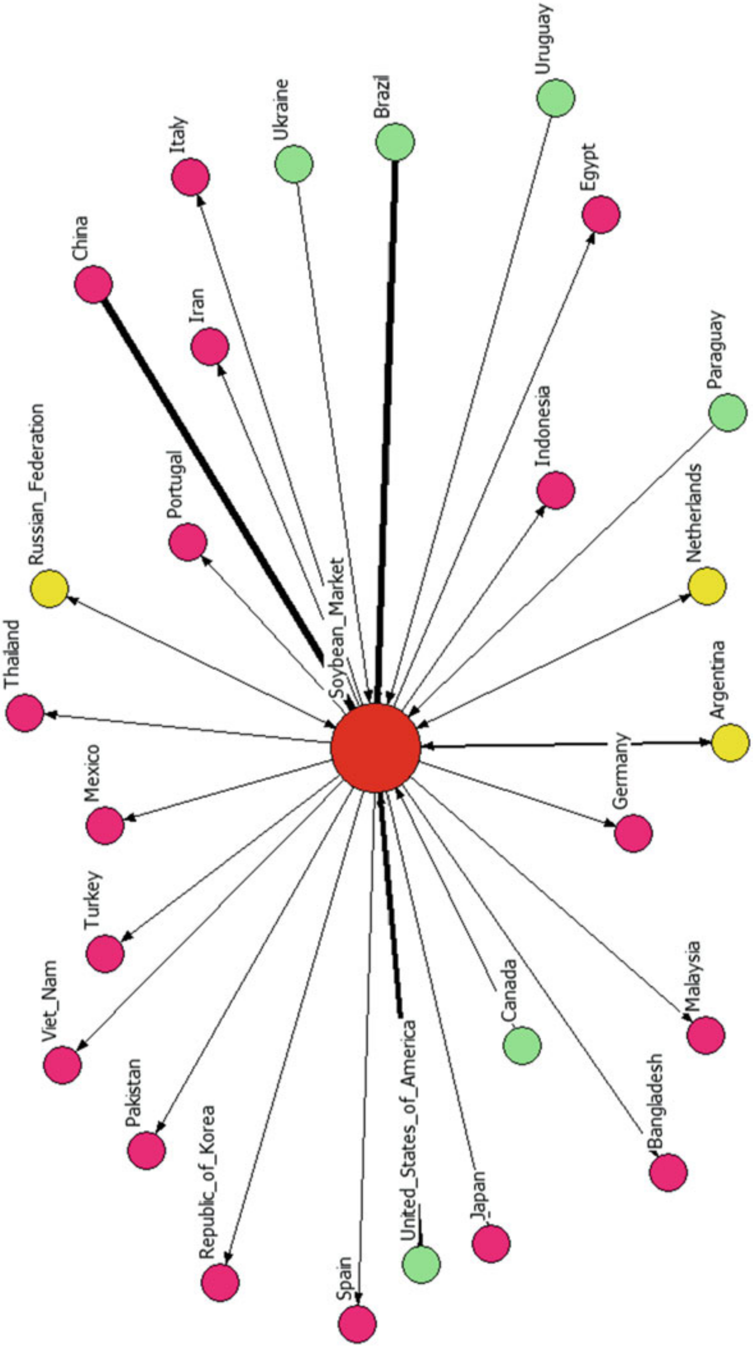


Fig. 2 Soybean market (Source: Adapted from FAO (2021) using database 2019 and NetDraw 2.138 and UCINET 6.698)

The first result indicated, on the contrary to our assumptions, the existence of few significant importing countries (1% or higher). We found 20 countries in this situation corresponding to 94% of the market in 2019 where China represents 60% of the total while Mexico, Argentina, Egypt, Netherlands, and Germany respond to only 3% each. The others represent between 1% and 2% each.

We confirmed that China is the main hub of soybean exports as presented by Sly (2017), unbalancing the bargaining power in the soybean market. On the other hand, the results show a large dependence on the Chinese market for Argentina, Brazil, and the United States. Strategically, this is a fragile situation to exporters being even harder for Argentina and Brazil where agriculture is a relevant part of Gross Domestic Product (GDP). Our results can represent a twist for exporting countries to create policies to motivate newcomers and potential markets.

Another result to be highlighted is the double role of Argentina, the Netherlands, and the Russian Federation. Only these three presented significant volumes of soybean (imported and exported) to appear in the network, but it is very prevalent for exporter countries also to buy the same commodity. It is a common ground for producers to sell for export based on the exchange rate and after the animal producers buy from international markets due to shortage of products in the internal market. Normally, the countries do not regulate these transactions. Finally, note that the Netherlands is not a producer, hence, import and exports are related to their ports that are important hubs to import products to Europe.

With these ideas in mind and in order to verify the logistics role, we chose the Brazilian scenario for three reasons: (1) the highest producer country with 33% of the total, (2) the biggest exporter with 48% of the total, and (3) the large data availability. Figure 3 presents Brazil's map by state.

Initially, using the SNA approach, we created a network considering state producer, port of exportation, and importing country, Fig. 4.

In Fig. 4, we can observe the importing countries, Brazilian states, and ports that move soybean to the international markets. The strength of the lines refers to the volume of soybean transported while the size of nodes represents the number of relationships according to Freeman's degree centrality (Hanneman & Riddle, 2014). Moreover, the proximity of nodes matters, and it means a similar behaviour among them.

The graphical analysis allowed us to identify that South and Southeast ports such as Santos, Paranaguá, São Francisco do Sul, and Rio Grande present the highest volume of soybean exports (strength of lines) forming a cluster. However, the ports of the North—Santarém, Belém—show the highest number of importing countries and connections amongst Brazilian states with a small volume of exports (size of nodes—Freman's degree centrality). Also, in the graphical analysis, we can infer that Mato Grosso and Paraná are the main producers but Bahia, Goiás, and Rio Grande established a cluster connected to many ports creating resilience to the soybean supply in these states. It is possible to verify the importance of China for the soybean transactions. Moreover, some groups of importing countries buy soybean from different states showing a great bond with Brazil such as China, the Netherlands, Thailand, and Spain. Based on these results, we tested the existence of

Fig. 3 Brazilian map (Source: Instituto Brasileiro de Geografia e Estatística [IBGE] (2021))

subgroups in the network. To do so, we plotted the *k*-core network as can be seen in Fig. 5.

The main network, that means these nodes have more connections inside the network, composed by Ports of Belém, Santarém, Paranaguá, Salvador, São Luíz, Vitória, Santos, Rio Grande, São Francisco do Sul, and Imbituba; Brazilian states of Mato Grosso, São Paulo, Goiás, Rio Grande do Sul, Minas Gerais, and Paraná; and countries such as Spain, China, Thailand, Pakistan and Netherlands. The idea of *k*-core is to verify what nodes have a strong relationship based on the number of connections (Hanneman & Riddle, 2014). The network indicates that even though the volume is highly exported by the South and Southeast ports, the Brazilian soybean network is highly connected creating a governmental issue of deciding how to invest and improve these routes.

Eventually, to analyse our results in detail, we established the quantitative analysis from degree centrality considering Freeman indegree (Table 2) and Freeman outdegree (Table 3).

Table 2 indicates that Paranaguá and Santos are the most important ports for soybean export (43.6%). Nevertheless, the North and Northeast ports (Belém, Santarém, Manaus, and São Luiz) are more connected and have been increasing

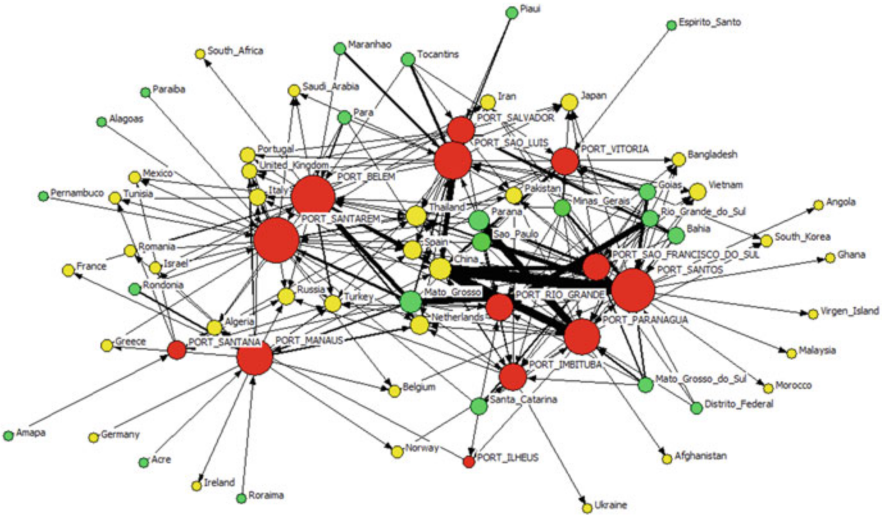


Fig. 4 Brazilian soybean export network (Source: Adapted from ME (2021) using NetDraw 2.138 and UCINET 6.698)

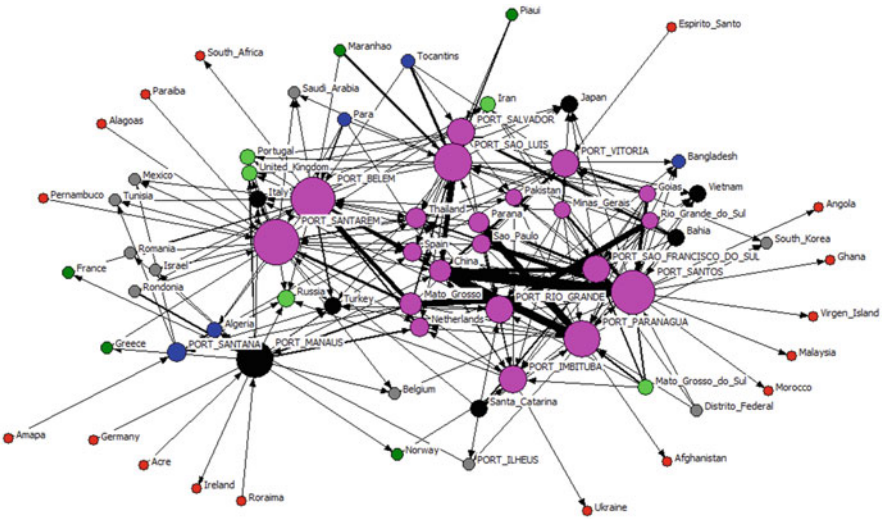


Fig. 5 Brazilian soybean export network *k*-core (Source: Adapted from ME (2021) using NetDraw 2.138 and UCINET 6.698)

the volume representing 27.2% already. Insofar as Brazilian production advances to those areas, the ports in the region raise their participation. At the same time, these ports have proximity with Europe and the Panama channel so that bigger producers in Brazil such as Mato Grosso commence sending the soybean through these corridors. These results confirm what was presented in the graphical analysis. The

Table 2 Freeman's indegree centrality

Node	Indegree Node	Indegree Volume	Volume (%)
Ports			
Belém	19.000	7260760.000	8.8
Santarém	19.000	3708139.000	4.5
Manaus	17.000	3119811.000	3.8
Santos	16.000	21735744.000	26.5
São Luís	14.000	8285500.000	10.1
Paranaguá	11.000	14050985.000	17.1
Salvador	8.000	3283623.000	4.0
Vitória	7.000	4391460.000	5.4
Rio Grande	6.000	9120933.000	11.1
Santana	6.000	187782.000	0.2
São Francisco do Sul	6.000	6095293.000	7.4
Imbituba	5.000	777784.000	0.9
Ilhéus	1.000	62878.000	0.1
Countries			
China	11.000	59961548.000	73.1
Thailand	9.000	2587752.000	3.2
Netherlands	8.000	4277712.000	5.2
Spain	8.000	2818986.000	3.4
Pakistan	7.000	1200107.000	1.5
Turkey	7.000	2134337.000	2.6
Italy	6.000	618225.000	0.8
Japan	6.000	448064.000	0.5
Russia	6.000	1071428.000	1.3
Vietnam	6.000	705219.000	0.9

Source: Authors using UCINET

countries' indegrees indicate the predominance of China (73.1%) which receives cargo from 11 out of 12 Brazilian Ports and all Brazilian states.

The outdegree results demonstrate that the ports of Santos and Paranaguá are the main corridors connected to 10 states. The main reasons for that are (1) The depth of the draft, (2) Port connections with roadways and railroad, and (3) Port capacity. But the results also confirm a network attempt to reconfigure for other routes and diminish the dependence on the principal corridors in the direction of called 'Arch North Ports'. The outdegree also indicates a predominance of Mato Grosso, Paraná, and Rio Grande do Sul; however, it does not necessarily mean a higher number of connections.

The second supply chain that we studied in this chapter was the beef supply chain. Brazil is also one of the most important world beef producers with more than ten million metric tons (FAO, 2021).

Table 3 Freeman's outdegree centrality

Node	Outdegree Node	Outdegree Volume	Volume (%)
Ports			
Santos	10.000	21736260.000	26.5
Paranaguá	10.000	14051403.000	17.1
Santarém	9.000	3706689.000	4.5
São Luís	9.000	8284538.000	10.1
São Francisco do Sul	9.000	6095293.000	7.4
Imbituba	9.000	777784.000	0.9
Belém	8.000	7260792.000	8.8
Rio Grande	8.000	9120933.000	11.1
Vitória	7.000	4390705.000	5.3
Salvador	6.000	3283071.000	4.0
Manaus	4.000	3119766.000	3.8
Santana	2.000	187782.000	0.2
Ilhéus	2.000	62298.000	0.1
Brazilian States			
Mato Grosso	11.000	22305256.000	27.2
Paraná	9.000	13932975.000	17.0
São Paulo	8.000	4910494.000	6.0
Bahia	8.000	3790383.000	4.6
Goiás	7.000	7851350.000	9.6
Minas Gerais	7.000	4607550.000	5.6
Rio Grande do Sul	7.000	8527981.000	10.4
Santa Catarina	7.000	1934017.000	2.4
Mato Grosso do Sul	5.000	4397782.000	5.4
Pará	4.000	2226544.000	2.7
Tocantins	4.000	2553876.000	3.1

Source: Authors using UCINET

To understand how the world production is divided, the same procedure of soybean was adopted, and we plotted Fig. 6 considering 2019 year using data from the Food and Agriculture Organisation from the United Nations (FAO).

In 2019, the main players were the United States (18%), Brazil (15%), China (9%), Argentina (5%), Australia (3%), and Mexico (3%). The United States, Brazil, and Argentina were responsible for 38% of production. If the dominance is not the same as soybean, we can realise that there is a huge connection between meat and soybean production. The participation of other producers is lower than 2%. Perhaps, the same could be seen in chicken and pork supply chains; however, this would require other individual studies.

We also plotted a network using SNA and data from FAO (2021), UCINET 6.698, and Netdraw 2.138, Fig. 7.

Yellow nodes (i.e., light grey in b/w) are exporters, pink nodes (i.e., dark grey in b/w) are importers, except for Madagascar that acts in both cases. Beef marketing is

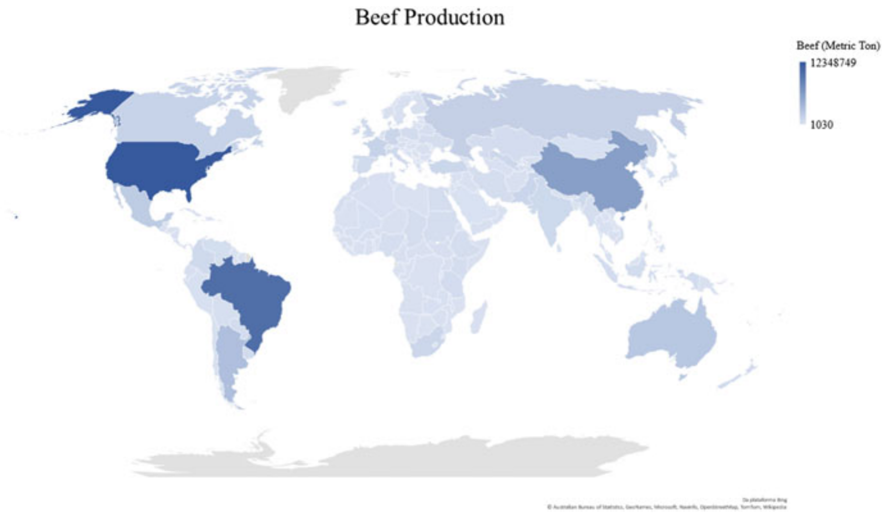


Fig. 6 Beef Producers (Source: Adapted from (FAO, 2021) using database 2019 and Microsoft Excel© 16)

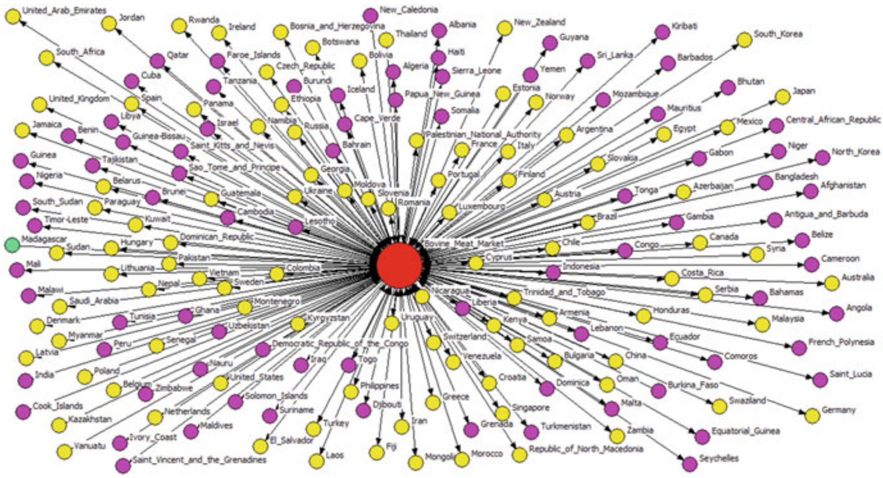


Fig. 7 Beef market (Bovine Meat Market) (Source: Adapted from (FAO, 2021) using database 2019 and NetDraw 2.138 and UCINET 6)

composed of many importers and exporters. Differently from soybean, there are not high-strength lines, and it is very diverse. While chicken and pork supply chains are more concentrated in some regions, beef marketing deals with many trades.

Brazil has the highest number of bovine cattle with almost 215 million heads. Of this total 32.5 million were slaughtered in 2019 (FAO, 2021). Beef is the most important protein produced by the country, Fig. 8.

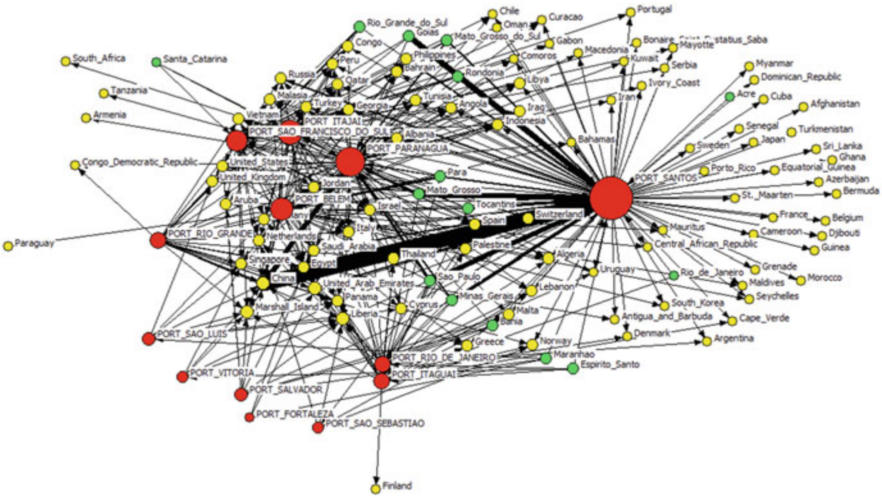


Fig. 8 Brazilian beef export network (Source: Adapted from ME (2021) using NetDraw 2.138 and UCINET 6.698)

In Fig. 8, we depicted Brazilian ports, Brazilian states and importing countries. The strength of the lines refers to export volume. The graphical analysis allowed us to identify that the Port of Santos has the highest number of connections and the highest volume of exports. The Port presents a dissociate behaviour in the network and many countries import only from Santos. This Port is the main foreign trade hub of Brazil with 46 terminals, an operational draught of 13.20 m which its movement diversity allows for more facilities and flexibility than other Brazilian ports where is possible to berth and operate 33 vessels simultaneously (Santos & Hilsdorf, 2019). Nevertheless, it maintains connections with countries that interact with other ports in the network. Ports of Rio Grande, São Francisco do Sul, Paranaguá, Belém, and Itajaí form a cluster. Except for Belém, the others are located in the south of the country. Besides that, the ports of Itaguaí and Rio de Janeiro present the same behaviour in the network because they are located together both in Rio de Janeiro state, around 100 km far.

Regarding the Brazilian states, they form two big clusters: first includes Pará, Mato Grosso, Tocantins, São Paulo, Minas Gerais, Bahia, Espírito Santo, and Maranhão; second Rio Grande do Sul, Goiás, Mato Grosso, and Rondônia. Finally, about the countries, it is possible to note the pivotal role of China as well as in the soybean supply chain.

We explore more these results by analysing sub-groups by the *k*-core network, Fig. 9.

Analysing our results in detail, we established the quantitative analysis from degree centrality considering Freeman indegree (Table 4) and Freeman outdegree (Table 5).

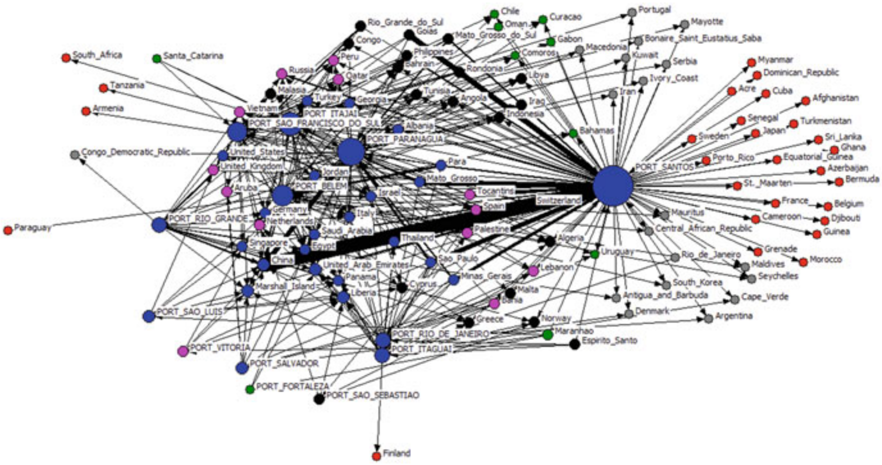


Fig. 9 Brazilian beef export network *k*-core (Source: Adapted from ME (2021) using NetDraw 2.138 and UCINET 6.698). The analysis of *k*-core shows the existence of the main network including 10 Brazilian Ports; the Brazilian states Pará, Mato Grosso, São Paulo, Minas Gerais; and countries like China, Egypt, Thailand, Germany, Albania, Turkey among others. Note that an important Brazilian state producer, Goiás, is in a secondary network due to the few connections.

Port of Santos confirms the results of the graphical analysis being responsible for 68.6% of Brazilian beef exports, receiving for all the states and exporting for 93.6% of importing countries. Besides that, Santos and Paranaguá move 87% of Brazilian beef production. Two aspects can be highlighted: both ports repeat the primary role of soybean, and both are in a more concentrated network.

The indegree result confirms the role of China being responsible for 67.4% of Brazilian meat acquisition. The second most important partner, Egypt, is far from that with 7.4%. The indegree also showed us that six states are the main Brazilian exports, São Paulo, Mato Grosso, Minas Gerais, Goiás, Rondônia, and Mato Grosso do Sul corresponding to 92% of the total altogether). Mato Grosso is the first place in beef as well as soybean showing to be one of the main agriculture states for the country. Mato Grosso's distance from the ports reflects the logistics bottlenecks of Brazil.

5 Discussion

Our findings are discussed in this section using different aspects revolving around the supply chains studied.

Table 4 Freeman's indegree

Node	Indegree node	Indegree volume	Volume (%)
Ports			
Santos	14.000	1091171200.000	68.6%
Paranaguá	10.000	294992256.000	18.5%
Belém	10.000	47580752.000	4.9%
Itajaí	9.000	2202584.000	3.0%
São Francisco do Sul	7.000	28294664.000	1.9%
Rio Grande	4.000	77669584.000	2.7%
Rio de Janeiro	4.000	2697212.000	0.2%
Itaguaí	2.000	2394791.000	0.1%
Salvador	1.000	42394012.000	0.2%
São Luís	1.000	37451.000	0.0%
Vitória	1.000	26189.000	0.0%
São Sebastião	1.000	6670.000	0.0%
Fortaleza	1.000	367700.000	0.0%
Countries			
China	10.000	1072652032.000	67.4%
Liberia	9.000	235247.000	0.0%
Marshall Island	9.000	52921.000	0.0%
United Arab Emirates	9.000	38518196.000	2.4%
Egypt	8.000	117599168.000	7.4%
Italy	8.000	23650258.000	1.5%
Panama	8.000	52300.000	0.0%
Germany	7.000	4199694.000	0.3%
Jordan	7.000	11598578.000	0.7%
Saudi Arabia	7.000	39582216.000	2.5%
Singapore	7.000	18884816.000	1.2%
Thailand	7.000	2613614.000	0.2%

Source: Authors using UCINET

5.1 *Marketing and Trade*

In a certain way, the results of the research prove that we cannot simplify the discussion of food production and demand considering local or non-local food, smallholders or large-scale farm producers, the influence of large corporations and governments, and non-government organisations. We need to behold past and present that mould our history and realise that supply chains are formed based on many factors that should not be analysed separately.

Globalisation in FSCs is supported by centuries of global trade. Since the Age of Discovery started, agriculture has been included in the centre of the question. The access to the new cultures provided new tastes, new discoveries, new fruits and vegetables, and changed the trade among countries forever. According to Paciarotti & Torregiani (2021), the influence of globalisation on food systems

Table 5 Freeman's outdegree

Node	Outdegree Node	Outdegree Volume	Volume (%)
Ports			
Santos	88.000	1091171200.000	68.6%
Paranaguá	54.000	294992256.000	18.5%
Itajaí	36.000	47580752.000	4.9%
Itaguaí	34.000	2202584.000	3.0%
São Francisco do Sul	30.000	28294664.000	1.9%
Belém	22.000	77669584.000	2.7%
Riod Janeiro	19.000	2697212.000	0.2%
Salvador	12.000	2394791.000	0.1%
Rio Grande	8.000	42394012.000	0.2%
São Luís	8.000	37451.000	0.0%
Vitória	5.000	26189.000	0.0%
São Sebastião	4.000	6670.000	0.0%
Fortaleza	2.000	367700.000	0.0%
Brazilian States			
São Paulo	11.000	310004352.000	19%
Mato Grosso	9.000	352084832.000	22%
Minas Gerais	8.000	168954032.000	11%
Para	8.000	110946632.000	7%
Bahia	7.000	2855023.000	0%
Tocantins	7.000	74858304.000	5%
Espírito Santo	7.000	2034968.000	0%
Goiás	7.000	248382672.000	16%
Maranhão	5.000	4407143.000	0%
Mato Grosso do Sul	4.000	130921088.000	8%
Rio Grande do Sul	4.000	43154064.000	3%
Rondônia	4.000	139797040.000	9%

Source: Authors using UCINET

rose in the second half of the nineteenth century, which is confirmed by the high movement in ports around the world. Furthermore, consumers' expectations for the availability of fresh food products encouraged the globalisation of food markets. With growing global competition and the greater distances between food production and consumption locations, there is increasing pressure for the integration of food production and distribution (Nagurney et al., 2013).

Among agriculture cultures in the American continent, we can emphasise sugar and cotton production which started in the sixteenth century (Cottonmill, 2021; Muhammad, 2019), coffee production in the nineteenth century (Watson & Achinelli, 2008), and most recently the soybean production in Argentina, Brazil, and the United States (Mendes dos Reis et al., 2020; Toloi, 2018). These are some examples of the role of global agriculture in world trade.

Many times, agriculture was supported by slavery (Wright, 2003), low-cost domestic and immigrant workforce (Lesser, 2013), children labour (International Labour Organization [ILO], 2021), and nowadays using cutting-edge equipment (Smith, 2020). Unfortunately, as a mark of shame to actual human society, many countries still face the challenges of slavery, children labour, and immigrant and local low-skills workforce (Campbell, 2008; ILO, 2021).

Currently, the great value of agriculture is connected to Trading Companies that control the marketing and logistics. In soybean production, for instance, Toloi (2018) found that prices in the field can be 30% less than Chicago Trade Board.

Interesting to note how the food production is unbalanced considering countries. Despite our work showing that there is a lot of players in soybean and beef supply chains, there is an enormous absorption of the Chinese market for both items. Chinese economic dynamism accelerates the process of pursuit raw materials, food, and natural resources overseas to sustain its industrial activity where South America represents an important source of supply (Sly, 2017).

Despite the world's move in this twenty-first century to nationalism, to create borders and implement barriers, it is unthinkable not to have larger agriculture/livestock producing countries. In Brazil's case, this is the country's vocation since the discovery by Portugal in 1500. Starting with sugar production going to the coffee production and most recently with many organised cultures pulled by the soybean production, these supply chains are responsible for more than 20% of GDP and 36% of Brazilian exports (Arias et al., 2017). After five centuries of supplying European commodity markets, the growth in Asia, mainly to China, has diverted Brazilian commodities to another direction. Our results showed the complete dependence of Brazil on China's purchasing (soybean 73% and beef 69%).

There is one effect not observed directly in the internal shortage. Trading Companies control the supply chain and exert a tremendous lobby over local governments to maintain their business running. Moreover, farmers are interested in the highest profitability possible since they are subject to the uncertainties of agriculture production such as marketing prices and climate risks. Therefore, many countries sell to external markets products needed in the domestic one.

In 2020, due to the Pandemic of COVID-19 crisis, Brazil sold great quantities of soybean to external markets creating a shortage of soybean for animal feedstuff and oil production. As a result, the animal protein and soybean oil prices exploded in the domestic market because it was necessary to import soybean at a moment that the Brazilian currency lost more than 30% of its face value before the US Dollar (Pamplona, 2020).

Having said that, our results allowed us to infer some marketing considerations: (1) there are many import players in the soybean and beef marketing, but the competition is not equal among them. As in many commodities markets, China has a lot of influence in this balance; (2) for many commodities it is not possible to only take local food into account. They require a lot of natural resources that directly impact food prices and hence it requires imports to alleviate internal prices; and (3) the exportations are supported by an extensive maritime network and low prices of freight.

5.2 Logistics Management Impact

Logistics management is vital to linking agricultural producers, the food industry, and the final consumer. The path between 'farm to fork' involves many transportation and storage operations that are essential to balance supply and demand and ensure world population food access.

Nevertheless, the agricultural network is composed of many developing countries that have seen in agriculture an opportunity to compete on international trade marketing and improve the welfare of their inhabitants. Analysing the main logistics aspects that affect international trade such as infrastructure, custom, logistics competence, among others, the developed countries have excellence in infrastructure while the developing ones depend on their competence to move cargo in a sheer deficient system (Arvis et al., 2018).

Brazil is recognised for its poor logistical infrastructure that affects domestic costs in comparison to its main competitors and reduces producers' profitability (Salin, 2019), presenting situations as in Fig. 10.

Nevertheless, the country presents some exceptions. Santos and Paranaguá ports, for instance, present a better infrastructure quality of roads and port capacity (Souza et al., 2017; Toloi et al., 2016) and this condition is the primary reason for using these corridors for agriculture commodities as indicated in our study (43.6 Soybean, and 87.1% Beef) as can be seen in Fig. 11.

Another important discussion regarding agriculture commodities is that what really moves the trade is the low cost of maritime freight. A metric ton of soybean from Brazil or the United States to China revolves around 50 US dollars (Salin, 2019) that are low cost in comparison to planting the grain, importing fertilisers,

Fig. 10 Road BR 158 that connects Pará to Rio Grande do Sul state (Source: Authors)

Fig. 11 Ports of Paranaguá and Santos (Source: Adapted from Agência Estadual de Notícias do Estado do Paraná [AEN], (2021) and Santos Port Authority (2021))

suffering from climate risks etc. Therefore, logistics management has a significant role in FSCs because connecting countries allows them to buy their food necessities or export the agriculture surplus.

On the other hand, all these logistics movements cause a lot of gas emissions, cost contamination, weakness of internal supply chains, and huge competition among countries for the same resources. When it is inexpensive to buy outside, why produce internally? Li & Hanafi (2013) confirmed that transportation is one of the main contributors to greenhouse gases and optimising logistics service with integrated economic and ecological objectives reduces the number of carbon emissions and improves operations efficiency. The author proposed strategies to diminish carbon emissions for UK FSCs.

5.3 Pandemic and New Ruptures in FSCs

The 2020 year brought another element to the discussion. The impact of diseases and risks of food contamination. The advent of COVID-19 made many countries raise their internal stocks concerned about a rupture of supply chains. In the same way, exporter countries did not limit the sales. As a result, it increased the imbalance in supply chains and created a recession in exporting countries that saw the internal prices skyrocket because of the shortage.

Coluccia et al. (2021) explain that fresh and perishable products, whose production or harvest took place during the first wave of COVID-19, have suffered price level effects, while storable products have not registered significant impacts due to the vulnerability of the harvest and production phases, which affected fresh and perishable products supply.

The use of antibiotics and agriculture defensives are a reality in FSCs and maybe represents a serious risk of provoking resistance to the bacteria, viruses, and generating all sort of diseases. Unfortunately, the COVID-19 is one of the crises but not the least. The concern in FSCs has been strengthened by several sector-wide crises such as the BSE crisis, the dioxin crisis, and the swine fever (van der Vorst, 2006). All of them seem to be generated by the FSCs and we need to have a concern about that. As it can be seen in our results, FSCs are connected, and diseases are overspread fast using the human body and the countless ships over the seas.

Despite FSCs being heavily regulated, it is possible to observe an increase in the spotlight for safety scares, recalls, and disruptions. FSCs are particularly fragile owing to the geographic, economic, and legislative spread of participating entities and any materialisation of risk quickly permeates across the different players (Dani & Deep, 2010).

The question is whether the human being in charge of the countries will use their animal instincts to save food and reduce access over the world increasing poverty and hungry or they will permit the resilience of supply chains recovering the economical balance.

6 Conclusions

The present chapter discussed the role of logistics management in FSCs using as an example two international large chains: Soybean and Beef. We used Social Network Analysis, the world volume of production, the Brazilian export network connections, and volumes to figure out the importance of these networks for food nutrition and how logistics impact these marketing creations and relationships.

Our results reveal that China is the most relevant commodity world import due to population size and economic growth. This fact is creating dependence on exporting countries, as exemplified by Brazil, and mining the participation in the marketing for other countries. Moreover, we realised that the low-cost land acquisition, productivity, and efficiency of the maritime network make it attractive to purchase agriculture products from big players rather than to adopt a local food strategy.

Despite the current pandemic crisis and other diseases and implications that put at risk international food network, the dependence on natural and financial resources will keep pushing the countries in the direction of suppliers that offer the best prices in connection with good quality. There is a trend to reconfigure these FSCs where resilience, information technology, low-cost logistics, and income prevail against the uncertainties and food security. The pandemic proves that food access is not only a question to have food production, reducing food losses and waste, but it is a financial condition. The countries that have more income and planning buy items to create stocks generating a shortage in the more fragile ones. Thus, they see their population implode to the incapacity of having enough resources for food acquisition. The unemployment rate explodes and countries that are offering aid for the population are better at reducing the impacts than others that decide to deny the reality. It is not a mistake only in developing countries, but great economic nations as the United States saw the no-welfare condition policy move millions to the poverty level.

This study provided many managerial implications as the necessity to improve the logistics process to be more competitive in international food markets and turn feasible any attempt to increase local food participation. To have natural resources is strategic to produce food; however, income is more important than that once a high income makes it possible to obtain food access. The FSCs seem resilient enough to avoid collapses but it is necessary to develop policies for food access and establish a minimum level of welfare. The SNA approach allows creating a bridge between theoretical analysis and assumptions to practice.

The research is not free of limitations; hence, it is essential to comprehend that empirical research of the players involved may contribute to a better understanding of FSCs issues. The database used is subject to eventual errors for organisations that collected them implying the obligation of seeing the results as exploratory instead of definitive. We only could analyse two supply chains with similar behaviour. However, despite all these remarks, the results presented in this chapter cannot be invalidated and can contribute to explaining a little bit more about FSCs behaviour and logistics implications.

Finally, as future research, we suggest analysing FSCs implications comparing local SFSCs with a large-scale international one to obtain a managerial understanding of how the world can create a hybrid food production system so that it can tackle any risk of shortage and food insecurity. Moreover, we believe that it is necessary to discuss FSCs issues not only considering economic and logistics aspects, urge to include sustainable development and the option of countries agriculture for monoculture. No doubt changes in the environment affect agriculture and the ability of soil to recover. The extractive industry model must to be rethink considering future generations. Furthermore, the connections among different food supply chains should be considered as well as the role of the consumer in generating food losses and wastes believing in a model with infinite resources.

References

Abraham, E. R., Mendes dos Reis, J. G., Vendrametto, O., de Oliveira Costa Neto, P. L., Carlo Toloi, R., de Souza, A. E., & de Oliveira Morais, M. (2020). Time Series Prediction with Artificial Neural Networks: An Analysis Using Brazilian Soybean Production. *Agriculture, 10*(10), 475. https://doi.org/10.3390/agriculture10100475

Agência Estadual de Notícias do Estado do Paraná, A. E. de N. do E. do. (2021). *Galeria de Imagens—Agência Estadual de Notícias*. http://www.aen.pr.gov.br/modules/galeria/detalhe.php?foto=405807&evento=60458

Agneessens, F., Borgatti, S. P., & Everett, M. G. (2017). Geodesic based centrality: Unifying the local and the global. *Social Networks, 49*, 12–26. https://doi.org/10.1016/j.socnet.2016.09.005

Arias, D., Vieira, P. A., Contini, E., Farinelli, B., & Morris, M. (2017). *Agriculture Productivity Growth in Brazil*. World Bank.

Arvis, J. F., Ojala, L., Wiederer, C., Shepherd, B., Raj, A., Dairabayeva, K., & Kiiski, T. (2018). Connecting to compete 2018: Trade logistics in the global economy. *World Bank*. https://openknowledge.worldbank.org/handle/10986/29971.

Behnke, K., & Janssen, M. F. W. H. A. (2020). Boundary conditions for traceability in food supply chains using blockchain technology. *International Journal of Information Management, 52*, 101969. https://doi.org/10.1016/j.ijinfomgt.2019.05.025

Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2013). Analyzing social networks. *SAGE*.

Britannica. (2021). Soybean. In *Encyclopedia Britannica*. https://www.britannica.com/plant/soybean

Campbell, J. (2008). A Growing Concern: Modern Slavery and Agricultural Production in Brazil and South Asia. *Human Rights & Human Welfare*, 131–141.

Chaboud, G., & Moustier, P. (2020). The role of diverse distribution channels in reducing food loss and waste: The case of the Cali tomato supply chain in Colombia. *Food Policy, 98*, 101881. https://doi.org/10.1016/j.foodpol.2020.101881

Chang, W.-S., Lee, H.-I., & Hungria, M. (2015). Soybean production in the Americas. In B. Lugtenberg (Ed.), *Principles of plant-microbe interactions* (pp. 393–400). Springer International Publishing. https://doi.org/10.1007/978-3-319-08575-3_41

Cheah, I., & Shimul, A. (2018). Consumer ethnocentrism, market mavenism and social network analysis. *Australasian Marketing Journal (AMJ)*. https://doi.org/10.1016/j.ausmj.2018.05.017

Christopher, M. (2011). *Logistics & supply chain management* (4th ed.). Financial Times Prentice Hall.

Coluccia, B., Agnusdei, G. P., Miglietta, P. P., & De Leo, F. (2021). Effects of COVID-19 on the Italian agri-food supply and value chains. *Food Control, 123*, 107839. https://doi.org/10.1016/j.foodcont.2020.107839

Cottonmill. (2021). *A brief history of cotton in America*. https://www.cottonmill.com/a-brief-history-of-cotton-in-america/

Dani, S. (2015). *Food supply chain management and logistics: From farm to fork*. Kogan Page.

Dani, S., & Deep, A. (2010). Fragile food supply chains: Reacting to risks. *International Journal of Logistics Research and Applications, 13*(5), 395–410. https://doi.org/10.1080/13675567.2010.518564

de Gorter, H., Drabik, D., Just, D. R., Reynolds, C., & Sethi, G. (2020). Analyzing the economics of food loss and waste reductions in a food supply chain. *Food Policy, 98*, 101953. https://doi.org/10.1016/j.foodpol.2020.101953

Despoudi, S. (2020). Challenges in reducing food losses at producers' level: The case of Greek agricultural supply chain producers. *Industrial Marketing Management.*https://doi.org/10.1016/j.indmarman.2020.09.022

dos Reis, J. G. M., Amorim, P., & Cabral, J. A. S. (2016). Effects of Price and Transportation Costs in Soybean Trade. In I. Nääs, O. Vendrametto, J. Mendes Reis, R. F. Gonçalves, M. T. Silva, G. von Cieminski, & D. Kiritsis (Eds.), *Advances in Production Management Systems. Initiatives for a Sustainable World* (Vol. 488, pp. 563–570). Springer International Publishing. https://doi.org/10.1007/978-3-319-51133-7_67

Everett, M. G., & Borgatti, S. P. (2012). Categorical attribute based centrality: E–I and G–F centrality. *Social Networks, 34*(4), 562–569. https://doi.org/10.1016/j.socnet.2012.06.002

FAO. (2017). *The future of food and agriculture: Trends and challenges*. Food and Agriculture Organization of the United Nations.

FAO. (2021). *FAOSTAT*. http://www.fao.org/faostat/en/

Foreign Agricultural Service/USDA. (2019). Oilseeds: World Market and Trade. https://downloads.usda.library.cornell.edu/usda-esmis/files/tx31qh68h/7d2797406/cj82kp34g/oilseeds.pdf.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks, 1*(3), 215–239. https://doi.org/10.1016/0378-8733(78)90021-7

Gholami-Zanjani, S. M., Klibi, W., Jabalameli, M. S., & Pishvaee, M. S. (2021). The design of resilient food supply chain networks prone to epidemic disruptions. International Journal of Production Economics, 233, 108001. https://doi.org/https://doi.org/10.1016/j.ijpe.2020.108001.

Glover, J., & Touboulic, A. (2020). Tales from the countryside: Unpacking "passing the environmental buck" as hypocritical practice in the food supply chain. *Journal of Business Research, 121*, 33–46. https://doi.org/10.1016/j.jbusres.2020.06.066

Hanneman, R. A., & Riddell, M. (2011). A Brief Introduction to Analyzing Social Network Data. In J. Scott & P. J. Carrington (Eds.), *The SAGE Handbook of Social Network Analysis*. SAGE.

Hanneman, R. A., & Riddle, M. (2014). Concepts and Measures for Basic Network Analysis. In I. J. Scott & P. Carrington (Eds.), *The SAGE Handbook of Social Network Analysis* (pp. 340–369). SAGE Publications Ltd.. https://doi.org/10.4135/9781446294413.n24

HighQuest Group. (2020). Soy facts. *Oilseed & Grain News.*. https://www.oilseedandgrain.com/soy-facts

Huang, M., Wang, Z., & Chen, T. (2019). Analysis on the theory and practice of industrial symbiosis based on bibliometrics and social network analysis. *Journal of Cleaner Production, 213*, 956–967. https://doi.org/10.1016/j.jclepro.2018.12.131

IBGE. (2021). *Brasil em síntese*. https://brasilemsintese.ibge.gov.br/territorio/divisao-politica.html

IGC. (2021). Grain Market Report. https://www.igc.int/en/gmr_summary.aspx

International Labour Organization. (2021). *Child labour in agriculture (IPEC)*. http://ilo.org/ipec/areas/Agriculture/lang--en/index.htm

Jones, P., Comfort, D., & Hillier, D. (2004). A case study of local food and its routes to market in the UK. *British Food Journal, 106*(4), 328–335. https://doi.org/10.1108/00070700410529582

Kim, Y., Choi, T. Y., Yan, T., & Dooley, K. (2011). Structural investigation of supply networks: A social network analysis approach. *Journal of Operations Management, 29*(3), 194–211. https://doi.org/10.1016/j.jom.2010.11.001

Krackhardt, D. (1990). Assessing the Political Landscape: Structure, Cognition, and Power in Organizations. *Administrative Science Quarterly, 35*(2), 342–369. https://doi.org/10.2307/2393394

Kukovic, D., Topolsek, D., Rosi, B., & Jereb, B. (2014). A comparative literature analysis of definitions for logistics:between general definition and definitions of subcategories. *Business Logistics in Modern Management, 14*, 111–122. https://ideas.repec.org/a/osi/bulimm/v14y2014p111-122.html.

Lesser, J. (2013). *Immigration, Ethnicity, and National Identity in Brazil, 1808 to the Present*. Cambridge University Press.

Li, D., & Hanafi, Z. (2013). A Study of Eco-Performance of Logistics Services in Food Supply Chains. In Z. Zhang, R. Zhang, & J. Zhang (Eds.), *LISS 2012* (pp. 223–228). Springer. https://doi.org/10.1007/978-3-642-32054-5_33

Loiseau, E., Colin, M., Alaphilippe, A., Coste, G., & Roux, P. (2020). To what extent are short food supply chains (SFSCs) environmentally friendly? Application to French apple distribution using Life Cycle Assessment. *Journal of Cleaner Production, 276*, 124166. https://doi.org/10.1016/j.jclepro.2020.124166

Loo, C. (2019). Access to Food. In D. M. Kaplan (Ed.), *Encyclopedia of Food and Agricultural Ethics* (pp. 9–15). Springer. https://doi.org/10.1007/978-94-024-1179-9

Marsden, P. V. (2002). Egocentric and sociocentric measures of network centrality. *Social Networks, 24*(4), 407–422. https://doi.org/10.1016/S0378-8733(02)00016-3

ME. (2021). *Comex Stat*. http://comexstat.mdic.gov.br/pt/home

Mendes dos Reis, J. G., Sanches Amorim, P., Pereira, S., Cabral, J. A., & Toloi, R. C. (2020). The Impact of Logistics Performance on Argentina, Brazil, and the US Soybean Exports from 2012 to 2018: A Gravity Model Approach. *Agriculture, 10*(8), 338. https://doi.org/10.3390/agriculture10080338

Messner, R., Johnson, H., & Richards, C. (2021). From surplus-to-waste: A study of systemic overproduction, surplus and food waste in horticultural supply chains. *Journal of Cleaner Production, 278*, 123952. https://doi.org/10.1016/j.jclepro.2020.123952

Mittal, A., Krejci, C. C., & Craven, T. J. (2018). Logistics Best Practices for Regional Food Systems: A Review. *Sustainability, 10*(1), 168. https://doi.org/10.3390/su10010168

Muhammad, K. G. (2019, August 14). The Barbaric History of Sugar in America (Published 2019). *The New York Times*. https://www.nytimes.com/interactive/2019/08/14/magazine/sugar-slave-trade-slavery.html, https://www.nytimes.com/interactive/2019/08/14/magazine/sugar-slave-trade-slavery.html

Nagurney, A., Yu, M., Masoumi, A. H., & Nagurney, L. S. (2013). *Networks Against Time*. Springer. https://doi.org/10.1007/978-1-4614-6277-4

Nieminen, U. J. (1973). On the centrality in a directed graph. *Social Science Research, 2*(4), 371–378. https://doi.org/10.1016/0049-089X(73)90010-0

Paciarotti, C., & Torregiani, F. (2021). The logistics of the short food supply chain: A literature review. *Sustainable Production and Consumption, 26*, 428–442. https://doi.org/10.1016/j.spc.2020.10.002

Palomo-Campesino, S., Ravera, F., González, J. A., & García-Llorente, M. (2018). Exploring Current and Future Situation of Mediterranean Silvopastoral Systems: Case Study in Southern Spain. *Rangeland Ecology & Management, 71*(5), 578–591. https://doi.org/10.1016/j.rama.2017.12.013

Pamplona, N. (2020). Brasil importa soja dos EUA para repor exportação para China. *Folha de São Paulo.*. https://www1.folha.uol.com.br/mercado/2020/11/brasil-importa-soja-dos-eua-para-repor-exportacao-para-china.shtml

Powell, W. W., Koput, K. W., & Smith-Doerr, L. (1996). Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. *Administrative Science Quarterly, 41*(1), 116–145. https://doi.org/10.2307/2393988

Salin, D. L. (2019). *Soybean Transportation Guide: BRAZIL 2018*. USDA.

Santos, M. C., & Hilsdorf, W. de C. (2019). Planning and organization of road port access: The case of the Port of Santos. *Transportation Research Part D: Transport and Environment, 75*, 236–248. https://doi.org/10.1016/j.trd.2019.08.030

Santos Port Authority. (2021). Santos Port Authority. http://www.portodesantos.com.br/conheca-o-porto/o-porto-de-santos/.

Sly, M. J. H. (2017). The Argentine portion of the soybean commodity chain. *Palgrave Communications, 3*(1), 1–11. https://doi.org/10.1057/palcomms.2017.95

Smith, H. P. (2020). *Farm Machinery and Equipment*. Read Books Ltd..

Souza, A., Reis, J., Abraham, E., & Machado, S. (2017). Brazilian Corn Exports: An Analysis of Cargo Flow in Santos and Paranagua Port. AICT-514, 105–112. https://doi.org/10.1007/978-3-319-66926-7_13.

Stranieri, S., Riccardi, F., Meuwissen, M. P. M., & Soregaroli, C. (2021). Exploring the impact of blockchain on the performance of agri-food supply chains. *Food Control, 119*, 107495. https://doi.org/10.1016/j.foodcont.2020.107495

Thapa Karki, S., Bennett, A. C. T., & Mishra, J. L. (2020). Reducing food waste and food insecurity in the UK: The architecture of surplus food distribution supply chain in addressing the sustainable development goals (Goal 2 and Goal 12.3) at a city level. *Industrial Marketing Management*. https://doi.org/10.1016/j.indmarman.2020.09.019

Thomé, K. M., Cappellesso, G., Ramos, E. L. A., & Duarte, S. C. de L. (2021). Food Supply Chains and Short Food Supply Chains: Coexistence conceptual framework. Journal of Cleaner Production, 278, 123207. https://doi.org/https://doi.org/10.1016/j.jclepro.2020.123207.

Toloi, R.C. (2018). *Análise da Cadeia de Suprimentos da Soja no Estado de Mato Grosso* [Universidade Paulista]. https://www.unip.br/presencial/ensino/pos_graduacao/strictosensu/eng_producao/ss_eng_producao_teses2018.aspx.

Toloi, Rodrigo Carlo, Reis, J. G. M., Vendrametto, O., Machado, S. T., & Rodrigues, E. F. (2016). How to improve the logistics issues during crop soybean in Mato Grosso State Brazil. *6th International Conference on Information Systems, Logistics and Supply Chain., Bordeaux. Anais... In: INTERNATIONAL CONFERENCE ON INFORMATION SYSTEMS, LOGISTICS AND SUPPLY CHAIN-ILS. Bordeaux: KEDGE Business School*.

van der Vorst, J. G. A. J. (2006). Product traceability in food-supply chains. *Accreditation and Quality Assurance, 11*(1–2), 33–37. https://doi.org/10.1007/s00769-005-0028-1

Wajszczuk, K. (2016). The Role and Importance of Logistics in Agri-Food Supply Chains: An Overview of Empirical Findings. *Logistics and Transport, 30*(2), 10.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.

Watson, K., & Achinelli, M. L. (2008). Context and contingency: The coffee crisis for conventional small-scale coffee farmers in Brazil. *The Geographical Journal, 174*(3), 223–234. https://doi.org/10.1111/j.1475-4959.2008.00277.x

Wichmann, B. K., & Kaufmann, L. (2016). Social network analysis in supply chain management research. *Int Jnl Phys Dist & Log Manage, 46*(8), 740–762. https://doi.org/10.1108/IJPDLM-05-2015-0122

Wright, G. (2003). Slavery and American Agricultural History. *Agricultural History, 77*(4), 527–552. JSTOR. https://www.jstor.org/stable/3744933

Yustiawan, Y., Maharani, W., & Gozali, A. A. (2015). Degree Centrality for Social Network with Opsahl Method. *Procedia Computer Science, 59*, 419–426. https://doi.org/10.1016/j.procs.2015.07.559

A New Paradigm of Transport in the Wake of a New Age?

Yücel Candemir

Abstract In the present process of development, connectivity as sine qua non of the contemporary global economy, it has risen to prominence to define the trends of the twenty-first century. As this concept has physical and digital dimensions, it takes place on a network basis in the present global economy and is an attribute of these networks and is a measure of the quality of their performance.

Our envisaged game depends on this mimesis, performed by the actions of real-world players like businesses, agencies of governmental and international agencies at both global and local levels on themes determined by the rules of the time. The rules are set by the institutional and technological setup. In the meantime, the background is subject to a permanent change.

The aim of our study is to investigate the working of the global economy structured around the framework we summarized.

Our methodology is rather theoretical based on a model which comprises the basic sectors of the economy and the interconnections between them.

We come to the conclusion that the world is passing through another revolutionary transformation period which will determine the future structure of the world economy. It is important to reduce the costs of a transformation process if and when we develop an analytic framework beneficial for workable policy making.

Keywords New Paradigm · Transformation · Connectivity · Unbundling · New age

Abbreviations

1IR	First Industrial Revolution
2IR	Second Industrial Revolution

Y. Candemir (✉)
Istanbul Technical University,İstanbul, Turkey
e-mail: candemiry@itu.edu.tr

Y. I. Topcu et al. (eds.), *New Perspectives in Operations Research and Management Science*, International Series in Operations Research & Management Science 326,
https://doi.org/10.1007/978-3-030-91851-4_22

GSC Global Supply Chains
GVC Global Value Chains
ICT Information and Communication Technology
MGI McKinsey Global Institute
OECD Organisation for Economic Cooperation and Development
WEF World Economic Forum
WTO World Trade Organization

1 Introduction

This chapter is going to be a shorter one than otherwise. In other words, we will not harbor on a thoroughgoing study here. This is certain as the first and last of keywords read: *New Paradigm* and *New Age*. Both inherit uncertainty which reflects some degree of obscurity. The author is neither so talented nor so equipped as to speculate into the depth of these pretty wet elements of a taxonomy for the future.

When the respectable British weekly, The Economist, has written in 1999 that "*In the summer of 1999, with a qualm or two about the Asian economies and the mess in Russia, he*[1] *still believes that "liberal democracy and a market-oriented economic order are the only viable options for modern society"—just the culmination of the politico-economic process that Hegel had in mind when he talked about the end of history after Napoleon's victory at Jena in 1806. Mr. Fukuyama has lately grown alarmed about what is happening in the world of biotechnology, the arrival of mood-altering drugs such as Prozac and Ritalin and the possibility that science may soon be able to create "a new kind of human being", which could herald the arrival of "post-human history*."[2]

Now, without going into the depths of a sophisticated analysis, we can claim that the world is in the midst of a similar but deeper interim period with a perhaps sharply different situation that is on a razor's edge. Then, apart from a malignant virus, the so-called COVID-19, a much greater question looms large in the offing: Climate change and the affiliated issues of environmental outlook.

Before this gloomy picture has been set in, a revolutionary process of transformation in the world economy was already in progress. Beginning in the 1970s in the twentieth century, an unprecedented development of the Information and Communication Technology, ICT, or Digital Technology, in short, has been in action to mold and transform every segment of human life among them economy.

The obscure nature of the future is not the only reason for the shortness of this study. The author has a random model in mind. In an attempt for a brief depiction, the character of the abundance of factors to be considered in an effort to predict the

[1] The Economist article refers to Francis Fukuyama and his notable article "*The End of History*" in the American magazine, *The National Interest*, in 1989.

[2] The Economist. "***A Survey of the New Geopolitics-The road to 2050***." The Economist July 31, 1999.

nature of the future should be deemed to play a dominant role in the new paradigm. In other words, the sequential nature of the variables of a sample (global) economy should leave this main linear quality in a real, actual new nonlinear structure. This is because they themselves will be transformed into new identities and can no longer be relied upon in a new analysis as a result of the new interactive model. We will be discussing this rather abstract approach below.

Our aim of composing this chapter is simply to draw the attention of the reader(s) to the gravity of the process of change and transformation our planet is passing through and the inevitable outcome of the change of ruling paradigms that are in the offing. As we know nothing about them yet, we had to resort to long quotations in a desperate attempt to underline the importance of our main theme.

One final word is on the extent of transport in this study. For reasons we put in our closing remarks below, we will not treat transport per se in our study in its full context. We know that the transport activity and the infrastructure it operates on are derived from *other sectors'* functioning and as we study these other sectors the transport sector appears to be studied in parallel as the offshoot of them.

The main theme of our study can be labelled "*world economy in transition*." This topical concept has been and is subject to change as the developments in the structure of the world economy themselves are subject to transform. Due to the importance of the topic, there is a wide range of studies in this field.

We can classify this theme into relevant chapters like economic epochs, historical sequencing categorization of them, geographical division, globality, and global shifts. There is a vast number of figures studying the world economy from different perspectives. These can be classified into individual authors, research organizations in universities and independent ones, and international organizations alike.

In the category of individual researchers, Angus Maddison of the Univ. of Groningen is the most prominent to divide the economic history into epochs with a geographical extension. Richard Baldwin and Peter Dicken are among the researchers to evaluate the historical changes with special attention to global supply and value chains. The names of Gary Gereffi, Marcel Timmer, Bart Los, and L. Tavasszy can be named a few of the many.

Within the private research organizations McKinsey Global Institute, World Economic Forum, and the Economist Intelligence Unit can be mentioned. International organizations like OECD, WTO, World Bank International Transport Forum have all paid serious attention to a wide range of topics. In search of the Literature, Park et al. (2013) Study for WTO and FGI is worth to be mentioned. Finally, and not unimportantly, the role of some respectable press and magazines should not be underestimated in terms of keeping attention on worldly affairs. After all, global issues can best be dealt with effectively if public opinion is considered a rewarding stakeholder in this regard.

2 World Economy in Transition

> *When the capital development of a country becomes a by-product of the activities of a casino, the job is likely to be ill-done*. J.M. Keynes. General Theory of Employment, Interest, and Money.

Nearly two decades ago, a group of scientists have made a remark on the then timely developments taking place in the world on a global scale: "*In the last decades we have seen a steady growth of international trade and international transport. Logistics chains are constantly changing, to facilitate these increasingly global movements. In qualitative terms, long term trends in logistics services indicate a growing degree of product customization and an increased responsiveness in order delivery. These developments impact on the development of technology and the growth of welfare in different world regions in different ways.*"[3] This was a clear-cut description of what was going on in the world then and till approximately a year ago. The world did not expect any disruption to this on a scale invented by a Pandemic, COVID-19 virus. As of today, a year after the eruption of this, in navigating a "*new normal*" supply chains issue is still at a focal point.

After the breaking point of Bretton Woods meeting in 1944, the world economy has experienced through a long period of prosperity for nearly half a century till late 1990s, with an interim era beginning in 1985 and named as "*Great Moderation*," a span of unparalleled macroeconomic stability and with prominence of institutions which epitomized the neo-liberal, free-market ideology until a beginner decade ending by 2008. Then, as the Lehman Brothers, symbolizing those institutions, collapsed, the critical question comes into effect: "Does the economic turmoil that began in 2008 herald 'the end of globalization'? Well, it all depends on what we mean by 'globalization.'"[4] The answer lies in the assessment of "the actual *structural* changes that are occurring in the way the global economy is organized and integrated."[5]

In fact, the paragraph above can perhaps be labelled as the most condensed summary of the last 50 years of the world economy. The world economic order which has been devised by the Bretton Woods could not meet the expectations of its architects. The way out to the answer above lies in the metaphor of history of the genesis of the world economy before and after the mid-twentieth century. Before then,[6] the world has been divided between the industrial powers of the *West and the rest*[7] on a basis of industrialization fed by a *core–periphery configuration*. The global economic map of this situation endured until WWII, which shattered the bases of this division, mainly by the appearance of Capitalist–Socialist Blocs

[3] Ruijgrok et al. (2002).

[4] Dicken (2011), p. 1.

[5] ibid.

[6] Especially before 1914.

[7] Here, though not among our main references, we deem it beneficial for those who might have an interest to refer: Ferguson (2011). *Civilisation: The West and the Rest*.

and the so-called "Third World." Two compelling features have characterized the global economy, namely the emblematic concept of *economic development* and the growing *interconnectedness*. Among them, we can note two interconnected fields, namely a steady growth of international trade and international transport. "*Logistics chains are constantly changing, to facilitate these increasingly global movements. In qualitative terms, long term trends in logistics services indicate a growing degree of product customization and an increased responsiveness in order delivery. These developments impact on the development of technology and the growth of welfare in different world regions in different ways.*"[8]

3 Emerging Sectors or Emerging New Paradigms

Global supply chains, GSC, have transformed the world economy over the last 30 years. By making collaboration across borders easier than ever before, innovations will support further growth in supply chains. The Internet of Things and developments in big data, improved communications technology will all boost supply chains. World trade and production are increasingly structured around global value chains, GVC. "The fact that they are increasingly spread over several countries explains why the value chain is regarded as 'global.'"[9]

According to Baldwin, it is a major misunderstanding that "globalization is often viewed as driven by the gradual lowering of natural and man-made trade costs."[10] He asserts that this is a serious mistake. He advances to claim that globalization has been driven by the developments in two connective fields, more than others: transport and transmission. These are the most effective sectors with which global connectivity blooms to carry trade further ahead. It is certainly not the growth of global trade itself that feeds the globalization process. "*It is a truism that nations have become more interdependent*[11] *through the flows of goods, services, and financial capital since the 1970s. The growing importance of export-oriented industrialization has made integration into the global economy virtually synonymous with development for a number of nations. Most recently, the projection of national production systems across borders through direct investment and international subcontracting has deepened the interdependence and functional integration of the world economy*."[12] This can be specified as "interdependence through growing connectivity" with the result of internationalization of supply chains. In fact, the last era following the 1970s can be characterized as a process of unbundling the world economic order through disruption of long-established geographies of production

[8] Ruijgrok et al. (2002).
[9] OECD (2012).
[10] Baldwin (2012–13).
[11] I underlined.
[12] Gereffi et al. (2001).

via supply chains. "*There has been a huge transformation in the* ***nature*** *and the* ***degree*** *of interconnection in the world economy and, especially, in the speed with which such connectivity occurs, involving both a* ***stretching*** *and an* ***intensification*** *of economic relationship. Without doubt, the world economy is a qualitatively different place from that of only 60 or 70 years ago, although it is not so much more open as* ***increasingly interconnected in rather different ways***."[13]

It is here that two keywords/concepts deserve attention: the "*nature*" of interconnection and the "*speed*" with which this connectivity occurs. Our interpretation of these brings us to the conclusion that this is a sort of "*revolution*." Klaus Schwab, the founder-director of the World Economic Forum, calls this "Fourth Industrial Revolution" which enjoyed widespread acceptance, with some variances in some cases, like Industry 4, or I4. We agree with the revolution side of the process. However, we have skepticism on the very concept of "industrial." Nevertheless, this digress is not directly connected with the theme of this chapter.

3.1 Connectivity

Connectivity, in a way, has been the impetus of human societies for thousands of years. But today it gained an unprecedented momentum to affect and change the forces of reengineering the world and investing huge resources into the transport, energy, and communications infrastructure and in linking the megacities flourishing in every corner of the planet, with revolutionary consequences in every part of human life, societal as well as individual levels. Parag Khanna goes on to say that "*Connectivity is our destiny*."[14] In a way, if we confine ourselves to our main theme, connectivity is the most efficient tool of logistics and transport sectors for tomorrow as well as today. This is much more so today than before and it will be the litmus paper for tomorrow not only in these sectors but the whole economy and the extremely intricate edifice raising above it. Paradoxical though it may be dubbed, this splendidly complex framework may (and probably will) be disrupted by the same forces which animated it.

3.2 Disruption

Richard Baldwin starts a paper writing "*Three eminent economists from Princeton University have recently argued that globalization has entered a new phase that requires a new paradigm understand*."[15] This very word of *paradigm* is one of the

[13] Dicken (2011). p. 23.

[14] Parag Khanna (2016).

[15] Baldwin (2006).

keywords of this paper and we can even allege that it symbolizes the main theme. Indeed, even before the outbreak of recent pandemic the author of this paper was (and still is) convinced that the world has come to the brink of a comprehensive revolutionary transformation of economic order, and with it the whole social system. In this, we fully agree with Baldwin that "*... new paradigm globalization differs from the old in that it is occurring at a much finer level of disaggregation*" and that "*deeper new paradigm globalization will seem quite unpredictable from the perspective of firms and sectors.*"[16]

Disruption has been and is being taking place within a process named as "unbundling" by several authors. Perhaps the first unbundling movement has ensued the beginning of Great Divergence between the West and the Rest of the World began around the turn of the nineteenth century (1820s, Industrial Revolution) when spotting factories and consumers first at different locations (first unbundling) and unpackage factories and various production stages far apart from each other (second unbundling). This was the first great disruption of the world economy, in terms of an industrial revolution. The second came to work following 1970s in the twentieth century, to set in motion forces to revolutionize the global economy not only in terms of physical production but in a phenomenal way: digital. With this second element, the trajectory of processes one following the other at a bewildering momentum, extends to transform not only the mode of production but the rational sphere of our conception as well to seek a new paradigmatic world.

3.3 Physical Versus Digital: To What Extent and How

It is clear that when we mention the very concept of "*digital*," this falls into the sphere of ICT which inscribes the role of technology of a matrix of a whole system of the time. When it was the ruling technology of industrial revolution (of the First, referring to Schwab, i.e., The First Industrial Revolution), as a sequence within a time span ranging from the end of Feudal Ages to the present, it reflects an exponentially rising growth curve unparalleled before.

Based on studies of Angus Maddison, we can see the first great divergence when Western Europe and North America diverged from the rest of the world (The Economist, 2013). We can critically analyze the First Great Divergence in terms of "physical" disruption where mechanization, steam power, waterpower, all serving to assembly lines, electricity for mass production in company with cars, airplanes, and all the rest serving to a consumer society.

Today, the world/global economy is again passing through another revolutionary epoch. In our understanding, the primary metaphor here is the fact of globalization. The global economy is interconnected at a rate not only unseen but unforeseen before. This can be attached widely to the development of digital technology

[16] ibid.

at unprecedented rates. Global economy has been disrupted earlier during the industrial revolution of the nineteenth century with the motive force technology on rather physical lines. Now, disruption is taking place more on digital lines at all levels: factory, services, including transport with all its ramifications (operational as well as infrastructural contexts) and the techniques to measure (even accounting methods at firm and macro levels). All these lead us to one focal point: a new paradigm. This will be taken later.

3.4 Transport and Logistics Sectors in Process

Economic activity has always been dependent on the movement of goods. The more developed economies are the more efficient and complex transport and logistics networks have to be and vice versa. In order to make a reliable analysis of the networks we have to peer over the level of a multidimensional connectivity. This covers a sort of connectivity map which consists of transport linkages, communications (to cover ICT global flows), trade, and even finance. This reflects a wider view of economic activity to gain more momentum and density as the development process of the world economy proceeds further in terms of sophistication.

A delicate subject here is the distinction to be made between urban transport and inter-cities international transport. Taking the growing importance of cities for the future into account, we cannot ignore the role of cities as synaptic points connecting all centers to each other (Dicken, 2011).

Just 15 years ago *The Economist* wrote "'*THE great manufacturers in the Yorkshire and Lancashire districts tell me that, under modern conditions, they have got into the habit of laying in supply not for a period of two to five months but they are dependent week by week on the importation of the raw material.' So Sir George Parkin described the alarming business practices found in Britain at the dawn of the twentieth century. As a leader of the Imperial Federation League, he sought to replace the British empire with a bigger group of trading partners, so as to guarantee supplies. A hundred years on, Sir George would have marveled at globalization, but been aghast that today's manufacturers measure their inventories in only a few hours of production.*

The great manufacturers now have amazingly lean operations. They have outsourced business to contractors that can do the work more efficiently, often in places where wages are lower. A huge logistics industry has sprung up to move stuff around the world at dazzling speed."[17]

Being before the present COVID-19 hit the world, this was the scene of a period of time stretching from the beginning of one revolutionary process to another within 100 years. Indeed, for long, "*the fastest-growing category of international trade was "intermediate" trade between the nodes in global supply chains, adding value*

[17] "***The logistic revolution. Chain reaction***". The Economist Jun 15, 2006.

incrementally. This generated higher levels of freight movement, and helped the value of international trade grow at least two or three times faster than global GDP."[18]

The relationship between international trade and international transport is one of the critical sides of our study where the nature of affiliation is also the reflection of interdependence between trade, logistics, transport, and even technology being more so when the impact of digital technology appears to be the dominant force.

In this regard, we can refer to Tavasszy et al.[19] which underlines the importance of three subjects:

1. The strategic implications of borderless supply chain management of the choice of alternative logistics structures in supply chains.
2. The possible impacts of the expected changes in supply chain processes upon regional economic activities.
3. The impacts of changes in global logistics on the transportation systems and, in turn, on the environment.

The referred study of Tavasszy and colleagues was one of the most notable at the time of its publication (2003) and is still a noteworthy analysis. However, especially after the outbreak of recent Viral crisis, some points of critical importance need shrewd attention. We think that the time is come to look at worldly affairs in a new approach not necessarily because of the Viral crisis, but because it has spurred a movement which has already been in progress. Fuzzy though it may seem, a global economic course typified by a high level of connectivity has begun to drive the structure of the world economic system to move and transform in the direction of a new mode. We are aware that opening the chapter of critical *transformation* problem within the context of this chapter can easily lead us to a state of conundrum. But we firmly contend that if there appears to be a problem, disregarding it leads us to nowhere in the solution of the problem. What we intend to do here is to mark the problem, especially if it has a serious potential to affect the context of our main theme. How will be the new world economy where the transport and logistics activities assume crucial roles? We leave the discussion of this issue, however, to Chapters "Partially Non-discretionary Measures for Green Transportation Corridors Performance Index: A DEA Approach" and "Competitiveness of Nations and Inequality-Adjusted Human Development: Evaluating the Efficiency of Nations using DEA and Random Forest Classification".

[18] "Supply Chain and Transport." WEF—https://toplink.weforum.org/knowledge/insight/a1Gb0000000pTDoEAM/explore/summary

[19] Tavasszy et al. (2003).

4 Global Supply Chains and Global Value Chains

In analyzing the transformation of the global economy from the state of hard-industrial economy pattern shaped by the Industrial Revolution of steam technology to the present-day economy pattern bred and staged by the ICT revolution, the global supply and value chains have the primary role to figure out the basic causes. Indeed, the foreword of a well-known work starts as "*Any discussion today of international trade and investment policy that fails to acknowledge the centrality of global value chains (GVC) would be considered outmoded and of questionable relevance. The idiom might vary—referring to trade in value-added, production sharing, supply chains, vertical integration, or fragmented production instead of GVCs—but the core notion of internationally joined-up production is the same. Every international agency dealing with economic affairs as well as many governments are working on various aspects of GVCs in order to understand better their various dimensions. The central concern from this quarter, of course, is what GVCs mean for trade policy and for international cooperation in trade-related matters*.

While the business, management, economics, and development literature on GVCs goes back at least two decades, attention from the international policy community is much more recent. It is interesting to consider the process through which GVCs became more mainstream in policy thinking."[20]

Before the recent structural transformation of global economy, global intra-industry trade among industrialized countries was in parallel with the fact that global supply chains have long existed among advanced economies. With the evolution (though at a revolutionary high rate of momentum) of ICT, the mode of production has begun to transform globally, in conformity with global trade—beginning in the late twentieth century. At this stage, digital technology enabled the coordination of spatially dispersed complex tasks within the production processes scattered over a wide range of geography. Just as the steam revolution, through empowering railways and steamships, thereby reducing trade costs dramatically and making it feasible to enlarge markets, roughly 200 years ago, a comparable process made it feasible to enlarge the global trade over the whole globe today. This unusual trade growth has paralleled with the production processes stretching over many countries each specializing in particular stages of production sequence. As a study[21] has put it "*... the production of pins in late Eighteenth century England is no different from today's production of tee-shirts, cars, computers or semiconductors. Today, however, production processes increasingly involve global supply chains spanning multiple countries, with each country specializing in particular stages of a good's production sequence, a phenomenon which Hummels et al. (*2001*) refer to as vertical specialization*."[22] Hummels et al. have documented a key aspect of these

[20] Elms and Low (2013).

[21] Costinot et al. (2013).

[22] I underlined.

vertical linkages and they called it ***vertical specialization*** (VS). Their contention (verified by their estimates) was that the impressive spread of VS has lied in the fact that trade barriers, in the form of tariffs and transport costs, have steadily kept on falling. Even small reductions in these barriers have led to extensive VS.

It is a truth that GSC have transformed the global economy. "*Global supply chains, however, are themselves rapidly evolving. The change is in part due to their own impact and in part due to rapid technological innovations in communication technology, computer integrated manufacturing and 3D printing*."[23]

What we aim to do, trying to draw conclusions out of comparing the current structural transformation with that of 200 years ago in an analogy to the alignment between what Baldwin calls "first and second unbundling(s)." For his first, described as "*globalization's first unbundling*," "*the separation of production and consumption increased the importance of proximity in the production process. By enabling international trade, the transport revolution provided an incentive for large-scale production, which involved bringing together goods, technology, people, training, investment, and information.*

Proximity lowers the costs and risks of coordinating such complexity and hence trade resulted in the bundling of all stages in individual factories, often clustered locally in industrial sectors."[24]

In the ongoing transformation process which Baldwin calls *Second Unbundling* (and in Schwab's terming third and fourth Industrial Revolution, 3IR and 4IR/I4.0) production stages, performed in close proximity in the first case, began to be dispersed geographically. In this ongoing stage (before the COVID affair?), GSCs are different from those that existed among advanced economies in the twentieth century. As of today, they are much more than extra trade in parts and components as well as for telecommunications, transport, and logistics. International supply chains merely became workable when advances in information and transport technology made it possible to extend production processes across countries and the world. "*The resulting structures have led to complex inter-linkages among numerous goods and services markets and the creation of networks that can only be understood in their entirety. This is why a multi-disciplinary approach to the study of GVCs is indispensable*."[25]

Although we mentioned above that "GSCs have transformed the world economy over the last 30 years," they are older than that. "*Trade and production networks are not new. Firms have been producing items with components sourced from around the globe for centuries. Businesses have continuously sought out new markets for their products. What have changed, however, are the speed, scale, depth and breadth of global interactions. Increasingly, new players have become active in what have*

[23] Baldwin (2012–13).

[24] *op cit*, p. 28.

[25] Park et al. 2013, p. 10.

come to be called global value chains or global supply chains. This process of organization has brought entirely new issues to the table for consideration."[26]

As we have mentioned before, Baldwin's "second unbundling" has evolved in close touch with the development of digital technology and is more functional for our purposes, as ours is an attempt to analyze and interpret the ongoing transformation process. In this second phase, ICT made it possible to disperse production stages geographically over wide distances unforeseen before. In other words, the production stages previously performed in close proximity as in the time of First Unbundling (1IR?) can now be implemented within very wide connectivity networks—global networks. This is revolutionary not only within the concept of production processes,[27] but in a multidimensional, abstract sphere as well. All these bring us to a new stage, a paradigmatic one.

5 Time to Change for the Old Paradigm: Competition Versus What?

As we look at either the First Unbundling or the Second what we see is a relative resemblance between the two, at the first sight:

- Both are associated with the size of a market, scaled globally.
- Both have had the pushing effect of a major force: Technology.[28]
- Both have enacted a global transformation in economic as well as sociological territories with far-reaching consequences.
- In both logistical (supply chains) networks first and transport networks in parallel did and do play a leading role to connect the world on a global scale.
- Both have the same dynamics where the factors affecting the future of supply chains have similar characteristics.
- Both have enjoyed the teaching of the same economic doctrine: *Liberal economy*. In a way, the same paradigm is the ruler—up until now. The pillars of this *indisputable* metaphor are the *competition* and the *invisible hand*.

This list can be prolonged further. Yet, these are relevant enough to enable us to set assumptions on questioning the aspects of a trajectory stretching toward the future. In line with similarities of these two successive chapters of global development, we can detect some adversaries as well. Along the way, we may detect some points where the two sketches diverge:

[26] Elms and Low (Eds) (2013) Introduction.

[27] Here, we would be very eager to associate the case with the concept of Cobb-Douglas Function, at least metaphorically. However, ours is not yet a mature postulate.

[28] Although all the notable transformative events throughout the history have enjoyed the reproductive role of some sort of technology of the time, the impetus of technology of Industrial Revolution (1820s) was of a novel sort.

- First and foremost, one important basic element: the issue of competition. This divine concept has been nibbled enough in the last century.[29] But, only in theory, not in practice.
- Although both have enjoyed the push of a forceful technological revolution, the exertion of the second is of a different disposition. The technology of the Industrial Revolution (1IR?) is a sequential event no matter its power was incomparably greater than those before. This time, the consequences of the ruling technology as a driving force appears to go beyond the conceptual framework of a system in which technology is an independent variable. This variable seems to have idiosyncratic symptoms to acquire powers to transform the mode of production together with a superstructure by all its ingredients including human breed being at the top of the list.
- Connected with this, the odd phenomenon of *fragmentation* which prevails not only in the context of production which has been one of the basic elements of the transformation process of the *First Unbundling*, but in the configuration of the market itself. Till the late twentieth century, the global market was in the hands of major powers of the capitalist West. But now, the dictum is being shared by an increasing number of partners: China, Russia, and India in the offing yet. The composition of "*ruling classes*" is in a process of transformation.
- An important discrepancy between the two is about the navigation of the global markets. In the nineteenth century, there were big companies involved in economic decision-making as well as today. And today there are governments guiding the world economy as well as 200 years ago. But the extent of their power in this play is sharply changed—at least in the Capitalist West not to mention China where the Communist Party takes the decisions on behalf of government-led companies. Either way, it is fair to say that the picture is not the same as it was in the nineteenth century.

Taking all these points into consideration, we can come to conclude that the factors handled in the prognosis of the future have been mutated if not changed for the most part. And it is much harder now to take action concerning the future of the planet—in the name of future of global economy if not for other (more) critical issues like climate change, environmental degradation, and machine–human interaction. And all these and many others have to be judged within the system of the future: in the hands of governments or private companies or machine / technology-based structures.

Most important of all, at least in the eyes of the author of these lines, should / can we make our judgments and analyze the issues with the paradigm basically depending on the one(s) constructed and evolved throughout the last 200 years?[30]

[29] Cf. Joan Robinson. **Economics of Imperfect Competition** and (1933) and Edward Chamberlin. **Theory of Monopolistic Competition** (1933) and a long list in the same context.

[30] It would not be wrong to say that the present base of the ruling paradigm has been evolved throughout the centuries long before the beginning of Industrial Revolution, which consolidated the basics of Capitalist Civilisation and cannot be overlooked.

A paradigm as such certainly cannot be written out simply by individual efforts; such a concept goes beyond the limits of a conceptualization sphere; this is not the intention of the author. What we are trying to do here is just to suggest that the efforts based on a dying schedule are doomed to fail, especially if it aims to extrapolate the future. Otherwise, the expectation of a paradigmatic shift will be carved into the future exclusively by a new system on the remains of the ceased.

Whatever the new world economic order may be, it is bound to be organized by some serious constraints. In close connection/interaction with others, the most imminent and most insurmountable is the formidable problem of climate change and environmental destruction. The formulation of a new paradigm is destined to involve these basics alongside with others like the ones expected to be paraphrased by technology and society interaction. As we look at either the First Unbundling or the Second what we see is a relative resemblance between the two, at the first sight:

- Both are associated with the size of a market, scaled globally.
- Both have had the pushing effect of a major force: Technology.[31]
- Both have enacted a global transformation in economic as well as sociological territories with far-reaching consequences.
- In both logistical (supply chains) networks first and transport networks in parallel did and do play a leading role to connect the world on a global scale.
- Both have the same dynamics where the factors affecting the future of supply chains have similar characteristics.
- Both have enjoyed the teaching of the same economic doctrine: *Liberal economy*. In a way, the same paradigm is the ruler—up until now. The pillars of this *indisputable* metaphor are the *competition* and the *invisible hand*.

This list can be prolonged further. Yet, these are relevant enough to enable us to set assumptions on questioning the aspects of a trajectory stretching toward the future. In line with similarities of these two successive chapters of global development, we can detect some adversaries as well. Along the way, we may detect some points where the two sketches diverge:

- First and foremost, one important basic element: the issue of competition. This divine concept has been nibbled enough in the last century.[32] But, only in theory, not in practice.
- Although both have enjoyed the push of a forceful technological revolution, the exertion of the second is of a different disposition. The technology of the Industrial Revolution (1IR?) is a sequential event no matter its power was incomparably greater than those before. This time the consequences of the ruling

[31] Although all the notable transformative events throughout the history have enjoyed the reproductive role of some sort of technology of the time, the impetus of technology of Industrial Revolution (1820s) was of a novel sort.

[32] Cf. Joan Robinson. **Economics of Imperfect Competition** and (1933) and Edward Chamberlin. **Theory of Monopolistic Competition** (1933) and a long list in the same context.

technology as a driving force appear to go beyond the conceptual framework of a system in which technology is an independent variable. This variable seems to have idiosyncratic symptoms to acquire powers to transform the mode of production together with a superstructure by all its ingredients including human breed being at the top of the list.

- Connected with this, the odd phenomenon of *fragmentation* which prevails not only in the context of production which has been one of the basic elements of the transformation process of the *First Unbundling*, but in the configuration of the market itself. Till the late twentieth century, the global market was in the hands of major powers of the capitalist West. But now, the dictum is being shared by an increasing number of partners: China, Russia, and India in the offing yet. The composition of "*ruling classes*" is in a process of transformation.
- An important discrepancy between the two is about the navigation of the global markets. In the nineteenth century, there were big companies involved in economic decision-making as well as today. And today there are governments guiding the world economy as well as 200 years ago. But the extent of their power in this play is sharply changed—at least in the Capitalist West not to mention China where the Communist Party takes the decisions on behalf of government-led companies. Either way, it is fair to say that the picture is not the same as it was in the nineteenth century.

Taking all these points into consideration, we can come to conclude that the factors handled in the prognosis of the future have been mutated if not changed for the most part. And it is much harder now to take action concerning the future of the planet—in the name of future of global economy if not for other (more) critical issues like climate change, environmental degradation, and machine–human interaction. And all these and many others have to be judged within the system of the future: in the hands of governments or private companies or machine / technology-based structures.

Most important of all, at least in the eyes of the author of these lines, should / can we make our judgements and analyze the issues with the paradigm basically depending on the one(s) constructed and evolved throughout the last 200 years?[33]

A paradigm as such certainly cannot be written out simply by individual efforts; such a concept goes beyond the limits of a conceptualization sphere; this is not the intention of the author. What we are trying to do here is just to suggest that the efforts based on a dying schedule are doomed to fail, especially if it aims to extrapolate the future. Otherwise, the expectation of a paradigmatic shift will be carved into the future exclusively by a new system on the remains of the ceased.

Whatever the new world economic order may be, it is bound to be organized by some serious constraints. In close connection/interaction with others, the most imminent and most insurmountable is the formidable problem of climate change and

[33] It would not be wrong to say that the present base of the ruling paradigm has been evolved throughout the centuries long before the beginning of Industrial Revolution which consolidated the basics of Capitalist Civilization and cannot be overlooked.

environmental destruction. The formulation of a new paradigm is destined to involve these basics alongside with others like the ones expected to be paraphrased by technology and society interaction. The destiny of GSCs and the transport networks, therefore, rests upon this group of determinants.

5.1 The Big Issue: Environment and Greening—or "Jacta Alea Est"[34]

Table 1 from MGI is strong enough to indicate the point of no return as Caesar has done 10 January 49 BD right before the birth of the Roman Empire with one difference: Caesar has succeeded to do it; ours is dubious whether we can save our planet. However, one point is very clear: We are at the edge of no return; there is no way of returning to the good *old*[35] system we enjoyed till this moment.

> *In the summer of 2010, record-high temperatures hit Moscow. At first it was just another heat wave, but the scorching heat that started in late June continued through mid-August. Western Russia was so hot and dry in early August that 300 or 400 new fires were starting every day. Millions of acres of forest burned. So did thousands of homes. Crops withered.*[36]
>
> *From deep-rooted racism to the Covid-19 pandemic, from extreme inequality to ecological collapse, our world is facing dire and deeply interconnected emergencies. But as much as the present moment painfully underscores the weaknesses of our economic system, it also gives us the rare opportunity to reimagine it. As we seek to rebuild our world, we can and must end the carbon economy.*
>
> *The carbon economy amplifies racial, social and economic inequities, creating a system that is fundamentally incompatible with a stable future.*[37]
>
> *IN JUNE 1988 scientists, environmental activists and politicians gathered in Toronto for a "World Conference on the Changing Atmosphere". The aspect of its changing that alarmed them most was the build-up of carbon dioxide, a greenhouse gas. In the late 1950s, when systematic monitoring of the atmosphere's carbon-dioxide level began, it stood at around 315 parts per million (ppm). By that summer, it had reached 350 ppm—and a heatwave was bringing record temperatures to much of North America.*
>
> *The week before the Toronto conference James Hansen, a climate scientist at NASA, had pointed to the heatwave when telling the US Senate that it was time "to stop waffling . . . and say that the evidence is pretty strong that the greenhouse effect is here". The Toronto conference took a similar view, calling for an international effort to reduce global carbon-dioxide emissions by 20% by 2005.*[38]

[34] ***Jacta Alea Est***—or "*let the die be cast*": The words when Julius Caesar declared when he stepped into the River Rubicon, which marks a point of no return.

[35] I choose the very word "old" intentionally to underline the fact that there is no way of return to it.

[36] Lester B. Brown (2011). **World on the Edge**.

[37] The Guardian (2020).

[38] The Economist (2020).

Table 1 Impacts of climate hazards (*adapted fromMcKinsey Global Institute (2020), p. 14*)

Year	Climate hazards	Impact
2003	European heat wave	$15 billion in losses
2010	Russian heat wave	~55,000 deaths attributable
2012	Hurricane Sandy	$62 billion in damage
2012	30-year record low Arctic Sea ice	Reduced albedo effect, amplifying warming
2013–2014	Australian heat wave	~$6 billion in productivity loss
2015	Southern Africa drought	Agriculture outputs declined by 15%
2016	Fort McMurray fire, Canada	$10 billion in damage, 1.5 million acres of forest burned
2017	East African drought	~800,000 people displaced in Somalia
2017	Flooding in China	$3.55 billion of direct economic loss, including severe infrastructure damage
2017	Hurricane Harvey	$125 billion in damage
2019	European heat wave	~1500 deaths in France
	Decline of Himalayan glaciers	Potential reduction in water supply for more than 240 million people
	Ocean warming	Up to 35% decline in North Atlantic fish yields

There is no need to amplify the samples; the samples are ample enough. The ones we chose are highly convincing that so long as the present system of doing things prevails, an iceberg frighteningly big is doomed to bring the global human society to a day of apocalypse. As the heading of *The Economist* article quoted reads, the situation is *without precedence*. Nevertheless, no matter how big our anxiety is, our schedule here is to approach the issue at the context of its interaction with GSCs and transport networks.

GSCs, alongside with other sectors of the global economy (and society), are plunged into an awkward position where a transformation is taking place and need still more so. The Coronavirus crisis has revealed the flaws in the supply chains prevailing till pre-COVID times.

5.2 COVID-19 and the Aftermath

The Coronavirus crisis revealed the flaws in the GSCs. In our opinion, the seemingly greatest strength of the present global system, connectivity, has begun to work adversely as it turned out the greatest threat immediately by the outbreak of plague. However, though it may seem satirical, as the resilience of the connectivity inherent in the system has been fabricated and evolved by the ICT, the way out of this stalemate lies in the digital technology itself as well. The crisis exposed the vulnerability of the complex supply chains system as the linkages between producers themselves as well as those between buyers (consumers) and sellers (companies) were all locked down. *"Despite rapid advances in technology, the relationship between buyers and suppliers remains predominantly paper-based. Digitizing the buyer-supplier relationship is a fundamental element for building sturdy supply chains and will make identifying and recruiting new suppliers far less time-consuming. With technologies like artificial intelligence and the Internet of Things, supply chains could quickly switch to alternative providers when regular suppliers face disruption.*

The current crisis is an opportunity to reset a system that has relied on outdated processes. Creating smart and nimble supply chains is the key to building a global trade and investment network that's capable of weathering future storms."[39]

The pandemic exposed the weaknesses in the global community as well alongside with the "*inbuilt imperfection*" of the capitalist economy. "*COVID-19 is an unforeseen global issue that is now being confronted with global action as we realize the weaknesses it's exposed. Today, it presents us with an opportunity to rebuild the supply chain of the future that we want to see and that's an opportunity we cannot miss*."

In order to be able to run this course effectively a collaborative action move between governments, academia, and the civil society is vitally important. Without going into the theoretical discussions in the Marxian "infrastructure—superstructure" domain, it is time to say that a rhetoric based on collaboration as such is incompatible with the conception of competition. Furthermore, it is time to approach the issue *en passan* with a passing paradigm of a passing civilization—The Capitalist Civilization. If we confine ourselves to the economic dimension of this ***old*** code, it is almost impossible to analyze the interim period we are passing through. We are to conclude this critical issue in the last (next) paragraph.

6 Resetting?

In the expectation of a new global shift, several questions arise out of a probably much longer list. This is one of the most intricate points to be decided in an overview of the rich collection of studies being carried out throughout the world. If we try to make a shortlist according to our approach in this chapter, we can sum up some below on an understanding of the following:

(a) There must be a selection of issues based on the taxonomy we specify.
(b) We can proceed on the lines clarified by the methodology of the taxonomic approach specified.

Based on our approach for the (near) future in which we come up with the perception that there will be a new paradigm to dominate the new world as a whole, no matter how less is known for this new world. What we only know for this new

[39] Lin and Christian Lanng (2020).

world is that "*the future is not what it used to be*."[40] Then what can/should be done to draw a roadmap for the future? Being conscious of our position not to turn it into a prophecy as such, all we can say is:

> No historic event can be explained by cause & effect relationships of a certain time by the parameters of another time if and when we resort to the fact that the happenings of a certain period do not necessarily pass through the similar course in a sequential line. In other words, a configuration of a certain point of time can and will probably subject to the determinants of the same period and not those belonging to that already passed (or passing) away.

Therefore, we are in a position to realize that the password for passing through this interim period of transformation is to try to look to the future with the paradigm of the next period—which is not known yet. Our analogy here is not necessarily ought to bring us to a point of little or no avail. What we have to do is to frame a model that comprises a selected bunch of factors forming an agenda for the future.

We can do this by formulating a traditional production function as below: [41,42]

$$O = f(K, L, N, T, S, \alpha, t)^t$$

For this model, we presume some basic elements that have characteristically been prevailed throughout the economic history though at different dimensions: Connectivity through several sections of an economy like trade, investment, migration, and other channels critical in the formation of macroeconomic linkages is the forerunner of these to form, for instance, Internet to be the climax of technological development. But the most important issue for our purposes here is the question of how much knowledge is transmitted is not the result of the overall level of connectivity, but also to whom a unit is connected, as well as how these connections complement each other.[43] Our reference here is the relativity of the functional interaction of all factors between themselves.

Furthermore, the model can hardly be usable taken at a linear path; a nonlinear bundle can be more suggestive. Assuming all the determinants in the equation above as separate vectors, every vector formulating the functional relationship within itself would bring us to a matrix, which represents a multidimensional connectivity among all individual constituents of the total sum. This can be turned into action within an

---

[40] Sneader and Singhal (2020).

[41] K: Capital, L: Labour, N: Nature, T: Technology, S: Socio-cultural Factor, α: Constraint, t: Time (dimension).

[42] This equation is a rude, simplistic functional written form of what we have in mind. We are aware that to turn this equation into a workable version necessitates a more sophisticated econometric manipulation. Then, we should go into a detailed analysis where we have to add a time dimension, dummy variables, etc. to write down a working set of equations. But our purpose in this chapter is to give an elementary metaphor and no more.

[43] Cf. Gold et al. (2018).

Input–Output context. This is not, in fact, our original idea/invention and it has been worked out several times before.[44]

Up to this point, we spared too little place for transport in its own merit as compared with the title of this chapter. This is of course not unintentional. The reason for this portioning is what we deem decisions concerning the transport as such either in operational or infrastructure depend on the fact that both decisions are up to focus on transport demand, which itself is derived from other decisions. And all the context of sections we tried to analyze are connected with transport in one way or the other, directly, and/or indirectly.

It is clear that transport activity, as part of the supply/value chains, is a vitally important element of the networks of connectivity and the key role it plays gains further importance. The latest pandemics brought this importance into light further. It will show itself in the facts that:

1. Supply chains will transform further, making it inevitable that transformative investments, including those related transport will have to be planned.
2. Supply chains are to be focused on resilience and digitization.
3. In the planning of future emerging trends in supply chains of the future occupy a special place on the agenda of planning.
4. Some points of importance will have to be taken into account seriously[45]:
   (a) Shocks that affect global production are growing more frequent and more severe.
   (b) Value chain disruptions cause substantial financial losses.
   (c) The interconnected nature of value chains limits the economic case for making large-scale changes in their physical location.
   (d) Building supply chain resilience can take many forms beyond relocating production.
5. In the presence of *Global Megatrends* (in demographics, technology, issues of individualistic matters, economic interconnectedness, urbanization, and most important of all, climate change and environmental issues) the posture of the future state.

A distinctively important aspect of our times is its extent of universal comprehensiveness. On the large periphery of the globe, the very extensive and comprehensive network of connectivity made it inevitable that the basic decisions concerning the future are to be explored within the context of a global network. And then transport research and policy constitutes and integral part of this frame.

[44] Cf. Timmer (2010) and Timmer et al. (2014) and B. Los et al. (2017) and good many others.
[45] Cf. Lund et al. (2020).

References

Baldwin, R. (2006, October). *Globalisation: The great unbundling(s) eu2006.fi* – Helsinki communique – Discussion paper.

Baldwin, R. (2012–13). *Global supply chains: Why they emerged, why they matter, and where they are going*. The Graduate Institute.

Brown, L. B. (2011). *World on the edge. How to Prevent Environmental and Economic Collapse. Earth policy institute*. W.W.Norton & Co.

Costinot, A., et al. (2013). An elementary theory of global supply chains. *Review of Economic Studies., 80*, 109–144.

Dicken, P. (2011). *Global shift. Mapping the changing contours of the world economy* (6th ed.). The Guilford Press.

Elms, D. K., & Low, P. (Eds.). (2013). *Global value chains in a changing world*. WTO and Fung Global Institute.

Ferguson, N. (2011). *Civilization: The west and the rest*. Penguin

Gereffi, G., et al. (2001). Globalisation, value chains and development. *Institute of Development Studies. IDS Bulletin, 32*, 3.

Gold, D., et al. (2018). *Multidimensional connectivity. Benefits, risks, and policy implications for Europe and Central Asia*. World Bank Group. Policy Research Working Paper 8438. May.

Hummels, D. L., et al. (2001). The nature and growth of vertical specialization in world trade. *Journal of International Economics, 54*, 75–96.

Khanna, P. (2016). *Connectography. Mapping the future of global civilization*. Random House.

The Guardian (2020) Letter from economists: to rebuild our world, we must end the carbon economy: The Guardian. 4 August 2020.

Lin, J., & Lanng, C. (2020). *Here's how global supply chains will change after COVID*-19. WEF 6 May.

Los, B., et al. (2017). *Analyzing global value chains using the world input-output database*. BBVA Foundation-Ivie workshop. Oct 30. Valencia.

Lund, S., et al. (2020). *Risk, resilience, and rebalancing in global value chains*. McKinsey G.I..

McKinsey Global Institute. (2020, January). *Climate risk and response. Physical hazards and socioeconomic impacts*.

OECD. (2012). *Mapping global value chains*.

Park, A., Nayyar, G., & Low, P. (2013). Supply chain perspectives and issues. A literature review. WTO & Fung Global Institute.

Ruijgrok, C. J., Tavasszy, L. A., & Thissen, M. J. P. M. (2002). *Emerging global logistics networks: Implications for transport systems and policies*. STELLA focus group 1 meeting "globalisation, e-economy and trade", Siena, 9–10 June 2002.

Sneader, K., & Singhal, S. (2020, April). The future is not what it used to be: Thoughts on the shape of the next normal. *McKinsey & Company*.

Tavasszy, L. A., Ruijgrok, C. J., & Thissen, M. (2003). *Emerging global logistics networks: Implications for transport systems and policies*. Article in Growth and Change.

The Economist. (2013). *What was the great divergence?* Sept. 2nd.

The Economist. (2020). *Why tackling global warming is a challenge without precedent.* April 25th edition.

Timmer, M. P. (2010). *Measuring global value Chains with the WIOD*. OECD Conference 21 Sept.

Timmer, M. P., et al. (2014, Spring). Slicing up global value chains. *Journal of Economic Perspectives, 28*(2), 99–118.