Cooperative Management

Evangelia Krassadaki
George Baourakis
Constantin Zopounidis
Nikolaos Matsatsinis *Editors*

Operational Research in Agriculture and Tourism

7th International Symposium and 29th National Conference on Operational Research, Chania, Greece, June 2018

EXTRAS ONLINE

Springer

Cooperative Management

Series Editors

Constantin Zopounidis, School of Production Engineering and Management, Technical University of Crete, Chania, Greece
George Baourakis, Department of Business Economics and Management, Mediterranean Agronomic Institute of Chania, Chania, Greece

The Book Series on Cooperative Management provides an invaluable forum for creative and scholarship work on cooperative economics, organizational, financial and marketing aspects of business cooperatives and development of cooperative communities throughout the Mediterranean region and worldwide. The main objectives of this book series are to advance knowledge related to cooperative entrepreneurship as well as to generate theoretical knowledge aiming to promoting research within various sectors wherein cooperatives operate (agriculture, banking, real estate, insurance, and other forms). Scholarly edited volumes and monographs should relate to one of these areas, should have a theoretical and/or empirical problem orientation, and should demonstrate innovation in theoretical and empirical analyses, methodologies, and applications. Analyses of cooperative economic problems and phenomena pertinent to managerial research, extension, and teaching (e.g., case studies) regarding cooperative entrepreneurship are equally encouraged.

More information about this series at http://www.springer.com/series/11891

Evangelia Krassadaki • George Baourakis •
Constantin Zopounidis • Nikolaos Matsatsinis
Editors

Operational Research in Agriculture and Tourism

7th International Symposium and 29th National Conference on Operational Research, Chania, Greece, June 2018

Springer

Editors
Evangelia Krassadaki
School of Production Engineering
and Management
Technical University of Crete
Chania, Greece

George Baourakis
Department of Business Economics
& Management
Mediterranean Agronomic Institute of Chania
Chania, Greece

Constantin Zopounidis
School of Production Engineering
and Management
Technical University of Crete
Chania, Greece

Audencia Business School
Nantes, France

Nikolaos Matsatsinis
School of Production Engineering
and Management
Technical University of Crete
Chania, Greece

ISSN 2364-401X ISSN 2364-4028 (electronic)
Cooperative Management
ISBN 978-3-030-38765-5 ISBN 978-3-030-38766-2 (eBook)
https://doi.org/10.1007/978-3-030-38766-2

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

Preface

The main activity of the Hellenic Operational Research Society (HELORS) is the promotion of operational research (OR) in Greece, in academia and industry. Since 1977, HELORS has organized 29 national conferences, while since 2012 the annual conference has been organized along with an international symposium, which provides a forum for exchanging ideas about the current and future trends in OR, among not only Greek researchers but also a wider audience.

The 7th International Symposium and 29th National Conference on Operational Research, was held in Chania, Greece, from June 14 to 16, 2018. The event was devoted to the "*Contribution of Operational Research, New Technologies and Innovation in Agriculture and Tourism*." Three invited speakers presented their recent work: Professor Markos Papageorgiou from the Technical University of Crete (Greece), Professor Gert van Dijk from Nyenrode University (the Netherlands), and Professor Yannis Siskos from the University of Piraeus (Greece). Fifteen sessions were organized and sixty-two papers were presented by researchers coming from eleven countries. The covered topics included all recent advances in operational research, new methodological developments, and applications in fields such as agriculture, food and beverage, tourism, transport, environment, energy, and marketing.

The current edited volume was prepared on the occasion of the above event. After a review process, 10 papers were selected for this book. The first two papers cover topics related to agriculture and the emphasis needed for the adoption of technological innovation in the field. The next two papers refer to the food and beverage (F&B) industry. The following four papers cover topics related to the tourism sector. Finally, the book closes with two papers: one about the most commonly used criteria for the evaluation of green suppliers and another one about a new algorithm for the orienteering problem.

The first paper, by Vartan Kesiz Abnousi, Konstantinos Karantemiris, and Andreas Doulis, presents issues related to the acceptance of technological innovation by agricultural cooperatives. The authors place emphasis on traceability and on

certification (molecular or otherwise) of plant propagation material as examples of technological innovation.

Christina Diakaki, Nikoleta Banani, and Evangelos Grigoroudis, in the second paper, document the need for the development of virtual incubators and online business tools for agro-food SMEs. The authors analyze the limited access to resources, experts, and integrated business tools for SMEs, which comprise the major part of the agro-food sector, after a survey they have undertaken in five countries.

The third paper, by Anastasios Liapakis, Theodore Tsiligiridis, and Constantine Yialouris, presents the design of a sentiment lexicon for the Greek F&B sector. The authors provide a literature review and the proposed design of a sentiment lexicon as a preliminary step to further analyze the customers' reviews of some leading companies.

The fourth paper, by Konstantina Panagiotakopoulou, Petros Kalantonis, and Panagiotis Kaldis, examines if the profitability of F&B firms was significantly affected in the period of Greek crisis. The authors further explore if debt and liquidity are determinant factors of profitability. The results indicate that firms' profitability has been affected by the economic crisis; however, their book and market value does not seem to be significantly affected.

Georgios Alexopoulos, Alexandros Apostolakis, Constantin Zopounidis, Alexandros Garefalakis, and Marianna Eskantar, in the fifth paper, examine the utilization of the fiscal multipliers under extremely uncertain conditions in the tourism sector. They explore how the impact of the economic strategy on performance can shift depending on whether the economy is in progress or recession.

The sixth paper, by Dimitrios I. Vortelinos, Konstantinos Gkillas, Christos Floros, and Lavrentios Vasiliadis, presents the performance of the *k* nearest neighbor (kNN) forecasts in the context of European tourism demand. The parameterization of the kNN model affects forecasting performance, while the inclusion of international stock indices significantly increases forecasting accuracy.

The seventh paper, by Basil Tselentis, Theodoros Pelagidis, and Sotirios Manologlou, focuses on the role of modern ports and marine protected areas in East Mediterranean countries' environmental sustainability. The authors discuss the provided policies and the applied practices, which are closely related to the quality of the environment, the tourism sector, and in general the socioeconomic well-being.

The next paper, by Manousos Rigakis, Dimitra Trachanatzi, Magdalene Marinaki, and Yannis Marinakis, presents a new algorithm to simulate a tourist trip design problem, namely the Hybrid Firefly Algorithm Based on Coordinates (FAC) for the Prize-Collecting Vehicle Routing problem (PCVR). The formulation of the problem is able to successfully simulate a tourist trip and it could be integrated in a tourist recommendation system.

The book closes with two papers: the first about the most commonly used criteria for green supplier evaluation and the last about a new algorithm for modeling and solving orienteering problems. In the first, Syrine Jemaa, Ahmed Alayidi, Athanasios Migdalas, George Baourakis and Periklis Drakos present an updated literature review (2012–2019) about the most cited green criteria for the evaluation

and selection of suppliers. In the last paper of the book, Eleftherios Tsakirakis, Magdalene Marinaki, and Yannis Marinakis present a new Similarity Hybrid Harmony Search (SHHS) algorithm for the solution of the orienteering problem, along with a number of benchmark results.

Closing this short preface, we would like to sincerely thank all the participants of the 7th International Symposium and 29th National Conference on Operational Research, who supported the event in Chania-Crete, and in particular the authors who contributed to this edited volume. We would also like to thank all those who devoted considerable time to review the submitted papers.

Chania, Greece

Evangelia Krassadaki
George Baourakis
Constantin Zopounidis
Nikolaos Matsatsinis

Contents

Agricultural Cooperatives and Acceptance of Technological Innovation

Vartan Kesiz Abnousi, Konstantinos Karantemiris, and Andreas G. Doulis

Abstract This chapter deals with issues related to the adoption of technological innovation by agricultural cooperatives. These cooperatives possess some special characteristics, by comparison to other types of primary producers and are considered as lagging or even resisting innovation adoption. Consequently, the first part of the chapter presents some general notions of progress, growth innovation, and challenges their mainstream definition by placing within a defined cultural and philosophical framework. The second part of the chapter deals with economic aspects of technology innovation and innovation diffusion, especially as they relate to agricultural cooperatives, and it presents a series of case studies derived from different countries representing diverse cultural, developmental and organizational characteristics by emphasizing traceability, and on certification (molecular or otherwise) of plant propagation material as examples of technological innovation.

Keywords Agricultural cooperatives · Innovation adoption · Innovation diffusion · Technological innovation · Traceability and certification of plant propagation material

V. Kesiz Abnousi (✉)
Department of Agricultural and Applied Economics, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
e-mail: vkesizab@vt.edu

K. Karantemiris · A. G. Doulis
Department of Viticulture, Floriculture, Vegetable Crops and Plant Protection, Laboratory of Plant Biotechnology and Genomic Resources, Hellenic Agricultural Organisation "DEMETER" (ex. NAGREF), Institute for Olive Tree, Subtropical Crops and Viticulture, Heraklion, Greece
e-mail: andreas.doulis@nagref-her.gr

E. Krassadaki et al. (eds.), *Operational Research in Agriculture and Tourism*, Cooperative Management, https://doi.org/10.1007/978-3-030-38766-2_1

1 Introduction

This chapter aims at providing some initial concepts related to the adoption of technological innovation by agricultural cooperatives. It is mostly addressed to students and professionals with little or no previous exposure to the special characteristics of such enterprises. Customary readers of books, similar as the present, are not expected to be familiar with notions related to the use of certified plant propagation material and their products (i.e. foods) by agricultural cooperatives. Such an adoption holds the promise for added benefits and opportunities. Simultaneously, it presents some further organizational and fiscal costs, occasionally leading to resistance built, mostly, by local stakeholders or wide heterogeneity in accepted norms. In our present show-case we discuss prerequisites of different cooperatives installed in different countries around the globe representing diverse regulatory, social and market frameworks. The technological underpinnings of material certification as well as traceability and quality assurance issues are presented.

Historically, the notion of progress has been central in the understanding, interpretation, and function of science and technology. Similarly, the notion of social progress has been applied within the context of various types of societies. Nevertheless, the definition of progress, its actual meaning, manifestations, and consequences for science, technology, and mainly societies is contested; 'elusive' as at least one author almost in the beginning of the last century suggests (Woods, 1907: 779). Other authors focus on the critical importance that the idea of progress has for Western civilization and in particular, the value that the so-called Western world and intellect has placed on the historical movement from past to present to future perceived to occur in a more or less continuous, gradual, cumulative and mostly uni-linear fashion (Nisbet, 2009: 4–5). Put another way, albeit simplified, it can be argued that to some extent "the idea of progress holds that mankind has advanced in the past—from some aboriginal condition of primitiveness, barbarism, or even nullity—is now advancing, and will continue to advance through the foreseeable future" (Nisbet, 2009: 4–5). A dominant assumption, at times implicit and, other times explicit, in this approach to progress, is the assumption that progress is a process that leads from an 'inferior' to a 'superior' state of affairs liberating humans from ignorance and fear and leading them to civilization and material and spiritual achievements; an assumption portrayed in Aeschylus' *Prometheus Bound*.

In the realm of social sciences the concept of progress was addressed by theorists such as August Comte who sees progress as the right balance between conservative and innovative forces (No author, 2011: 389), Karl Marx who envisions progress through revolution and the overcoming of human alienation, the so-called Social Darwinists or social evolutionists who linked progress to evolution (Sklair, 2001: xiii) and later on, Max Weber and Emile Durkheim who although not concerned with progress per se they view societies as progressively adopting more complex, individualistic forms of organization with increased levels of specialization (Sklair, 2001: xiii).

In societies such as the above, however, critical is the belief in rationality and the quantification of natural and social phenomena which in turn underline the promise

and the deriving expectations that science and technology perceived to be *value free*, ideologically neutral, can continue delivering a better future; deliver progress (Kosellek, 2004). Viewed in this way, then, progress is a modern phenomenon (Borup, Brown, Konrad, & van Lente, 2006).

Scientific and technological neutrality, however, along with the value-laden belief in progress, came under the scrutiny and criticism of various schools of social thought with the Frankfurt School of critical studies being these schools of thought flagship. Adorno and Horkheimer's (1992 [1944]) position on the drawbacks of social and technological progress, for example, and Foucault's (1972, 1980) break of utopias mark a significant shift in the study of progress, science, technology, rationality and, neutrality away from the Enlightenment paradigm.

Frequently associated and linked with notions of 'development', 'change', 'advancement' the notion of progress came under observation and criticism by philosophers and social scientists for being a value-laden term imposing normative understandings of progress as a dogma; a dogma that characterized Western-type culture, ideology, science and technology and propagated in the rest of the world (Nisbet, 2009: 9) suggesting the future as having a utopian end. Social scientists tend to be 'uncomfortable' with the notion of progress as a history-driving force since as Best argues "…progress involves making judgments that violate the principle of cultural relativism" (2001: 2) or better obfuscates and devalues cultural, ethnic, historical, religious, economic differences between societies as well as social groups; failing to take these differences into account may potential lead to inequalities (Best, 2001: 2). In more than one settings, such inequalities can entail imposing cultural and ideological normative constructs developed in one setting (i.e., Western industrialized, individualistic oriented societies) into another (i.e., non-Western, agricultural, community-oriented societies).

Criticism and skepticism notwithstanding, the term progress gradually receded. While other terms have been used to replace it (or maintain it as the case might have been), the term that dominates contemporary discourses on the advancement and overall betterment of social and individual lives (the pillar concept in progress) is the term innovation. The centrality and importance of innovation in contemporary societies can be seen in that "when societal changes are discussed. The debate is most likely framed through the lens of 'innovation'" (Skjølsvold, 2012: 1). Alternatively, we tend to use the notion of innovation when we discuss "…how to improve technology, organizations, business, policy, nature and even in some cases religion" (Skjølsvold, 2012: 1). As Michel Callon puts it (Callon, 1986), innovation has become an obligatory passage point.

Similarly, to the notion of progress, innovation tends to be understood as a linear, gradual, cumulative process. Even though the fact that 'the linear model of innovation' has been discredited academically (Godin, 2006: 659; Rosenberg, 1994: 139), it continues to underline policy recommendations and commercialization processes which emphasize how research and development occurs and is disseminated (Smith, 2008: 13).

The need, however, to put innovation in a socially, culturally as well as the politically sensitive context has been widely acknowledged in Europe, and this is further

translated into the European Commission's policies and procedures. It can be argued that for the EC

> Innovation provides **real benefits** for us as citizens, consumers, and workers. It speeds up and improves the way we conceive, develop, produce and access new products, industrial processes and services. It is the key not only to **creating more jobs, building a greener society and improving our quality of life**, but also to maintaining our **competitiveness** in the global market[1] (Innovation Union).

For the EC, it is imperative to gear the innovation process to societal needs (EC, 2013: 11). To that effect, it can be argued that within Europe, or at least for the EC, innovation is understood as social progress contextualized in notions of equity for the largest possible number of people while allowing for cultural differences to be taken into account.

2 Economics of Innovation in Agriculture: A Quick Overview

Simon Kuznets (1962) stated that "the greatest barrier in understanding the role of innovation in economic processes has been the lack of meaningful measures of innovative inputs and outputs." Schumpeter (1942), in his Capitalism, Socialism, and Democracy, also probed into the subject, concluding that innovations are generated by established, large firms with monopoly power because only such firms had the ability control and commanded the resources for research and development. Investment, innovation, and diffusion are closely interrelated through multiple layers. The following sections begin with the role of investment and its relation to innovation. The returns, the role of different stakeholders, and the control that each of them exercise to investment are briefly discussed. A definition of innovation is provided. Different classifications of innovation are presented as well as the way each of them is differentiated in terms of adoption and investment. Finally, different measures of diffusion are presented, including theoretical and empirical models. Diffusion is analyzed with respect to its influence by the size of the firm, risk, uncertainty, geography, access to credit, land ownership, and other characteristics.

3 Investment in Agriculture: Economic and Policy Issues

Investment in agricultural innovation technology can be either public or private. Public investment in agricultural research and extensions has high rates of returns that, according to some researchers, see Huffman (1998) and Alston, Norton, and

[1] Innovation Union. A Europe 2020 Initiative. http://ec.europa.eu/research/innovation-union/index_en.cfm?pg=intro

Pardey (1995), which could be as high as 20% or more. This could indicate that there is considerable underinvestment. Studies evaluating returns of agriculture to society are based on a partial equilibrium analysis. Moreover, even though the social gain from the research is positive, there can be cases where the producers might be in a worse position than before because the demand is sufficiently inelastic in that industry (Griliches, 1957). Subsequently, different stakeholders have different interests and gains from research. Mechanical innovation, while benefiting society overall might be harmful to workers, a well-studied example is the introduction of the tomato harvester in California (Schmitz & Seckler, 1970). de Gorter and Zilberman (1990) provide a mathematical model that analyzes the good public inputs in agriculture, concluding that underinvestment is likely to occur when producers control the level of investment and finance it. The less elastic is the demand for the final product, the less optimal it is to invest if they invest at all. This has also spawned several political implications that producer groups can lobby for an underinvestment in research. Gordon and Zusman (1991) use a cooperative game theoretic framework to illustrate how the political system decides on water quality and pricing matters. This can also be extended in the framework of agricultural research, where the government decides to put more weight on the vested interests of agricultural producers than the taxpayers unless they are compensated. This analysis is insightful in describing the process for public support in agricultural research and development, and ultimately, innovations.

Therefore, the public sector has been very important in supporting and funding agricultural research and development, especially non-shielded disembodied or embodied innovations (Sunding & Zilberman, 2001).

Cooperatives, like any firm, invest in technologies that provide the best-expected benefits. The term expected is used because investments, by definition, entail a level of uncertainty. The benefits and the costs can differ across cooperatives depending on their size, their specialization and the willingness, and commitment of the members. On a farm level, they might differ depending on the size, the human capital, and the land quality, among other factors.

4 Innovation: Definition and Classifications in Agriculture

Innovations are the basic element of technological change. In the current framework, innovations are defined as new methods, customs, or devices used to perform new tasks. There are several categories of innovations. We can distinguish them between innovations that are embodied in capital goods or products (e.g., improved varieties of seeds, tractors, fertilizers) and those that are disembodied (e.g., integrated pest management schemes). Embodied innovations are more likely to be generated by private parties. However, it is necessary that intellectual property rights are protected for that to be possible. On the other hand, disembodied innovations are less likely to receive private funding because of the difficulties in marketing and selling the end product. Therefore, it usually becomes an area of public action. Classifications of

innovation can also be done according to their form. Subsequently, they can be categorized as biological innovations (seed), chemical innovations (fertilizers and pesticides), mechanical innovations (tractors), agronomic innovations (new management practices), biotechnological and informational innovations (computerization of agricultural cooperatives).

Innovations can also be distinguished based on their economic impact. Henceforth, there are yield-increasing, cost reducing, quality-enhancing, risk-reducing, environmental-protection increasing, and shelf-life enhancing innovations. Most innovations might fall to one or more of these categories. The importance of categorization stems from the fact that it renders the analysis of the forces and factors that lead to the adoption of new technologies more clear. Hence agricultural cooperatives that introduce new pesticides adopt a yield-increasing, risk reducing, environmental-protection increasing, and shelf-life enhancing innovation (Sunding & Zilberman, 2001). Subsequently, we should also take into account all the factors, such as the degree of risk-aversion, when we want to analyze the forces behind the adoption of new technologies.

5 Diffusion: Theoretical and Empirical Models, Characteristics and Factors Related to Agriculture and Agricultural Cooperatives

Understanding the adoption and diffusion of innovations is an integral part of our analysis. The adoption of a particular innovation can be understood as the utilization of new technology. The underlying factors of adoptions depend on one or many factors. The timing and extent of adoption also vary. On the other hand, diffusion is a measure of the aggregate adoption of innovation for industry. There are various measures of diffusion. For instance, adoption can be measured by the percentage of cooperatives, agricultural enterprises, or farmers that adopt a particular technology. It can also be quantified by the percentage of land share on which the innovation is utilized.

Diffusion is the process by which an innovation is communicated through certain channels over time among the members of a social system (Rogers, 2003). An innovation is "an idea, practice, or object that is perceived to be new by an individual or another unit of adoption." "Communication is a process in which participants create and share information with one another to reach a mutual understanding" (Rogers, 2003). Diffusion research can be traced as far back as French Sociologist Gabriel Tarde, 1903, who pioneered the S-shaped diffusion curve. According to Rogers (2003), most innovations have an S-shaped rate of adoption. The fundamental research paradigm for the diffusion of innovations can be traced to the Iowa study of hybrid seed corn. Although its advantages compared to traditional seed, there were some barriers to the adoption of this new technique. Iowa state administrators were seemingly frustrated on the reasons why, despite the clear advantages, hybrid

corn seed was not adopted immediately. Bryce Ryan and Neal C. Gross (1943) who investigated the diffusion employed quantitative techniques and laid the foundations for the modern perspective on the diffusion of innovations in agriculture.

Given that decisions are not authoritative or collective, each member of the social system faces his/her own innovation-decision that follows a five-step process:

1. Knowledge—person becomes aware of an innovation and has some idea of how it functions,
2. Persuasion—person forms a favorable or unfavorable attitude towards the innovation,
3. Decision—person engages in activities that lead to a choice to adopt or reject the innovation,
4. Implementation—person puts an innovation into use,
5. Confirmation—person evaluates the results of an innovation-decision already made.

Rogers argued that innovation consists of four stages: invention, diffusion (or communication) through the social system, time, and consequences. Information regarding the innovations is conveyed through social networks. The likelihood that the innovation will be adopted depends on the ability of the opinion leaders to communicate the new technology, as well as the nature of the social networks. Agricultural extensions can be considered such networks. Furthermore, Rogers suggested (2003) five types of adopter categories, which are: (1) innovators, (2) early adopters, (3) early majority, (4) late majority, and (5) laggards. The distribution of adopters follows a normally distributed bell-shaped curve, the derivative of the S-shaped diffusion curve, over time.

The first group, the "innovators," adopts the innovation before the rest. They test the innovations as they are less risk-averse. Then, the second group of "early adopters" after assessing whether the innovation was successful, will either adopt or ignore the innovation. Since they use data and empirical observations for the assessment, they tend to make judicial decisions, and they are well-respected by the society. This is the "tipping point" that motivates the widespread adoption of innovations. The "opinion leaders" are primarily located in the group "early adopters". Hence the social system uses the decisions of the opinion leaders as an indicator of whether they should adopt a particular innovation or not.

There are various reasons why individual cooperatives, and in general, businesses might be lagging in the adoption of innovations. In Rogers' framework, the last adopters, the laggards, are regarded as more traditional or isolated agents, in terms of the social system. Moreover, uncertainty regarding the outcome is one of the reasons for late adoption.

Statistical models estimating diffusion models use the logistic equation of the form:

$$Y_t = K\left[1 + e^{-(a+bt)}\right]^{-1}$$

where Y_t is the diffusion at time t, b is a measure of the pace of diffusion, K is the long-run upper limit of diffusion, α reflects diffusion at the beginning of the estimation period, and b is a measure of the pace of diffusion. Empirical research has shown that these diffusion parameters are affected by profitability and other economic factors (Sunding & Zilberman, 2001). Furthermore, Griliches (1957) who examined the case of hybrid seed corn adoption in the United States, finds that the diffusion parameters (K, b, α) can be explained based on of varying profitability of entry, which in turn is a function of market density, innovation and market cost.

Kennedy and Thirlwall (1972) found no indication that larger firms and monopolistic industries are a necessary condition for technological change. Mansfield (1963) explored the factors that affect the rate of adoption of a technology once it is adopted by a firm, the intra-firm diffusion rate. He used the example of locomotives when they switched from steam to diesel power to conclude that the profitability of an investment opportunity stimulates the speed of adopting an innovation. In addition, he found that smaller firms substitute old technology to new just as fast as large firms. The relationship between the size of the firms and the adoption of new technologies has been a matter of interest among researchers. Acs and Audretsch (1990) argued about the importance of small firms in generating technological innovations. Furthermore, they found that innovative activity was hindered in highly concentrated markets and that such activity tends to be more prevalent in industries consisting of larger firms. On the other hand, when larger firms dominate an industry, small firms have an advantage in innovative activity. In industries that are utilizing skilled labor and are characterized by many innovations, small firms have an innovative advantage. In contrast, large firms had an innovative advantage in capital and advertising intensive industries that are not concentrated. In general, most research on innovative activity has been concentrated on the role of large firms, Acs and Audretsch (1988) demonstrated that firms with fewer than 500 workers contribute half of the innovations. Nooteboom (1989) conducted empirical research that explored diffusion, firm size, and uncertainty. He stressed that it is expected that smaller firms will lag in adopting technological innovations since expected returns are proportional to the size of the firm whereas risk is independent of size. Furthermore, because of the S-curve, the gap between large and small firms is wide at first. David (1969) argued that farms adopt new technologies if their size is beyond a specific "threshold," as he writes. Moreover, he provided an analytical framework to calculate the "threshold farm size," beyond which adoption of the mechanical reaping process is more profitable than labor intensive methods. The requirement was that the total savings in labor costs are more than the additional fixed costs that the farmer incurs yearly with the introduction of agriculture machines. Davies (1979) found that the speed of diffusion depends on the profitability, the number of the firms, the size, the labor intensity, and the rate of growth of the industry. Fewer firms, smaller inequalities in the sizes, and high labor intensity in industry result in a faster diffusion of innovations. However, Olmstead and Rhode (1993) found that small

farms cooperated by jointly purchasing the new equipment rendering the cost of the machinery cheaper. Furthermore, a heterogeneity that stems from differences in land quality and human capital can also be a reason that some technologies are adopted. Caswell and Zilberman (1986) concluded that modern irrigation techniques are more likely to be adopted in lands that have low quality, and water is expensive. Many researchers (Rothwell & Zegveld, 1982) found empirical indications that the innovations of small and large are different. Large firms generate innovations that require highly specialized labor, capital intensive, and risky, whereas small firms generate innovations that require close interactions with customers and require flexibility. Rothwell (1985) underlies that the advantages of large firms are mainly material while those of the small firms behavioral. Subsequently, the notion of "dynamic complementarity" was developed, where small and large firms play a different role in an economy, depending on the time and the industry.

Risk considerations have also been an important factor in innovation diffusion. Farmers initially allocated a portion of their land to new higher yield seed varieties to see their performance before fully adopting them. As Roumasset (1976) puts, a higher expected yield is associated with higher risk.

In addition, agricultural policy, through price support policies, reduces the level of uncertainty, which in turn can accelerate the diffusion of innovations (McGuirk & Mundlak, 1991). Cochrane (1979) also argued that commodity programs in the United States were pivotal in the adoption of agricultural innovations, mainly mechanical and chemical. On the other hand, taxation on agricultural output had a negative effect on the adoption of technological innovation. Cavallo and Mundlak (1982) argued that taxation of output was one of the primary reasons for the low growth in the agricultural sector in Argentina. International trade regulations also affect technological innovations. However, the impact of such changes in trade regimes depends on the relative comparative advantage of each country.

There are also geographical considerations for adopting new technologies. Innovations are more likely to be adopted if they are closer to the market centers. Support for new equipment is comparatively more accessible than a remote location.

Access to credit also affects diffusion (Hoff, Braverman, & Stiglitz, 1993). Large scale investments are difficult when cooperatives face credit constraints. Smaller farms generally have more difficulties in gaining access to credit. This may be another factor why smaller farms adopt new technologies later. High interest rates also can hinder the diffusions of innovation.

Land ownership considerations are also important. There is a distinction between land ownership and management. A considerable portion of farmland might be operated by individuals who do not own the land. Therefore, farmland is rented through contracts. The time-frame of such contracts is important if the contracts are short term, the managers will avoid making large scale investments and thus adopt new technologies. Subsequently, the time horizon of an investment is an additional crucial factor of diffusion.

Wahab, Rose, and Uli (2009) highlight the relationship between knowledge, technology recipients, technology suppliers, relationship characteristics, and the

degree of inter-firm technology transfer. They underlined that the degree of technology transfer depends on five characteristics. The first characteristic was knowledge, which in turn depends on the tacitness of the knowledge i.e., non-verbalized intuitive knowledge that is not easily transferable, the complexity of the knowledge, and the specificity. In addition, technology recipients' characteristics have been suggested to be an essential factor in knowledge transfer. The absorptive capacity and the recipients' collaborativeness were suggested to have a positive correlation with knowledge transfer. Technology supplier characteristics, such as partner protectiveness and transfer capacity, also facilitate inter-firm technology transfer. Finally, relationship characteristics such as the quality of the inter-firm relationship and the mutual trust between them have a significant impact on knowledge diffusion.

6 Agricultural Cooperatives: Innovation Strategies and Diverse Case Studies

6.1 Definition of Agricultural Cooperatives

An agricultural cooperative is defined by the International Cooperative Alliance (ICA) as: "an autonomous association of persons united voluntarily to meet their common economic, social, and cultural needs and aspirations through a jointly-owned and democratically-controlled enterprise" (International Co-operative Alliance, 1995). The agricultural sector has heterogeneous and challenging technological demands. Therefore, if we are concerned about which type of innovation system is desirable, it is difficult to assign a single promising and efficient archetype which fits the role of cooperatives (Fronzaglia, Guedes, & Santos, 2008).

6.2 Innovation Strategies Adopted by Cooperatives: A Case Study from Spain

Agriculture, as a production sector, has many limitations; a major one being its fragmented structure. This becomes immediately apparent when a traceability system is attempted to be introduced. Such limitations could be easily overcome by the collaborative structures used by cooperatives or unions of cooperatives which, desire traceability or guaranteed quality attributes.

Several agricultural cooperatives in the Region of Valencia, Spain (where the cooperative sector is deeply rooted) have adopted innovative strategies to increase the added-value of their products (Ortiz-Miranda, Moreno-Pérez, & Moragues-Faus, 2010). Traceability, quality standards, and food safety requirements are being implemented during the last years as a response to increasing concerns about food safety and animal health. These innovation strategies can be classified as new

configurations of AFNs (alternative food networks). In that case, individual cooperatives formed a cooperation framework to create a common trademark (Ortiz-Miranda et al., 2010). The network of cooperatives centralizes transaction costs while it deals with the certification authority and bureaucracy. This framework emphasizes (Ortiz-Miranda et al., 2010) locally based (and organic) characteristics of products (for example, production of oil exclusively from autochthonous varieties or from 1000-year-old trees). Of course, there is a demand for organic certification as well as for traceability of the oil. New processing lines have evolved to separate organic from conventional production. Territory-based labeling of olive oil from the Mountain of Castello (Valencia, Spain) is a promising example where traceability can be coupled with molecular techniques. In the future, cooperatives have the opportunity to implement molecular traceability techniques for olive oil using DNA extracted from oil and polymerase chain reaction (PCR) amplification of simple sequence repeat (SSR) markers (Corrado, Imperato, La Mura, Perri, & Rao, 2011; Raieta, Muccillo, & Colantuoni, 2015; Rotondi, Beghè, Fabbri, & Ganino, 2011). If technical support (experimentation and diffusion of information) can be coupled with convincing potential technology adopters, then the implementation of traceability techniques will add value to the product. On the other hand, environmentally friendly farming practices already enable the cooperatives to claim quality certification for their "organic" production. In addition to increased shelf price, food quality improvement can be further linked with tourism and the culinary sector.

6.3 Production and Commercialization of Novel Products by Collaboration of Cooperatives with Research Centers: A Case Study from Brazil

In 1996, a federation of Brazilian cooperatives (***Aurora Alimentos***) signed a contract with ***Embrapa Swine and Poultry*** research center, buying ***Embrapa*** improved hog line MS58 boars and sows for commercial breeding and distribution of progeny (Fronzaglia et al., 2008). ***Aurora Alimentos*** is a federation of 17 member cooperatives (77,500 farmer members) and one of the largest federations of cooperatives in meat processing technology in Brazil. The **E*mbrapa Swine and Poultry*** research center is an Embrapa's research unity with expertise on poultry and hog genetic improvement, **based on quantitative genetics**. The ***Embrapa Swine and Poultry*** provides technical and scientific expertise, training, courses, equipment, installations, laboratory-based diagnosis and software required for swine, and poultry raising (Fronzaglia et al., 2008).

In this technology transfer process, **Embrapa** offered technical assistance to ***Aurora Alimentos*** for the analysis of MS58 breeding performance. ***Aurora Alimentos*** offered breeding facilities and the commitment to sell the piglets grown by its affiliated cooperatives farmer members (Fronzaglia et al., 2008). ***Aurora Alimentos*** pays royalties to **E*mbrapa*** as specified by the contract.

Interaction of ***Aurora Alimentos*** with ***Embrapa Swine and Poultry***'s research center made possible the development of an even more improved hog line (MS60) and the substitution of MS58 with MS60 (Fronzaglia et al., 2008). Both lines [**MS58** (first to be achieved) and **MS60** (known as "Light Pork," substitution of **MS58**)] are swine lines derived from a genetic enhancement effort which included crossing of pure races as part of a Research and Development program to enhance meat quantity and quality[2].

6.4 Innovation Process in Agricultural Production

By definition, an agricultural value chain refers to all collaborators (from farmers and their cooperatives to final distributors and consumers) and their activities in input supply, production, processing, transport, distribution, marketing and purchasing of a particular final product or a group of closely related products (Anandajayasekeram, 2011). The addition of value in the product through its route, from producers to consumers, is of particular interest. Value addition to a product is the outcome of diverse activities (cleaning, bulking, packaging, and transporting) at each stage in the value chain mediated by different groups of people.

Innovation process in agricultural production may be divided into four major steps (Anandajayasekeram, 2011):

1. Invention (it includes the scientific contribution to priority problems solving)
2. Transformation of scientific results into novel technological processes or products
3. Commercialization of processes or products
4. Adoption of novel products by consumers and validation of socioeconomic benefits of the new technology.

6.5 Drawbacks in the Innovation Process

An innovation platform is a physical or virtual environment for thinking, talking, sharing ideas, discussing problems, listening, learning, and collaborating to innovate. There were many drawbacks to the innovation process in the past (Dons & Bino, 2008). Science was mainly curiosity-driven and the importance of its applications was sidelined. The commercialization/industrialization of science was not seen as a crucial factor, and was even considered a threat to the independence of the universities. Finally, establishing relationships with the private sector (industry) was an underestimated factor for the success of the innovation platforms. Moreover, a major constraint in agricultural development is that agricultural research is solely under the supervision of the Ministry of Agriculture in most developing countries (Anandajayasekeram, 2011). It is proposed by the same authors that if the

[2]By improving feed conversion efficiency and decreasing fat fraction.

responsible agency were the Ministry of Science, many constraints could be bypassed. The higher education system is focused on teaching and thesis supervision rather than Research and Development (Anandajayasekeram, 2011). The orientation of agricultural research and related priorities must also be reconsidered. For example, a shift from the "prescriptive" tradition towards participatory and action-oriented research in developing countries is necessary (Anandajayasekeram, 2011).

6.6 The Role of Cooperatives into Agricultural Innovation Systems

Cooperatives must be incorporated into collaborative Agricultural Innovation Systems, which include all the organizations actively participating in technological change in agriculture. An Agricultural Innovation System (Anandajayasekeram, 2011) is an arrangement of institutions and persons which includes all participants ranging from agricultural research institutes to farmer's organizations as well as their collaborations with the private sector (local, national, and multinationals agro-industrial firms). Each agricultural technological innovation must acquire acceptance of consumers, pressure groups, and their organizations and be compatible with regulations, laws, norms, customs and beliefs that affect the development, and diffusion of the innovation process. It is an absolute priority to harmonize the Research and Development planning, and implementation in the context of the economic, political, and cultural interrelationships in which it takes place. The innovative dynamics of a country depends not only on the dynamics of each participant (universities, research institutes, cooperatives, private sector) but also on their collaboration, the relationships, and interactions between these components as elements of a collective system and their interaction with the society.

6.7 The Three-Legged Stool Model for the Role of Research in the Innovation Process

Blackswan (2010) proposed an illustration of a three-legged stool for the role of research in the innovation process and the diffusion/commercialization of knowledge (Blackswan, 2010). The first leg (critical element) of the stool is knowledge generated through basic research. The second leg is the ability to transform ideas and experimental findings into real products and services through adaptive applied research. The third leg is the ability for marketing and commercialization of the ideas. These three legs in the innovation process must be of the same length to achieve a stable, effective structure. Customers (end users) are the necessary platform to hold the legs and to support the stool (institutes, institutions, government, and society). The stability of the overall system depends on the stability of each part.

6.8 Agricultural Research-for-Development (AR4D) Projects: A Case Study from Dutch Agricultural Innovation System

The Dutch agricultural innovation system can be used as a case study for the reliability of the three-legged stool model (Dons & Bino, 2008). During the previous years, significant changes have taken place in the Dutch agricultural innovation system for the reorganization of the knowledge system and re-orientation of the interactions of research organizations with the private sector. The classical linear scheme of information flow (from basic/fundamental university research towards implementation/commercialization through the intermediate step of strategic and applied research at governmental research institutes and experimental stations) has been replaced with more flexible forms of collaboration. We are in the middle of a transition phase, where the classical model is replaced by more open forms of co-innovation and collaboration between research centers and industries (Dons & Bino, 2008). This co-innovation model [the so-called public–private partnership (PPP)] is based on the close collaboration of various stakeholders in a more open and dynamic system. In the Netherlands, the agro-food research and agro-food industrial complex exist in close interaction for many decades, and this is the key to the expansion of the agricultural innovation system (Dons & Bino, 2008).

This model radically transforms collaboration in Research and Development as well as in education. Education, as well as academic research and industries, all work together to form a network system, which in turn establishes effective education programs.

Nowadays, the Dutch agricultural innovation system is considered a critical success factor for the strong development of Dutch horticulture.

6.9 Agricultural Research-for-Development (AR4D) Projects: A Case Study from Nigeria

The RIU-Nigeria program is an AR4D project focused on the generation, evaluation, accumulation and, sharing of evidence for the reorientation of innovation and development strategies in agricultural research (Ugbe, 2010). This task involves providing a number of activities including the identification of possible collaborations of farmers' associations and co-operations with scientists from research institutes, technology producers, intermediaries, post-harvest processors and traders, private sector companies, policy-makers from related public agencies and government officials responsible for agriculture. The first step is to convince farmer cooperatives and other participants that they are all parts of the same agricultural production chain and that collaboration among them will be highly beneficial (directly or indirectly) for all. Cooperatives and other participants are involved in discussions and negotiations concerning innovation challenges, evaluation of strategies (on-farm productivity, post-harvest value addition, market development, and

policy enhancement), bottlenecks, or even policy issues. The formation of an innovation platform between cooperatives and other participants is strengthened by their connection to responsible state agencies and collaboration with international relevant development assistance agencies working in specific agricultural product value chains. The organizational capacity of co-operations and diverse local intermediaries must be complemented by a combination of strategies in attempting to improve development processes and agricultural innovation. Fundamental issues for the success of such programs are management of research and research outputs, investment in research as well as intended-user access to relevant information. Towards that goal a centralized Council (the central Agricultural Research Council of Nigeria) coordinated the efforts of the 15 NARIs—National Agricultural Research Institutes (Ugbe, 2010). Reorientation of NARIs to address key innovation challenges relating to local capacity to produce high-quality, collation and compilation of research outputs from all institutes (by also including national and international research outputs), formulation and implementation of innovation platforms, formulating quality-control measures, collaborating in agricultural development programs are the main challenges in these type of agricultural research-for-development projects (Ugbe, 2010).

6.10 Introduced Technologies Versus Locally Developed Innovation and the Role of Cooperatives

Introduced technologies are frequently less appropriate for addressing challenges than locally developed approaches. Locally developed innovation is adapted to local conditions within which it has been developed (Scogings, Ngubane-Shezi, & Shezi, 2009) mainly because farmers and farmer cooperatives can acquire an active role in contributing ideas, initiating research activities or re-orienting research to harmonize with their priorities and motivations. The next step in this more balanced partnership is development programs driven by farmers' ideas. This is a truly 'participatory' approach where farmers and cooperatives play an active role in the process that, in general, involves a very complex set of participants. Another critical factor for a successful agricultural innovation system is its adeptness in addressing challenges that local farmers and innovators are experiencing in each separate country (Scogings et al., 2009). One additional challenge is to support local innovation facilities while making these facilities available to farmers. This, for example, is the aim of a program coordinated by the Prolinnova (**PRO**moting **L**ocal **INNO-VA**tion) network. Prolinnova is a network operational in 20 countries in Asia, Africa and South America which is coordinated by a Non-Governmental Organization (NGO) based in the Netherlands. Farmer's organizations and cooperatives form a network with educational institutions, government departments, and NGOs focused on promoting ecologically oriented agriculture and participatory R&D approaches. Local Innovation Support Facilities supported by the Prolinnova network are

engaged in enhancing the impact of local innovation by supporting access to financial resources. Such facilities have already been established on a pilot basis and in several communities due to a Prolinnova supported initiative FAIR (Farmer Access to Innovation Resources). Enhancing farmers' innovative capacity is achieved by establishing locally managed funds to purchase materials and inputs needed for experiments and bringing the necessary expertise and skills which are necessary for the experiment or investigation (Scogings et al., 2009). The benefit from such programs is the opportunity to adapt research to local needs and to propose solutions harmonized with local cultures and familiar to local farmers rather than applying externally (and often inappropriate) derived ideas.

6.11 *The Role of Agricultural Research Information Systems in Cooperatives Innovation*

An Agricultural Research Information System (ARIS) can be defined "as a system that enables **digital connectivity** among **institutions** and **stakeholders** engaged in agricultural research within countries for sharing **scientific**, **technical** and **research management** information; and supports, in full or part, the **infrastructure** and **services** for electronic messaging and **communication among participating institutions**" (Maru, 2002). By definition, an ARIS is a common information system with the ability to provide an organizational backbone to spread and diffuse information among the national agricultural research institutions (Sadovaskaya, 1999; Singh, 1998). This network of agricultural research institutions is based on the implementation of new Information and Communication Technologies (ICT) and is a critical factor for agricultural development, especially in under-developed or developing countries in Africa, Asia, South and Central America (Maru, 2002). The success of an ARIS is dependent on the ability of the national agricultural innovation system to share information competitively. An ARIS must be structured with the flexibility to share appropriate information with its clients and partners (for example, co-operations) in agricultural innovation. The next level of ARIS development is the orientation towards supporting Research Institutes to seek new partners at a national, regional, and global level (Maru, 2002).

6.12 *Role of Cooperatives in Adopting Innovation*

Diederen, van Meijl, and Wolters (2003) shares empirical evidence concerning the determinants of innovation adoption behavior in Dutch agriculture during the years 1995–1998 (Diederen et al., 2003). There is a positive and statistically significant impact of "cooperation" on the easiness of adopting innovations, while "cooperation" can be viewed as an indicator for adaptation of innovation (Diederen et al.,

2003). Furthermore, there is a positive correlation between the number of farmers involved in agricultural cooperative networks and the easiness of adopting innovations early. Larger cooperatives have members who can adopt innovations earlier (Diederen et al., 2003). Innovation adoption is also positively correlated (Diederen et al., 2003) with: (1) farmer's access to external sources of information—in which case networking can play a significant role (2) market position if the market of the agricultural product permits product differentiation and (3) available technical skills. The benefits of innovation can be determined by the improvement in performance between the old and new technology, by the cost-effectiveness of the investment in the new technology and by characteristics of the business environment. For example, strict market regulations may hinder the acceptance of novel ideas. On the other hand, once someone becomes an innovator, then there is an increased probability of being an innovator in the future.

6.13 The Role of Innovation Intermediaries in Collaboration Between Research Institutes and Cooperatives

'Intermediary firms' are firms that help in the process of adaptation of specialized solutions on the market to the requirements of individual "user firms."

Innovation intermediaries should act as 'intelligent interfaces' between their clients and their 'task environment' in relation to analytical, environmental, and testing matters. Sometimes they are referred to as "techno-rockers". The short-term role of Innovation intermediaries is to provide immediate, 'one-off' intermediary services to their clients (technology donors or technology recipients). It is preferable if Innovation intermediaries can offer longer-term, 'relational' innovation capabilities to their clients as well (Howells, 2006). Information scanning (and gathering) function, communication function, knowledge processing, scanning and information processing, testing and validation, gate-keeping, brokering, and commercialization are the main functions associated with intermediation (Blaikie, 2009). Of course, the primary role of intermediaries is to provide information scanning and exchange functions but a broader role of intermediaries as evaluators of technology *after* it has been transferred should not be underestimated. The advisory role of Innovation intermediaries includes advice about how the client company should develop in the future, how it should respond to the changing regulatory environment and, which kind of technical improvements should adapt (Howells, 2006). Innovation intermediaries should provide additional functions (for example, contract research, testing, or training work) within an innovation system (Howells, 2006). Building linkages with external knowledge providers, scanning for new sources of knowledge, and developing and implementing innovation strategies have also been proposed (Bessant & Rush, 1995). The contribution of innovation intermediaries may expand to activities such as acting as a mediator between researchers and co-operations while providing information to research institutes and co-operations about potential

collaborators. If the collaboration has already been established, funding and supporting the innovation outcomes of such collaborations, helping co-operations find advice or brokering a transaction between the two parties when they are already collaborating. Stankiewicz (1995) recognized the pivotal role that 'bridging institutions' may play (as another type of innovation intermediaries) in helping to link players within a technological system (Stankiewicz, 1995). Howells (2006) highlights the pivotal role of Innovation Intermediaries as essential players in the UK innovation system and their contribution to the growth and development of biotechnology in the UK (Howells, 2006). By contrast, significant gaps remain between the private and public spheres of research and innovation in the French innovation system, which might be potentially bridged by innovation intermediaries (Lallement & Paillard, 2003).

6.14 A Discrete Innovation Case: Traceability

Traceability is defined as "the capability for reconstruction of (1) the origin of materials and components used during the production, (2) the history of the processes used during distribution and (3) the identification of the location of the product after delivery within the supply chain using documented identifications" (Giacomini & Mancini, 2001; Giacomini, Mancini, & Mora, 2002). These identifications are related to the flows of material and each traceability system must be extended throughout the whole supply chain. The ISO 8402 standard refers to "company" traceability as the "capability to retrace the history, use or location of an entity using recorded identification data". The term "traceability system" is indicative of two different processes: (1) tracking (the process which the product follows throughout the whole supply chain), and (2) the reverse process (tracing) which allows the product to be traced back up the chain). The existence of a traceability system provides the involved producers and processing cooperatives (or companies) with a powerful competitive tool/competition instrument. It provides them the opportunity (1) to establish improved production of products, (2) to manage stock more efficiently through the identification process, (3) the standardization of the production procession and finally (4) it contributes to the increase of the product value and passage of additional costs to the (willing to pay for this "superior" product) purchaser (Giacomini & Mancini, 2001; Giacomini et al., 2002).

6.15 Implementing Traceability Solutions by Cooperatives: A Case Study from Emilia Romagna Region (Italy)

A traceability model system has been adopted (Giacomini & Mancini, 2001; Giacomini et al., 2002) by an Italian fruit and vegetable farmers cooperative (about

500 small and medium-size farms) located in the Emilia Romagna Region (province of Modena, Italy). This cooperative replied to growing consumer demand for product traceability and safe food products in terms of hygiene and quality by establishing a robust traceability system as the only way to become competitive in terms of quality and safety of the products. Various EU directives and regulations have been activated since the early 1980s in order to systematically monitor the residue levels of pesticides in foodstuffs and to encourage the spread of phytopathological control agents with low-environmental-impact by implementing organic, supervised and integrated techniques in the vegetable and fruit sector.

The implementation of a traceability system enforced the reorganization of the Emilia Romagna Region co-operative (Giacomini & Mancini, 2001; Giacomini et al., 2002). The implementation of traceability services was accompanied by an increase in production costs (costs for the purchase of the computers, hardware, and software, costs for the purchase of the optical readers and costs of phytosanitary data production and management) and redefinition of contractual relations with their suppliers. The Emilia Romagna Region co-operative managed to recover the greatest part of the costs by selling the software packages which were developed by the cooperative to the union of Italian cooperatives (CONERPO consortium) for installation, in all the cooperatives associated with CONERPO (Giacomini & Mancini, 2001; Giacomini et al., 2002). Furthermore, the added value (provided by the traceability possess) to its products enabled the cooperative to penetrate new markets such as Casino and Carrefour, which require complete product traceability from their suppliers and increased its competitiveness with respect to other private sector competitors in general. Furthermore, it was straightforward to adopt their specifications, given the fact that it already had all the organizational skills and capacities required.

6.16 *Implementing Traceability Solutions by Cooperatives: The French Paradigm*

Large French cooperatives are implementing traceability solutions in both animal production and arable crops as well as decision support systems initially developed by Research and Development institutes (such as ARVALIS). For example, INVIVO (the French cooperatives union) contacted a recent agreement with a major provider of traceability solutions (ISAGRI—www.isagri.fr, the French leader of ICT for Agriculture). AMI, like ISAGRI, is a provider of traceability solutions not only in crop production but also in many other sectors. Arcade Conseil (arcadeconseil.fr), CDER Informatique, and maferme.com (www.cder.fr), Agreen Tech (agreentech.com), GIE TRACENORD (prestilem.fr) are other major challengers of ISAGRI. They are companies dedicated to traceability solutions of agricultural production and offer a wide range of traceability tools to farmers and

cooperatives to harmonize and simplify traceability systems and to minimize traceability costs for cooperatives and farms (Waksman, Escriou, & Gentilleau, 2003).

6.17 *Implementation of Traceability Systems by Collaboration of Cooperatives with the Private Sector: A Case Study from Brazil*

IMCOPA is a Brazilian family-owned company specialized in non-GM soybean which collaborates with about 11 cooperatives. These cooperatives provide 80% of the volume of non-GM soybean bought by the company. IMCOPA, which is well positioned geographically near large non-GM soybeans supplies, made a decision for raceability system implementation in 1998 (during the initial phase of non-GM soybean certification system implementation) and gained large profits (Pelaez, Aquino, Hofmann, Melo, & Goldsmith, 2010). The testing, purchase and distribution of non-GM soybean seeds to cooperatives is monitored by IMCOPA and re-tested by the cooperatives. GM testing is also implemented at the final product stage (Pelaez et al., 2010). Commercial strategies were also implemented by several other cooperatives in Brazil which seek to implement their own systems for identification, preservation and certification of non-GM soybean and to create a new market for these products. Inspired from IMCOPA example, three other cooperatives implemented their own certification and traceability systems for non-GM soybean in Paraná from 2004 to 2006 (Pelaez et al., 2010).

6.18 *Can Cooperatives Be Involved in Other Types of Traceability Systems? Geographical Indications Is a Promising Example*

The EU (European Union) protects high-quality agricultural products based on geographical location by using the designations of GIs (geographic indications) since 1992 (Babcock, 2015; Babcock & Clemens, 2004). EU and other countries are trying to expand this type of protection through geographical indications and negotiate about this issue with WTO (World Trade Organization). On the other side, a certification mark is a type of trademark which can be obtained by U.S. processors, producers, and provides, similarly to geographic indications, protection with the limitation that the protection of the products is not applicable outside the United States (Babcock, 2015; Babcock & Clemens, 2004). To ensure the high quality status of products Protected by Geographical Indications, it is required that each producer or cooperative maintains detailed records that include the origin and quality of the product, production data, quantities produced, control procedures for the finished product, and information to permit traceability. Agricultural cooperatives

with experience in the implementation of traceability systems, in general, can easily expand to adopting a traceability system that would also include geographical indications.

6.19 The Promise of Implementing a Traceability System for Plant Propagation Material by Agricultural Cooperatives

Quality requirements regarding health, purity and identity of agricultural products as well as post-control and comparative tests and labeling increase the demand of implementing traceability tools in agricultural practice. Harmful organisms can slip through and cause devastating outbreaks as a result of the high volumes of international commerce with countries where harmful pests are widespread in plant propagation material. Plant Propagation (reproductive) material is a fundamental factor for the quality of agricultural food products and public health as well as productivity and biological diversity. If farmers, cooperatives, and even governments will not harmonize to a widely accepted plant health regime, large-scale outbreaks caused by harmful organisms can lead to an increase in production losses and costs. On the other hand, if cooperatives can adopt a (molecular) traceability system dedicated with surveillance of plant reproductive material, many outbreaks of harmful pests can be detected in plant propagation material and successfully eradicated. In addition, such a system can foremost safeguard true-to-typeness of reproductive material itself. Of course, there is a requirement for scientific support for the development of traceability systems (molecular, logistic, or otherwise) and collaboration of cooperatives with diagnostic laboratories. Molecular biology methods involve a series of discoveries such as restriction enzymes for DNA molecules fragmentation, development of techniques such as PCR for amplification of DNA fragments, and lately introduction of second generation massive DNA sequencing. These discoveries create a new innovative toolbox for plant breeders to unravel the genetic variation in the plant germplasm (Dons & Bino, 2008). For example, Amplified Fragment Length Polymorphisms (AFLPs) were developed as a molecular marker technology in 1990 by Keygene (Vos et al., 1995). It became a widespread standardized technology for breeders across many plant and animal species and enabled them to use genotyping instead of phenotyping in the selection process and reach a clear insight into the genetic variation of their breeding lines. Such DNA fingerprinting technologies, similar to SSR markers have revolutionized plant breeding and are used routinely by breeding companies for many vegetable crops. Currently, an array of more efficient molecular marker technologies are routinely used by research laboratories and hold the promise for standardization, dissemination, and commercial end use.

References

Acs, Z., & Audretsch, D. (1988). Innovation in large and small firms: An empirical analysis. *The American Economic Review, 78*(4), 678–690.

Acs, Z., & Audretsch, D. (1990). *Innovation and small firms*. Cambridge: MIT.

Adorno, T. W., & Horkheimer, M. (1992 [1944]). *Dialectic of enlightenment*. London: Verso.

Alston, J., Norton, G., & Pardey, P. (1995). *Science under scarcity: Principles and practice for agricultural research evaluation and priority setting*. Ithaca, NY: Cornell University Press.

Anandajayasekeram, P. (2011). *The role of agricultural R&D within the agricultural innovation systems framework* (In report prepared for the ASTI/IFPRI-FARA conference).

Babcock, B. A. (2015). Geographical indications, property rights, and value-added agriculture. *Iowa Ag Review, 9*, 1.

Babcock, B. A., & Clemens, R. L. (2004). *Geographical indications and property rights: Protecting value-added agricultural products* (MATRIC Briefing Paper 04-MBP 7).

Bessant, J., & Rush, H. (1995). Building bridges for innovation: The role of consultants in technology transfer. *Research Policy, 24*, 97–114.

Best, J. (2001). *How claims spread : cross-national diffusion of social problems*. New York: Aldine de Gruyter.

Blackswan. (2010). *The 3-legged stool of innovation: Driving innovation success*. Blackswan whitepaper series.

Blaikie, N. (2009). *Designing social research*. Cambridge: Polity.

Borup, M., Brown, N., Konrad, K., & van Lente, H. (2006). The sociology of expectations in science and technology. *Technology Analysis and Strategic Management, 18*(3/4), 285–298.

Callon, M. (1986). Elements of a sociology of translation: Domestication of the scallops and the fishermen of St Brieuc Bay. In J. Law (Ed.), *Power, action and belief: A new sociology of knowledge* (pp. 196–233). London: Routledge.

Caswell, M., & Zilberman, D. (1986). The effects of well depth and land quality on the choice of irrigation. *American Journal of Agricultural Economics, 4*, 798–811.

Cavallo, D., & Mundlak, Y. (1982). *Agriculture and economic growth in an open economy: The case of Argentina (Research report 36)*. Washington, DC: International Food Policy Research Institute.

Cochrane, W. (1979). *The development of American agriculture: A historical analysis*. Minneapolis, MN: University of Minnesota Press.

Corrado, G., Imperato, A., La Mura, M., Perri, E., & Rao, R. (2011). Genetic diversity among olive varieties of Southern Italy and the traceability of olive oil using SSR markers. *Journal of Horticultural Science and Biotechnology, 86*, 461.

David, P. (1969). *A contribution to the theory of diffusion*. Stanford, CA: Stanford Centre of Economic Research in Economic Growth.

Davies, S. (1979). *The diffusion of process innovations*. Cambridge: University Press.

de Gorter, H., & Zilberman, D. (1990). On the political economy of public good inputs in agriculture. *American Journal of Agricultural Economics, 72*(1), 131–137.

Diederen, P., van Meijl, H., & Wolters, A. (2003). Modernisation in agriculture: What makes a farmer adopt an innovation? *International Journal of Agricultural Resources, Governance and Ecology, 2*, 328–342.

Dons, H. J., & Bino, R. J. (2008). Innovation and knowledge transfer in the Dutch horticultural system. In W. Hulsink & H. Dons (Eds.), *Pathways to high-tech valleys and research triangles: Innovative entrepreneurship, knowledge transfer and cluster formation in Europe and the United States* (pp. 119–137). Dordrecht: Springer.

European Commission. (2013). *EUR 25766 options for strengthening responsible research and innovation*. https://doi.org/10.2777/46253. Accessed January 21, 2016, from https://ec.europa.eu/research/swafs/pdf/pub_public_engagement/options-for-strengthening_en.pdf

Foucault, M. (1972). *The archaeology of knowledge*. London: Tavistock.

Foucault, M. (1980). *Power/knowledge. Selected interviews and other writings 1972–1977*. Brighton: The Harvester Press.
Fronzaglia, T., Guedes, V. G. F., & Santos, E. (2008). *The role of agricultural cooperatives interaction with public research on technological change in Brazil*. Co VEdPL-ad, operativismo (Eds.). São Paulo, Brasil.
Giacomini, C., & Mancini, M. C. (2001). A study case about a traceability system in the fruit and vegetable chain. In *IV International Symposium Perspective of the Agri-Food System in the New Millennium*, 5–8 September 2001, Bologna.
Giacomini, C., Mancini, M., & Mora, C. (2002). Case study on the traceability systems in the fruit and vegetable sector. In *17th Symposium of the International Farming Systems Association* (pp. 17–20). Lake Buena Vista, FL.
Godin, B. (2006). The lineal model of innovation. *Science, Technology and Human Values, 31*(6), 639–667. https://doi.org/10.1177/0162243906291865.
Gordon, R., & Zusman, P. (1991). *Political-economic analysis: Explanation and prescription*. New York: Cambridge University Press.
Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica, 25*(4), 501–522.
Hoff, K., Braverman, A., & Stiglitz, J. (1993). *The economics of rural organization: Theory, practice and policy*. New York: Oxford University Press.
Howells, J. (2006). Intermediation and the role of intermediaries in innovation. *Research Policy, 35*, 715–728.
Huffman, W. (1998). *Finance, organization and impacts of U.S. Agricultural Research: Future prospects 21st century*. Paper Prepared for Conference "Knowledge Generation and Transfer: Implications for Agriculture in the 21st Century". Berkeley: University of California.
International Co-operative Alliance. (1995). *Statement on the Co-operative Identity*. Retrieved from https://www.ica.coop/en/cooperatives/cooperative-identity
Kennedy, C., & Thirlwall, A. (1972). Technical progress: A survey. *Economic Journal, 82*(325), 11–72.
Kosellek, R. (2004). *Futures past—On the semantics of historical time*. Columbia, NY: Columbia University Press.
Kuznets, S. (1962). Inventive activity: Problems of definition and measurement. NBER Chapters. In *The rate and direction of inventive activity: Economic and social factors*. Princeton, NJ: National Bureau of Economic Research.
Lallement, R., & Paillard, S. (2003). *The French innovation system in the knowledge-based economy*. Paris: Commisariat General du Plan.
Mansfield, E. (1963). The speed of response of firms to new techniques. *Quarterly Journal of Economics, 77*(2), 290–311.
Maru, A. (2002, October). A normative model for agricultural research information systems. In *Proceedings of the Third Asian Conference for Information Technology in Agriculture. Asian Federation for Information Technology in Agriculture (AFITA)-Chinese Academy of Agricultural Sciences (CAAS)* (pp. 26–28).
McGuirk, A., & Mundlak, Y. (1991). *Incentives and constraints in the transformation of Punjab agriculture*. Research Report 87. Washington, DC: International Food Policy Research Institute.
Nisbet, R. A. (2009). *History of the idea of progress 4th printing*. New Brunswick: Transaction. https://books.google.gr/books?hl=el&lr=&id=QDRWfZ9Ydw0C&oi=fnd&pg=PR9&dq=the+concept+of+progress+in+sociology&ots=WRMY3fNPrI&sig=Y-h98FK4ReRtOP3FLPZnWUs92Uo&redir_esc=y#v=onepage&q=the%20concept%20of%20progress%20in%20sociology&f=false
No Author. (2011). Auguste comte on the natural progress of human society. *Population and Development Review, 37*, 389–394. https://doi.org/10.1111/j.1728-4457.2011.00416.x.
Nooteboom, B. (1989). Diffusion, uncertainty and firm size. *International Journal of Research in Marketing, 6*, 109–128.

Olmstead, A., & Rhode, E. (1993). Induced innovation in American agriculture: A reconsideration. *Journal of Political Economy, 101*(1), 100–118.
Ortiz-Miranda, D., Moreno-Pérez, O. M., & Moragues-Faus, A. M. (2010). Innovative strategies of agricultural cooperatives in the framework of the new rural development paradigms: The case of the Region of Valencia (Spain). *Environment and Planning A, 42*, 661.
Pelaez, V., Aquino, D., Hofmann, R., Melo, M., & Goldsmith, P. D. (2010). Implementation of a traceability and certification system for non-genetically modified soybeans: The experience of Imcopa co. in Brazil. *International Food and Agribusiness Management Review, 13*, 27–44.
Raieta, K., Muccillo, L., & Colantuoni, V. (2015). A novel reliable method of DNA extraction from olive oil suitable for molecular traceability. *Food Chemistry, 172*, 596–602.
Rogers, E. (2003). *Diffusion of innovations*. New York: Free Press of Glencoe.
Rosenberg, N. (1994). *Exploring the black box. Technology, economics and history*. New York: Cambridge University Press.
Rothwell, R., & Zegveld, W. (1982). *Innovation and the small and medium sized firm*. London: Frances Pinter.
Rothwell, S. (1985). Company employment policies and new technology. *Industrial Relations Journal, 16*, 43–51. https://doi.org/10.1111/j.1468-2338.1985.tb00524.x.
Rotondi, A., Beghè, D., Fabbri, A., & Ganino, T. (2011). Olive oil traceability by means of chemical and sensory analyses: A comparison with SSR biomolecular profiles. *Food Chemistry, 129*, 1825–1831.
Roumasset, E. (1976). *Rice and risk: Decision-making among low-income farmers*. Amsterdam: North-Holland.
Ryan, B., & Gross, N. (1943). The diffusion of hybrid seed corn in two Iowa communities. *Rural Sociology, 8*(15), 16–24.
Sadovaskaya, L. (1999). ARIS: Structure, information resources and services, place and role in the unified information field of the Russian Federation. *IAALD Quarterly Bulletin, XLIV*, 141–146.
Schmitz, A., & Seckler, D. (1970). Mechanized agriculture and social welfare: The case of the tomato harvester. *American Journal of Agricultural Economics, 52*(4), 569–577.
Schumpeter, J. (1942). *Capitalism, socialism and democracy*. New York: Harper and Row.
Scogings, B., Ngubane-Shezi, N., & Shezi, Z. (2009). Supporting or enhancing local innovation as a tool for ensuring the sustainable use of natural resources. In *The Fourth Biennial National Land Care Conference*. Polokwane, Limpopo Province.
Singh, G. (1998). Strategies for development of agricultural research information system: ICAR experience. *The Journal of Higher Education, 2*, 171–192.
Skjølsvold, T. M. (2012). *Towards a new sociology of innovation. The case of bioenergy in Norway and Sweden*. PhD dissertation. Department of Interdisciplinary Studies of Culture, Faculty of Humanities, Norwegian University of Science and Technology, Trondheim, Norway. Accessed January 21, 2016, from http://www.diva-portal.org/smash/get/diva2:549380/FULLTEXT02.pdf
Sklair, L. (2001 [1970]). *The sociology of progress*. London: Routledge.
Smith, K. (2008). *Climate change and radical energy innovation: Policy issues*. Oslo: Centre for Technology and Culture.
Stankiewicz, R. (1995). The role of the science and technology infrastructure in the development and diffusion of industrial automation in Sweden. In B. Carlsson (Ed.), *Technological systems and economic performance: The case of factory automation* (pp. 165–210). Dordrecht: Springer.
Sunding, D., & Zilberman, D. (2001). The agricultural innovation process: Research and technology adoption in a changing agricultural sector. *Handbook of Agricultural Economics, 1*, 207–261.
Ugbe, U. P. (2010). *It may take a little while: Insights on agricultural research for innovation and development in Nigeria*. NR International.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., Friters, A., Pot, J., Paleman, J., & Kuiper, M. (1995). AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research, 23*, 4407–4414.

Wahab, S., Rose, R., & Uli, J. (2009). Relationships between knowledge, technology recipient, technology supplier, relationship characteristics and degree of inter-firm technology transfer. *European Journal of Social Sciences, 11*(1), 86–102.

Waksman, G., Escriou, H., & Gentilleau, G. (2003). The situation of ICT in the French agriculture. In *European Scientific Association (EFITA) 2003 Conference* (pp. 5–9).

Woods, E. B. (1907). Progress as a sociological concept. *American Journal of Sociology, 12*(6), 779–821.

Virtual Incubators and Online Business Tools for Agro-Food SMEs

Christina Diakaki, Nikoleta Banani, and Evangelos Grigoroudis

Abstract Due to their limited access to resources, experts, and integrated business tools, the SMEs, which comprise the major part of the agro-food sector, struggle to start and develop their business in the highly competitive environment of our days. A recent survey suggested a need for support with respect to information, learning, networking and use of business tools. Such support can be offered via online tools and services, and has been the aim of the work presented in this chapter to search and provide a relevant review. Since such services and tools are typically provided by business support agents such as incubators, accelerators and innovation/entrepreneurship centers, a literature review of these agents and the business services they offer is also provided. The reviews reveal that there is a number of online platforms offering services and tools along the aforementioned four axes of interest, but no platform exists to fully cover all the expressed needs, being also specialized in the agro-food sector. This lack reveals a hidden potential for the sector, which can be elaborated through the development of online business platforms and tools tailored to its special requirements and needs.

Keywords Agro-food SMEs · Incubators · Business accelerators · Online business services · Online business tools

1 Introduction

The agro-food sector comprises mainly of Small Medium Enterprises (SMEs) (García-Álvarez de Perea, Ramírez-García, & Del Cubo-Molina, 2019), which, due to their limited access to resources, experts, and integrated business tools, struggle to start and develop their business in the highly competitive environment that prevails nowadays. According to a recent, interview-based survey

C. Diakaki (✉) · N. Banani · E. Grigoroudis
School of Production Engineering and Management, Technical University of Crete, Chania, Greece
e-mail: cdiakaki@isc.tuc.gr; nbanani@isc.tuc.gr; vangelis@ergasya.tuc.gr

E. Krassadaki et al. (eds.), *Operational Research in Agriculture and Tourism*, Cooperative Management, https://doi.org/10.1007/978-3-030-38766-2_2

(INNOVAGRO, 2019), undertaken within the frame of the INTERREG-ADRION project INNOVAGRO among representatives of agro-food companies from Greece, Slovenia, Italy, Albania, and Serbia, most entrepreneurs in this sector decide to set up a new business, mainly by using their instinct and less frequently by doing a concrete business calculation or a business plan. At the very early business steps, entrepreneurs often seek advice from their close environment (i.e., family, friends, and work). Initial funds are usually given by the owners. However, while the company is growing, more funds are required and the company has to present a justified business plan to get additional funds. In many cases, a business plan is prepared by the owner and/or the company's staff and the most difficult issue refers to the market estimations and investment returns.

The most important problems that entrepreneurs face in the first steps of their business are funding, bureaucratic procedures (generally and particularly for getting the initial operating licenses), and strategy development. Later, the frequent change of the food legislation, tax policy, starting exports and finding partners are also issues of concern.

Most agro-food companies target their product to the local or the national market and only few of them export their products, mainly to countries of the European Union (EU), but also in other countries. The first problem that they have to overcome in order to go abroad is their low capacity, but secondly they have to face the complicated procedures and probably the additional certifications required. However, before aiming on a foreign market, a company needs to have information about demographic data of the country, eating and drinking habits, as well as any special legislation concerning food imports and distribution. However, although exports is an opportunity for growth, only few companies have conducted a market research for a foreign country and most of them do not even have a written marketing plan. Furthermore, most companies do not update their business plans in a regular basis.

Regarding networking, most companies participate in national and international fairs in order to develop collaborations, while business to business (B2B) meetings organized by chambers are also helpful. However, the participation cost in this kind of events (as exhibitors) is quite high and it is thus difficult for a small company to afford it. The most important available online information that may be useful for agro-food companies includes friendly farming practices and legislation, while the most useful business tools are related to marketing plan development and effectiveness evaluation of a promotional campaign.

The aforementioned survey findings highlight the need of agro-food SME for support of their effort towards a viable and sustainable development in the following four axes (INNOVAGRO, 2019):

1. *Information* with respect to case studies and success stories, good practices, new/existing agro-food clusters and clusters policy, funding opportunities, relevant cooperation networks, exports/imports, regulations and laws, environmental-friendly farming practices, etc.
2. *Learning* of financial and management issues or other topics such as how to start a new business.

3. *Networking* with exporters and importers, educational organizations and research centers, as well as entities and clusters, which may foster collaboration and mutual activities among SMEs.
4. *Use of business tools* facilitating development by enabling in an easy manner the self-assessment of the companies with respect to their innovation and extroversion potential, the preparation of business plans, the financial management, etc.

Given their typically limited resources, the best way to support SMEs with respect to their aforementioned needs, is through the provision of online tools and services. It has been thus the aim of this chapter to search and provide a structured review of readily available online business tools focusing on entrepreneurship support and growth.

Given that services and tools, which support business development and growth are typically provided by business support agents such as incubators, accelerators and innovation/entrepreneurship centers, the chapter starts with a literature review of these agents and the business services they offer. The business support agents review is followed by a review of existing business platforms and websites, and a presentation and discussion of specific online tools that can be used to address the needs of agro-food SMEs along the axes of information, learning, networking and use of business tools. The chapter concludes with a brief overview of the major findings and conclusions.

2 Business Support Agents

2.1 Incubators

Helping startups to grow and turn their ideas to viable ventures was always a promising way to enhance the entire economy, reduce unemployment and improve the living standards of the society. Business incubators first appeared in the 50s providing infrastructure facilities and several services to the incubated companies in order to help them grow.

Initially, incubators were a means to revitalize declining manufacturing areas, and they served as a tool for reconversion. They offered services to all kinds of enterprises from low-tech, to no-tech, including manufacturing and services. The oldest one is situated in the United States of America (USA). It was created in the 1950s in response to plant closures in Batavia and New York. In Europe, one of the first incubators was set up in the United Kingdom (UK) in 1975, when British steel formed a subsidiary called the British Steel Industry to create jobs in steel closure areas. Both in the USA and Europe, step by step, the concept evolved (Aernoudt, 2004). Actually, physical business incubators took off during the 90s. Thus, the number of business incubators was about 200 at the beginning of the 1980s and it became more than 3000 in the beginning of the new century (UNECE, 2000). Carayannis and von Zedtwitz (2005) notice that "*incubation has experienced*

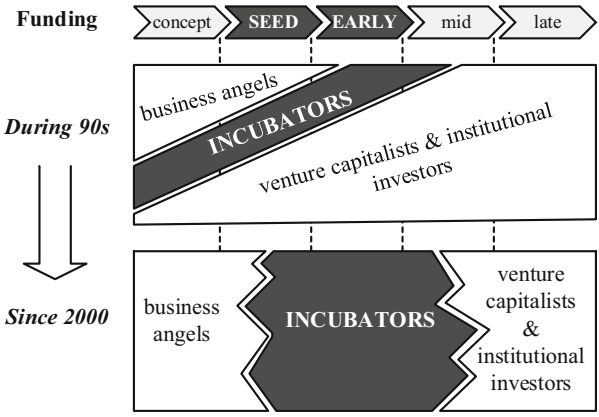


Fig. 1 Incubators evolution through the years

increased attention as a model of startup facilitation. Venture capitalists see incubators as a means to diversify risky investment portfolios, while would-be entrepreneurs approach incubators for startup support. Incubators are faced with the challenge and the opportunity of managing both investment risks, as well as entrepreneurial risks".

Through the years, incubators tended to target seed and early-stage startups, thus their "primary market" considered to be the intermediary stages of startup evolution, as described by von Zedtwitz (2003) and displayed in Fig. 1.

At the beginning of the new century, having the internet developed, virtual types of incubators were also born. The incubated companies could still enjoy the provided services without being located in a particular site, in the incubator's site. This way, virtual business incubators were developed aiming at supporting new businesses, while the incubated companies were hosted online. The obvious advantages of virtual incubators are their low operational cost and their high flexibility, in comparison to physical incubators. Today, many different types of business incubators can be found online, aiming on startups in a wide range of business sectors, regions, etc.

Many scientists have tried to categorize incubators. Von Zedtwitz (2003) suggests the following categories:

- Regional business incubators;
- University incubators;
- Independent commercial incubators;
- Company-internal incubators;
- Virtual incubators.

This categorization is based on the competitive scope of an incubator (industry, geography and segment), its strategic objective, according to its profit orientation (profit and not-for-profit incubators), and its service package. With respect to virtual incubators, von Zedtwitz (2003) notes that they "*offer no physical workspace or office support. Instead, they offer online access to a network of entrepreneurs,*

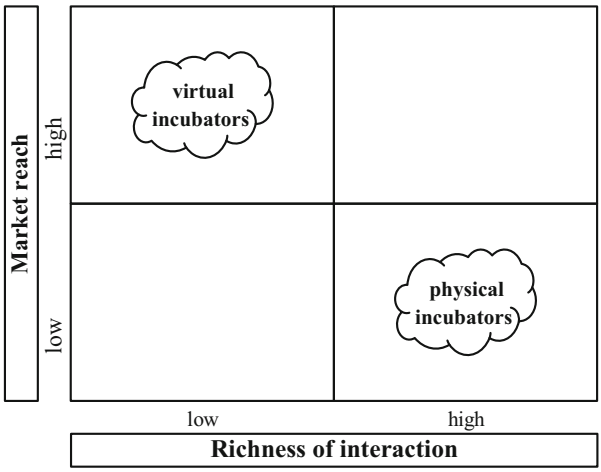


Fig. 2 Differences between virtual and physical incubators in terms of market reach and interaction richness

investors and advisors, as well as support to help match other entrepreneurial needs to professional advice" (see Fig. 2, for differences between virtual and physical incubators).

Von Zedtwitz (2003) also notices that "*virtual incubators do not offer the positive effects of local synergy between similar startup companies obtained through face-to-face networking and problem-solution sharing. Also, startups do not have a running start to their business life, with secretarial or infrastructure support. However, virtual incubators are able to offer a greater advisory network to their incubatees, better matching supply and demand of management and technical talent. This is often left to the initiative of the entrepreneur—the incubator merely provides the platform and the network*". In addition, he observes two functions of virtual incubators: online matchmaking and service aggregation. The online matchmaker provides a communication and news platform for entrepreneurs and startups, and organizes conferences and seminars. Matchmakers also design online learning groups around special interests, which include advice seeking from professionals as well as experience sharing among peers. As Von Zedtwitz (2003) pinpoints, "*although there was little hands-on coaching from the incubator management, there was certainly a lot of exchange of advice and best practices within the startup community*".

The main three incubator model components, according to which different incubators may be distinguished, include (Bergek & Norrman, 2008):

- selection;
- business support; and
- mediation.

With respect to selection, a four-field matrix consisting of the following strategies is used: picking-the-winners and idea, picking-the-winners and entrepreneur, survival-of-the-fittest and idea and survival-of-the-fittest and entrepreneur. Also,

business support strategies may be positioned on a scale from strong intervention to laissez-faire.

The following five services provided by incubators are considered crucial (von Zedtwitz, 2003):

1. *Access to physical resources*: offices, furniture, computer network and other physical infrastructure.
2. *Office support*: secretarial and reception services, mail handling, fax and copying services, computer network support, book-keeping, etc.
3. *Access to financial resources*: venture capitalists, business angels, local institutions and companies, other private funds, etc.
4. *Entrepreneurial startup support*: support for defining a business plan, providing accounting and legal advice for incorporation and taxation issues, formulating ownership etc.
5. *Access to networks*: incubators have a great network after years of operation and they can help new entrepreneurs to find customers, professionals, an interested venture capitalist and others (see also Hansen, Chesbrough, Nohria, and Sull (2000), for the importance of networks in incubation success).

The actual service mix of each incubator depends on the focus of the incubator, as well as the needs and preferences of the entrepreneur (von Zedtwitz, 2003). However, an incubator offering all five of the above mentioned services is an incubator in the strong sense of the term. Organizations that offer only four services are considered incubators in the weak sense of the term, while organizations that offer fewer than four of these services should be considered to be in the domain of the accelerators, technology-transfer offices or entrepreneur-in-residence programs of consulting and accounting firms.

Apart from the provided services, some basic operational characteristics of incubators are the followings (Bone, Allen, & Alley, 2017; Molnar, Grimes, & Edelstein, 1997):

- Open-ended duration of incubation program for the incubatee company (exit usually based on the stage of the company, rather than a specific time frame);
- Typically rent/fee-based;
- Admissions on ad-hoc basis (not cohort-based);
- Selective admission (but typically less strict than the procedure followed by accelerators);
- Often provide technical facilities such as laboratory equipment (not true for virtual incubators);
- Focus on physical space over services (not true for virtual incubators).

Most importantly, graduating startups, which had grown up under the "protection" of an incubator, have much more chances to survive and grow than startups in a similar stage, which have been strangling independently. Actually, the survival rate of startups coming from an incubator has been estimated to be about 80–95% (the higher percentage is for the for-profit incubators), while for other startups the survival rate is about 20% (Molnar, Grimes, & Edelstein, 1997; see also Aernoudt,

2004, for similar results regarding startups graduated from USA incubators). In addition, incubation may be particularly helpful in less-developed economies, where incubators can help bridging knowledge, digital, socio-political and even cultural divides and help increase the availability, awareness, accessibility and affordability of financial, human, intellectual, and even social capital, the key ingredients of entrepreneurial success (Carayannis & von Zedtwitz, 2005). However, incubators should be run like a business and the future of incubation rests on the readiness of sponsoring organizations (either for-profit or not-for-profit) to support their incubation vehicles in the long term (von Zedtwitz, 2003).

Keeping von Zedtwitz's (2003) remark about the number of services an incubator should provide, there are many organizations on the web offering only one or two of the aforementioned services, such as funding, mentoring and networking. These organizations are usually characterized as accelerators or innovation/entrepreneurship centers, and are discussed in the following section.

2.2 *Business Accelerators and Innovation/Entrepreneurship Centers*

Business accelerators appeared in 2005 in the USA and since then a variety of business accelerator programs have emerged as a new tool to support startups (Miller & Bound, 2011). Especially after 2011, hundreds of accelerators have been developed all around the world and if the growth rate is maintained, accelerators will be more than incubators soon. Business accelerators are short-duration programs that help new ventures in early stages of development by providing support services such as office spaces, coaching and mentoring, small amount of financial support and set of education programs. One of the most crucial elements of business accelerators is the mentorship provided by high quality mentors. Accelerators also provide a networking opportunity that consists of different events such as "Demo Days", designed to connect startups with investors (Miller & Bound, 2011).

Based on the above, it seems that accelerators and incubators have a lot in common, but they also have some important differences. Generally, accelerators assist startups that are more mature and ready for external financing (see also Fig. 3).

Some of the most typical characteristics of accelerators are the following (Miller & Bound, 2011):

- Fixed duration programme (usually between 3 and 12 months);
- Often provide seed funding;
- Receiving equity in exchange for the provision of funding to the startups (rather than fees);
- Focus on services over physical space (e.g. mentorship, entrepreneurial training);
- Admission in cohorts;
- Highly selective selection process.

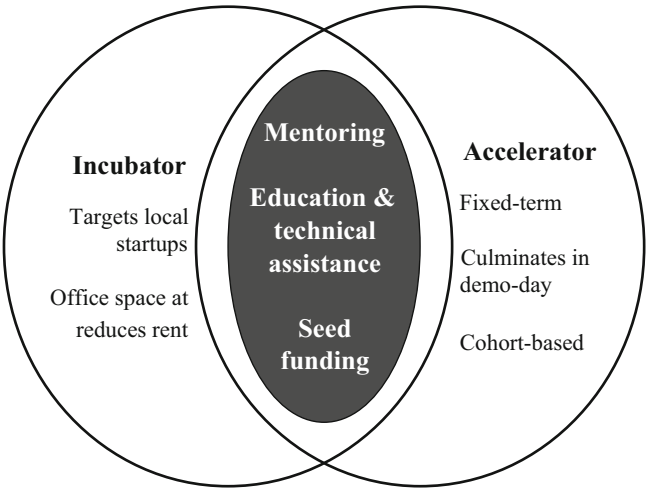


Fig. 3 Overlapping features between incubators and accelerators (adopted by Dempwolf, Auer, and D'Ippolito (2014))

The critical success factors of business accelerators, in the same way as for business incubators, are the selection process and criteria, the business support services and networks provided (Clarysse & Yusubova, 2014).

Concluding, many accelerators provide some similar to incubators services, like mentorship, training, connections, but typically they focus on providing seed funding to startups, usually by taking equity of them. The selection of the startups is usually a very competitive procedure, while the supporting period is generally short and it rarely exceeds 1-year time.

On the other hand, innovation and/or entrepreneurship centers are often connected to a university or a particular firm and provide special training programmes to potential entrepreneurs/university students. They sometimes provide services to SMEs concerning skills development, professional training and industry partnerships establishment, networking, mentoring etc. Some of them organize "Business plans competitions", while others support their members throughout the patent protection procedure.

2.3 Overview of Online Business Services

Many of the agents mentioned in the previous sections have a presence on the web through business platforms and websites, where they present or even provide their services to the interested entrepreneurs.

Online offered business services typically include, but are not limited to (INNOVAGRO, 2019):

- Mentorship
  - Business planning
  - Marketing assistant
  - Legal and accounting advice
- Training
- Funding
- Patent application support, intellectual property rights (IPR) protection
- Access to business resources (articles, business tools, business calculators, templates, etc.)
- Networking
- Business tools

The following section reviews existing online platforms and websites of agencies providing business support to agro-food SMEs and beyond.

3 Review of Existing Online Business Platforms and Websites of Business Development Support Agencies

Despite the need of the agro-food sector SMEs for business development support, an extensive internet-based search revealed a limited presence of platforms and websites of business development support agencies relevant to the agriculture and related sectors (INNOVAGRO, 2019). Their main characteristics are summarized in Table 1. Most of them, like AGROPOLIS in Belgium, Agricultural Incubator Foundation (AIF) in Ohio and AgrOnov in France, are based on physical incubators and offer, beyond any online support and tools, plots, offices, experimental equipment and other services to startups. On the other hand, virtual incubators/accelerators focusing on the agricultural sector exist, like Agro Innovation Lab, which, however, offer limited free material online.

The aforementioned internet-based search revealed also that there are several other online platforms for business support, such as Bplans, SME Toolkits and Re-rise, which do not focus on a particular sector. Table 2 summarizes the main characteristics of some representative such platforms (see INNOVAGRO, 2019, for the complete record of the review results), which, unlike those of Table 1, may be exploited by any business in need.

The outcome of the extensive research on e-business platforms summarized in Tables 1 and 2, showed that, despite the need and importance, there are not many online platforms specialized in the agro-food sector.

This finding may be justified by the perceived low innovation capacity of the agribusiness sector, the very small size of business, or the shallow entrepreneurship culture. This finding may however reveal a hidden potential for the agribusiness sector that can be elaborated through the development of business platforms and tools tailored to the special needs of this particular sector. Thus, deepening the search

Table 1 Entrepreneurship support e-platforms and websites of agriculture-related agencies

Name	Location	Services	Web address
Agricultural Incubator Foundation (AIF)	Ohio, USA	Business development services Physical spaces Networking Training	https://www.agincubator.org/
Agro innovation lab	Vienna, Austria and Munich, Germany	Conventional training Funding Mentoring	https://www.agroinnovationlab.com/
AgrOnov	Bretenière, France	Physical spaces Networking Funding	https://agronov.com/
AGROPOLIS	Kinrooi, Belgium	Plots Offices Business units Support and guidance Networking	https://www.agropolis-kinrooi.be/en/content/incubator
F2F Genetics Network	Virtual (concerns USA)	Information Networking	https://www.farmersbusinessnetwork.com/
MarketMaker	Virtual (concerns USA)	Networking	https://national.foodmarketmaker.com/
StartLife	Wageningen Campus, The Netherlands	Incubation program including virtual place, business plan development tools, access to facilities and experts, network, contacts, funding, mentoring, etc.	https://start-life.nl/
WellFOOD	Virtual (concerns Adriatic countries)	Information Networking	http://www.wellfoodhub.eu/

of the existing online platforms, the following section presents and discusses specific online tools that can be used to address the needs of agro-food SMEs along the axes of information, learning, networking and use of business tools, which have been identified as crucial for their development and growth.

4 Online Business Support Development Tools

4.1 Information

Information may be provided online in the form of e-library modules. An e-library module may be extremely broad, it typically, however, includes material such as (INNOVAGRO, 2019):

Table 2 Entrepreneurship support e-platforms of general interest

Name	Services	Web address
Bplans	E-library Information Business tools	https://articles.bplans.com/
Business Queensland	Information Business tools Networking	https://www.business.qld.gov.au/
Free management library	E-library	https://managementhelp.org/businessplanning/
morebusiness.com	Information Business tools Training	https://www.morebusiness.com/
RE-rise	Incubation program including mentoring, information, training, networking, workspace, business tools	https://www.rerise.in/
SCORE	E-library Mentoring Training Networking	https://www.score.org/
SPARK	Training Mentoring	https://www.spark-online.org/
Startups.co	E-learning Information Mentoring Training Business tools Funding	https://www.startups.com/

- Guides
- Templates
- Case studies/Success Stories/Good practices
- Checklists
- Articles
- Videos

As far as the guides are concerned, several types have been found online, such as a guide on "how to start a farming business" or "how to get your business online" or even a guide for conducting a business or a marketing plan. Such guides are offered as online materials or downloadable files, containing general information concerning specific business topics.

Usually, guides regarding starting online farming businesses provide information about the following issues (see for example https://www.score.org/resource/how-get-your-business-online):

Fig. 4 Example of marketing plan template (Source: https://www.morebusiness.com/marketing-plan-example/)

- *Registration of a domain name*: The domain name plays a critical role, since it is the main identity of a business. It can be used in several cases, including a website address, a company-branded email or connecting company's web address to its social media page. The domain name may describe the product and services offered to customers or highlight a specialty niche or attribute, support online marketing campaigns and protect the brand.
- *Setting up a company-branded email*: When registering a domain name, providers give the option to create a business email. A professional email address is also a significant part of a business identity, given its role in communication with customers.
- *Creating online space*: A website may maximize the benefits of online businesses, since it offers more control, credibility and flexibility in marketing activities. It can also be linked to the social media page of the business, usually, however, customers prefer to get information from a business's website rather than its social media page, and most consumers believe a website makes business appear more credible.

Several templates are also available online for conducting a business and/or a marketing plan, while financial projections templates may be also found (see Fig. 4).

Generally, templates have more specific information than guides and are usually downloadable so that the user can edit them offline. Business and marketing planning templates are typically in the form of editable documents, while spreadsheets are provided for financial projections. However, special tools have also been developed, which allow the user to develop his/her own business plan (including financial projections) by using a simple step-by-step guide. The final output seems to be a well formatted portable document format (pdf) file, with charts and tables. Nevertheless, even though there are many templates online, business report templates specialized in agro-businesses are rather rare.

Regarding online business plans or business plan templates, it should be noted that they are usually offered as a guide only, since it is necessary, in many cases, to modify the suggested structure. A business plan is a document that details how the important aspects of the business are handled, covering key areas like market research, marketing planning, legal and risk management planning, operational planning, human resources planning, and financial planning.

A business plan usually consists of the following main parts (see for example https://www.publications.qld.gov.au/dataset/business-plan-template):

1. *Executive summary* (brief information about business vision, legal structure, offered products and services, customers, competitors, market, operation, financial projections, owners, etc.)
2. *Product/service and market analysis* (key features, benefits and limitations of products/services, cost and sale price, substitutes, description of the market in local/regional/national level, seasonal influences, price ranges, etc.)
3. *Marketing plan* (customer profile regarding age, gender, location, income, education, buying patterns, and motivation, pricing, after sales service and warranties, unique selling proposition, promotional activities, online marketing strategy, sales analysis and forecast, etc.)
4. *Legal matters and risk management*: (legal structure, contracts, licenses and agreements, intellectual property, insurance, employment-related legislation and obligations, risk management, legal and ethical trading, etc.)
5. *Operating plan* (location requirements such as space, type of building, zoning, access, parking, production processes regarding costs, documentation quality control, new product development, benchmarking, customer management, plant and equipment, supplies, inventory control, etc.)
6. *Human resources plan* (organizational structure, skills and experience of key personnel, job descriptions and specifications, performance assessment processes, training, recruitment, human resource policies, workforce planning, etc.)
7. *Financial plan* (start-up or expansion costs, profit and loss forecast, cash flow forecast, balance sheet, financial ratios, etc.)
8. *Action plan* (coordination and time frame of actions)

In addition to guides and templates, business plan samples, case studies and good practices can be found online. For example Bplans.com (https://www.bplans.com/farm_and_food_production_business_plan_templates.php) provides a list of sample business plans for farms, food growers, food production facilities, and other

Fig. 5 Checklist example (Source: http://www.humanmetrics.com/entrepreneur/quiz)

agriculture-related businesses, including fruit farms, brewery, coffee export, feed and farm supply, vegetables, garden nursery, hydroponics farms, pasta manufacturing, juices wholesale food manufacturing.

Success stories are also available, where specific businesses are presented in articles, videos, etc. Some of them are only presented online, others can be downloaded, while pictures, tables, charts, and videos are used to make the presentation more attractive.

Moreover, several online checklists are available. Checklists usually concern the assessment of a new/potential entrepreneur/project, a business startup awareness evaluation or an entrepreneurs' skills assessment (see Fig. 5). They are either in document format or presented as online questionnaires.

Last not least, information can be provided in the form of articles and videos. A vast number of articles on different business topics is readily available online. Some of them are categorized by topic, while in other cases a search engine helps the visitor to spot the articles he/she is more interested in, by using filters (see Fig. 6). Several videos are also available online mainly for presenting business cases, while others have a motivational aim, as they try to promote entrepreneurship and inspire potential entrepreneurs.

Summarizing, an e-library section of a business support platform may include several information in the form of articles, links, guides, templates, case studies, success stories, good practices, checklists/quizzes or even videos. Studies and reports may also be included, if available. All of them can be presented in several ways (online, downloadable files, files connecting with tools/spreadsheets etc.).

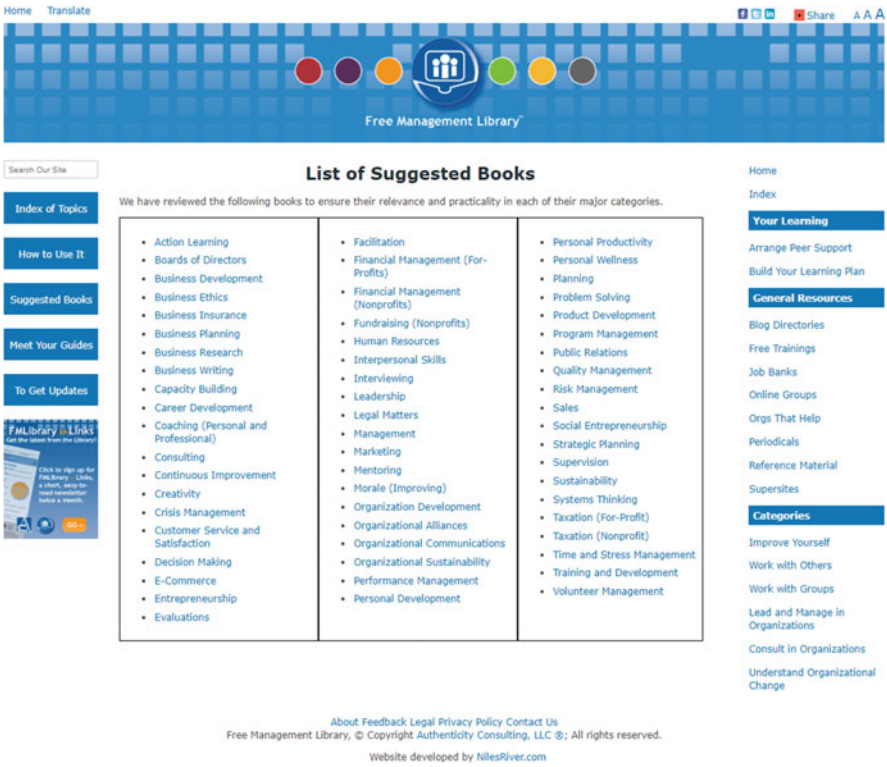


Fig. 6 Online library example (Source: https://managementhelp.org/books/list-of-books.htm)

Moreover, in order to help users, some platforms have search engines, while others try to categorize available material. Also, different platforms have different ways of enriching their resources. In some cases, the resources are "open" to all users; in others a registration is prerequisite, while in some others a fee must be paid.

4.2 Learning

Compared to information, learning is much narrower. E-learning tools usually include interactive videos (e.g., for financial management issues) or webinars in many different topics (e.g., how to start a new business) (INNOVAGRO, 2019).

Webinars refer to seminars available online. Webinars with business and agricultural topics may be found on several websites. Some of them are offered live, so visitors can participate and impose questions to the presenter, while others are recorded, so visitors can only access the educational material that has already been produced. Attending a live webinar, usually requires subscription.

For example, the Food and Agricultural Organization (FAO) of the United Nations provide in its e-agriculture platform a series of webinars (http://www.fao.org/e-agriculture/), including topics on spur agricultural technological innovation, drones for agriculture, satellite data and machine learning, eco-farming, airborne agriculture analytics, data visualizations, open data, remote controls, water data collection, etc.

FAO's e-agriculture platform may also be considered as a global community of practice that facilitates dialogue, information exchange and sharing of ideas related to the use of information and communication technologies (ICTs) for sustainable agriculture and rural development.

4.3 Networking

Online tools for networking are the least developed, possibly because several questions need to be answered prior their design, such as:

- What types of businesses/sectors will be included?
- How will the companies be presented? Map/list/other?
- How will the presentation of a particular company be? What kind of information will be included?
- Is it going to include a search engine? What kind of filters will be available?
- How is the database going to be enriched/expanded?
- Will it be an open resource database or will a registration be required?

As a consequence, there are some networking platforms, but most of them are not free, as they have been developed from specific private companies, which make them available only to registered users (INNOVAGRO, 2019). There is, of course, the yellow pages, a well-known free networking platform, but the information provided therein is limited (address, telephones, e-mails, website, map location, working hours, etc.).

Summarizing, as far as online platforms offering B2B networking information are concerned, only few cases can be found and even less are specialized in agribusiness. For example, AgroNetwork (http://www.agronetwork.com/global/) provides networking opportunities for agriculture business. A company may be registered in this network, in order to develop and customize its own webpage (with info about products and activities), search useful agriculture related sites worldwide, share ideas with other members in forum, find agribusiness partners and products, and have access to news, weather, etc. This networking platform provides a large list of companies (almost 800 entries) operating globally in the agriculture sector. These companies are specialized in particular products or services (fruits, grain, nuts, agricultural seeds, foods, beers, beverages, coffee, agricultural chemicals and fertilizers, agricultural machinery, etc.), having different roles in the agro-food chain (e.g., import, export, manufacture, distribution). An important feature of AgroNetwork is the sell and buy market, where users can find and pose offers for

agricultural related products. These offers are presented based on standard information (e.g., product title, product detail description, price for offer, terms of validity, product image).

4.4 *Use of Business Tools*

Several different types of e-tools for SMEs are available online (see examples in Figs. 7 and 8), covering many different topics, like accounting, human resources management, social-media management, e-mail marketing, online survey tools, etc. Such tools enable in an easy manner the self-assessment of the companies with respect to their innovation and extroversion potential, the preparation of business plans, the financial management, etc. Some of these tools are totally free (at least under specific limitations), while others are available after paying a fee.

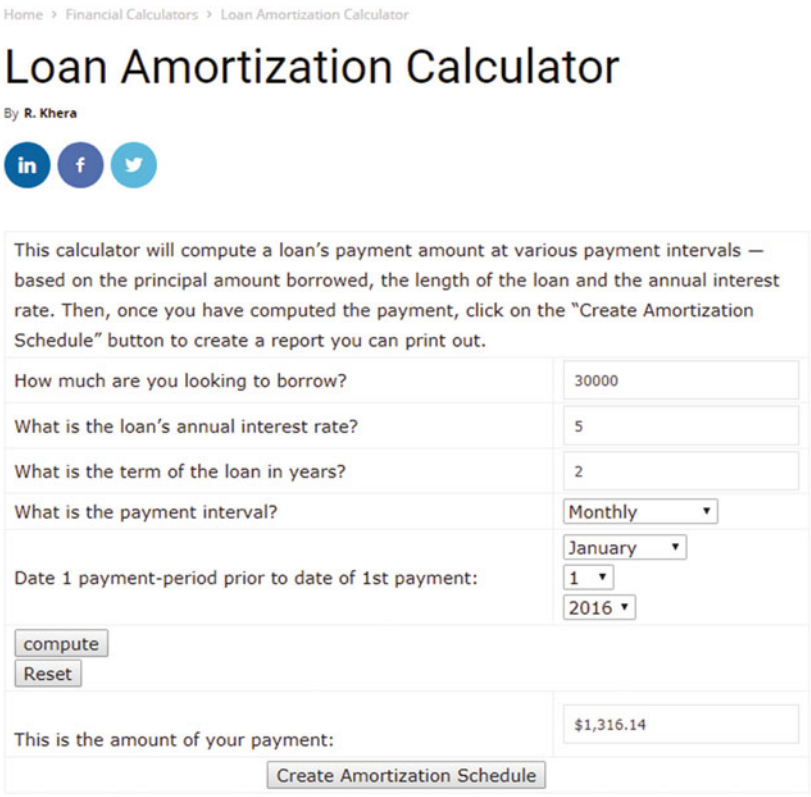


Fig. 7 Loan amortization calculator example (Source: https://www.morebusiness.com/amortization-calculator)

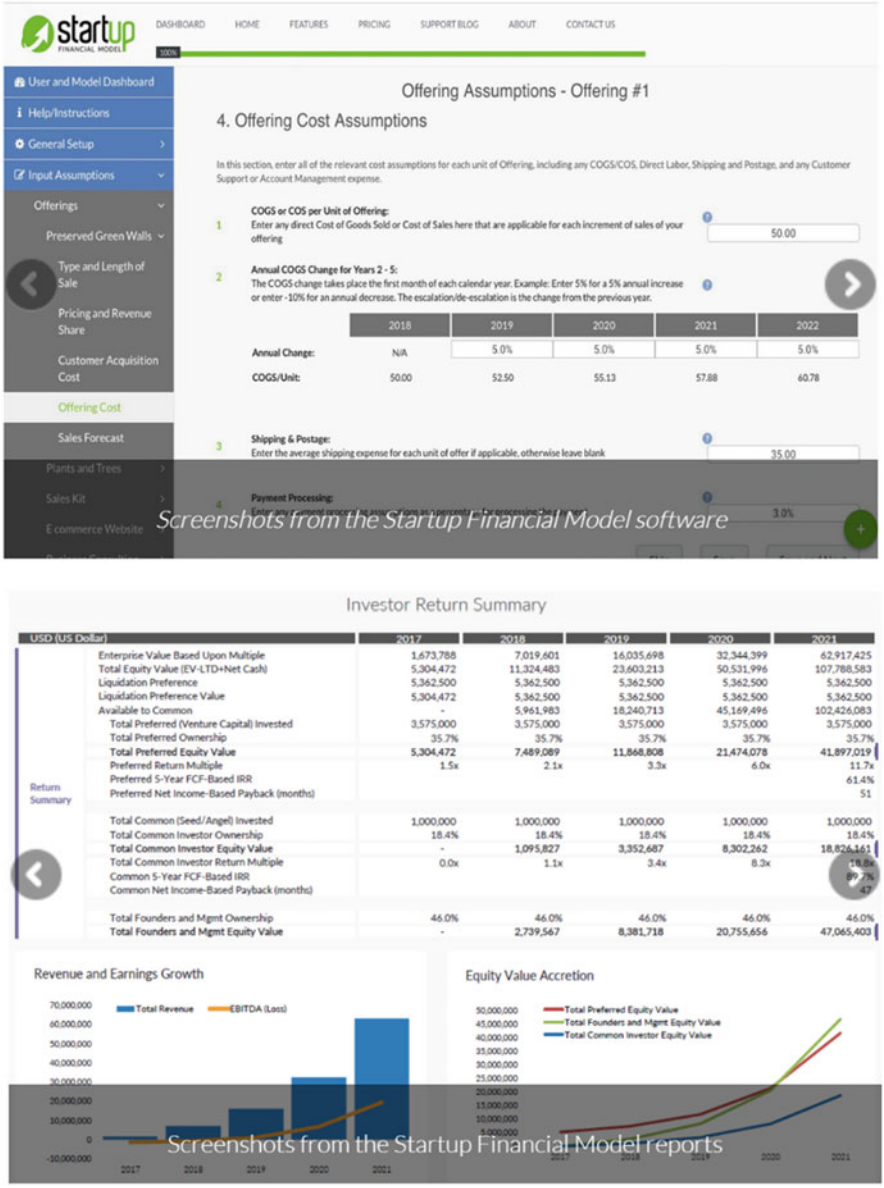


Fig. 8 Financial model application example (Source: https://www.startupfinancialmodel.com/)

For example Bplans, a platform owned and operated by Palo Alto Software, Inc., provides a large set of business tools for new and existing entrepreneurs. An indicative list of Bplans' business calculators includes the following (https://www.bplans.com/business-calculators):

- *Cash flow calculator*: it shows how business-to-business sales, carrying inventory, and rapid growth can absorb a business' money.
- *Discounted cash flow*: it is similar to the previous tool, but it takes into account the time value of money.
- *Starting costs calculator*: it estimates the cost to start a new business.
- *Break-even calculator*: it shows how much revenue is needed in order to cover both fixed and variable costs.
- *Conversion rate calculator*: it shows the impact on total online sales in case of increasing the website conversion rate.
- *Investment offering calculator*: it estimates the risk and the potential gain of an investor based on the calculation of the Net Present Value and the Internal Rate of Return.
- *Email marketing ROI calculator*: it calculates the Return on Investment (ROI) for email campaigns, based on expected results (e.g., response rate, conversion rate, average buyer purchase) and marketing campaign data (e.g., audience size, campaign costs).
- *Direct mail and pay-per-click ROI calculators*: using similar data, it calculates the ROI for direct mail campaigns or pay-per-click advertising campaigns.

Focusing on new businesses, e-tools that can support business development and growth include (INNOVAGRO, 2019):

- Business calculators (e.g. for cash flows, starting costs, conversions, etc.);
- Quizzes for entrepreneurs; and
- Logo creation tools.

Finally, there are several other multi-purpose online tools that may be useful also for agro-food companies. Some of them, such as the following, are very popular and free:

- *Google Forms*: for designing an online questionnaire and conducting a survey, including the analysis of survey results (https://www.google.com/forms/about/).
- *Dropbox*: for sharing files with others or between different personal computers (https://www.dropbox.com/).
- *Mailchimp*: tools for marketing (e-mail campaigns, ads, signup forms, etc.) provided by The Rocket Science Group LLC (https://mailchimp.com/); the service is free for up to 2000 contacts or 12,000 emails per month.
- *Hactchful*: for designing a logo in just a few minutes; the tool is powered by Shopify Inc. and it is free for all users (https://hatchful.shopify.com/).
- *TeamViewer*: for remotely connecting to desktop computer from anywhere and at any time. The service is free and provided by TeamViewer GmbH (https://www.teamviewer.com/).
- *Skype*: for communicating with other people using a computer or a mobile and also for taking place in a videoconference. Skype is a provided by Microsoft (https://www.skype.com/).

Mailchimp is a characteristic example of an integrated online marketing platform for small businesses (https://mailchimp.com). It provides CRM (Customer Relationship Management) tools that can integrate and synchronize customer data. It can also help business organizations in building their own marketing campaign (e.g., customer segmentation, prediction of demographics, personalization of promotional emails). Brand building tools are also an important part of Mailchimp, providing custom domains for website and landing pages, as well as editing tools and templates for designing promotional material and websites. Moreover, the platform offers several online tools to build, launch, and measure campaigns across different channels (emails, social media, landing pages, digital ads, postcards, etc.).

Other online business tools may concern to a specific part of marketing activities, such as the creation of a business logo. For example, Shopify Logo Creator (https://hatchful.shopify.com/) provides free services to potential users, regardless if someone is a Shopify merchant or not.

5 Major Findings and Conclusions

Due to their limited access to resources, experts, and integrated business tools, the SMEs, which comprise the major part of the agro-food sector, struggle to start and develop their business in the highly competitive environment of our days. A recent, interview-based survey suggested a need for support in four axes: (1) information with respect to case studies and success stories, good practices, new/existing agro-food clusters and clusters policy, funding opportunities, relevant cooperation networks, exports/imports, regulations and laws, etc.; (2) learning of financial and management issues or other topics; (3) networking with exporters and importers, educational organizations and research centers, etc.; (4) use of business tools facilitating activities such as self-assessment, preparation of business plans, financial management, etc. Given their typically limited resources, the best way to get support with respect to the aforementioned needs is through the provision of online tools and services. It has been thus the aim of this chapter to search and provide a structured review of readily available online business tools focusing on entrepreneurship support and growth.

The review revealed that there is a number of online platforms offering services and tools along the aforementioned four axes of interest, but no platform has been identified to fully cover all the needs, being also specialized in the agro-food sector.

A business development support platform should include an information, a learning, a networking and a business tools section. For the information section an e-library may be used including information in the form of articles, links, guides, templates, case studies, success stories, good practices, checklists/quizzes or even videos. Studies and reports may also be included, if they are available. All of them can be presented in several ways: online, downloadable files, files connecting with tools/spreadsheets etc. The learning section may include training videos (e.g. presentations with or without voice) and/or webinars (recorded or live), while

the networking can include information about exporters, importers and generally information about agro-food businesses, trying to provide easy business linkages. Finally, regarding business tools, several tools can be considered for inclusion that should enable in an easy manner the self-assessment of the companies with respect to their innovation and extroversion potential, the preparation of business plans, the financial management, etc. All sections should be provided under the umbrella of a user-friendly interface that will allow and assist easy access.

The lack of an online platform that covers all the aforementioned needs focusing on the agro-food sector reveals a hidden potential for the sector, which can be elaborated through the development of online business platforms and tools tailored to its special requirements and needs.

Acknowledgments The research leading to these results took place within the frame of the Interreg V-B Adriatic-Ionian Cooperation Programme, project INNOVAGRO (ADRION-613). This document has been produced with the financial assistance of the European Union. The content of the document is the sole responsibility of "Technical University of Crete" and can under no circumstances be regarded as reflecting the position of the European Union and/or ADRION programme authorities.

References

Aernoudt, R. (2004). Incubators: Tool for entrepreneurship? *Small Business Economics, 23*(2), 127–135.

Bergek, A., & Norrman, C. (2008). Incubator best practice: A framework. *Technovation, 28*(1/2), 20–28.

Bone, J., Allen, O., & Alley, C. (2017). *Business incubators and accelerators: The national picture*. BEIS Research Paper 7, UK Department for Business, Energy, and Industrial Strategy, London.

Carayannis, E. G., & von Zedtwitz, M. (2005). Architecting gloCal (global–local), real-virtual incubator networks (G-RVINs) as catalysts and accelerators of entrepreneurship in transitioning and developing economies: Lessons learned and best practices from current development and business incubation practices. *Technovation, 25*(2), 95–110.

Clarysse, B., & Yusubova, A. (2014). Success factors of business accelerators. In *Proceedings of Technology Business Incubation Mechanisms and Sustainable Regional Development*. Accessed September 11, 2019, from https://biblio.ugent.be/publication/6842877/file/6843687.pdf

Dempwolf, C. S., Auer, J., & D'Ippolito, M. (2014). *Innovation accelerators: Defining characteristics among startup assistance organizations*. Washington, DC: Small Business Administration.

García-Álvarez de Perea, J., Ramírez-García, C., & Del Cubo-Molina, A. (2019). Internationalization business models and patterns of SMEs and MNEs: A qualitative multi-case study in the agrifood sector. *Sustainability, 11*(10), 2755.

Hansen, M., Chesbrough, H., Nohria, N., & Sull, D. (2000). Networked incubators: Hothouses of the new economy. *Harvard Business Review, 78*(5), 75–83.

INNOVAGRO. (2019). *Collection of primary and secondary data*. Deliverable T2.1.1 of the Interreg V-B Adriatic-Ionian Cooperation Programme, project "Development of an innovative network for the promotion of extroversion of agro-food companies in Adriatic – Ionian Area (INNOVAGRO)" (ADRION-613). Accessed September 11, 2019, from https://innovagro.adrioninterreg.eu/wp-content/uploads/2019/07/Collection-of-primary-and-secondary-data_Final.pdf

Miller, P., & Bound, K. (2011). *The startup factories: The rise of accelerator programmes to support new technology ventures*. London: Nesta Innovation Foundation.
Molnar, L., Grimes, D., & Edelstein, J. (1997). *Business incubation works*. Ohio, OH: NBIA Publications.
UNECE. (2000). *Best practice in business incubation*. Geneva: Economic Commission for Europe, United Nations.
von Zedtwitz, M. (2003). Classification and management of incubators: Aligning strategic objectives and competitive scope for new business facilitation. *International Journal of Entrepreneurship and Innovation Management, 3*(1/2), 176–196.

Design of a Sentiment Lexicon for the Greek Food and Beverage Sector

Anastasios Liapakis, Theodore Tsiligiridis, and Constantine Yialouris

Abstract Sentiment Analysis is a computational method aiming to extract opinions/evaluations of individuals for an entity such as a product, a service etc. In social media networks and other online sources, as for example the food websites, sentiment analysis is able to identify all possible terms, such as simple words, combinations of words, or phrases (pre-processing stage) that can be used to express the feelings of a user for a specific entity. Then, by considering the characteristics of these media, such as time sensitivity, text size limitation and unstructured expressions, converts them giving a positive or negative significance. For the analysis, a set of linguistic, statistical and machine learning techniques are usually considered to structure the information contained in text sources. The main purpose of this paper is twofold. First is to provide a literature review of the sentiment analysis techniques and second is to design a sentiment lexicon as a preliminary step to further analyze the customers' reviews of some leading companies in the Greek Food and Beverage industry as they uploaded in the most common Opinion Social Networks in Greece (fb). Existing research has focused mainly on the recognition on English characters, while to our knowledge, limited research papers have been published so far concerning the Greek language, concentrating mainly on the banking and financial sector, neglecting contributions on food industry. Note that since significant portion of online text-based Greek communications ignore the rules of spelling and grammar the study takes into account this trend and improves the calculation of a sentiment score accordingly. As appears, the findings are expected to contribute in the design issues of a sentiment lexicon particularly devoted to the Greek food and beverage industry and to be used for further analysis.

Keywords Sentiment analysis · Modern Greek · Food and beverage sector

A. Liapakis (✉) · T. Tsiligiridis · C. Yialouris
Informatics Laboratory, Department of Agricultural Economics and Rural Development, School of Applied Economics and Social Sciences, Agricultural University of Athens, Athens, Greece
e-mail: liapakisanastasios@aua.gr; tsili@aua.gr; yialouris@aua.gr

E. Krassadaki et al. (eds.), *Operational Research in Agriculture and Tourism*, Cooperative Management, https://doi.org/10.1007/978-3-030-38766-2_3

1 Introduction

Social Media networks are usually changed consumers' behavior. Nowadays, more and more companies use social media marketing in order to attract more customers. This modifies consumers' attitudes and companies cannot detect these modifications due to the big volume and the diversity of the produced information. According to surveys (comScore/the Kelsey group, 2007; Horrigan, 2008), 81% of the Internet users in USA have done online research on a product at least once, the vast majority of questioned are willing to pay from 20 to 99% more for a five-star-rated item or a service than a low-star-rated item and interestingly, Rainie and Hiltin (2004), report that "Individuals who have rated something online are also more skeptical of the information that is available on the Web". Other surveys have shown that the majority of Internet users, usually do researches on opinion networks for products that they are willing to buy and claim that reviews influence their purchase decision. Obviously, the variety of social media networks creates multimedia data as text, audio, images and videos files which are difficult to be edited or sorted by the companies or other users.

Facebook is the most popular social media platform all over the world and in Greece too. Companies are using it for reading what other users say about their products or services, responding to users' messages and distributing content about their product or services. Table 1, shows according the Eurostat, the most common social media (Hu & Liu, 2012) tools that are used from the Greek companies and individuals too.

In the case of Greek Food and Beverage (F&B) industry, the use of social media networks is very high to multinational and large companies. This is due to the fact that this sector, is especially prone to problems in sustainability given its high impact and dependence on natural, human and physical resources (Liapakis, Costopoulou, Tsiligiridis, & Sideridis, 2017). The challenges that face are numerous, including environmental sustainability (usage of natural resources, animal welfare etc.), social sustainability (labor and work conditions, food safety etc.) and economic sustainability (energy usage, waste management etc.) (Genier, Stamp, & Pfitzer, 2009). Every day, various campaigns regarding the previous aspects are shown in social media networks by the companies in order to eliminate the attacks concerning the good reputation of their brand names. In this framework the quality of services, marketing and maximization of sales will be enforced by considering the textual content that is generated by Internet users. Sentiment analysis (subjective analysis, opinion mining, and appraisal extraction with some connections to affective

Table 1 Categorization of the most common social media tools in Greece

Category	Platforms
Blogging	Blogger
Micro-blogging	Twitter
Opinion and reviews	Trip advisor, efood
Media sharing	YouTube, Instagram
Social networking	Facebook, LinkedIn

computing are also alternative names) can approach methodologically the problem by identifying relevant information from the huge communication of stakeholders over the Internet. The data created has resulted in the development of web opinion mining, as a concept in web intelligence focusing on extracting, analyzing and combining web data about user thoughts. The sentiment analysis is significant since the users' opinions provide information of how people feel about a topic of interest and understand how this was acknowledged by the market.

Sentiment analysis can be seen as a classification process divided into three main levels; document-level, sentence-level, and aspect-level. Document-level aims to classify a document by expressing a positive or negative opinion (sentiment). It assumes that the whole document is a basic information unit and it refers about one topic only. Sentence-level classifies opinion expressed in each sentence. Only subjective sentences need to be considered and, in such cases, sentence-level will determine whether the sentence expresses positive or negative opinions. Since sentences are short documents there is no substantial difference between document and sentence level classifications. However, to provide detail opinions on all aspects of the entity which is needed in many applications, we have to classify the sentiment with respect to the specific aspects of entities, namely to move to the aspect level sentiment analysis. This level of analysis requires in the first place to identify the entities and their aspects. The sentiment holders may provide various opinions for different aspects of the same entity like, for example: "the food quality of this restaurant is very good, but the service time is long".

Sentiment analysis is, in fact, a sentiment classification problem and therefore it is necessary to extract and select text features. According to Aggarwal and Zhai (2012), some of these features are the terms presence and frequency which are individual words or word n-grams and their frequency counts. It either gives the words binary weighting (zero if the word appears, or one if otherwise) or uses term frequency weights to indicate the relative importance of features (Yelena & Padmini, 2011); the Parts of Speech (POS), tags the words of a sentence as verb, adverb, adjective and noun based on some POS tagging rules. The following example could help us to understand the method precisely. Assuming the sentence: "The place is clean and the food is very good." Then in POS method, the review should be written as: The/***DT***, place/***NN***, is/***VB***, clean/***JJ***, and/***CC***, the/***DT***, food/***NN***, is/***VB***, very/***RB***, good/***JJ*** where, **DT**: Determiner, **NN**: Noun, **VB**: Verb, **JJ**: Adjective, **CC**: Conjunction, **VB**: Verb in present tense, and **RB**: Adverb.

Finally, negations, which are negative words that may change the opinion orientation, like 'not good' an alternative to 'bad'. Feature selection approaches can be lexicon-based that usually start with a small set of seed words, and extend this set through synonym detection, online resources and or by using a large corpus of documents from a single domain in order to obtain a larger lexicon or statistical-based automatic methods that are used more frequently. The feature selection techniques treat the documents either a group of words (Bag of Words (BoWs) or as a string which retains the sequence of words in the document. BoW is used more often because of its simplicity for the classification process. The most common feature selection step is the removal of stop-words and stemming (returning the word

to its stem or root i.e. 'flies' to 'fly'). Another popular term-weighting numerical statistic used in this context is the Term Frequency–Inverse Document Frequency (TF-IDF), which reflects the importance of a word in a document of a collection or corpus. The value of the TF-IDF statistic increases proportionally to the number of times a word appears in the document and is balanced by the number of documents in the corpus that contain the word. This helps to adjust for the fact that some words appear more frequently in general. Frequently used statistical methods in feature selection are the Point-wise Mutual Information (PMI) (Cover & Thomas, 1991), the Chi-square $\chi 2$ (Fan & Chang, 2012) and the Latin Semantic Indexing (LSI) (Deerwester, Dumais, Landauer, Furnas, & Harshman, 1990). Other methods are the Gini index (Yelena & Padmini, 2011), the Hidden Markov Model (HMM) and the Latent Dirichlet Allocation that simultaneously model topics and syntactic structures in a collection of documents (Duric & Song, 2012; Griffiths, Mark, Blei, & Tenenbaum, 2005) as well as an approach to detect irony in customer reviews (Reyes & Rosso, 2012). Some discussion on the methods above can be seen in Medhat, Hassan, and Korashy (2008) and Feldman (2013).

Sentiment classification analysis relies on three types of techniques, i.e., machine learning-based and lexicon-based techniques as shows in Fig. 1.

Machine learning based techniques apply some well-known machine learning algorithms and use linguistic features (Singh, Singh, & Singh, 2016). To describe the corresponding text classification problem, we assume that each record of the training set of records is labeled to a class. Then, for a given instance of an unknown class, a classification model is used to predict a class label for it. Obviously, the model is related to the features in the underlying record to one of the class labels. A label can be assigned to an instance either directly or indirectly, by selecting it from a set of labels in accordance to some predefined probability. The text classification machine

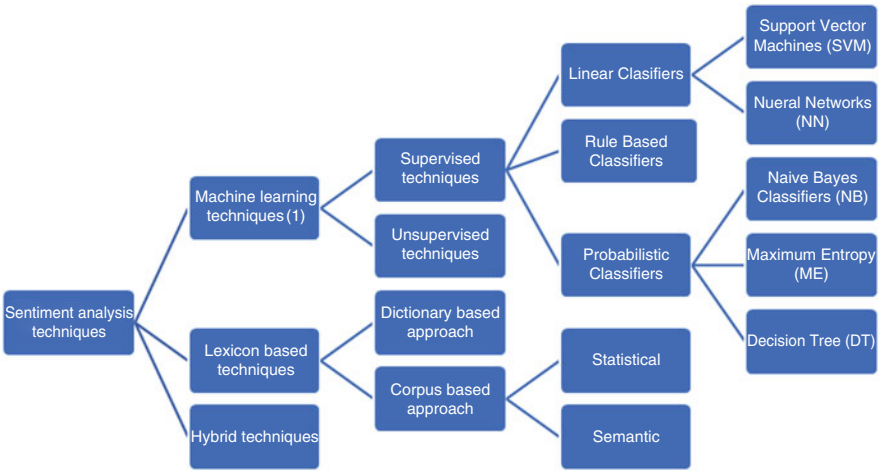


Fig. 1 Types of sentiment analysis classification techniques

learning methods, can be distinguished by the approach used; namely supervised, semi-supervised and unsupervised.

The supervised methods make use of a large number of labeled training documents and include four categories of classifiers; the decision tree, the linear, the rule-based and the probabilistic. Insight on these categories subdivides them further; however, the most interesting cases include the linear classifiers which are split into Support Vector Machines (SVM) and Neural Networks (NN) as well as the probabilistic classifiers into Naive Bayes (NB), Bayesian Network (BN), Decision Tree (DT), and Maximum Entropy (ME).

As for the unsupervised text classification methods, which are used when it is difficult to find the labeled training documents and consequently it is difficult to classify documents into categories, alternative methods are used (Ko & Seo, 2000; Read & Carroll, 2009; Turney, 2002; Xianghua, Guo, Yanyan, & Zhiqiang, 2013). Details for the semi-supervised classification can be seen in He and Zhou (2011).

The lexicon-based approach (2) relies on sentiment or opinion lexicon, a collection of known and precompiled sentiment terms. It is divided into manual, dictionary-based and corpus-based approaches. The first one is quite consuming and it is used only in particular cases. The dictionary-based approach depends on finding opinion seed words and then searches the dictionary of their synonyms and antonyms. The corpus-based approach begins with a seed list of opinion words and then finds other opinion words in a large corpus to help in finding opinion words with context-specific orientations. This could be done by using statistical or semantic methods.

The remaining two use statistical or semantic methods to find sentiment polarity. Corpus-based techniques are based on decision trees such as *k*-Nearest Neighbors (*k*-NN), Conditional Random Field (CRF), Hidden Markov Model (HMM), Single Dimensional Classification (SDC), Sequential Minimal Optimization (SMO), and related to methodologies of sentiment classification. The hybrid approach combines both approaches and is very common with sentiment lexicons playing a key role in the majority of methods. Other approaches to sentiment analysis may also be found in the literature. Some of the most characteristic studies include (Gräbner, Zanker, Fliedl, & Fuchs, 2012) in which sentiment analysis does what every user is required to do after writing a product review: to quantify the opinion represented by the text with percentages or stars; Yu, Zhou, Zhang, and Cao (2017) in which it aims to automatically classify the text of written reviews from customers into positive or negative opinions; and Vinodhini and Chandrasekaran (2012) in which the type of natural language processing for tracking the mood of the public about a particular product or topic.

To conclude, sentiment analysis is contextual mining of text which identifies and extracts information from online reviews, websites, and social media networks. It is applied in many fields such as consumer information, marketing, books and websites (Griffiths et al. 2005). It helps businesses to compute the social sentiment of their brand, products or services while was examining online conversations. To facilitate processing and storage in databases, an opinion/review of a user about an entity can be analyzed as a five-dimensional vector of the form:

Opinion (*Entity Name, Entity Part, Polarity, User, Date*).

In the case of the Food and Beverage industry and for the purposes of this study we suggest the following form:

Opinion (*Name*, *Function*, *Polarity*, *User*, *Date*)

where, *Name* is the restaurant's or cafeteria's name; *Function* could be the image of the store (decoration etc.), the service facilities (location, working hours etc.), the staff (availability, information offered etc.) and the products (quality, variety etc.) taking values *positive* or *negative*; *Polarity* could be *positive* or *negative*; *User* is the user's name; and *Date* is the date that the review was written. For the better understanding, a review for a known Athens restaurant (named: *X*) on Tripadvisor could be used: "The place is clean and the food is very good. We like the atmosphere and the way its look. Very good coffee and breakfast"—Reviewed July 18, 2018 by *Name A*. According to the suggested form and for further analysis, the review could be shown as follows:

Opinion1 = {*X*, image of the store: *positive*, *A*, *July 18, 2018*}
Opinion 1.1 = {*X*, products: *positive*, *A*, *July 18, 2018*}

The structure of the article is as follows: The second section presents a literature review of the most important supervised techniques that are used in sentiment analysis as well as an overview of the related to our study articles. Section 3 presents the methodology used for identifying the most common Greek adjectives in social media networks (written in modern Greek and Greeklish too). The final section ends with a discussion on the findings of this study.

2 Literature Review

2.1 Sentiment Analysis Techniques

As it has been pointed out in sentiment analysis, lexicon-based and machine learning are two techniques that are used in common; in lexicon-based schemes, a dictionary with selected sentiment words is proposed, whereas, in machine learning techniques various methods are used for sentiment classification (supervised and unsupervised algorithms) (Haseena & Tanvir, 2014).

BoW and TF-IDF are two of the most popular feature selection schemes used so far and will be described in the sequel. In addition, for the machine learning approach the interesting is focused on the super-vised classification techniques, consisting of three consecutive phases, namely, the determination of the data set, the selection of the appropriate classifier and finally the feature selection. The selection classifiers to be considered in this section are the *Supporting Vector Machines* (SVM) (Joachims, 2001), the *Naïve Bayes* (NB) (Lewis, 1998), the *Maximum Entropy* (ME) (Pang, Lee, & Vaithyanathan, 2002) and the *K-nearest neighbors' algorithm (*Adeniyi, Wei, & Yongquan, 2016*)*.

The *Bag of Words* (BoW) method (Gabryel, Damaševičius, & Przybyszewski, 2018) is the most known algorithm in SVM. It aims to deal document with linear algebraic operator by transforming a text into sparse numeric vectors. Terms are stored in dictionaries and can be represented by simple words, (1-gram) or composed words (2, 3, . . ., n-gram) that occur in various documents. In n-grams dictionaries the priority of the words composing the term (phrase) is very important as it changes the semantic, e.g. for the 4-gram term "The food was beautiful" or "Was the food beautiful?" Each term is used as an attribute of the data set represented in the attribute-value form. Thus, in BoW model, a term is represented as a separate variable having numeric weight of varying importance. In particular, after preprocessing the set of documents d_i; $i = 1, 2, \ldots, n$, can be represented in a matrix of the form $[u_{ij}]$, with u_{ij}; $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, m$ be a count of the number of occurrences of term t_j in document d_i, where n documents are represented and each document consists of m terms. Each document d_i can be seen as a histogram $d_i = [u_{i1}, u_{i2}, \ldots, u_{im}]$, with u_{ij} referring to the value of the j-th term of the i-th document. This results in a numerical vector which can be utilized as inputs in the various machine learning algorithms in order to classify documents into topics, to be used for many other machine learning applications where text is the initial input. This simple model is proven very flexible and effective in various scenarios.

The *Term Frequency—Inverse Document Frequency* (TFIDF or *tf-idf*) (Gabryel et al., 2018) is another most popular term-weighting scheme used today in information retrieval, text mining, and user modelling. This weight is a statistic showing how important a term is to a document in a collection or corpus of examined online reviews. It is the product of two factors: the first computes the normalized *Term Frequency* (TF), e.g., the number of times a term appears in a document, divided by the total number of terms in that document; the second term is the *Inverse Document Frequency* (IDF), computed as the logarithm of the number of the documents in the collection or corpus divided by the number of documents where the specific term appears. The simplest choice is of course to use the raw count of a term in a document, i.e., the number of times that term t occurs in document d. Thus,

$$tf - idf = tf \times idf = f^1_{t,d} \times f^2_{t,D} = f^1_{t,d} \times \log\left(\frac{N}{n_t}\right).$$

where: $f^1_{t,d}$, and $f^2_{t,d}$ are the *tf* and *idf* respectively, with $t = 1, 2, \ldots, T$; $d = 1, 2, \ldots, N$; $N = |D|$; and $n_t = |\{d \in D : t \in d\}|$. In the above T is the total number of terms appeared in the corpus N is the total number of documents the corpus contains and n_t is the number of documents containing the term t. Note that in the usual case, when the term is a simple word (1-gram) then the above formula becomes:

$$tf - idf = tf \times idf = f^1_{w,d} \times f^2_{w,D} = f^1_{w,d} \times \log\left(\frac{N}{n_w}\right)$$

Now *tf-idf* is the relative weight of the feature in the vector, *tf* presents the number of words occurrences in total of reviews, n_w is the number of documents in the

corpus containing the word, N is the number of reviews in the corpus. The *tf–idf* value increases proportionally to the number of times a word appears in the document but is offset by the number of documents in the corpus that contain the word. This helps to adjust for the fact that some words appear more frequently in general.

As a working example consider a document containing 300 words wherein the words 'lovely' and 'badly' appear 15 and 45 times, respectively. The term (word) frequencies, i.e., $tf = f^1_{w,d}$ for $w =$ 'lovely' and $w =$ 'badly' is then $15/300 = 0.05$ and $45/300 = 0.15$, respectively. Now, assume we have two million documents and the word 'lovely' and 'badly' appear in 50 and 450 thousands of these, respectively. Then, the inverse document frequency, i.e., $idf = f^2_{w,D}$ is calculated as $\log(2{,}000{,}000/50{,}000) = 1.6$ and $\log(2{,}000{,}000/400{,}000) = 0.699$ and the *tf-idf* weight is the product of these quantities: $0.05 \times 1.6 = 0.08$ and $0.15 \times 0.699 = 0.105$ respectively.

Variations of this scheme are often used by search engines as a central tool in scoring and ranking a document's relevance given a user query. A simplest ranking function is computed by summing the *tf-idf* for each query term, however, many more sophisticated rank functions are variants of this simple model. The advantages are that the *tf-idf* metric, as well as, the similarity between two documents are computed easily, and in addition, it extracts in a flexible way the most descriptive terms in a document. In contrast, *tf-idf* is based on BoW model, and therefore it is only useful as a lexical level feature and cannot capture position in text, semantics, co-occurrences in different documents, etc.

The *Support Vector Machine* (SVM) classifier is a supervised machine learning algorithm or kernel methods that can be used for binary classification or regression. It constructs an optimal hyperplane as a decision surface such that the margin of separation between the two classes in the data is maximized. To introduce SVM (Joachims, 2001) we consider a training dataset $S = \{(\boldsymbol{x}, \boldsymbol{y})\} = \{(x_i, y_i)\}_{i=1}^m$. where $x_i \in R^n$ is the input vector, and $y_i \in \{+1, -1\}$ is the target value. SVMs map these input vectors into a high dimensional reproducing kernel space, where a linear machine is constructed by minimizing a regularized functional. The linear machine is of the form $f(\boldsymbol{x}) = <\boldsymbol{w} \cdot \varphi(\boldsymbol{x})> + \mathrm{b}$, where $\varphi(\cdot)$ the mapping function, b is the bias, and the dot product $<\varphi(\boldsymbol{x}) \cdot \varphi(\boldsymbol{x}^T)>$ is also reproducing kernel $\mathrm{K}(\boldsymbol{x}, \boldsymbol{x}^T)$. Then, the standard SVM can be stated as the following objective function to be minimized:

$$\min_{w \in R^n} \frac{\lambda}{2} \|\boldsymbol{w}\|^2 + \frac{1}{m} \sum_{(x,y) \in S} \max \{0.1 - \boldsymbol{y} < \boldsymbol{w}, \boldsymbol{x} >\}$$

According to Shalev-Shwartz, Singer, Srebro, and Cotter (2011), Pegasos algorithm can be used to solve the SVM since it was proved to be faster, having higher computational efficiency than the standard SVM. Then the modified objective SVM function above becomes:

$$\min_{w \in R^n} \frac{\lambda}{2} \|\boldsymbol{w}\|^2 + \frac{1}{m} \sum_{i=1}^{m} \max \{0.1 - y_i \boldsymbol{w}^T x_i\}$$

In the above, λ is the regularization term, $\boldsymbol{w}$ is the estimated vector, indicating a score for each word, m is the number of reviews, y_i is 1 if the review is positive or -1 if the review is negative, and x_i is the feature vector of each review. Pegasos is a stochastic sub-gradient descent method with varied step size with the following pseudocode:

Input $\lambda > 0$
Choose $w_1 = 0, t = 0$
While epoch $<$ max_epochs
  For $j = 1, \ldots, m$
    t = t + 1
    $n_t = 1/(t * \lambda)$
    If $y_j * w_t^T < 1$ *then*
      $w_{t+1} = (1 - \eta_t * \lambda) * w_t + \eta_t * y_j * x_j$
    Else
      $w_{t+1} = (1 - \eta_t * \lambda) * w_t$
    End if
  Next j

In the proposed model, each word in a review treats as an individual feature and a sparse feature matrix with very high dimensions would be generated. To avoid numerous zeroes in the list of words that is generated, a dictionary would be set up for each review, in order to provide information only for the words appear in the review. The sentiment score of each word in a sample of reviews, is multiplied by its average frequency among all reviews, namely:

$$\textit{Sentiment score of a word } (t) = \textit{score}(t) \times \frac{\textit{total frequency } (t)}{\textit{total number of reviews}}$$

In the above, *score*(*t*) is the word score calculated from the model, and *total frequency*(*t*) is the total frequency of word t in all the reviews.

The Naïve Bayes (NB) classifier is based on the so-called Bayesian theorem and is particularly suited when the dimensionality of the inputs is high. Despite its simplicity, Naive Bayes can often outperform more sophisticated classification methods. It computes the posterior probability of a class, based on the probability of a class, say *Pclass*, and the conditional probability of a class, say *P*(*w*/*class*) in a review, is computed as *Pclass* = *Nclass*/*N* where, *Nclass* is the probability of a class (positive or negative) *N* is the total count of class in the training set (Sharma & Dey, 2012):

$$P(w|class) = \frac{count(w, class) + 1}{count(class) + |V|}$$

In the above w is the word attribute, c is the class, $count(w, class)$ is the total count of word attribute occurs in class, + 1 is the Laplace smoothing, $count(class)$ is the total count of word in a particular class occurs in the data set, and $|V|$ is the vocabulary T.

The Maximum Entropy (ME) is a probabilistic classifier which belongs to the class of exponential models that is used to classify text documents. Unlike the NB classifier the ME does not assume that the features are conditionally independent of each other. From all the models that fit the training data, the ME selects the one which has the largest entropy. It can be used to solve a large variety of text classification problems such as language detection, topic classification, sentiment analysis and more. According to Pang et al. (2002) it outperforms NB in standard text classification tasks. It estimates the conditional probability of a class by the following form:

$$P(class|w) = \frac{1}{Z(w)} \exp\left(\sum_i \lambda_{i,c} F_{i,c}(w, c)\right)$$

In the above $Z(w)$ is a normalization function, $\lambda_{i,\ c}{}'s$ are the feature $-$ weight parameters, and $F_{i,\ c}$ is defined as follows:

$$F_{i,c}(w, c) = \begin{cases} 1, & n_i(d) > 0 \\ 0, & otherwise \end{cases}$$

The parameter values (Reyes & Rosso, 2012) are set in a way that the maximum entropy classifiers maximize the entropy of the induced distribution while maintaining the constraints enforced by the training data.

The K-nearest neighbors' algorithm (KNN) (Shalev-Shwartz et al., 2011), is a similarity-based learning algorithm that has been shown to be very effective for a variety of problem domains including text categorization. It compares the unknown data with the training data comparing the distance between them. The distance could be computed by the following form:

$$d(x_i - x_t) = \sqrt{(x_{i1} - x_{t1})^2 + (x_{i2} - x_{t2})^2 + \cdots + (x_{ip} - x_{tp})^2}$$

In the above, x_i is an input of reviews with p features $x_i = (x_{i1},\ \ x_{i2}, \cdots, x_{ip})$, x_t is the other input of reviews with p features $x_t = (x_{t1}, x_{t2}, \cdots, x_{tp})$, p the total number of inputs $j = 1, 2, 3, \ldots, p$ and n is the total number of inputs $i = 1, 2, 3, \ldots, n$. Given a test document, the KNN algorithm finds the K-nearest neighbors among the training documents, and uses the categories of the K neighbors to weight the category candidates. The similarity score of each neighbor document to the test document is used as the weight of the categories of the neighbor document.

According to Feldman (2013), the most important resource for the sentiment analysis techniques is the design and development of a sentiment lexicon. There are three approaches that are used for the acquisition of the sentiment lexicon. The first is

the manual approach in which the coding of the proposed lexicon is done by hand. The second is the dictionary-based approach in which a set of words is used by utilizing external resources as for example the WordNet dictionary, and the third is the corpus-based approach in which a large set of documents is examined in order to propose a dictionary.

In the following section, we propose a lexicon in Greek and Greeklish (Greek words written with English letters) language after the analysis of reviews in the data set. This lexicon should be used in the lexicon-based techniques and especially, in the sentence-level sentiment analysis techniques.

2.2 Related Articles

For the purpose of this work, we examined a number of related articles presented in Table 2. Since our interest is for the food and beverage of the Greek sector, we restrict the review in English and Greek language. The first column indicates the methodology adopted by each article and it is divided into two categories (Liu, 2010), namely, the unsupervised learning methods (methods that perform classifications based on some fixed syntactic phrases which are likely to be used to express opinions), and the supervised learning methods (methods that should be readily applied to sentiment classification), such as Naive Bayesian (NB) and Support Vector Machines (SVM). The second column specifies the language that analyzed in each article. The third column determines the data source of each case. The sources could be online reviews, forecast etc. The fourth column specifies the part of speech detected, namely, nouns, pronouns, adjectives, verbs, participles and adverbs. The fifth column shows the industry the corresponding article is referred and the final column specifies the examined study. We should note that all the following articles use the binary polarity (positive or negative).

The main result that comes out from the above summary is that existing research has focused mainly on the recognition of English characters and that very limited research work has been published on the Greek language and Greeklish idioms so far. Most of the work focuses on the technology and movie sector, while there was a very limited work on food industry.

3 Design of Greek Sentiment Lexicon

3.1 Data Set

We created a corpus of Greek and Greeklish (Latin letters) reviews by retrieving individuals reviews from Facebook and Twitter social media platforms. Specifically, we examine 275 customer reviews of 10 Greek leading companies in the Greek food industry. For a proper training set, the reviews were manually checked. The small

Table 2 Related articles

Methods used	Language	Sources	Part of speech	Industry	References
Unsupervised	English	Reviews	Adj and Adv	N/A	Liu (2010)
Super/ unsupervised	English	Forecasts	All	N/A	Kang, Yoo, and Han (2012)
Supervised	English	Restaurant reviews	All	Fast food	Li and Wu (2010)
Supervised	English	Product reputation	All	Mobile phones	Morinaga, Yamanishi, Tateishi, and Fukushima (2002)
Unsupervised	English	Reviews	Adj and Adv	Movies	Turney (2002)
Supervised	Greek	Reviews	All	–	Kalamatianos, Mallis, Symeonidis, and Arampatzis (2015)
Super/ unsupervised	Greek	Reviews	All	–	Adam (2018)
Supervised	Greek	Reviews	All	Movies and Tech.	Giatsoglou et al. (2017)
Supervised	Cantonese	Restaurant reviews	All	F&B	Zhang, Ye, Zhang, and Li (2011)
Supervised	English	Hotel reviews	All	Hotels	Duan, Cao, Yu, and Levy (2013)
Supervised	English	Hotel reviews	All	Hotels	Gräbner et al. (2012)
Supervised	English	Restaurant reviews	All	F&B	Vinodhini and Chandrasekaran (2012)
Supervised	English	Reviews	All	F&B	Yu et al. (2017)
Lexicon based	Greek	Reviews	All	Technology	Agathangelou, Katakis, Kokkoras, and Ntonas (2014)
–	All	Reviews	All	All	Kokkoras, Ntonas, and Bassiliades (2013)

volume of the examined reviews comes as a result of the General Data Protection Regulation (GDPR). This regulation may cause problems in the mining of published online reviews from companies' social media accounts.

3.2 Results

Greek is a typical high inflection language due to its complex grammatical and syntax rules. It is a particularly challenging language for Neuro-linguistic programming and specifically for sentiment analysis. For instance, according to Giatsoglou

et al. (2017), in English, there are only 4 forms of the regular verb ask (ask, asks, asked and asking), while there are 93 different forms corresponding to the regular Greek verb "ρωτώ". According to the data set, we found the most used positive adjectives and we present them in Tables 3 and 4.

The most used word is the adjective Exypiretiko which means helpful. However, we should note that this adjective has also negative meaning if for example precedes the word not (δεν) as presented in the two following examined reviews:

The restaurant's staff was very helpful (To proswpiko tou estiatoriou itan poly exypiretiko)
The restaurant's staff was not so helpful (To proswpiko tou estiatorioy den itan poly exypiretiko)

In the case of Greek language, there are many adjectives whose semantic orientations depend on contexts in which they appear (fifth column). This problem can be resolved with the negation rules (e.g.: Not helpful/Oxi exypiretiko). The negation words or phrases usually reverses the opinion expresses in a sentence such as no, not and never. The basic rule that is applied for negation is: Negation Positive transforms to negative.

We should also note that each of the adjective that is shown in Table 3, may describe different *Function* (Facilities, Staff, Products and Image of the Store) of a restaurant (Table 5).

In order to solve this restriction, the proposed lexicon should include word sequences. The basic idea of this method is to compute sentiment scores in

Table 3 Proposed Greek adjectives

Adjective (Greek)	Pronunciation (in English)	Adjective (Greeklish)	Possible similar writing
Καλός/η/ο	kalos	Kalo (Good)	–
Ωραίος/α/ο	oreos	Oraios (Nice)	Oreo
Νόστιμος/η/ο	nostimos	Nostimo (Tasty)	Nosthmo, Nostymo, Nosteimo,Nostoimo
Τέλειος/α/ο	telios	Teleio (Perfect)	Telio
Φιλικός/η/ο	filikos	Filiko (Friendly)	–
Αγαπημένος/η/ο	agapimenos	Agapimeno (Favourite)	–
Εξαιρετικός/η/ο	ekseretikos	Exairetiko (Brilliant)	Exeretiko
Υπέροχος/η/ο	iperohos	Iperoho (Fabulous)	Iperoxo, Yperoxo, Yperoho
Εξυπηρετικός/η/ο	ekseepiretikos	Exypiretiko (Helpful)	Exipiretiko
Οικονομικός/η/ο	ikonomikos	Oikonomiko (Cheap)	Ikonomiko
Όμορφος/η/ο	Omorfos	Omorfo (Nice)	–
Θαυμάσιος/α/ο	θaumasios	Thaumasio (Terrific)	Thaymasio
Καταπληκτικός/η/ο	katapliktikos	Kataplitiko (Exceptional)	–
Γρήγορος/η/ο	γriγoros	Grigoro (Fast)	Grhgoro
Ευγενικός/η/ο	evjenikos	Evgeniko (Polite)	Ebgeniko

Table 4 Statistics of proposed Greek adjectives

Adjective	Frequency positive meaning	Relative frequency positive meaning	Frequency negative meaning	Relative frequency negative meaning	tf-idf positive meaning	tf-idf negative meaning
Καλός/η/ο	37.0	0.13	3.0	0.14	0.117	0.021
Ωραίος/α/ο	40.0	0.15	–	–	0.122	–
Νόστιμος/η/ο	18.0	0.07	–	–	0.078	–
Τέλειος/α/ο	4.0	0.01	–	–	0.027	–
Φιλικός/η/ο	5.0	0.02	4.0	0.18	0.032	0.027
Αγαπημένος/η/ο	6.0	0.02	–	–	0.036	–
Εξαιρετικός/η/ο	26.0	0.09	–	–	0.097	–
Υπέροχος/η/ο	6.0	0.02	–	–	0.036	–
Εξυπηρετικός/η/ο	50.0	0.18	6.0	0.27	0.135	0.036
Οικονομικός/η/ο	10.0	0.04	3.0	0.14	0.052	0.021
Όμορφος/η/ο	40.0	0.15	2.0	0.09	0.122	0.016
Θαυμάσιος/α/ο	5.0	0.02	–	–	0.032	–
Καταπληκτικός/η/ο	5.0	0.02	–	–	0.032	–
Γρήγορος/η/ο	8.0	0.03	2.0	0.09	0.045	0.016
Ευγενικός/η/ο	15.0	0.05	2.0	0.09	0.069	0.016

Table 5 Alternative functions that each adjective describes

Adjective	Function that referred most	Alternative function 1	Alternative function 2
Kalo (Good)	Products	Image of the store	Staff
Oraio (Nice)	Products	Image of the store	–
Nostimo (Tasty)	Products	–	–
Teleio (Perfect)	Products	Image of the store	Staff
Filiko (Friendly)	Staff	–	–
Agapimeno (Favourite)	Products	Staff	Image of the store
Exairetiko (Brilliant)	Products	Staff	Image of the store
Iperoho (Fabulous)	Products	Staff	Image of the store
Exypiretiko (Helpful)	Staff	–	–
Oikonomiko (Cheap)	Products	–	–
Omorfo (Nice)	Products	Image of the store	–
Thaumasio (Terrific)	Products	Image of the store	Staff
Kataplitiko (Exceptional)	Products	Image of the store	Staff
Grigoro (Fast)	Staff	–	–
Evgeniko (Polite)	Staff	–	–

sequences of words depending the function that describes. Another obstacle with the Greek language is that the vast majority of online text-based communications ignore the rules of spelling and grammar. After the analysis of the examined reviews, are identified alternative ways to write the same adjective (column 4, Table 3). For

Table 6 Transcription of Greeklish into Latin letters

Greek letters	Latin letters	Greeklish
Αι, αι	Ee	Ai/ai, Ee
Αυ/αυ	Av/av, Au/au	Av, au, Ay, Af
Ββ	Vv	Bb, Vv
Γγ	Gg, Yy	Gg, Yy
Εε	Ee	Ee
Ει, ει	Ii	Ii, Ei/ei
Ευ, ευ	Ev/ev, Eu/eu	Ev, Eu, Ey, Ef
Ηη	Ii	Ii, Hh
Ιι	Ii	Ii
Κκ	Kk, Cc	Kk, Cc
Ξξ	Xx	Xx, Jj, Ks/ks
Οο	Oo	Oo
Οι, οι	Oi, oi	Oi, oi
Ου, ου	Ou, ou	Ou, Oy
Υυ	Ii	IiYi
Υι, υι	Ii/ii	Ii, Ui, Yi
Φφ	Ff	Ff, Ph/ph
Χχ	Xx	Xx, Ch/ch
Ωω	Oo	Oo, Ww

instance, the adjective tasty was presented with the following five different ways: Nostimo, Nosthmo, Nostymo, Nosteimo, Nostoimo. This is an outcome of the complexity of the Greek alphabet. Table 6, cites an equivalent standard for transcription of Greek letters into Latin letters (Pedersen, 2009) and into Greeklish from the examined reviews.

Adapting the abbreviation Greek/Greeklish there are two ways to write the letter e (ε/e and αι/ai), four ways to write the letter i (ι/i, η/h, υ/y or υ/u and ει/ei), two ways to write the letter o (ο/o, ω/o or ω/w), and two ways to write the letter v (β/b and β/v), since in each one of the above cases the corresponding Greek letters are pronounced identically. All the above could help us to suggest a sentiment dictionary with Greek sentiment adjectives which may be used in lexicon-based techniques.

According to the proposed analysis, the transcription of Greek letters into Greeklish enables us to use the Greek dictionary in any recommended program that has already used in sentiment analysis in English language. The following procedure should be ideal. Firstly, the proposed lexicon should be saved in a .csv file (greeklexicon.csv). The alternative ways to write the same adjective must be registered as different (column 1 and 2, Table 3). Each row of the first column has to contain an adjective. Secondly, the data set of the examined reviews should be saved also as a .csv file (examinedreviews.csv). Each row has to contain a review that has to be written with Latin letters (Table 4). In order to avoid the double binary semantic orientation, all neutral and negative sentiments should be removed. Fourthly, from

each row in the file examined reviews.csv the frequencies of each adjective have to be computed. Finally, the reviews which contain one of the adjectives in the greeklexicon.csv should be saved in a new file as follows: (Company's Name, Adjective, Positive, User, and Date of review).

4 Conclusions

In Greece, more and more people use the social media networking to make a decision about purchasing a product or a service. On the other hand, the vast majority of companies use the social media marketing to improve their brands. Due to the large and continuous volume of data that produced users face difficulties to identify all the necessary information. The problem can be approached with the method of sentiment analysis. Sentiment analysis has a wide variety of techniques (supervised, unsupervised and lexicon based) to detect users' positive or negative sentiments. While a lot have been written and researched about sentiment analysis in various languages, Greek has not drawn researcher's attention. After the analysis of 275 customer reviews of 10 leader companies in Greece, a mini vocabulary of 15 positive adjectives and a transcription of Greek into Latin letters (and Greeklish idioms) was suggested. The proposed vocabulary should be used in any recommended technique to detect comments that are written in Greek language (or Greeklish) in social media networks in case of food and beverage sector. Future work will focus, primarily, on the extension of this study regarding all the parts of speech of Greek language (verbs, nouns, adverbs) and secondly, on the restrictions as presented in Sect. 3 (negation rules and spelling mistakes). Especially, a bigger sample of online reviews will help us to predict all the alternative ways to write Greek words as presented in Greeklish idioms in the case of social media networks.

Acknowledgments This research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme "Human Resources Development, Education and Lifelong Learning 2014–2020" in the context of the project "Strengthening Human Resources Research Potential via Doctorate Research—2nd Cycle" (MIS 5000432).

References

Adam, T. *Greek sentiment lexicon*. Accessed May 23, 2018., from http://socialsensor.eu/results/datasets/147-greek-sentiment-lexicon

Adeniyi, D., Wei, Z., & Yongquan, Y. (2016). Automated web usage data mining and recommendation system using K-nearest neighbor (KNN) classification method. *Applied Computing and Informatics, 12*(1), 90–108.

Agathangelou, P., Katakis, I., Kokkoras, F., & Ntonas, K. (2014). Mining domain-specific dictionaries of opinion words. In *International conference on web information systems engineering* (pp. 47–62). Cham: Springer.

Aggarwal, C. C., & Zhai, C. X. (2012). *Mining text data*. New York: Springer Science.
comScore/the Kelsey group. (2007). *Online consumer-generated reviews have significant impact on offline purchase behavior*. Press Release, Accessed May 23, 2018, from http://www.comscore.com/press/release.asp?press=1928
Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory*. New York: Wiley.
Deerwester, S., Dumais, S., Landauer, T., Furnas, G., & Harshman, R. (1990). Indexing by latent semantic analysis. *JASIS, 41*, 391–407.
Duan, W., Cao, Q., Yu, Y., & Levy, S. (2013). Mining online user-generated content: Using sentiment analysis technique to study hotel service quality. In *System Sciences (HICSS), 46th Hawaii International Conference* (pp. 3119–3128).
Duric, A., & Song, F. (2012). Feature selection for sentiment analysis based on content and syntax models. *Decision Support Systems, 53*, 704–711.
Fan, T., & Chang, C. (2012). Blogger-centric contextual advertising. *Expert Systems, 38*, 1777–1788.
Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM, 56*, 82–89.
Gabryel, M., Damaševičius, R., & Przybyszewski, K. (2018, October). Application of the bag-of-words algorithm in classification the quality of sales leads application. In *ICAISC 2018, LNAI 10841* (pp. 615–622). New York: Springer International.
Genier, C., Stamp, M., & Pfitzer, M. (2009). Corporate social responsibility for agro-industries development. In C. Da Silva, D. Baker, A. Shepard, C. Jenane, & S. Miranda-da-Cruz (Eds.), *Agro-industries for development* (pp. 223–251). Oxfordshire, UK: CABI.
Giatsoglou, M., Vozalis, M. G., Diamantaras, K., Vakali, A., Sarigiannidis, G., & Chatzisavvas, K. C. (2017). Sentiment analysis leveraging emotions and word embeddings. *Expert Systems with Applications, 69*, 214–224.
Gräbner, D., Zanker, M., Fliedl, G., & Fuchs, M. (2012). Classification of customer reviews based on sentiment analysis. In M. Fuchs, F. Ricci, & L. Cantoni (Eds.), *Information and communication technologies in tourism* (pp. 460–470). New York: Springer.
Griffiths, T. L., Mark, S., Blei, D. M., & Tenenbaum, J. B. (2005). Integrating topics and syntax. *Advances in Neural Information Processing Systems, 17*, 537–544.
Haseena, R., & Tanvir, A. (2014). Opinion mining and sentiment analysis - challenges and applications. *International Journal of Application or Innovation in Engineering & Management (IJAIEM), 17*, 25–29.
He, Y., & Zhou, D. (2011). Self-training from labeled features for sentiment analysis. *Information Processing and Management, 47*, 606–616.
Horrigan, J. A. (2008). *Online shopping pew internet & American life project report*. https://www.pewinternet.org/2008/02/13/online-shopping/
Hu, X., & Liu, H. (2012). Text analytics in social media. In *Mining text data* (pp. 385–414). Boston: Springer.
Joachims, T. (2001). A statistical learning model of text classification for support vector machines. In *Proceedings of SIGIR-01, 24th ACM International Conference on Research and Development in Information Retrieval* (pp. 128–136). New York: ACM.
Kalamatianos, G., Mallis, D., Symeonidis, S., & Arampatzis, A. (2015) Sentiment analysis of Greek tweets and hashtags using a sentiment lexicon. In *Proceedings of the 19th Panhellenic Conference on Informatics, Athens* (pp. 63–68). New York: ACM.
Kang, H., Yoo, S. J., & Han, D. (2012). Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews. *Expert Systems with Applications, 39*, 6000–6010.
Ko, Y., & Seo, J. (2000). Automatic text categorization by unsupervised learning. In *Proceedings of the 18th Conference on Computational Linguistics* (pp. 453–459). Stroudsburg, PA: Association for Computational Linguistics.
Kokkoras, F., Ntonas, K., & Bassiliades, N. (2013). DEiXTo: A web data extraction suite. In *Proceedings of the 6th Balkan Conference in Informatics, Thessaloniki* (pp. 9–12). New York: ACM.

Lewis, D. D. (1998). Naïve (Bayes) at forty: The independent assumption in information retrieval. In *Proceedings of ECML-98, 10th European Conference on Machine Learning* (pp. 4–15). Berlin: Springer.

Li, N., & Wu, D. D. (2010). Using text mining and sentiment analysis for online forums hotspot detection and forecast. *Decision Support Systems, 48*, 354–368.

Liapakis, A., Costopoulou, C., Tsiligiridis, T., & Sideridis, A. (2017). Studying corporate social responsibility activities in the agri-food sector: The Greek case. *International Journal of Agricultural and Environmental Information Systems (IJAEIS), 8*, 1–13.

Liu, B. (2010). Sentiment analysis and subjectivity. In N. Indurkhya & F. J. Damerau (Eds.), *Handbook of natural language processing* (Vol. 2, pp. 627–666). Boca Raton: Chapman and Hall/CRC.

Medhat, W., Hassan, A., & Korashy, H. (2008). Combined algorithm for data mining using association rules. *Ain Shams Journal of Electrical Engineering, 1*(1), 1–12.

Morinaga, S., Yamanishi, K., Tateishi, K., & Fukushima, T. (2002). Mining product reputations on the web. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Edmonton* (pp. 341–349). New York: ACM.

Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing* (Vol. 10, pp. 79–86). Stroudsburg, PA: Association for Computational Linguistics.

Pedersen, T. T. (2009). *Transliteration of non-roman scripts*. Accessed May 23, 2018.

Rainie, L., & Hitlin, P. (2004). *The use of online reputation and rating systems*. Washington, DC: Pew Internet and American Life Project.

Read, J., & Carroll, J. (2009). Weakly supervised techniques for domain in dependent sentiment classification. In *Proceeding of the 1st International CIKM Workshop on Topic-Sentiment Analysis for Mass Opinion* (pp. 45–52). New York: ACM.

Reyes, A., & Rosso, P. (2012). Making objective decisions from subjective data: Detecting irony in customer reviews. *Decision Support Systems, 537*, 754–760.

Shalev-Shwartz, S., Singer, Y., Srebro, N., & Cotter, A. (2011). Pegasos: Primal estimated sub-gradient solver for SVM. *Mathematical Programming, 127*, 3–30.

Sharma, A., & Dey, S. (2012). A comparative study of feature selection and machine learning techniques for sentiment analysis. In *Proceedings of the 2012 ACM Research in Applied Computation Symposium* (pp. 1–7). Raleigh, NC.

Singh, J., Singh, G., & Singh, R. (2016). A review of sentiment analysis techniques for opinionated web text. *CSI Transaction on ICT, 4*, 241–247.

Turney, P. D. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics* (pp. 417–424). Philadelphia, PA.

Vinodhini, G., & Chandrasekaran, R. M. (2012). Sentiment analysis and opinion mining: A survey. *International Journal of Advanced Research in Computer Science and Software Engineering, 2*, 282–292.

Xianghua, F., Guo, L., Yanyan, G., & Zhiqiang, W. (2013). Multi-aspect sentiment analysis for Chinese online social reviews based on topic modeling and how net lexicon. *Knowledge-Based Systems, 37*, 186–195.

Yelena, M., & Padmini, S. (2011). Exploring feature definition and selection for sentiment classifiers. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*. Barcelona, Catalonia.

Yu, B., Zhou, J., Zhang, Y., & Cao, Y. (2017). Identifying restaurant features via sentiment analysis on yelp reviews. *arXiv preprint arXiv:1709.08698*.

Zhang, Z., Ye, Q., Zhang, Z., & Li, Y. (2011). Sentiment classification of internet restaurant reviews written in Cantonese. *Expert Systems with Applications, 38*, 7674–7768.

Firms' Profitability in the Period of Crisis: Evidence from Greek Food and Beverage Listed Firms in the Athens Stock Exchange

Konstantina Panagiotakopoulou, Petros Kalantonis, and Panagiotis Kaldis

Abstract Food and beverage industry has a fundamental role in the Greek manufacturing sector and also for the Greek economy. Even in the period of economic crisis the food and beverage industry has demonstrated competitiveness extroversion and dynamic development in Greece and also in the global market. This study investigates the impact of economic crisis on F&B firms' profitability and value. We explored the listed F&B firms in the Athens Stock Exchange for the sub periods before and during the crisis. We found evidence that F&B firms' profitability has been affected by the economic crisis, however, their book and market value doesn't seem to be significantly affected. Also impact of crisis on the relationship between the debt and profitability has been detected.

Keywords Food and beverage · ROA · ROE · Tobin's Q · MAVPS · EPS · Financial statements analysis

1 Introduction

Food and Beverage (F&B) industry is one of the most developed business sectors in Greece. There are more than 16,000 small, medium and big F&B enterprises in Greece. These firms constitute the 24% of the Greek industrial sector. The revenue of the F&B firms equals to the one third of that of the Greek manufacturing sector. In addition, more than 100,000 employees constitute almost the 36% of the total

K. Panagiotakopoulou
Department of Home Economics and Ecology, Harokopio University, Athens, Greece

P. Kalantonis (✉)
Department of Tourism Management, University of West Attica, Egaleo, Greece

P. Kaldis
Department of Vine, Wine and Beverage Sciences, University of West Attica, Egaleo, Greece
e-mail: pkaldis@uniwa.gr

E. Krassadaki et al. (eds.), *Operational Research in Agriculture and Tourism*, Cooperative Management, https://doi.org/10.1007/978-3-030-38766-2_4

manufacturing sector's employment.[1] Nevertheless, the 1% of Greek F&B firms produces the 63% of the total sector's revenue (see footnote 1).

During the years of financial crisis and specifically since Greek economy has been stated under financial probation—May, 2010—Greek F&B firms performed better than those of other industry sectors. According to a Pricewaterhouse Coopers, PWC (2018) report for the Food and Beverage industry in Greece, the annual revenue of F&B firms appeared a mean increase by 1.86% in the period 2009–2016. The fact that F&B firms seem to be less affected by the financial crisis—at least in terms of revenues—could attract the interest of investors for the F&B firms in Greece. On the other hand there is not adequate information for the impact of the recent crisis on firms' financial performance.

The main purpose of this paper is to examine the effect of crisis on firms' book and market value and also on their profitability which constitute the main financial performance components.[2] Prior and recent relevant studies used only reported accounting information or macroeconomic variables to explore the impact of crisis on business firms of different industrial sectors. In this study we focus on the listed F&B firms which are more interesting to investors. In addition to the prior research, we introduce the market value of the listed firms to measure the financial performance of F&B firms.

Relevant studies, which explore the macroeconomic and the microeconomic effects of the global economic crisis set the years 2007 or 2008 as the starting point for the global crisis. The year 2007 has been related with the Housing bubble in United States. However the bankruptcy of Lehman Brothers in 2008 has been also recognized as the beginning of the market crash.

Nevertheless, although signs of the contagion of crisis in Greek economy have been appeared since 2007 and 2008, no serious effects of the global crisis were appeared in the Greek market. The most significant consequences became after the May, 2010 when Greek Economy was placed in the status under financial probation. In the entire research we compare the financial performance of the F&B firms in two sub periods. The first sub period has been extended from 2007 to 2009 and the second one from 2010 to 2014. Consequently, we recognize the year 2010 as the beginning year of the crisis, differentiating our approach for the beginning of the crisis from that of the previous studies.

This study has been structured as follows: In the following section we include the literature review for the performance of the food and beverage industry in Greece and the impact of prior crises on firms' profitability. The research hypotheses and the search methodology are presented in the third section. We test the hypotheses and discuss our findings in the fourth section. The conclusions, the limitations and proposals for future research have been explained in the last section.

[1] https://www.pwc.com/gr/en/publications/greek-thought-leadership/the-food%20&%20beverage-sector-in-greece(en).pdf

[2] https://www.investopedia.com/terms/f/financialperformance.asp; http://www.businessdictionary.com/definition/financial-performance.html

2 Literature Review

Food and Beverage business sector in Greece and its prospects have been analyzed in professional and research studies (European Commission 2013, IOBE 2015, National Documentation Center 2015a, b, etc). Christodoulakis and Stathis (2014) used SWOT analysis to examine the F&B industry in Greece. They specified as strengths the well-organized sales network of food and beverage products, the positive impact of tourism in the increased seasonal consumption of these products and the fact that Greek household give priority to the consumption of Greek products. On the other hand the big number of small sized enterprises with limited access to the big retail network and the low export ability were identified as a significant weak point of Greek F&B industry. Moreover the absence of a well organized and efficient policy for the promotion of the Greek F&B products in the Global market and the increase of Value Added Tax—which is a barrier to the industrial development—have been also pointed as weak points of the F&B manufacturing sector in Greece. The increased demand for Mediterranean and organic products, the promotion of the Greek traditional F&B products in the global market and the interest in new emerging markets, were characterized as opportunities. However the decrease of the demand due to the crisis, the uncertainty of the Greek Economy Framework and the rising prices of the F&B products and their raw materials were reported as the threats to the development of the sector.

The impact of crisis on manufacturing firms was explored in prior and recent literature. Claessens, Djankov, and Lang (2000) compared the growth and corporations financial measures to the years before and after the beginning of East Asian Economic Crisis. According to their findings the effect of crisis on firms' performance has been related with the economy of the country. Also they examined firms' financial and non financial characteristics and they indicated that firms, with higher performance, better before the beginning of crisis, have been less affected by the crisis. The size of firms was not identified as a significant factor for the firms' performance in those two examined periods. Larger firms didn't perform better than smaller ones in the era of East Asian crisis.

Negative effect of crisis on firms' profitability has been observed also by Iraizoz, Bardaji, and Rapun (2005) who explored the technical efficiency and profitability of Spanish beef firms during the B.S.E crisis. Of course B.S.E crisis was not a financial crisis but it had serious economic consequences to food industry. Lien (2010) supported that the effect of crisis on firms performance could be related with the business sector. Some sectors—due to their activity—tend to be less sensitive than others in the period of crisis. The elasticity of the demand of their products could be a good reason for that.

In a more recent study Sujova (2015) analyzed the reported financial information to ascertain the incidence of crisis in Wood Industry firms' performance in Czech Republic. Specifically she used financial indicators, as return on equity, return on sales, return on added value per employee, labor productivity indices and investment

rates. In her findings, global crisis affect negatively sales, investments, profitability and labor productivity.

Alternatively, El Bilali et al. (2012) made a macroeconomic approach. They investigated the effect of recession on Serbia's Agrofood Industry, based on macroeconomic variables, such as salaries, GDP, unemployment rates and imports to exports ratio.

However, in Greece, the debt and the low liquidity became the most important problems for the Greek Firms, since the banks couldn't finance them for a long time period, especially in the early years of the economic crisis. Moreover it is known that the financial crisis affected negatively the profitability and the market value of the listed firms in the Athens Stock Exchange.

The relationship between the liquidity and profitability has been examined in previous research studies. Nevertheless contradictory findings have been observed. According to liquidity—profitability trade off approach (Saluja & Kumar, 2012), the "desire" of firms to increase their current assets and—at the same time—to decrease their current liabilities has negative effect on their profitability. Deloof (2003) and Lazaridis and Tryfonidis (2006) observed that profitability related negatively with working capital and cash conversion cycle. In a more recent study, Bhunia (2013) focused on 100 small-medium firms of the steel industry in India and concluded that liquidity and profitability were negatively related.

On the other hand, Owolabi et al. (2011) found that liquidity and profitability on Nigerian firms have been positively associated. The positive relationship between liquidity and profitability in Nigerian firms has been confirmed by Egbide, Uwuigbe, and Uwalomwa (2013), who found positive effect of the listed Nigerian firms' current liquidity ratio on their return on their capital employed ratio. Enqvist, Graham, and Nikkinen (2014), investigated the role of business cycles in the relationship between working capital and profitability and stated that the management of inventories, the decrease of accounts receivables and payables periods could lead to the profitability increase. Significant positive impact of liquid ratio on Return On Assets (ROA) ratio has been observed by Saleem and Rehman (2011) who analyzed firms listed on Karachi Stock Exchange. However no significant effect on Return On Equity (ROE) and on Return On Investments (ROI) has been detected for these firms.

The relationship between the debt and profitability has been also investigated in previous and recent studies. Positive relationship between the percentage growth of total assets and, the total debt ratio, has been found by Abdullah (2005). In addition to those findings, debt ratio affects negatively but not significantly the ROA of commercial banks in Kenya according to Kajirwa (2015).

Looking into the findings of the above studies we observed that the relation of profitability with debt and liquidity is not the same in the all the examined countries. We also estimate that the status of the economy may affect this relationship.

In our study—which is focused on the Greek listed F&B firms—we investigate the role of crisis on the effect of liquidity and debt on the profitability in terms of Return on Assets (ROA) and Tobin's Q. ROA is a most use indicator for the accounting profitability and Q ratio reflects the market performance. Q ratio has

been used as firms' market performance indicator in prior and recent research studies (Fu, Singhal, & Parkash, 2016; Wolfe & Sauaia, 2003).

Regarding the impact of financial crisis on business firms, Notta and Vlachvei (2014, 2018) examined if firms' performance has been affected by the Global Financial Crisis. According to the results of their research (2014) market share, liquidity, and beverage affect significantly the profits and therefore they explain profitability differences among the firms. Moreover they stated in their more recent study (2018) that the older and more efficient firms are more profitable in the period of crisis.

In addition to the literature for the performance of F&B firms before and after the beginning of the recent Financial Crisis, the research study of Kontogeorgos, Pendaraki, and Chatzitheodoridis (2017) focused on the profitability of the Greek Cheese industry. Financial data of the cheese industry have been analyzed for period 2006–2011. The profitability was the dependent variable of the OLS regression model and the firms' size, the liquidity, capital structure and activity indices were the independent variables and therefore determinant factors of the firms' performance. The results of the study showed that in the period of crisis manufacturing firms are less profitable than commercial firms and especially the cooperatives performed better during the crisis.

Beyond the age and the previous performance of firms, the activity sector of firms and its future perspective seem to have a significant role in their future performance. Castaner (2009) explained that the future growth will be based on the food consumption in the developing countries—which has been estimated to one billion consumers—and to the growth of high value added markets in developing countries.

3 Methodology

Then main purpose of this research is to investigate the effect of the economic crisis on F&B stock market sector. We first examine if the profitability of F&B firms was significantly affected in the period of crisis. We next explore if the debt and the liquidity are determinant factors on the F&B firms' profitability.

The research hypotheses which will be tested will be the following:

H_1 The profitability and firms' value of food and beverage listed firms has a significant difference between the period after the beginning of the economic crisis in Greece and the period before the crisis.

H_2 The debt and liquidity affect significantly on the profitability within the periods before and after the beginning of Greek's economy financial probation.

For this purpose, we selected all the listed F&B firms in the Athens stock exchange. Data from 22 firms which constituted this sector were analyzed. The firms of our sample have been distributed per subsector as follows (Table 1):

Table 1 Distribution of sample per sub sector

Sub sector	Number of firms
Food	13
Agriculture and fishery	6
Beverage (alcohol)	2
Beverage (non-alcohol)	1

The study period begins 3 years before the probation of the Greek Economy due to the contamination of the Global financial crisis in Greece. It ends the year 2014, before the obligatory adoption of Capital Controls from the Greek Government on year 2015. We splited the examined period in two sub periods. The first one is the period in which, Greek Economy has not been directly affected by the crisis and the other is the period since the significant political changes in Greece and the implementation capital controls. It is commonly accepted that the Greek economy has been affected by the consequences of the economic crisis since May 2010, when Greek Economy has been placed under financial probation. For that reason, we consider as the period of crisis the years from 2011 to 2014 and as the per recession period the years from 2006 to 2009.

3.1 Variables

The main purpose of this study was to evaluate the profitability of F&B firms in the sub period of Crisis in (2007–2014) and detect for any significant difference in the profitability of F&B listed firms in Athens Stock Exchange for the periods before and after the beginning of probation period for the Greek Economy and the adoption of the first *economic* adjustment program.

The financial indicators which were used to test the first hypothesis focused on the evaluation of the profitability and firms' market and book value. These indicators were examined in previous and recent studies for this purpose. Moreover these indicators have been also investigated in financial analysts reports. All data have been retrieved from the annual financial reports of the listed firms in the Athens Stock Exchange for the period 2007–2014. Based on the accounting literature (Courtis, 1978; Palepu, Healy, & Bernard, 2004; Penman, 2007) and also on professional sites specialized in accounting and financial analysis and practice (e.g. financial times FT lexicon, www.investopedia.com, www.financeformulas.net) we adopt in our research study the following indices in order to test our first research hypothesis:

Net Asset Value per share (NAVPS) is an indicator which reflects the firm's value per share. It is also introduced in the literature as book value per share. The formula for its calculation is:

$$Net\ Asset\ Value\ per\ share = Net\ Asset\ Value/total\ shares.$$

Return on Assets (ROA) index used to estimate the profitability of firms comparing it with total assets. Specifically, it is used to evaluate the efficiency of management to create earnings for a company. Dividing the net income with total assets, we calculate this index. In other words, Return on Assets reflects the ability of the administration of a firm to manage the invested capital in order to maximize the annual profits.

Return on Equity (ROE) demonstrates the portion of net income comparing it with shareholders equity. Furthermore, it is used for the evaluation of the profitability which has been generated from shareholders' equity. Investors can evaluate the profitability of their investment and compare it with alternative investments' return. For its calculation, we need to divide the net profit with shareholders equity.

Moreover, for measuring the return of an individual share, based on the share's book value, Earnings per Share (EPS) is the most appropriate indicator. It must be reported in the Annual Financial report of listed Firms according to the Accounting Framework of International Accounting Standards. Earnings per share indicator are mathematically expressed as the portion of the net income of the total shares number of business firms.

A key indicator which shows how the stock market evaluates a firm is a price to book ratio (MV/BV). It reflects the goodwill that has been given to the firm by the shareholders and investors in a globalized stock market. It is also a measure to evaluate the price of a share and recognize investment opportunities. The calculation of this indicator is:

$$Market\ to\ Book\ value = stock\ price\ of\ a\ firm\ or\ share/book\ value\ of\ a\ firm\ or\ share.$$

If the market value is lower than book value there is evidence that the market evaluates firms less than it is reported in their financial statements. On the other hand, if the market value to book value ratio of a firm equals more than one, then the stock market seems to believe that the firm will increase either its efficiency or competitiveness or profitability in the future. In other words market has more prospects from this share.

An alternative indicator for the valuation of the stock price of a firm and also for the evaluation of share price is Tobin's Q[3] ratio. It compares the market value of a company with total Asset Value. The denominator constitutes the difference from price to book ratio we stated above. Specifically, this ratio measures if the replacement cost of Total Assets of a company is greater than its stock price. Considering that if this ratio equals less than 1 it could mean that the firm is undervalued from the market and then it could attract the interest of investors, Tobin's Q could be used for the detection of profitable future investments.

[3] Also known as Q ratio.

Return on assets (ROA), Return on Equity (ROE), Earnings per share (EPS), were used for the evaluation of the profitability of food and beverage firms. Heikal, Khaddafi, and Ummah (2014) tested in a recent research on the effectiveness of ROA and ROE for the reflection of corporate profit growth. Earnings per share are also a financial indicator which is used to evaluate the return on investment and risk of a firm. IAS 33% and determine EPS and how it has to be disclosed in financial statements.

Market value to book value indicator has been used in this study to examine how the market evaluated F&B firms. Specifically, we estimate if there is a significant change in the market's sense for the value of F&B firms. Canibano, Ayuso, and Sanchez (2000) and Kalantonis, Gaganis, and Zopounidis (2014) used this indicator to investigate if the market gives a premium to innovative firms and to those which demonstrate research and development activity. A similar indicator which has been adopted in financial analysis research for testing the relationship between the market value of a firm and total asset value will be also used in order to appraise firms' value. Tobin's Q has been used in prior research from Lang, Stultz, and Walking (1989) as an estimator for managerial performance and recently by Suzuki and Chida (2017) as measure of physical capital investment for R&D intensive firms.

We tested all the variables for normality with Kolmogorov Smirnov's normally test. No significance evidence for normality was observed. The non-parametric test of Kruskal and Wallis was implemented for the detection of significant differences or financial indicator values of F&B listed companies for the periods before and after the economic crisis in Greece.

Regarding the second hypothesis we developed OLS regression models setting the profitability and indicators for the firms' market value as dependent variables. Specifically we test debt and liquidity which have a significant effect on profitability and market value of F&B firms. The formulas for the debt and liquidity ratios were[4,5]:

$$Debt\ ratio = Total\ debt / Total\ assets$$

$$Liquidity\ (current)\ ratio = Current\ assets / Current\ liabilities$$

The structure of the equations was as follows:

Dependent variables	Independent variables
$ROA=$	$a + \boldsymbol{b_1}\ debt\ ratio + \boldsymbol{b_2}\ liquidity + e$
$ROE=$	
$MV/BV=$	
$Q_{ratio}=$	
$EPS=$	
$NAVPS=$	

[4] https://www.investopedia.com/terms/d/debtratio.asp

[5] https://www.investopedia.com/terms/l/liquidityratios.asp

4 Analysis and Discussion

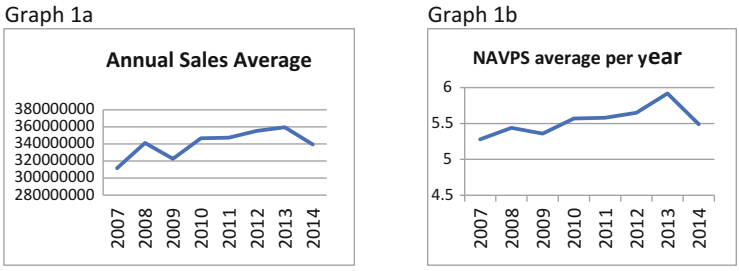
In this section, we present graphically the mean sales and the mean profitability—based on accounting and stock market indicators—of the F&B listed firms in the Athens Stock Exchange. In addition, we present descriptive statistics of the examined variables. Our observations concern the two sub periods individually (before the financial probation and during the probation period).

Descriptive Statistics

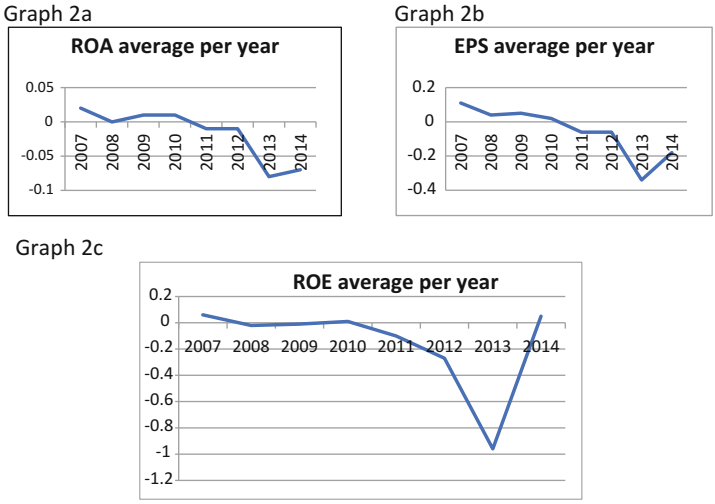
Looking into the Graph 1a, b, no negative effect from the recession could be observed on the Sales and the Net Asset Value of the Food and Beverage firms. The exports orientation of these firms and the increase of their export activity, so as the fact that food and beverage products cover fundamental needs of the people could prevent the affection of financial crisis on F&B enterprises. Of course, this does not mean that the financial performance of the F&B sector firms have not been affected by the recession, especially since the Greek Economy has become under financial probation.

Specifically we observe a downward trend (Graph 2a–c) in the profitability of the F&B firms. The decrease of profitability seems to be more significant in the year period 2010–2013, however 2013 could be characterized as the profitability worst year for the F&B listed Firms in the Athens Stock Exchange. The unstable political status—continuously governmental changes—and the uncertainty concerning the prospects of Greek economy could be related with the negative performance of firms profitability.

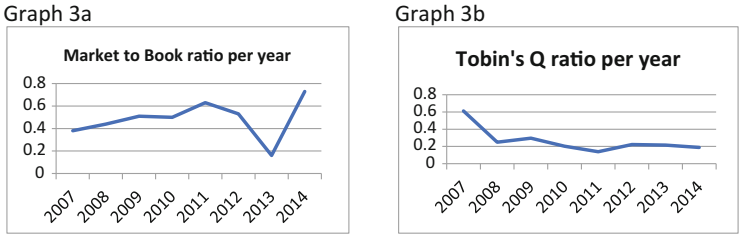
Moreover, the Market to Book ratio appears relevant view with the profitability ratios. The F&B firms' market value has been decreased since 2011, when Greek economy became under financial probation. However, we have to mention that Market to Book ratio was below 1 in the period after the beginning of the recession in the US economy. An optimist change of the Market to Book ratio was observed in the year 2014. On the other hand, Q ratio becomes lower within the period of crisis. The shrinkage of firms' total value—due to the low or negative profitability—could be a reason for the relevant reaction of the stock market (Graph 3).



Graph 1 Annual sales and NAVPS average



Graph 2 Return on Assets, Earnings per Share and Return on Equity



Graph 3 Market value to book value & Tobin's Q

Table 2 Descriptive statistics and normality test

Indices	Mean	St. dev.	KS[a]	P value
EPS	−0.058	0.341	0.246	0.01
ROE	−0.1421	0.895	0.346	0.01
ROA	−0.004	0.126	0.204	0.01
MV/BV	0.478	0.683	0.244	0.01
Tobin's Q	0.297	0.309	0.206	0.01
NAVPS	4.677	2.896	0.154	0.01

[a]Kolmogorov-Smirnov test

Secondly, in order to test our first hypothesis we searched for significant differences in profitability and firms' value between the two examined subperiods. Kolmogorov and Smirnov's Normality test has shown that none of the examined variables was normally distributed. For this reason, we implemented a non-parametric test for the detection of significant differences (Table 2). Also, high volatility of the examined variables was observed. Since the non parametric

Table 3 Descriptive statistics for the examined period

	Before crisis		After crisis	
Indices	Mean	St. dev.	Mean	St. dev.
EPS	0.05	0.149	−0.183	0.453
ROE	0.016	0.185	−0.341	1.305
ROA	0.033	0.113	−0.044	0.140
MV/BV	0.439	0.309	0.512	0.963
Tobin's Q	0.405	0.351	0.165	0.207
NAVPS	4.378	2.442	4.950	3.38

Table 4 Kruskal Wallis test for differences in profitability and value of F&B firms for the periods before and after the financial crisis

	Pre-crisis	Crisis period		
Indices	Median	Median	H	P value
EPS	−0.035	0.051	16.55	0.000
ROE	−0.015	0.031	12.85	0.000
ROA	−0.009	0.018	20.14	0.000
MV/BV	0.3355	0.3575	0.00	0.974
Tobin's Q	0.100	0.400	48.85	0.000
NAVPS	4.220	3.680	1.02	0.312

tests for the detection of differences of means are less sensitive to high volatility and are more appropriate for non-normal distributed variables we used the non parametric test of Kruskal and Wallis.

Looking at the mean values of EPS, ROA and ROE (Table 3) we can see that the profitability of food and beverage firms after the crisis became lower than the period before the beginning of the crisis. However, MV/BV, Tobin's Q and NAVPS seem to be increased during the period of financial crisis of the Greek economy. According to these findings, shareholders and investors did not seem to lose their interest in food and beverage firms. Also, though firms reported losses in their financial statements their net asset value per share was increased during the period of economic crisis.

Hypothesis 1

Implementing the test of Kruskal and Wallis (1952)[6] (Table 4) we can argue,[7] that reported earnings of F&B firms in the economic crisis period are significantly lower than in the period before the crisis. These findings are relevant with those of sectorial studies which have shown that economic crisis affected negatively the profitability of F&B firms. In contrast, no significant difference in the MV/BV and NAVPS indicators has been observed in a significance level of 5% (Table 4). Hence, we have found no evidence for significant effect of the economic crisis on firms' value as it has been reported in their financial statements on the one hand.

[6] Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one criterion variance analysis. *Journal of the American Statistical Association, 47*, 583–621.

[7] At a level of significance less than 1%.

Table 5 Effect of liquidity and debt ratio on ROA and Q ratio

		ROA				Q ratio		
	Coef	P value	R^2	VIF	Coef	P value	R^2	VIF
Before crisis								
Liquidity	0.017	**0.031***	0.116	1.07	−0.045	0.337	0.021	1.11
Debt ratio	−0.049	0.324			−0.211	0.436		
During crisis								
Liquidity	0.002	0.738	0.593	1.21	0.0296	**0.057** **	0.193	1.21
Debt ratio	−0.3060	**0.000** *			−0.2142	**0.001***		

Level of significance 5%* and 10%**

Furthermore, shareholder and investors keep their interest in these firms. However, Q ratio an MV/BV are lower enough than 1 which means that F&B firms were undervalued before the crisis too. Also, we found that Tobin's Q ratio of F&B firms before the crisis is significantly higher than during the period of crisis. These findings are not to the same direction with those of MV/BV ratio but they could be an alarm for the impact of the crisis on the managerial performance of F&B firms.

Overall, we could agree with the results of the sectorial studies which have also been based on reported financial data. According to these studies, economic crisis had a negative impact on F&B business sector but not too much as in other sectors. Hence, we could not claim that the economic crisis had a significant negative effect, either on food and beverage firms' book value or on their market value.

Hypothesis 2

In the second hypothesis of this study we explored how the liquidity and debt of F&B firms has been related with profitability. Specifically we answered to the question "Did the crisis affect on that relationship?" Our prospect was not the development of a forecasting model but the detection of significant effects within the examined variables.

In the Table 5 we observed that ROA is significantly[8] positively affected by the liquidity and negatively but insignificantly by the debt in the sub period before the beginning of crisis in Greece. However only the 11.6% of the ROA's variation has been explained by the regression equation. Nevertheless in the period of crisis the debt of F&B firms was significantly negatively related with their profitability and with their Q ratio. These findings could mean that firms which did not controlled their debt, during the first years of the crisis, became less profitable and their market value has been expected to be negatively affected. Moreover we found evidence the financial crisis affected the significance of the observed relationship between the examined variables (Table 5).Level of signicance 5%

[8]At a level of significance 5%.

Table 6 Effect of liquidity and debt ratio on ROE and MV/BV

		ROE				MV/BV		
	Coef	P value	R^2	VIF	Coef	P value	R^2	VIF
Before crisis								
Liquidity	0.022	0.409	0.056	1.11	−0.065	0.124	0.132	1.11
Debt ratio	−0.293	**0.073****			0.566	**0.025***		
During crisis								
Liquidity	0.019	0.839	0.071	1.21	−0.106	0.157	0.002	1.21
Debt ratio	−1.060	**0.008***			−0.281	0.352		

Level of significance 5%* and 10%**

Table 7 Effect of liquidity and debt ratio on NAVPS and EPS

		NAVPS				EPS		
	Coef	P value	R^2	VIF	Coef	P value	R^2	VIF
Before crisis								
Liquidity	0.083	0.79	0.06	1.11	0.027	0.198	0.084	1.11
Debt ratio	4.31	**0.021***			−0.229	**0.066****		
During crisis								
Liquidity	−0.459	0.094	0.019	1.21	0.008	0.759	0.478	1.21
Debt ratio	0.22	0.838			−0.897	**0.000***		

Level of significance 5%* and 10%**

Looking into the Table 6, ROE and debt have been significantly negatively related before and during the two examined sub periods. On the other hand, Market value to Book value used to be significantly positively related with the debt in the sub period before the beginning of crisis. Impact of financial crisis on the significance of the relationship between the MV/BV and debt has been also detected. The negative significant relationship between the debt and profitability—expressed by the Earnings per Share (EPS)—has been noticed in both of the two studied sub periods (Table 7). Last but not least we mention that although the debt ratio was significant related with the examined accounting and market profitability indicators at least in the one of the two searched periods, the liquidity did not seem to be significantly related with the F&B firms' profitability.

5 Conclusion

The main purpose of our research was to detect if the economic crisis in Greece had a negative impact on food and beverage listed companies on the Athens Stock Exchange. Our interest in this business sector comes from the fact that the food and beverage industry in Greece is one of the most important business sectors of the Greek economy. We found evidence that the profitability of listed food and beverage firms has been affected negatively by the economic crisis. On the other hand, there is

no evidence that food and beverage firms lost significantly their market value or their book value. Moreover, a short increase in their sales has been observed within the examined period.

Our findings are not relevant with those of Sujova (2015) who focused on the performance of Wood Industry. However the findings of this study for the listed firms are similar with those of the sectoral studies in which non-listed firms are also examined. We believe that the increased export activity of Greek firms contributed positively to the profitability performance of F&B firms. The sales performance of the listed F&B firms in the Athens Stock Exchange according to the results of this study follows more the corresponding performance of the multinational firms in Zhao, Jiang, and Li (2015) study. Furthermore in this study, we observed that debt and profitability were negatively related but there was not significant relationship between liquidity and profitability. Financial crisis affected the significance of the relationship between the debt and profitability.

A limitation of our study could be the short number of firms of our sample. However, these firms are all the listed F&B firms in the Athens Stock Exchange. The fact that no market value is available for non listed firms could be a barrier for the profitability analysis with indices such as Tobin's Q and Market Value to Book Value. A relevant comparative study within more business sectors is proposed in order to be analyzed sectors which became more sensitive to the financial crisis. Also the analysis of the period of the imposition of capital controls in the Greek banking system constitutes an extension of this study. Our sample's size could be a limitation on our research. We propose that a further analysis of the managerial and financial performance could be an extension of this study which will not only based on reported financial information. The discussion and analysis of the CEO's and financial managers' opinion on the impact of the crisis on F&B firms could explain a part of the variability which has been observed during this research study. Mylonas (2015) proposed—in a sectoral study of the National Bank of Greece—that the promotion of the Greek food products from the category of perishable commodities to high priced consumer goods should be a strategic goal of Greek food production firms. A future research could focus on the profitability of firms which produce high priced food and beverage products and also on the effect of global economic crisis on these firms.

References

Abdullah, M. A. (2005). Capital structure and debt maturity: Evidence from listed companies in Sandi Arabia. *Journal of Business & Economics, 11*, 15–33.

Bhunia, A. (2013). Importance of liquidity management on profitability. *Business Perspectives, 4*, 43–53.

Canibano, L., Ayuso, M. G., & Sanchez, M. P. (2000). Shortcomings in the measurement of innovation: Implications for accounting standards setting. *Journal of Management and Governance, 4*, 319–342.

Castaner, F. (2009). The food market in times of crisis. *Paradigmes* (2), 35–42.

Christodoulakis, I., & Stathis, E. (2014). The competitiveness of dynamic sectors of Greek economy. In M. Masourakis & C. H. Gortsos (Eds.), *Competitiveness for the development: Policy proposals*. Athens: Hellenic Bank Association.

Claessens, S., Djankov, S., & Lang, H. P. L. (2000). The separation of ownership and control in East Asian Corporations. *Journal of Financial Economics, 58*(1–2), 81–112.

Courtis, J. K. (1978). Modelling a financial ratios categoric framework. *Journal of Business Finance & Accounting, 5*, 371–386.

Deloof, M. (2003). Does working capital management affect profitability of Belgian firms? *Journal of Business Finance and Accounting, 30*(3 & 4), 573–588.

Egbide, B. C., Uwuigbe, O., & Uwalomwa, U. (2013). Liquidity management and profitability of manufacturing companies in Nigeria. *Journal of Business and Management, 9*(1), 13–21.

El Bilali, H., Panin, B., Berjan, S., Driouech, N., Despotovic, A., & Zucaro, N. M. (2012). Impacts of the global financial and economic crisis on the agro-food industry and rural livelihoods in Serbia. *International Journal APSTRACT – Applied Studies in Agribusiness and Commerce, 6* (1–2), 113–118. Retrieved from http://ageconsearch.umn.edu/bitstream/138096/2/15_Bilali_Apstract.pdf

Enqvist, J., Graham, M., & Nikkinen, J. (2014). The impact of working capital management on firm profitability in different business cycles: Evidence from Finland. *Research in International Business and Finance, 32*, 36–49.

European Commission. (2013, July). *European Economy: The second economic adjustment programme for Greece.* Third Review.

Foundation for Economic and Industrial Research "IOBE". (2015). *Food and beverage industry: Facts and figures*. Annual sectoral report. Retrieved from http://iobe.gr/docs/research/RES_01_7042016_REP_GR.pdf

Fu, L., Singhal, R., & Parkash, M. (2016). Tobin's Q ratio and firm performance. *International Research Journal of Applied Finance, VII*(4), 1–10.

Heikal, M., Khaddafi, M., & Ummah, A. (2014). Influence analysis of return on assets (ROA), return on equity (ROE), net profit margin (NPM), debt to equity ratio (DER), and current ratio (CR), against corporate profit growth in automotive in Indonesia stock exchange. *International Journal of Academic Research in Business and Social Sciences, 4*(12), 101–114.

Iraizoz, B., Bardaji, I., & Rapun, M. (2005). The Spanish beef sector in the 1990s: Impact of the BSE crisis on efficiency and profitability. *Applied Economics, 37*(4), 473–484. https://doi.org/10.1080/0003684042000295359.

Kajirwa, H. I. (2015). Effects of debt on firm performance: A survey of commercial banks listed on Nairobi securities exchange. *Journal of Advanced Research, 2*(6), 1025–1029.

Kalantonis, P., Gaganis, C., & Zopounidis, C. (2014). The role of financial statements in the prediction of innovating firms: Empirical evidence from Greece. *Operational Research – An International Journal, 14*(3), 439–451.

Kontogeorgos, A., Pendaraki, K., & Chatzitheodoridis, F. (2017). Economic crisis and firms' performance: Empirical evidence for the Greek cheese industry. *Revista Galega de Economia, 26*(1), 73–82.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association, 47*(260), 583–621. https://doi.org/10.1080/01621459.1952.10483441.

Lang, L. H. P., Stultz, R. M., & Walking, R. A. (1989). Managerial performance, Tobin's Q and the gains from successful tender offers. *Journal of Financial Economics, 24*, 137–154.

Lazaridis, J., & Tryfonidis, D. (2006). Relationship between working capital management and profitability of listed companies in the Athens stock exchange. *Journal of Financial Management and Analysis, 19*, 26–35.

Lien, L. B. (2010). *Recessions across industries: A survey*. Working paper no. 16/10. Bergen, SNF project no 1306 Crisis, Restructuring and Growth.

Mylonas, P. (2015). *Unlocking the potential of Greek agro-food industry*. National Bank of Greece, Sectoral report. Retrieved from https://www.nbg.gr/greek/the-group/press-office/e-spot/reports/Documents/Sectoral%20Report_Agriculture%202015.pdf

National Documentation Centre. (2015a). *Research & development expenditure and personnel in Greece in 2013*. Retrieved from https://metrics.ekt.gr/en/publications/95

National Documentation Centre. (2015b). *Innovation in Greek enterprises 2010–2012, sector of food and beverage* (Original title in Greek: "Καινοτομία στις ελληνικές επιχειρήσεις 2010–2012, κλάδος Τροφίμων & Ποτών" Εθνικό Κέντρο Τεκμηρίωσης, 2015). Retrieved from https://metrics.ekt.gr/sites/metrics-ekt/files/ekdoseis-pdf/2019/CIS_2010-2012_Greece_FoodBeverages_el.pdf

Notta, O., & Vlachvei, A. (2014). The impact of financial crisis on firm performance in case of Greek food manufacturing firms. *Procedia Economics and Finance, 14*, 454–460.

Notta, O., & Vlachvei, A. (2018). Effects of the Greek financial crisis to the food manufacturing firms. In N. Tsounis & A. Vlachvei (Eds.), *International Conference on Applied Economics* (pp. 547–560). Cham: Springer.

Owolabi, S. A., Obiakor, R. T., & Okwu, A. T. (2011). Investigating liquidity-profitability relationship in business organizations: A study of selected quoted companies in Nigeria. *British Journal of Economics, Finance and Management Sciences, 1*(2), 11–29.

Palepu, K., Healy, P., & Bernard, V. (2004). *Business analysis and valuation using financial statement*. Mason, OH: Thomson.

Penman, S. H. (2007). *Financial statement analysis and security valuation*. Sydney: McGraw-Hill Higher Education.

Pricewaterhouse Coopers (PWC). (2018, June). *Food and beverage industry. On the verge of change*. www.pwc.com. Retrieved from https://www.pwc.com/gr/en/publications/greek-thought-leadership/the-food%20&%20beverage-sector-in-greece(en).pdf

Saleem, Q., & Rehman, U. R. (2011). Impacts of liquidity ratios on profitability (case of oil and gas companies of Pakistan). *Interdisciplinary Journal of Research in Business, 1*(7), 95–98.

Saluja, P., & Kumar, P. (2012). Liquidity and profitability trade off (a study on Airtel Bharti Limited). *International Journal of Advanced Research in Management and Social Sciences, 1*(3), 77–84.

Sujova, A. (2015). Influence of the economic crisis in 2008 on the performance of companies in wood-processing industry. *Procedia Economics and Finance, 34*, 581–586.

Suzuki, K., & Chida, R. (2017). Contribution of R&D capital to differences in Tobin's q among Japanese manufacturing firms: Evidence from an investment-based asset pricing model. *Journal of the Japanese and International Economies, 43*(C), 38–58.

Wolfe, J., & Sauaia, A. C. A. (2003). The Tobin Q as a company performance indicator. *Developments in Business Simulation and Experiential Learning, 30*, 155–159.

www.businessdictionary.com

Zhao, X., Jiang, X., & Li, Z. (2015). The impact of the economic crisis on the financial performance of multinational corporations. *International Review of Economics & Finance, 37*(C), 55–68.

Fiscal Multipliers Under Extreme Uncertainty: Case of Greek Tourism Economy

Georgios Alexopoulos, Alexandros Apostolakis, Constantin Zopounidis, Alexandros Garefalakis, and Marianna Eskantar

Abstract This article explores how the impact of the economic strategy on performance can shift depending on whether the economy is in progress or recession. This particular strategy seems to affect the tourism sector especially in the central regions of Greek territory and not in the areas bordering sea tourism destinations. Extensions and bends are characterized by the indication of the performance slit (positive and negative, separately). The choice to use the yield segment as a marginal variable is promoted by a few elements, one of which is that under a negative performance gap, free from the GDP growth rate, abundance limits are predictable for the economy, reducing the decline in private business measures after a stunning administration.

Keywords GDP growth rate · Tourism sector · Performance slit · Economy strategy

1 Introduction

Theoretically, economics is the name earning and consuming the money in a professional, ethical and effective manner. Economics is the study of how to earn the money and how to consume the same in an ethical manner. There are two main parts of Economics are there known as Microeconomics and Macroeconomics. Microeconomics is the branch of economics that deals with the demand and supply factor of the individuals in particular, while macroeconomics is a branch of economics that specifically deals with the economy as a whole.

G. Alexopoulos (✉)
Department of Social and Educational Policy, University of Peloponnese, Korinthos, Greece
e-mail: alexopoulosga@upatras.gr

A. Apostolakis · A. Garefalakis (✉)
Department of Business Administration and Tourism, Hellenic Mediterranean University (HMU), Crete, Greece

C. Zopounidis · M. Eskantar
Financial Engineering Laboratory, Technical University of Crete, Chania, Greece
e-mail: kostas@dpem.tuc.gr

E. Krassadaki et al. (eds.), *Operational Research in Agriculture and Tourism*, Cooperative Management, https://doi.org/10.1007/978-3-030-38766-2_5

It is more than essential for an economy to maintain their attentiveness in the market, and maximise the potential of the economy through the proper management of the macroeconomic variables and indicators. Gross Domestic Product (GDP) is an important macroeconomic variable that associated with the total products that have been manufactured in a specific financial year, and then selling the same in the domestic market. There are numerous elements that found interactive and highly potential for the economies to maximise their effectiveness in the market.

One of the major elements that affect positively over the income multiplier and GDP growth of a country is the monetary and fiscal policies. Monetary Multipliers have been referred as the most important and highly efficient multipliers that used by the economies to maximise their potential in the market. However, its utilisation is based on the criteria to create marginal economic benefits to a country. There are different uncertainties that found in the consortium of an economy, and among them multiplier effect is one of them.

The main perspective of this assignment is to examine the utilisation of the multipliers under extremely uncertain condition, such as economic crisis. This particular assignment is likely to undertake the situation of Greece. The main rationale behind the selection of Greece is simple, as it is one of those European based economy that affected adversely and seriously during the recent economic downturn. However, the central bank and the head headers of the country took some serious actions to overcome on the crisis to make the economy resilient again to gain its previous momentum.

2 What Are Fiscal Multipliers and How Large Are they?

Fiscal multipliers are regularly characterized as the proportion of an adjustment in yield to an exogenous and brief change in the monetary shortage as for their individual baselines (Spilimbergo, Symansky, & Schindler, 2009). Notwithstanding a broad research, there is still no accord with respect to size of monetary multipliers. They have a tendency to be littler in more open economies and in nations with bigger programmed stabilizers, however as the hypothetical and observational research propose, they contrast broadly over nations.

Monetary multipliers vary crosswise over nations on the grounds that the structure and conduct of economies are different. They additionally vary inside nations, contingent upon variables, for example, the monetary instrument actualized, the approach reaction to financial advancements, and desire arrangement by financial specialists. A great part of the late research on financial multipliers likewise proposes that the span of the multiplier may likewise rely on upon the condition of the economy (Auerbach & Gorodnichenko, 2012a, 2012b, 2012c; DeLong & Summers, 2012). Others have concentrated on recognizing joins between the financial position and the hazard premium on government lending, which is of specific significance in

the euro zone (Arghyrou & Kontonikas, 2011; Bernoth & Erdogan, 2012; Corsetti, Kuester, Meier, & Meuller, 2012; De & Ji, 2012; Schuknecht & Wolswijk, 2010).

The research on the measure of monetary multipliers is broad and questionable, in both hypothetical and experimental grounds. The debate originates from the way that there is no such thing as "a financial multiplier." In exact studies, the measure of the multiplier relies upon the identification technique, the nation or gathering of nations considered, the time skyline, and the observational model. Though creators that discover low monetary multipliers underscore the part of Ricardian operators or market fulfillment, those that support bigger financial multipliers push the part of dependable guideline or hand-to-mouth family units, non-Ricardian conduct, limited lifetimes, or ostensible rigidities. An incredible number of empirical analysis (Barro, 1981; Barro & Redlick, 2011; Hall, 1986; Perotti, 2005, 2012) ordinarily puts the administration spending multiplier inside the 0.5–1 territory, however different studies (Beetsma, Massimo, & Klaassen, 2008; Blanchard & Roberto, 2002; Ramey, 2011; Ramey & Shapiro, 1998). propose a higher interim, between 0.8–1.5.3 Tax multipliers are tended to, entomb alia, (Favero & Francesco, 2012; Perotti, 2012), the previous closing for a duty multiplier beneath one, and the last for an assessment multiplier in the 1–1.5 territory. These outcomes appear differently in relation to those in (Romer & Romer, 2010), who evaluate a duty multiplier involved somewhere around 2.5 and 3 (Chahrour, Schmitt-Grohe, & Uribe, 2012). Propose that such differentiating results are clarified by unmistakable estimation procedures, as various models recognize diverse expense stuns (Ilzetzki, Mendoza, & Vegh, 2010). Find that financial multipliers have a tendency to be little in the short run; however they are significantly bigger, and conceivably well above one, in the medium and long run. In a late article (Blanchard & Daniel, 2013; Blanchard & Roberto, 2002) contend that financial multipliers connected with arranged monetary combinations amid the Great Recession are around 0.7–1 rate focuses bigger than those epitomized in consistent estimating practices by approach foundations, however discover no proof of deliberate figure blunders for the pre-emergency period. This study, close by with the substantial contractionary impacts of financial solidification over the globe that are being watched, set off the open deliberation about the genuine size of monetary multipliers in times of emergency, recommending that they can along rely upon the business position of the economy. Despite the fact that DSGE models are for the most part state free, and in this manner not able to produce endogenously monetary multipliers that are state unforeseen, late hypothetical work (Christiano, Eichenbaum, & Rebelo, 2011; Woodford, 2011) has advanced a potential clarification for the reliance of financial multipliers on the repetitive position of the economy. On the off chance that the ostensible loan cost ties at the zero bound, the swarming out impacts of government lending over private utilization and private venture are lessened or even dispensed with, and the monetary multiplier can in this manner take bigger qualities, perhaps bigger than two. This compares and a multiplier beneath one when money related approach is represented by a Taylor control at positive premium rates. Six (Gali, David Lopez-Salido, & Valles, 2007) were likewise ready to produce government spending multipliers as high as two just if

the share of general guideline customers is adequately extensive and business is request decided.

Job market establishments assume a key part in this result, in particular by setting least ostensible wage levels or requiring ostensible wage slices to be liable to a common assertion amongst bosses and representatives. Aggregate wage understandings, especially in social orders where unions assume a vital part in wage setting, are likewise a wellspring of descending compensation unbending nature (Baker, Bloom, & Davis, 2016; Druant et al., 2012; Messina, Duarte, Izquierdo, Caju, & Hansen, 2010). Costs may likewise respond unevenly to stuns, as hypothesized (Laurence & Mankiw, 1994). Positive stuns make a bigger crevice amongst fancied and real relative costs than negative stuns do, since pattern swelling causes the relative costs to decrease between modifications. Subsequently, when confronting a negative stun, firms may basically let drift swelling modify relative costs downwards, hence keeping away from to pay modification costs and the value changes. By complexity, firms are all the more ready to pay the alteration cost and change costs in case of a positive stun, as the craved relative cost increments yet the real relative cost is declining because of pattern expansion. Thus, positive value alterations that happen are bigger than negative value modification (Gilchrist, Shoenle, Sim, & Zakrajsek, 2013). Set forward an alternate contention, proposing that value changes might be contingent on firms' monetary positions. While firms with more grounded money related positions can bring down costs, firms with weaker monetary records may need to expand costs saving in mind aiming to raise enough inner assets to respect the obligation benefit. Consequently, value elements get to be constricted in light of contractionary request stuns.

3 The Fiscal Multiplier

The term financial multiplier alludes to the proportion of an adjustment in yield (ΔY) to an exogenous change in the monetary adjust (ΔG if the last alludes to an adjustment in government spending or ΔT on the off chance that it identifies with an adjustment in government revenue). Two depending on the time skyline considered, there are a few pertinent proportions that fit the term of fiscal multiplier:

The impact multiplier, characterized as the proportion of a contemporaneous change in yield (at time t_0) to an exogenous change in the financial adjust at time t_0 ($\equiv \Delta Y(t_0)/\Delta G(t_0)$)

$$\equiv \frac{\Delta Y(t_0)}{\Delta G(t_0)}$$

The multiplier at some future point in time (say, N period from now), characterized as the proportion of an adjustment in yield at time $t_0 + N$ to an exogenous change in the monetary adjust at time t_0

$$\equiv \frac{\Delta Y(t_0 + N)}{\Delta \mathrm{G}(t_0)}$$

The cumulative multiplier, characterized as the proportion of the total change in yield over an exogenous change in the financial adjust over a period skyline of N periods

$$\sum \frac{\Delta Y(t_0 + N)}{\Delta \mathrm{G}(t_0)} \text{ with i} = 0, 1, \ldots, \mathrm{N}$$

The maximum or peak multiplier, characterized as the proportion of the biggest change in yield over at whatever time skyline N to an exogenous change in the financial adjust at time t_0.

Then, the formulation of the problem is the following:

$$(\equiv \max \Delta Y(t_0 + \mathrm{N})/\Delta G(t_0), \text{for each N}).$$

$$\equiv max \frac{\Delta Y(t_0 + N)}{\Delta \mathrm{G}(t_0)} \text{ for each N}$$

4 The Principles of the Fiscal Multiplier

Earlier hypothetical and experimental work on the reaction of primary macroeconomic totals to exogenous financial stuns has demonstrated that the size (and, in specific occurrences, the sign) of the monetary multiplier can be nation, time, estimation strategy, and particular financial conditions. Generally, it creates the impression that very various perspectives keep on existing among expert financial specialists and approach producers as respects both the quantitative and subjective impacts of monetary strategy (Ilzetzki et al., 2010). Give a very telling case of the continuous differences in the financial matters calling with respect to the span of the monetary multiplier. The writers alluded to a January 2009 Wall Street Journal piece, in which (Barro, 1981) contended that peacetime monetary multipliers were basically zero, while, on the other hand, Chair of President Obama's Council of Economic Advisers. Christina Romer, utilized multipliers as high as 1.6 in assessing the occupation picks up that would be produced by the $787 billion boost bundle endorsed by Congress in February 2009.

5 Do Multipliers Differ in Downturns and Expansions

Albeit most studies don't recognize multipliers in various basic conditions of the economy, the impacts of financial approach stuns on monetary action are likely nonlinear. Multipliers could be essentially bigger in downturns than in developments. In times of a negative yield crevice, the customary swarming out contention—that higher government spending uproots private spending—is for the most part less material since abundance limits are accessible in the economy. Moreover, the extent of credit-obliged families and firms, which modify spending in light of an adjustment in extra cash, is higher. Shockingly few studies have attempted to recognize multipliers in downturns and developments. These have generally centered around a solitary nation (Germany: (Baum & Koester, 2011) and the United States: (Auerbach & Gorodnichenko, 2012)), or utilized a board information approach, consequently giving normal multipliers crosswise over nations, which may veil critical heterogeneities in the estimation procedure (Auerbach & Gorodnichenko, 2012a, 2012b). A late study that is close to this paper and that recognizes multipliers on a nation by-nation premise is the work by (Batini, Callegari, & Melina, 2012). Utilizing administration exchanging VARs with yield development as the edge variable, the paper concentrates on collaborations amongst financial and money related arrangements.

6 Short-Term Impact Multipliers

One of the researches (Barro, 1981) present that multipliers are time and state subordinate. Financial multipliers contrast crosswise over nations on the grounds that the structure and conduct of economies vary. They additionally contrast inside nations, contingent upon elements, for example, the monetary instrument executed, the approach reaction to financial advancements, and desire arrangement by financial agents. Thus there is no single "multiplier" that can be credited to a given economy, as the effect of a financial development on GDP relies upon an extensive variety of components. In this area they deteriorate the effect of a few of these elements by transforming each one every time.

7 Economy's Structure

Nation size is a vital recognizing variable crosswise over nation multipliers, as the long haul fall in genuine loan fees that is created by solidifications is a universal marvel. At the point when capital moves unreservedly between nations, genuine financing costs are resolved generally by the harmony between worldwide sparing and worldwide venture, and vast nations, for example, the United States have a great

deal more effect than little ones, for example, Greece. What's more, the underlying financing cost reaction will be littler in nations in EMU in light of the fact that the ECB reacts to euro range swelling. Multipliers have a tendency to be littler in more open economies, in light of the fact that the more open an economy is the to a greater extent a stun will spread into different nations through imports, and little open economies, for example, Belgium have little multipliers. Another organizing component is the level of reliance of utilization on current wage. This is frequently identified with liquidity requirements, with a higher current-wage versatility more normal in monetarily unliberalised economies, for example, Greece than in Belgium or the United States.

8 Analysis and Discussion

Greece with an official name of Hellenic Republic is basically a country located in South Eastern Europe. The overall population of Greece is comprises of 10.9 million people. Athens is the known as the largest country of the world. Strategically, Greece is located into two different locations which are Asian and Europe. It is developed country of the world which affected heavily during the recent economic downturn. In the year 2016, the GDP of the country went on a level of US$ 287.11 billion with per capita income amounted to US$ 26,606 in the same year. Greece is one of those European countries of the world that affected adversely and negatively during the recent economic crisis, and the management of the country did a fantastic and outstanding job as far as overcoming on the potential problem is concerned. Greece with high Health and Development Index (HDI) is doing an exceptional job in brining economic prosperity towards the economy. The uncertain situation related with Greece stated from the year 2007 in which the economic crisis started to reflect over the screen in different parts of the world. Uncertainty in the economic well-being and inefficiency in the management of the economic functions are some of the major reasons that induced the country and its head headers to take effective monetary policy actions. Before going in the details about the monetary multipliers and monetary actions, it is imperative to get an idea about the generated and recognised GDP of the country.

The tourist reasons that prevent people from traveling to specific destinations were investigated in cases of countries threatened by a financial crisis. Several studies investigate the impact of natural, financial and political crises in the tourism sector (Garefalakis, Lemonakis, Alexopoulos, & Tabouratzi, 2017) informing us that crises lead to a reduction in international tourist demand and have a negative impact on tourist areas (Lemonakis, Garefalakis, Georgios, & Haritaki, 2018; Lemonakis, Garefalakis, Giannarakis, Tabouratzi, & Zopounidis, 2017), identifying issues of economic security and postponement of tourist travels. On the other hand, they ignore the fact that tourists perceive the danger in different situations and to different tourist destinations (Gkillas, Vortelinos, Floros, Garefalakis, & Sariannidis, 2019) secondly, studies focusing on the impact of economic crises on trips explore the

potential reactions of tourists immediately after a financial crisis (Lemonakis et al., 2018). Nevertheless, the response mechanisms facing an economic crisis can achieve a rapid recovery, especially at the level of Greek Tourist layout. However, this may not be the case in all the central regions that are undergoing continuous economic rebounds, but this is mainly the case in areas bordering on maritime tourist destinations.

9 GDP Analysis of Greece

GDP is basically an acronym of Gross Domestic Product, and has been referred the most important element that stride under the umbrella of macroeconomics. It is an important and highly intact tool that consider by the companies to analyse its financial and economic position. The countries which are having higher GDP are the one that can maximise their potential in this competitive market for a long span of time. There are different economies are there which have been striving in different parts of the world to maximise their income potential. Likewise other countries, the same is applicable with Greece as well. It should be undertaken by the researcher that Greece is the country that affected heavily due to the recent economic crisis. This particular part of the assignment is likely to undertake an economic analysis pertaining to the GDP Growth of Greece, related to both, pre and post crisis period. Mentioned below graph is showing the same thing in details (Fig. 1).

The aforementioned chart is clearly showing that the growth of the GDP of Greece started to lose its momentum even from the fiscal year 2006. Greece was one of the fastest growing economies of the world before the arrival of the recent economic downturn, which started lose its momentum from the commencement of the recent economic collapse. By the midst of the year 2009, the GDP growth of the country went to a level of −3%, which further decreased to a level of −10% in the

Fig. 1 Greece GDP annual growth rate

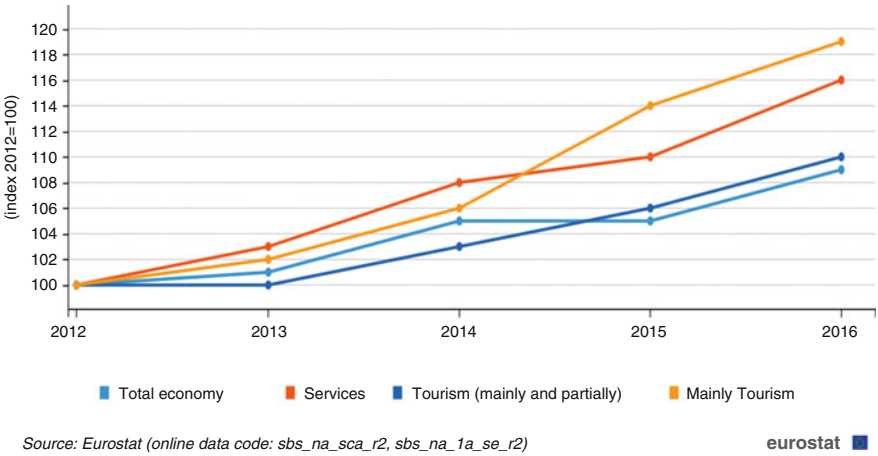


Fig. 2 Relation between GDP and GDP of tourism economy

financial year 2011. The power of resilience of the economy is not efficient enough due to which they showed their reluctance to make over it again. The country hadn't received any positivity in their economy by the end of the year 2014. After high amount of Bail out Packages and Assisting Grants from the members of the European Union, the country managed to extract out from the negative growth of the GDP to be covering up in the positivity after a year. This particular factor can be of high potential for the country in particular. One thing is for sure from this entire analysis is that the method of management in Greece is not suitable along with the economic capabilities that should have been located on a superior position. This particular factor is the one that induced the head headers of the country to regain the momentum in the market, and bring new monetary multiples into the consideration.

Taxes related to tourism, such as hotel room taxes. There is a relationship of taxes with the tourism sector, the local government directly affecting the central regions. In particular, the ability of room taxes to be exported by residents and supplying national, state and local governments is crucial to ensuring their use, they do not necessarily improve those areas. Hoteliers of island regions may be more burdened by other businesses if consumers are price-sensitive especially in the event of a financial crisis. Greek hoteliers cannot adjust the offer of hotel rooms, if the tax reduces the number of tourists, the cost remains largely variable depending on the Greek regions (Fig. 2).

Although a small tax increase will increase total tax collection for a local public sector, the percentage reduction in tax revenue cannot lead to a reduction in hotel consumption. Reduced consumption is burdened by hotel owners especially in the central areas (Fig. 3).

It is likely that hotels will be able to reduce the need for a lower-capacity workforce, but this would mean that the burden is greatly passed on to workers residing in the area and losing their jobs (Fig. 4).

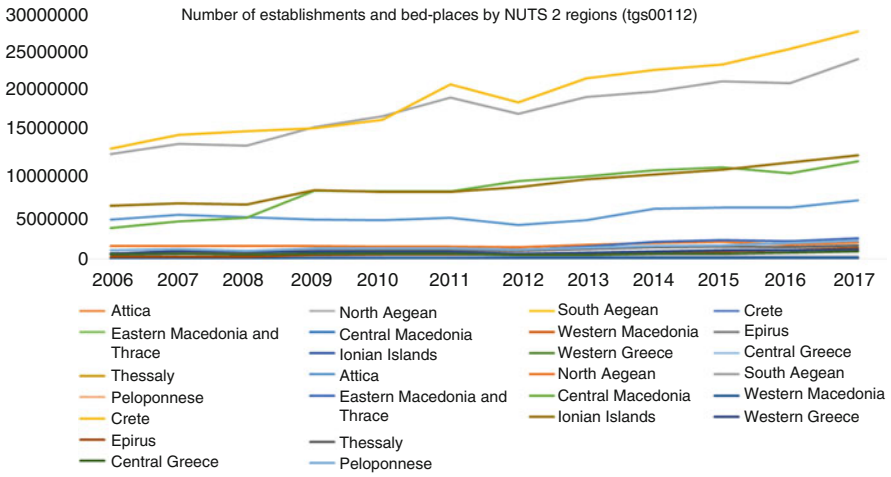


Fig. 3 Numbers of establishments and bed-places by Nuts2 regions

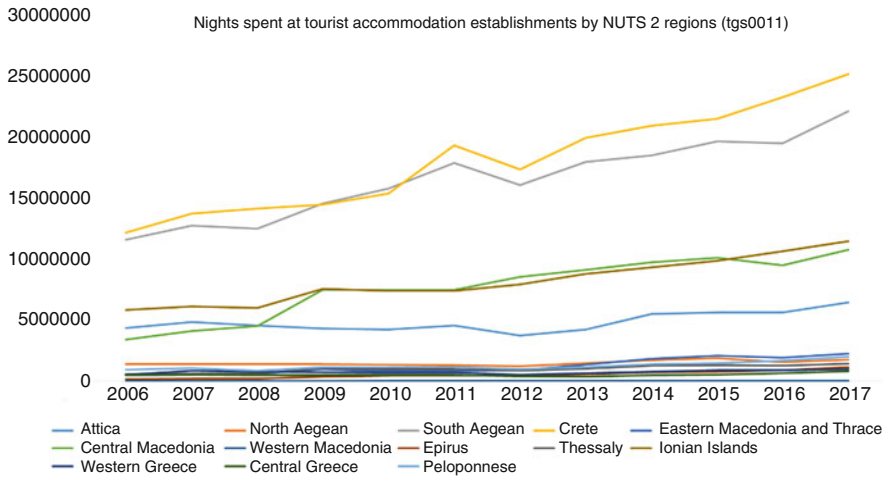


Fig. 4 Nights spent at tourist accommodation establishments by Nuts2 regions

10 Inadequacies of the Greek Adjustment

The Adjustment Program for Greece (APG) included broad and all inclusive cuts in use, a few expense climbs on family units and property, and an aggressive arrangement of market changes and privatizations to draw in new ventures and advance development. In this procedure an excessive number of things turned out badly and, a long way from containing monetary awkward nature in time, most targets were missed and the obligation load kept on rising. Despite the fact that specific change activities, connected in the Social Security framework, the Health part and in some

budgetary strategies, went in the right heading, different endeavors, for example, the progression of professional licenses in different divisions neglected to goad development as retreat was repressing potential new participants. In improving open substances or privatization arranges, the disappointment was much more broad disregarding costly arrangements and political costs (Christodoulakis, 2011). Overlooking the notices that profound slices in wages would prompt profound retreat, the powers added salt to the injury by receiving positive input strategies; at whatever point a financial target was missed, the measures—as opposed to being reevaluated—were re-escalated. Such strategies depleted the expendable wages of family units and affected incomes.

11 Debt Crisis in Greece

Since the first troika arrangement, every quarter of Greece missed its financial targets and the Greek government has declared extra severity activities. None of these activities tempered the ascent in risk. An October 2011 EU summit tried to quiet budgetary markets by reasserting part nations' duties to practical monetary strategies. Summit pioneers approached Private Sector Involvement (PSI) to help Greece achieve a 120% obligation GDP proportion. Under PSI, private financial specialists would consent to a 50% cut on Greek bond possessions, while euro zone nations would give 30 billion euro to the PSI bundle and add to recapitalizing Greek banks (European Union, 2011). Inside a month of the summit, premia had risen more than 500 premise focuses. In early November 2011, Lucas Papademos, previous legislative head of the Bank of Greece and Vice President of the ECB, was called to be the leader of a caretaking coalition government. On February 13, 2012 the new government endorsed the terms of the second troika bailout. Conditions incorporated a 22% cut in the lowest pay permitted by law, extensive diminishments out in the open work, and significant cuts in benefits, wellbeing, and barrier spending. Among the concessions the assertion accomplished were: an expansion to 53.5% for the cut taken by private investors, a timetable for coupon installments on Greek securities through 2042, a diminishment of financing costs of the Greek Loan Facility, and a guarantee by national bank holders of Greek bonds to pass profit from those bonds back to Greece (European Union, 2012). So far in 2012, the risk premium has balanced out at around 30 rate focuses.

12 Methods and Policy Details

We use estimates of the tax multiplier to study the influence of fiscal policy action on economic output through its influence on overall demand for goods and services. Part of the research on the magnitude of tax multipliers is centered on multipliers linked to increased government spending and large-scale tax cuts to analyze the

economic impact of a wide range of possible fiscal policy changes. To exhaust fluctuations, we use an estimation method that can use different policy details. The tax multiplication of the method results from two separate results. In particular, it boosts the multiplier to a direct effect (the effect of one euro on transfers or costs on the demand for goods and services) and an indirect effect (the impact on output that results when the direct impact is passed on to the economy as a whole). Using this method, the contribution to producing a change in euro area fiscal policy can be written as follows: (1 euro reduction from the cost budget) × [Tax multiplier] = (Change of cost budget) × [(Direct effect on demand) (Indirect effect on demand)] = Change in expenses. The direct effect of a different input of variables into fiscal policy depends on the details of this policy. In the case of a rise (decrease) in state markets, the direct result is one, because demand is rising (decreasing) by one euro. In the case of a rise in (reduction) in taxes or dollar transfers, the direct effect may differ significantly between fiscal policy provisions (mainly because they affect people with different characteristics and can be considered as more or less persistent, resulting in different reactions, as discussed above). For example, CRA's analysis is based on different estimates of the direct impact on eight major tax and transferable provisions of this legislation (Central Bank of Greece Policy) This study analyzes the effects of contributions to central fiscal policies in the Greek economy in the short and long term. In the short term, changes in fiscal policies affect economic output mainly by affecting demand for goods and services. In the long run, however, we assume that real output is close to rising output, changes in fiscal policies affect production primarily by changing incentives for work, saving and investing individuals and businesses. If the labor market is declining, changes in work incentives may also affect production and employment in the short term, but this result was usually small in our analyzes.

The indirect effect of a contribution on Greek policies offsets or strengthens the direct effect. For example, the direct effects of lower taxes or higher government spending are magnified when increased demand for goods and services leads businesses to increase investment and hire more workers than otherwise. On the other hand, direct effects are extinguished when, for example, the slowest government borrowing caused by tax cuts or cost increases leads to higher "crowd out" rates for investment and durable goods such as car interest rates increase the cost of borrowing from households and businesses. The consequence of this policy strengthens the overall change in euro-area output in the direct impact of demand. Given a small number of multiplication estimates and uncertainty about the economic relationships that form the basis of their assessment, this study uses several estimates in its analyzes. This range includes a broad range of researchers' views on underlying economic relationships and is updated by the results of macroeconomic and monetary models, time series models and DSGE models. Specifically, the higher segments of the study series are mainly based on macroeconomic models developed by Macroeconomic Advisers and IHS Global Insight. The lower edges are mainly based on bibliography using time series models. And research is based on the DSGE models, mainly to help understand the economic and behavioral mechanisms that form the basis of the empirical literature and to point out that changes in the behavior

Table 1 Ranges for GR fiscal multipliers

Type of activity	Estimated multipliers	
	Low estimate	High estimate
Purchases of goods and services by the Federal Government	0.02	1.2
Transfer payments to state and local governments for infrastructure	0.04	1.1
Transfer payments to state and local governments for other purposes type	1	2
Transfer payments to individuals	0.6	1.4
One-time payments to retirees	0.4	0.5
Two-year tax cuts for lower- and middle-income people	0.05	0.7
One-year tax cut for higher-income people	0.9	1.9
Extension of first-time homebuyer credit	0.5	0.9
Corporate tax provisions primarily affecting cash flow	1	2.3

Source: European central bank (ECB)
Note: The estimates above were produced for CBO's analysis of European central bank (ECB) of 2017

of businesses and consumers may negatively or positively affect multipliers. The ARRA analysis shows the results that can be obtained by combining estimates of direct and indirect monetary contributions. In particular, the ARRA method presents a series of estimates of fiscal multipliers for each major provision of this European legislation (see Table 1). In this survey, purchases of goods and services from the European Central Bank Government had the lowest estimated fiscal multiplier (estimates ranging from 0.02 to 1.2) and a set of corporate tax provisions had the highest estimated multiplier (estimates ranged from 1 to 2.3).

13 Economic Conditions and Confidence Reflections

Our assessment of the literature considers that the sizes of the multiplier used by the body vary according to the economic conditions. For example, when output is far below its potential and the European Central Bank's response to changes in fiscal policies is likely to be limited (as in recent years) when unemployment was high, inflation was high, the capacity of the European Central Bank (see Table 2) to reduce interest rates has failed to meet that multiplier ratios are greater than when production is close to or above its potential and the Greek Reserve it was unable to absorb the effects of the changes in the European directives on fiscal policies. Private spending could not be absorbed through other channels except interest rate rises. In 2017 the activities promoted by reinforcing fiscal policies could reduce the production elsewhere in the economy if they used rare materials or workers with specific skills and thus created bottlenecks that hamper the production of others. As with absorption due to the increase in short-term interest rates, the exclusion stemming from output growth points is likely to be much lower in Greece's last

Table 2 The effect of a 1€ increase in aggregate demand over eight quarters

	When output is well below		When output is close to	
Quarter	Low estimate	High estimate	Low estimate	High estimate
1	0.07	1,3	0.02	0.9
2	0.04	0.1	0.8	1.1
3	0	0.05	−0.01	1.3
4	0.01	0.16	−0.05	1.09
5	0	0.4	−0.34	−0.1
6	−0.07	0.1	−0.01	0.06
7	0.04	1.3	−0.09	1.1
8	0.005	0.2	1.09	1.29
Cumulative effect after eight quarters	0.095	3.61	1.41	6.74

recession and slow recovery (due to high unemployment and a small number of unused capital) during other periods. An important observation is worth mentioning that households will respond to a fiscal stimulus by reducing their spending in anticipation of higher taxes in the future, something that has not happened in the Greek case.

Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018) created a model in which uncertainty is time-varying and influences the volatility of technological vibrations, and businesses are heterogeneous and face the reduction cost of capital and labor adaptation. Tax policy is expressed as a wage subsidy. Practical implementation of this policy is smaller when uncertainty perceives the economy but is slightly less when the policy is executed a year later. The increasing set of reports (Baker et al., 2016; Bloom, 2009) suggests that uncertainty has a negative impact on economic activity. However, we have strong evidence on how uncertainty negatively affects the fiscal policy of Greece. This question could have important policy implications in terms of the extent to which a small intervention may be appropriate in a period of strong and stable recession. The present analysis of the repercussions and absorption of recessions of Greek fiscal policy proposing an alternative definition of crises strictly based on modern monetary and tax information. With this method, we avoid ambiguities in the bibliography and especially statistical errors that can cause the results. In addition, we mention the process and the way in which uncertainty can affect the responses of macroeconomic variables to Greek state spending and suggests an economic mechanism that could absorb the elements of the ongoing recession. Determination of HU periods given that there is no official Greek market for the Greek market, we are building the VASE-20 DTM for the corresponding period that we are looking at the other European indicators. For its construction, we follow the general VI X construction methodology with some modifications proposed by (Jiang & Tian, 2005, 2007) in order to reduce the numerical calculation errors resulting from the fact that there is a finite number of options being traded.

In some HU periods we use the methodology and data described in (Bloom, 2009; Bloom et al., 2018) manufactures a monthly measure of uncertainty using the VXO of implied volatility from 1986 onwards, in the current survey we use the actual monthly returns of the FTSE/ASE-20 index between 1970 and 2018. Given that for the Greek market there is no official DTM, we are building the VASE-20 DTM for the relevant period. For the construction of the index we use the VI X construction methodology with some modifications proposed by (Jiang & Tian, 2005, 2007) to reduce the statistical calculation errors that result from the fact that there is no finite number of options being traded. We extend these estimates until 1947. The major events of uncertainty are selected as those months with stock market volatility of 1.65 standard deviations above the trend of Hodrick-Prescott (with a normalization factor of 129,600). Since our sample has a quarterly frequency, we consider that the HU periods are the periods containing any of the monthly events described above.

14 The LU Periods Are Set to Three Months

Bloom (2009), in terms of uncertainty, did not find any absorption from macroeconomic developments. However, a potential research question may arise as to whether the measure of uncertainty should be determined taking into account other Greek government or European macroeconomic variables. While we cannot formally control the exogenity of the finders via HU, we can observe that they can be predicted from previous (or current) financial information. We are adapting the (Mertens & Ravn, 2012; Schuknecht et al., 2010) methodologies to implement two predictive tests over the 3-month periods. The first is a linear estimate of time HU for the lags of the different variables (i.e. a Granger causality test). The second method is a non-linear estimation using a probit model. In all cases, we test the value of explanatory variables as an analysis of the predictability of HU time periods. All of our interpretative variables include four GDP lags, government spending, tax revenues, consumption spending, employment, wages, taxed corporate earnings, prices, interest rates and oil prices. The results of the tests for different groups of these variables are presented in columns 1–5 of Panel A of Table 3. The Greek case considers the assumption that the previous values of the examined variables can help to predict the occurrence of the high uncertainty episodes, is clearly discarded in existing data, both in linear and non-linear models. In Panel B of Table 2, the results have similar estimates when adding the current values of explanatory variables. The statistical findings seem to suggest that even the values of the variables considered can be of considerable help in predicting the contribution of fiscal policy to the duration of the HU.

Table 3 Tests of predictability of uncertainty

	(1)	(2)	(3)	(4)	(5)	(6)
	GDP	GDP tourist sector	Fiscal	Macro	Nominal	All
Panel A						
Linear estim. (F test)	0.15	0.05	0.13	0.23	2.01	0.77
	(0.95)	(0.95)	(0.75)	(0.85)	(0.85)	(0.85)
Non-linear estim. (LR test)	0.01	0.01	1.69	7.73	5.52	74.09
	(0.95)	(0.95)	(0.85)	(0.85)	(0.85)	(0.85)
Panel B						
Linear estim. (F test)	0.30	0.08	0.17	0.99	0.55	0.26
	(0.85)	(0.95)	(0.85)	(0.95)	(0.85)	(0.85)
Non-linear estim. (LR test)	7.31	2.87	5.44	3.14	67.01	31.72
	(0.85)	(0.95)	(0.85)	(0.85)	(0.95)	(0.85)

Note: This table shows the results of predictive predictability of the virtual uncertainty variable using information from different sets of variables (column presentation). Panel A presents trials using four regressions of explanatory and dependent variables and panel B also includes modern regressors. Linear estimation reports the results of an F test that measures the relevance of the sets of variables 1–5 to the prediction of uncertainty episodes (Granger causality tests). The non-linear estimate presents the results of probit probability probability models involving the explanatory variables in columns 1–5 against the model model used which does not include them. The set of variables in columns 1–5 includes delays (and concurrent values in the case of Table B) of GDP (column 1) of government expenditure and tax revenue (column 2), consumption, investment, employment, wages and income from enterprises (column 3). Repayment of GDP, interest rates (3-month bonds) and oil prices (calculated on the basis of the median of the rate (column 4) and all variables together (column 5). All explanatory variables are expressed in logarithms and in real terms per head (except for the nominal variables in column 4) All regressions include a constant and linear and square voltage Each cell in the table shows the F/Likelihood Ratio value The value ρ associated with each structure (the higher values indicate the failure to reject the hypothesis that the variables in each regression do not contain information to explain uncertainty episodes)

15 Results

Table 3 shows the responses to government spending shocks identified using the news on future defense spending in both LU (Upper Panel) and HU (bottom). The increase in government expenditure during the LU causes a positive response to the product (although estimates are associated with relatively high levels of confidence). The cumulative multiplier of 1 year is estimated to be about 0.5. When the same shock occurs during the HU periods, the production area is negative by the tenth quarter. The cumulative multiplier of 1 year is −1.8. During the LU periods, a shock to government spending reduces output over the whole horizon. This result is significantly different from zero during the first year, with a multiplier of 0.7 during this period. On the other hand, when the shock that occurs during the HU periods, production is in line with the increase in government expenditure, it actually concludes contracts after the second year.

The objective of this study is to: (a) check that tourists expressing greater emotional solidarity with specific tourist areas; and (b) see if tourists repeat the

degree of financial contribution to Greek regions affected by the financial crisis. These specific differences in emotional solidarity and financial contribution reinforce it the island regions according to statistical data and not the central Greek regions. The study found reliable data on the positive relationship between the multipliers and the positive impact on the Greek tourism sector. First of all, this study highlights the valuable perspective of government tourism representatives to re-evaluate their tourism strategies according to the Greek regions. Benefits of the community and reorganize their attitude towards the Ministry of Tourism. In the same rationale, the economic environment can explain why there was a difference between the central tourist areas and the areas that are accompanied by maritime destinations. The impact of multipliers has directly affected the central areas, significantly reducing the financial contribution of tourism. However, tour operator managers should not consider that tourists can also have their own perceptions but constantly raise new destinations especially in the central regions in order to gradually increase their contribution to tourism gross product.

Tourists' behaviors usually act as a reaction to the behavior of financial crises (Su & Swanson, 2017), tourists are less likely to return if what they see is mainly their positive impact on the country of the tourist destination. As with many other forms of tourism, the dedication of the tourist destination varies according to the qualities of each Greek region. (e.g. custom traits and traditional tasting) (Kaplanidou, Jordan, Funk, & Ridinger, 2012). Experiences and perceptions of tourists may worsen if they feel less welcome or emotionally segregated from particular regions, especially if they notice negative results from their tourist visits. Given the pioneering nature of this research, it would be useful to compare how residents and tourists think about the impact of tourism and tourism development on specific areas affected by multiplier effects. Taking into account the facilities and services required for domestic tourism for both residents and tourists, the two have a different view of tourist destinations. As GDP from tourism appears to have a greater impact on the impact of fiscal multipliers, different analytical approaches may be useful for all Greek regions. For example, the analysis of the gross national tourist product has been more influenced by the multiplier effect and has affected the standard of living of the central regional tourist areas. A more sophisticated approach examines how the experience of tourists is directly affected by making them stop if they are of tourist interest in the central tourist regions. Although this study could have made such a tourist effort, its effort was great due to the large number of data coming from all Greek regions through the respective tourist statistical organizations. Finally, tourism scholars may need to start looking at the tax impact of multipliers on the quality of tourist accommodation throughout the Greek region. In fact, tourists would provide a good financial measure as tourist tourists often share collective beliefs and monetary behaviors that are known to encourage the local and economic development of specific areas of the Greek tourist region.

16 Conclusions and Future Research

The great recession of Greece has caused great interest in the economic impact of fiscal policy. Recent interest is reflected in an ongoing discussion about the magnitude of the tax multiplier, especially in cases of uncertainty. This research article looked at the models that economists use to assess the multiplier and how economists can use these estimates to analyze cautiously economic policy in the Greek case. Three types of models are often used to generate estimates of the fiscal multiplier: macroeconomic forecasting models, time series models and DSGE models. One basic finding is that the variance of multiplier estimates can not be fully explained by the use of various types of models by economists. The findings vary largely because analysts use a variety of valuation methods and because estimates may depend on the details of fiscal policy, the nature of economic conditions, and how fiscal policy affects confidence in economic activity. Our study changes the multiplicative series for different economic conditions. This particular strategy seems to affect the tourism sector especially in the central regions of Greek territory and not in the areas bordering sea tourism destinations. The approach of this research to assessing. The future steps of our research will be to test the tax multiplier wich involves distinguishing between the direct and indirect effects of changes in fiscal policy and then a combination of estimates of both results.

Appendix

Nights spent at tourist accommodation establishments by NUTS 2 regions (tgs00111)	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Attica	4,495,136	5,017,437	4,722,681	4,454,770	4,372,391	4,708,288	3,868,098	4,381,506	5,674,801	5,830,865	5,835,717	6,663,424
North Aegean	1,449,605	1,449,605	1,449,605	1,449,605	1,411,209	1,373,423	1,283,169	1,561,612	1,784,069	1,945,552	1,607,530	1,841,493
South Aegean	11,936,728	13,144,237	12,910,020	15,014,228	16,265,432	18,425,579	16,576,204	18,504,776	19,071,358	20,269,721	20,080,508	22,785,701
Crete	12,558,011	14,158,819	14,573,794	14,910,702	15,855,489	19,914,142	17,872,690	20,573,561	21,552,910	22,154,000	23,998,896	25,927,972
Eastern Macedonia and Thrace	522,339	579,291	631,898	1,038,884	884,571	884,571	943,008	1,407,725	1,939,169	2,158,765	2,006,680	2,349,007
Central Macedonia	3,511,475	4,240,108	4,673,721	7,742,605	7,715,904	7,715,904	8,843,134	9,429,032	10,071,849	10,426,966	9,782,234	11,126,503
Western Macedonia	37,516	37,516	37,516	37,516	39,759	39,759	43,695	50,171	46,206	42,989	42,728	64,013
Epirus	186,218	236,045	234,031	411,606	470,329	470,329	474,207	597,190	768,105	801,590	941,544	1,209,470
Thessaly	626,617	665,793	664,865	1,114,707	1,024,195	1,024,195	927,649	1,083,109	1,321,748	1,386,662	1,324,434	1,503,980
Ionian Islands	6,047,509	6,328,611	6,200,661	7,826,465	7,660,023	7,660,023	8,189,418	9,085,498	9,608,194	10,189,154	10,993,653	11,835,299
Western Greece	564,541	909,960	834,330	764,463	734,884	734,884	586,885	707,829	832,090	936,392	947,151	1,051,029
Central Greece	512,457	570,750	510,874	541,541	531,145	531,145	423,121	395,888	533,557	583,305	698,892	847,396
Peloponnese	971,102	1,110,592	908,198	1,208,322	1,182,438	1,182,438	1,023,105	1,214,743	1,471,100	1,528,563	1,765,789	2,079,099

Number of establishments and bed-places by NUTS 2 regions (tgs00112)	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Attica	63,749	64,518	64,798	77,729	77,613	77,788	73,949	73,400	76,493	74,287	73,958	73,608
North Aegean	171,268	172,140	174,982	282,249	295,337	301,795	302,067	304,247	309,495	313,551	310,391	313,345
South Aegean	111,923	110,732	111,073	173,864	176,060	172,768	167,419	165,933	170,856	172,152	171,326	173,966

(continued)

Number of establishments and bed-places by NUTS 2 regions (tgs00112)	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Crete	89,540	91,647	92,370	137,759	139,185	140,958	137,893	138,856	139,752	141,827	141,419	153,052
Eastern Macedonia and Thrace	1962	1948	1971	9048	9108	8999	8955	8993	9046	9112	9056	8845
Central Macedonia	1537	1525	1538	4507	4635	4253	4201	4228	4654	4659	4671	4751
Western Macedonia	365	385	392	1117	1128	1132	1122	1121	1141	1141	1142	1433
Epirus	1211	1238	1249	5096	5215	4795	4762	4764	4843	4870	4848	4935
Thessaly	276	276	276	276	269	248	241	238	248	249	243	224
Ionian Islands	291	310	338	1423	1541	1597	1599	1611	1678	1734	1786	1667
Western Greece	572	586	601	2514	2441	2280	2253	2253	2272	2269	2272	2291
Central Greece	918	924	927	3841	3821	3793	3773	3780	3779	3813	3825	4680
Peloponnese	274	275	284	508	544	526	501	506	533	534	536	512

References

Arghyrou, M. & Kontonikas, A. (2011). The EMU sovereign-debt crisis: Fundamentals, expectations and contagion. *European Commission economic paper, no. 436*.

Auerbach, A. J., & Gorodnichenko, Y. (2012a). Fiscal multipliers in recession and expansion. *American Economic Journal: Economic Policy, 4*(2), 1–27.

Auerbach, A. J., & Gorodnichenko, Y. (2012b). Measuring the output responses to fiscal policy. *American Economic Journal: Economic Policy, 4*, 1–27.

Auerbach, A. J., & Gorodnichenko, Y. (2012c). Fiscal multipliers in recession and expansion. In A. Alesina & F. Giavazzi (Eds.), *Fiscal policy after the financial crisis*. Chicago, IL: University of Chicago Press.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics, 131*(4), 1593–1636.

Barro, R. J. (1981). Output effects of government purchases. *Journal of Political Economy, 89*(6), 1086–1121.

Barro, R. J., & Redlick, C. J. (2011). Macroeconomic effects from government purchases and taxes. *The Quarterly Journal of Economics, 126*(1), 51–102.

Batini, N., Callegari, G., & Melina, G. (2012). *Successful austerity in the United States, Europe and Japan, IMF working paper 12/190*. Washington: International Monetary Fund.

Baum, A., & Koester, G. B. (2011). *The impact of fiscal policy on economic activity over the business cycle-evidence from a threshold VAR analysis, Bundesbank discussion paper, (03/2011)*. Frankfurt: Deutsche Bundesbank.

Beetsma, R., Massimo, G., & Klaassen, F. (2008). The effects of public spending shocks on trade balances and budget deficits in the European Union. *Journal of the European Economic Association, 6*(2–3), 414–423.

Bernoth, K., & Erdogan, B. (2012). Sovereign bond yield spreads: A time-varying coefficient approach. *Journal of International Money and Finance, 31*, 639–656.

Blanchard, O., & Daniel L. (2013). Growth forecast errors and fiscal multipliers. *Working paper 18779*. Cambridge, MA: National Bureau of Economic Research.

Blanchard, O., & Roberto, P. (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *The Quarterly Journal of Economics, 117* (4), 1329–1368.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica, 77*(3), 623–685.

Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I., & Terry, S. J. (2018). Really uncertain business cycles. *Econometrica, 86*(3), 1031–1065.

Chahrour, R., Schmitt-Grohe, S., & Uribe, M. (2012). A model-based evaluation of the debate on the size of the tax multiplier. *American Economic Journal: Economic Policy, 4*(2), 28–45.

Christiano, L., Eichenbaum, M., & Rebelo, S. (2011). When is the government spending multiplier large. *Journal of Political Economy, 119*(1), 782121.

Christodoulakis, N. (2011). From indecision to fast-track privatisations: Can Greece still do it? *National Institute Economic Review, 217*(1), R60–R74.

Corsetti, G., Kuester, K., Meier, A., & Meuller, G. (2012). Sovereign risk, fiscal policy and macroeconomic stability. *IMF working paper, 12/33*.

De Grauwe, P., & Ji, Y. (2012). Mispricing of sovereign risk and multiple Equilibria in the Eurozone. *CEPS working document, no. 361*.

DeLong, J. B., & Summers, L. H. (2012). Fiscal policy in a depressed economy. *Brookings Papers on Economic Activity*, 233–297.

Druant, M., Silvia, F., Gabor, K., Lamo, A., Martins, F., & Sabbatini, R. (2012). Firms' price and wage adjustment in Europe: Survey evidence on nominal stickiness. *Labour Economics, 19*(5), 772–782.

European Union. (2011). *Euro summit statement*. Brussels, October 26.

European Union. (2012). *Eurogroup statement*. Brussels, February 21.

Favero, C., & Francesco, G. (2012). Measuring tax multipliers: The narrative method in fiscal VARs. *American Economic Journal: Economic Policy, 4*(2), 69–94.

Gali, J., David Lopez-Salido, J., & Valles, J. (2007). Understanding the effects of government spending on consumption. *Journal of the European Economic Association, 5*(1), 227–270.

Garefalakis, A., Lemonakis, C., Alexopoulos, G., & Tabouratzi, E. (2017). History of Greece's debt crisis and the banking policy. In *The Greek debt crisis* (pp. 177–187). Cham: Macmillan.

Gilchrist, S., Shoenle, R., Sim, J. W., & Zakrajsek, E. (2013). Inflation dynamics during the financial crisis. *2013 meeting papers 826*, Society for Economic Dynamics.

Gkillas, K., Vortelinos, D., Floros, C., Garefalakis, A., & Sariannidis, N. (2019). Greek sovereign crisis and European exchange rates: Effects of news releases and their providers. *Annals of Operations Research*, 1–22.

Hall, R. E. (1986). The role of consumption in economic fluctuations. In *The American business cycle: Continuity and change* (pp. 237–266). Chicago, IL: University of Chicago Press.

Ilzetzki, E., Mendoza, E. G., & Vegh, C. A.. (2010). How big (small?) are fiscal multipliers? *Working paper 16479*, National Bureau of Economic Research.

Jiang, G. J., & Tian, Y. S. (2005). The model-free implied volatility and its information content. *Review of Financial Studies, 18*, 1305–1342.

Jiang, G. J., & Tian, Y. S. (2007). Extracting model-free volatility from option rices: An examination of the VIX index. *Journal of Derivatives, 14*, 1–26.

Kaplanidou, K., Jordan, J. S., Funk, D., & Ridinger, L. L. (2012). Recurring sport events and destination image perceptions: Impact on active sport tourist behavioral intentions and place attachment. *Journal of Sport Management, 26*(3), 237–248.

Laurence, B., & Mankiw, G. (1994). Asymmetric price adjustment and economic fluctuations. *Economic Journal, 104*(423), 247–261.

Lemonakis, C., Garefalakis, A., Georgios, X., & Haritaki, H. (2018). A study of the Banks' efficiency in crisis: Empirical evidence from Eastern Europe, Balkans and Turkey. *Journal of Governance and Regulation, 7*(3), 8–12.

Lemonakis, C., Garefalakis, A., Giannarakis, G., Tabouratzi, E., & Zopounidis, E. (2017). Innovation and SMEs financial distress during the crisis period: The greek paradigm. In *The Greek debt crisis* (pp. 285–308). Cham: Macmillan.

Mertens, K., & Ravn, M. O. (2012). Empirical evidence on the aggregate effects of anticipated and unanticipated US tax policy shocks. *American Economic Journal: Economic Policy, 4*(2), 145–181.

Messina, J., Duarte, C. F., Izquierdo, M., Caju, P., & Hansen, N. L. (2010). The incidence of nominal and real wage rigidity: An individual-based sectoral approach. *Journal of the European Economic Association, 8*(2–3), 4872496.

Perotti, R. (2005). Estimating the effects of fiscal policy in OECD countries. *Discussion paper 4842*, Centre for Economic Policy Research.

Perotti, R. (2012). The effects of tax shocks on output: Not so large, but not small either. *American Economic Journal: Economic Policy, 4*(2), 214–237.

Ramey, V. A. (2011). Identifying government spending shocks: It's all in the timing. *The Quarterly Journal of Economics, 126*(1), 1–50.

Ramey, V. A., & Shapiro, M. D. (1998). Costly capital reallocation and the effects of government spending. In *Carnegie-Rochester conference series on public policy* (Vol. 48, pp. 145–194). Amsterdam: Elsevier.

Romer, C. D., & Romer, D. H. (2010). The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review, 100*(3), 763–801.

Schuknecht, L., von Hagen, J., & Wolswijk, G. (2010). Government bond risk premiums in the EU revisited. The impact of the financial crisis, *European Central Bank working paper, no. 1152*.

Spilimbergo, A., Symansky, S., & Schindler, M. (2009). Fiscal multipliers. *IMF staff position note, SPN/09/11*, May 2009. Washington: International Monetary Fund.

Su, L., & Swanson, S. R. (2017). The effect of destination social responsibility on tourist environmentally responsible behavior: Compared analysis of first-time and repeat tourists. *Tourism Management, 60*, 308–321.

Woodford, M. (2011). Simple analytics of the government expenditure multiplier. *American Economic Journal: Macroeconomics, 3*(1), 1–35.

Forecasting Tourism Demand in Europe

Dimitrios I. Vortelinos, Konstantinos Gkillas, Christos Floros, and Lavrentios Vasiliadis

Abstract We study the performance of the k nearest neighbor (*k*NN) forecasts in the context of European tourism demand. The forecasting performance of neural networks is examined across different parameterizations of the *k*NN model. The selection of the most appropriate *k*NN parametrization can produce more accurate forecasts. Tourism demand is forecast monthly for 20 European countries. Tourism demand is measured via seven variables for the reason of consistency in results. *k*NNs better forecast tourism demand in shorter horizons; in specific, 1 month ahead. The parametrization of the kNN model affects forecasting performance. More sophisticated parameterizations perform better than either an ARIMA model or a naive *k*NN parametrization. The inclusion of international stock indices significantly increases forecasting accuracy. The more explanatory variables employed, the higher forecasting accuracy is retrieved. However, there is not a specific group of stock markets affecting more the *k*NN model's forecasting accuracy. The forecasting accuracy of *k*NNs differs between three (Western, Eastern and Southern) European regions.

Keywords Nearest neighbor · Forecast · Europe · Tourism demand

JEL Classifications L83 · C53 · E37

D. I. Vortelinos
Lincoln Business School, University of Lincoln, Lincoln, UK
e-mail: dvortelinos@lincoln.ac.uk

K. Gkillas (✉)
Department of Business Administration, University of Patras, Patras, Greece
e-mail: gillask@upatras.gr

C. Floros
Department of Accounting and Finance, Hellenic Mediterranean University, Crete, Greece
e-mail: cfloros@hmu.gr

L. Vasiliadis
Department of Economics, University of Patras, Patras, Greece
e-mail: lvasil@upatras.gr

E. Krassadaki et al. (eds.), *Operational Research in Agriculture and Tourism*, Cooperative Management, https://doi.org/10.1007/978-3-030-38766-2_6

1 Introduction

It is widely acknowledged that tourism is among the most important economic activities of the modern world that contributes significantly in local, regional and national economies. Currently, tourism is considered to be an integral part of most developed and less developed economies and as such has been subject to numerous macroeconomic studies that try to explain and measure the determinants of tourism attractiveness of a destination through supply-and-demand indicators. For such an approach, the world is technically divided into a blurring framework of tourist-generating and tourist-receiving countries.

Although diverse niche markets are constantly emerging as a result of new tourist products and experiences generated by destinations, the production and consumption of tourism goods and services still functions in such a way that an origin-demand side- and a destination-supply side-exist.

We investigate the performance of the k nearest neighbor (*k*NN) forecasts in the context of European tourism demand. The forecasting performance of neural networks is examined across different parameterizations of the *k*NN model. The selection of the most appropriate *k*NN parametrization can produce more accurate forecasts. We forecast tourism demand on a monthly basis for 20 European countries. We measure tourism demand via seven variables for the reason of consistency in results. *k*NNs better forecast tourism demand in shorter horizons; in particular, 1 month ahead. The parametrization of the kNN model affects forecasting performance. More sophisticated parameterizations perform better either an ARIMA model or a naive *k*NN parametrization. The inclusion of international stock indices significantly increases forecasting accuracy. The more explanatory variables are employed, the higher forecasting accuracy is retrieved. However, there is not a specific group of stock markets affecting more the *k*NN model's forecasting accuracy. The forecasting accuracy of *k*NNs differs among three (Western, Eastern and Southern) European regions.

The remainder of this paper is organized as follows. Section 2 presents the literature review with regard to demand theory. Section 3 describes the methodology used in this paper. Section 4 presents the empirical evidence. Finally, Sect. 5 provides the concluding remarks.

2 Literature Review

The role of demand theory in exploring the growth of international tourism can be measured quantitatively by means of statistical measurements and qualitatively by analyzing the incentives that determine the specific demand. As such, the main purpose of demand theory is to interpret the international specialization of a country

through the study of these conditions that shape domestic, regional and international demand.

The growth of tourism encouraged this field of research, which focuses largely on demand and its determinants. Within this framework, there is a rich literature examining demand side factors related to the socio-economic characteristics of tourists. More precisely, tourism is a system, which is a function of supply and demand interaction according to which measuring one's destination tourist performance is a consideration of many factors related to both the characteristics of location as well as the perceptions, socio economic and political conditions that exist in the tourists' countries.

In this respect, tourism planning in either a company or regional level is assisted by accurate demand forecasts. However, it was mostly after 2000 that literature started to systematically research tourism demand forecasting. Li, Song, and Witt (2005) reviewed studies in tourism demand forecasting between 2000 and 2004. Song and Li (2008) reviewed tourism demand forecasting literature between 2004 and 2008. Many papers forecast tourism demand via linear models, like ARIMA (see Smeral & Wuger, 2005). Tourism demand was also forecast by autoregressive distributed log model (ADLM) (see Song, Wong, & Chon, 2003), error correction model (ECM) (see Kulendran & Witt, 2003), vector autoregressive model (VAR) (see Song & Witt, 2006), or time varying parameter (TVP) (see Li, Wong, Song, & Witt, 2006) models. The tourism demand was forecast also nonlinearly via GARCH modelling in Chan, Lim, and McAleer (2005). Neural networks (NN) have been employed in tourism demand forecasting only recently (Kon & Turner, 2005; Palmer, Montano, & Sese, 2006; among others). Combining forecasting methods are nonlinear forecasting methods that have relatively recently employed in tourism. These methods combine either linear or nonlinear forecasts from other models. Wong, Song, Witt and Wu employed the variance-covariance, and the discounted mean square error combining methods. Recently, Shen, Li, and Song (2011) employed the forecasting power of the simple average, the variance-covariance, the discounted-mean-square-error, the shrinkage error, and the time-varying parameter combining methods in forecasting UK outbound tourism demand.

Another nonlinear forecasting method is the *k*-Nearest Neighbor approach. *k*NN models, with their ability to adapt to data-imperfection, nonlinearity and arbiter function mapping, are a useful alternative to the static regression forecasting models. The *k*NN forecasting method is based on the weighted average of the *k* nearest neighbors of the item predicted. Weights are proportional to the proximity of the neighbor to the predicted item. The forecasting accuracy depends on the choice of the number of neighbors (*k*) and the distance metric (Kulkarni, Lugosi, & Venkatesh, 1998).

The recent literature puts emphasis on the selection of the number of the *k* nearest neighbors. This selection is empirically examined internationally for industrial production. Literature has proposed different ways to select the number of *k* nearest neighbors. If the *k* value is too small eliminates important information; if

it is too large results in noise that deteriorates forecasting accuracy.[1] One of the first papers dealing with the appropriate number of k nearest neighbors is Stone (1977). In this present paper, RMSE and MAE loss functions are employed as distance metrics. Regarding the distance metric, see Basseville (1989). The k-nearest neighbor regression method was recently empirically analyzed by Ahmed, Atiya, Gayar, and El-Shishiny (2010) in a comparative forecasting framework with other machine learning models. Wong, Xia, and Chu (2010) and Kulesh, Holschneider, and Kurennaya (2008) adopt an adaptive k-nearest neighbors (AKN) method. More recently, a locally weighted regression was introduced by Li, Shen, and Xiong (2012).

Nearest neighbor predictions have been employed and analyzed by both practitioners and academic literature. Fernandez-Rodriguez, Sosvilla-Rivero, and Andrada-Felix (1999) extend nearest-neighbor predictors to allow for a wider information set. Gencay (1999) proposed the use of technical trading rules to in-sample select the nearest neighbors. The kNN algorithm is a simple yet effective and hence commonly used classification algorithm in industry and research. Some recent studies are: Maggini, Giles, and Horne (1997), Yang and Shahabi (2007) and Park and Samworth (2008). The kNN method is also used to forecast histogram time series (Arroyo & Mate, 2009). A most recent study on k-nearest neighbor forecasting is Zhang, Jank, and Shmueli (2010) with an application to online auctions. Wichard (2011) proposed a combined forecast of a neural network, a nearest neighbor model and a nearest trajectory model as well for the NN5 time series competition data set. kNN can also be used for correcting imbalanced samples; a recent application to tourism is Li and Sun (2012).

3 Methodology

In this study, we do not concentrate on a forecasting competition. Forecasting competition's studies have three limitations: (a) there is not a clear answer to which models should be examined instead of others; (b) in many cases, the forecasting method is assumed to be ordinary least squares (OLS); and (c) more heavily parameterized models are employed only when there are limited datasets available. Thus, we empirically evaluates different model parameterizations of kNNs. kNNs are compared to the benchmark ARIMA model in terms of forecasting accuracy. The inclusion of the major international stock market indices as explanatory variables tries to raise the forecasting accuracy of the methods employed. It is also examined whether groups of regional stock markets, across the world and per continent,

[1] A large part of the forecasting literature is volatility forecasting. An influential paper from the latter is Lux and Kaizoji (2007). They compared the forecasting performance of long-memory models to GARCH and ARMA models upon the predictability of both volatility and volume for a large sample of Japanese stocks.

increase the forecasting accuracy of *k*NNs. The best parametrization of the *k*NN method is revealed across 20 European countries and seven tourism demand measures. Three forecasting evaluation criteria provide consistency in results. European tourism demand in measured and forecast in a monthly frequency.

The main *k*NN modeling issues: number of nodes, hidden layers, hidden nodes, output nodes, activation function, training algorithm, training sample, test sample and performance measures (see, Choudhary & Haider, 2012; Palmer et al., 2006). The number of nodes, hidden layers, hidden nodes and output nodes are empirically selected. We employ a locally adjusted linear autoregressive model as activation function and retrieve forecasts at either a univariate or multivariate level. The out-of-sample significance of inclusion of the explanatory variables is examined.[2] The Levenberg-Marquardt backpropagation is the training algorithm employed.[3] Five hundred iterations are employed for optimization. Optimization is based on either root mean square error (RMSE) or mean absolute error (MAE). The 12 months of the year 2003 was used as input. The remaining of the dataset is split into training-, validation- and test-sets. The validation set corresponds to the period from January 2011 to June 2013. The test-set corresponds to the last 12 months, from June 2013 to June 2014. The out-of-sample performance of either ARIMA or *k*NNs is evaluated via the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) criteria. Both criteria have been extensively employed in the tourism demand forecasting literature.

One-month-ahead (monthly) forecasts are retrieved. The forecasting performance of a *k*NN model is examined across different parameterizations. Moreover, ARIMA model is employed as a benchmark. ARIMA was proposed by Box and Jenkins (1970). Smeral and Wuger (2005) found that ARIMA was a bad performer. ARIMA is selected as a benchmark forecasting model, since it is the most accurate forecasting model in the tourism demand forecasting literature (see, Claveria & Torra, 2014; among others).

The methodology section focuses on prediction by nearest neighbor methods. The target of *k*NN is to find the *k*-nearest neighbors of a test pattern x ($x \in \omega$) in T based on a dissimilarity measure d(., .), and then classify the pattern x with the same label as the majority voting of nearest patterns in the training set. The *k*NN algorithm was introduced in Cover and Hart (1967). Let $x_t (t = 1, \ldots, n)$ be a finite time series. Vectors $x_t^{m,T}$ of m observations are sampled from the original time series at intervals of $\tau \in N$ units.[4]

[2] Athanasopoulos and Hyndman (2008) employed explanatory variables, such as economic variables and dummy variables, in the regression function of the models used.

[3] Law (2000) and Pattie and Snyder (1996) were from the first to adopt the back-propagation neural networks model. A more recent study on tourism demand forecasting via a back-propagation neural network is that of Chen et al. (2012).

[4] We refer to the studies of Casdagli and Weigend (1994) and Hastie, Tibshirani, and Friedman (2001), for further details on estimation methods.

$$x_t^{m,T} = \left(x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}\right)$$
$$t = 1 + \tau(m-1), \ldots, n \tag{1}$$

where m is the embedding dimensions and τ the delay parameter. According to Takens (1981), $\tau = 1$ and $x_t^{m,1} = x_t^{m,3}$. Local predictions are generated by analyzing the historical paths of the vectors around the last available vector:

$$x_n^m = \left(x_n, x_{n-1}, \ldots, x_{n-(m-1)}\right) \tag{2}$$

The closest k vectors (Eq. 3) are selected to maximise the function:

$$p\left(x_t^m, x_n^m\right) \tag{3}$$

Next, the predictors of the future evolution of x_n^m are simply estimates of x_{n+1} for the next observation. The prediction x_{n+1}^f of x_{n+1} can be obtained using some extrapolation of observations:

$$x_{i1+1}, x_{i2+1}, \ldots, x_{ik+1} \tag{4}$$

subsequent to the k NNs m-histories.

$$x_{n+1}^f = F(x_{i1+1}, x_{i2+1}, \ldots, x_{ik+1}) \tag{5}$$

Locally adjusted linear autoregressive predictions are employed for weighting:

$$x_{n+1}^f = \widehat{\alpha}_0 x_n + \widehat{\alpha}_1 x_{n-1} + \cdots + \widehat{\alpha}_{m-1} x_{n-(m-1)} + \widehat{\alpha}_m \tag{6}$$

Coefficients are fitted by Ordinary Least Squares. The regression is between the k NNs as in Eq. (3) on their future evolution. The $\hat{a}_i$ are the values that minimize either (a) Mean Square Errors or (b) Mean Absolute Errors. For example, MSE:

$$\sum_{r=i}^{k} \left(x_{i_r+1} - \alpha_0 x_{i_r-1} - \cdots - \alpha_{m-1} x_{i_r-(m-1)} - \alpha_m\right)^2 \tag{7}$$

4 Data and Descriptive Analysis

4.1 Data

Tourism demand can be measured by various variables, such as tourism revenues (Akal, 2004), tourism employment (Witt, Song, & Wanhill, 2004), tourism expenditure in the destination (Li et al., 2006) and overnight stays (Claveria & Datzira, 2010). Some studies in tourism demand forecasting filter data for removing trends and seasonalities.

However, there is part of the literature showing inconclusive evidence as to whether seasonality should be treated as stochastic or deterministic, and whether imposing seasonal unit roots may lead to more accurate forecasts (see, Alleyne, 2006; Coshall, 2005). In this paper, the null hypotheses of a unit root and nonstationarity are rejected in a 5% significance level for all European countries and tourism demand measures. Moreover, as evidenced in Claveria and Torra (2014), there is a trade-off between the degree of pre-processing and the accuracy of the forecasts obtained with neural networks, which are especially suited to deal with nonlinear data. Here, data series are de-trended; however, no filter for seasonalities is employed. This is because when data series are filtered heavily for accounting for the presence of seasonality and eliminating the existing outliers, *k*NNs produce less accurate forecasts. The broader economics literature has extensively studied the recent financial crisis, starting from the housing sector in the USA. A plausible crisis effect is not irrelevant to tourism demand forecasting. Min (2005) examined the effect of SARS on tourism demand in Taiwan. However, the limitation for such impact analysis is in periods of successive crises. This is the case since 2008 for Europe. All European countries, for which tourism demand forecasts are produced, are still affected by the consequences of the financial crisis which started in 2007. For this reason, in this paper, it is not necessary to consider the crises' effects in *k*NN forecasting.

This paper examines the most out-of-sample accurate parametrization of a *k*-nearest neighbor model for seven tourism demand variables in 20 European countries as destinations and not as regions of origin (disaggregate data). In particular, we measure tourism demand through seven different tourism accommodation variables.[5] The seven tourism demand variables are: (1) the number of nights spent by both residents and non-residents at any tourist demand establishments (*ALL–ALL*), (2) the number of nights spent by both residents and non-residents at any hotel demand (*ALL–HOT*), (3) the number of nights spent by residents at any tourist demand establishments (*RES–ALL*), (4) the number of nights spent by residents at any hotel demand (*RES–HOT*), (5) the number of nights spent by non-residents at any tourist demand establishments (*NRES–ALL*), (6) the number of nights spent by non-residents at any hotel demand (*NRES–HOT*) and (7) the Net occupancy rate of

[5]The monthly number of nights spent in hotels forecast has been examined by Thawornwong and Enke (2004) and Teixeira and Fernandes (2012).

bed-places and bedrooms in hotels and similar accommodation (%) (*NORBP*). The 20 European countries, for which tourism demand forecasts are provided, are: Austria, Belgium, Cyprus, Czech Republic, Denmark, Finland, Germany, Greece, Hungary, Italy, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and the UK. Since, in most cases, tourism policy-makers have to make decisions in the short run (in less than a year)—examples are short-term business planning, staffing, stock management etc.—monthly data series are employed and forecasts are retrieved 1 month ahead. Data for tourism demand come from Eurostat.

The tourism demand of Cyprus as a destination was researched in Bicak, Altinay, and Jenkins (2005) (in an annual frequency). The tourism demand of Greece as a destination was researched in Petropoulos, Patelis, Metaxiotis, Nikolopoulos, and Assimakopoulos (2003), Dritsakis (2004), Patsouratis, Frangouli, and Anastasopoulos (2005), and Petropoulos, Nikolopoulos, Patelis, and Assimakopoulos (2005). All these studies concerned annual data. The tourism demand of Denmark as a destination was researched in an annual frequency both in Song, Witt, and Jensen (2003) and Witt et al. (2004). The tourism demand of Portugal as a destination was researched in Rodrigues and Gouveia (2004) and Gouveia and Rodrigues (2005). Both studies concern a monthly frequency. The tourism demand of Spain as a destination was researched in Gil-Alana, Gracia, and Cunado (2004) (in a quarterly frequency) and Roget and Gonzalez (2006) (in an annual frequency). Finally, the tourism demand of Sweden as a destination was researched in Gustavsson and Nordstrom (2001) and Salman (2003). Both studies concern monthly measurements.

We also examine whether four groups of 28 equity indices (main countries' indices) from around the World (*Europe*, *Americas*, *Asia*, and *World*) provide a significant additional explanatory power to the most accurate parametrization of a *k*-nearest neighbor model. The first group concerns 12 European equity markets (*Europe*): Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Spain, Switzerland and the UK. The second group concerns four American equity markets (*Americas*): Argentina, Brazil, Mexico and the USA. The third group concerns nine Asian equity markets (*Asia*): Australia, Hong Kong, India, Indonesia, Japan, Malaysia, Singapore, South Korea and Taiwan. The fourth group concerns all international equity markets (*World*) including the *Europe*, *Americas*, *Asia* stock indices, as well as the ones from Israel, Russia, and South Africa. The explanatory variables were de-trend. Then, ADF-values revealed no unit root at a 5% significance level and thus without further trend. The data series for all tourism demand measures and all stock (equity) market indices are available in a monthly sampling frequency for a decade starting in January 2003 and ending in June 2014. Data for international equity indices come from Data-stream.

4.2 Descriptive Analysis

The 20 European countries for the seven tourism demand measures are split into three groups: *North* (Austria, Belgium, Denmark, Finland, Germany, Netherlands, Norway, Sweden and the United Kingdom), *East* (Czech Republic, Hungary, Poland, Romania, Slovakia and Slovenia) and *South* (Cyprus, Greece, Italy, Portugal and Spain).

We used sojourn densities to indicate whether the distribution of a series is the stationary density of the Markov chain. The sojourn density summarizes the long-term behavior of the hidden-states: it quantifies the proportion of time spent by the system at each point in state-space (the so-called "sojourn time"). Sojourn densities characterize the time spent by the series in local neighborhoods and can be estimated by nonparametric methods (see Phillips, 2001a, 2005; Rust, 2005). Bulla and Bulla (2006) showed that the sojourn time distribution is the difference between a hidden Markov Model (HMM) and a hidden semi-Markov model (HSMM). A coherent analysis of sojourn densities is provided in Phillips (2001b). If there is irregularity in the sojourn density, there is a unit root; so, the data is not equally spaced in time. In the present paper, the discussion about the distributional properties of the tourism demand and equity indices is derived by sojourn densities.[6] There is evidence of irregularity (not bell-shaped curves) in sojourn densities for tourism demand.

Moreover, the presence of unit root is a common distributional characteristic between European tourism demand and international stock indices. All other distributional properties between these two groups of data series differ. This is the reason why, international stock indices have been employed to explain the European tourism demand. The problem of unit root is treated via the first differences in all data series. The European tourism demand is forecast either without or with the inclusion of explanatory variables in any *k*NN method. Witt and Song (2000) suggested the inclusion of variables indicative to tourism demand. The present paper examines whether the inclusion of international equity indices improves European tourism demand forecasting. It is also examined whether a group of stock indices increases forecasting performance in the *k*NN methods. The following groups of stock indices are considered: *Europe*, *Americas*, *Asia*, and *World*. Each group of the first three concern the corresponding continents; whereas, the last group concerns all international stock indices. Table 1 analytically deploys the countries for which their major stock indices are the explanatory variables. Forecasting concerns a 1-month ahead forecasting horizon. Forecasting takes place in either a univariate or multivariate level for any method employed.

[6]More analytical results are available upon request.

Table 1 Classification of dependent and explanatory data time series

Dependent-variables groups			Explanatory-variables groups			
North	East	South	Europe	Americas	Asia	World
Austria	Czech Republic	Cyprus	Austria	Argentina	Australia	Israel
Belgium	Hungary	Greece	Belgium	Brazil	China	Russia
Denmark	Poland	Italy	Finland	Mexico	Hong Kong	South Africa
Finland	Romania	Portugal	France	USA	Japan	
Germany	Slovakia	Spain	Germany		India	
Netherlands	Slovenia		Greece		Indonesia	
Norway			Ireland		Malaysia	
Sweden			Italy		Singapore	
United Kingdom			Netherlands		South Korea	
			Spain		Taiwan	
			Switzerland			
			United Kingdom			

Note: This table reports the classifications of dependent (country tourism demand) and explanatory (stock market indices) variables to three and four groups accordingly. Country tourism demand is measured via seven measures: (1) the number of nights spent by both residents and non-residents at any tourist establishments (*ALL–ALL*), (2) the number of nights spent by both residents and non-residents at hotels (*ALL–HOT*), (3) the number of nights spent by residents at any tourist establishment (*RES–ALL*), (4) the number of nights spent by residents at hotels (*RES–HOT*), (5) the number of nights spent by non-residents at any tourist establishment (*NRES–ALL*), (6) the number of nights spent by non-residents at hotels (*NRES–HOT*), and (7) the net occupancy rate of bed-places and bedrooms in hotels and similar tourist accommodation (%) (*NORBP*)

5 Empirical Findings

5.1 Research Questions

Analytical empirical findings are depicted in Tables 1, 2, 3, 4, 5, 6, 7, 8 and 9. Three evaluation criteria (MAE, RMSE, and DIR) have been employed. MAE is the mean absolute error. RMSE is the root mean square error. DIR is the percentage (%) at which the forecasts accurately predict the direction actual figures. Table 2 reports the forecast evaluation of seven tourism demand series for 20 European countries, via the Mean absolute error (MAE) criterion. Table 3 reports the forecast evaluation of seven tourism demand series for 20 European countries, via the Root mean square error (RMSE) criterion. Table 4 reports the forecast evaluation of seven tourism demand series for 20 European countries, via the directional (DIR) criterion. Table 5 reports the groups of international stock indices (if any), for which the most accurate tourism demand forecasts are retrieved for the Norther, Easter, Southern and Total European countries. The groups of international stock indices which are indicated as Europe, Asia, Americas and World, concern the European, Asian, American and

Table 2 Mean absolute error (MAE) evaluation of *k*-nearest neighbor European tourism demand forecasts

	ALL–ALL	ALL–HOT	RES–ALL	RES–HOT	NRES–ALL	NRES–HOT	NORBP	Average
Panel A. Northern European countries (*North*)								
Austria	0.1703	0.4370	0.0123	0.0575	0.1202	6.27e-4	0.0361	0.1191
Belgium	0.0129	0.0699	0.0628	0.0017	0.0288	0.0087	0.0359	0.0315
Denmark	0.3291	0.0116	0.0040	0.0066	0.0643	0.0123	0.0847	0.0732
Finland	0.0211	0.1173	0.0208	0.0056	0.0103	0.0028	0.0235	0.0288
Germany	0.3753	0.1326	0.2841	0.0844	0.1699	8.39e-4	0.0573	0.1578
Netherlands	0.8046	0.0010	0.2978	0.0020	0.4837	0.0269	0.0204	0.2338
Norway	0.0412	0.0082	0.0394	0.0135	0.0229	0.0059	0.0765	0.0297
Sweden	0.1052	0.3093	0.0934	0.0770	0.0019	0.0066	0.1389	0.1046
United Kingdom	0.0961	0.8597	0.6425	0.1449	0.2994	0.1257	0.0727	0.3201
Panel B. Eastern European countries (*East*)								
Czech Republic	0.2752	0.1800	0.0945	0.0161	0.0075	0.0451	0.1157	0.1049
Hungary	0.1163	0.1591	0.0438	0.0338	0.0494	0.0152	0.3158	0.1048
Poland	0.1415	0.0044	0.0404	0.0099	0.0087	0.0051	0.0860	0.0423
Romania	0.0335	0.2688	0.0033	0.0190	0.0021	3.85e-4	0.0410	0.0526
Slovakia	0.1227	0.0028	0.0113	0.0027	0.0498	0.0246	0.0234	0.0339
Slovenia	0.2649	0.0400	0.0356	0.0018	0.3004	0.0113	0.6777	0.1902
Panel C. Southern European countries (*South*)								
Cyprus	0.0361	0.0135	–	–	0.0150	0.0095	0.1442	0.0437
Greece	0.1339	0.1731	0.0020	0.0131	0.1276	0.1170	0.1414	0.1012
Italy	0.9931	–	0.4546	0.2035	0.5005	0.0615	0.0087	0.3703
Portugal	0.1060	0.7129	0.0629	0.3031	0.0305	0.0177	0.8165	0.2928
Spain	4.24	7.84	4.04	1.1137	0.5837	0.0397	0.4968	2.6226
Average	0.4209	0.5971	0.3287	0.1111	0.1438	0.0269	0.17066	0.2570

Note: This table reports the forecast evaluation of seven tourism demand series for 20 European countries, via the Mean absolute error (MAE) criterion. Each of the reported values concerns the best forecasting evaluation across all *k*NN parameterizations, and across all selections of stock indices as explanatory variables. The forecast tourism demand series are: the number of nights spent by both residents and non-residents at any tourist demand establishments series (*ALL–ALL*), the number of nights spent by both residents and non-residents at any hotel demand series (*ALL–HOT*), the number of nights spent by residents at any tourist demand establishments series (*RES–ALL*), the number of nights spent by residents at any hotel demand series (*RES–HOT*), the number of nights spent by non-residents at any tourist demand establishments series (*NRES–ALL*), the number of nights spent by non-residents at any hotel demand series (*NRES–HOT*), and the Net occupancy rate of bed-places and bedrooms in hotels and similar demand (%) (*NORBP*). Moreover, an average measure (*Average*) indicates the average result across all tourism demand series forecast

World-wide (all) stock indices. Table 6 reports whether the most accurate forecasts come from univariate or multivariate kNN forecasts. Table 7 reports whether the most accurate forecasts come from univariate or multivariate kNN forecasts. Table 8 reports whether the most accurate forecasts come from univariate or multivariate

Table 3 Root mean square error (RMSE) evaluation of *k*-nearest neighbor European tourism demand forecasts

	ALL–ALL	ALL–HOT	RES–ALL	RES–HOT	NRES–ALL	NRES–HOT	NORBP	Average
Panel A. Northern European countries (*North*)								
Austria	6.99	2.47	2.90	2.51	8.46	6.18	6.09	5.09
Belgium	1.72	1.42	2.02	0.5457	1.63	1.04	2.99	1.62
Denmark	4.22	0.8490	3.29	0.4359	2.27	0.5939	3.08	2.10
Finland	1.75	1.23	1.57	1.06	0.8633	0.5795	2.53	1.37
Germany	32.34	13.14	25.11	12.19	4.69	2.92	2.90	13.33
Netherlands	11.87	2.15	4.33	1.27	4.99	1.24	3.11	4.14
Norway	3.23	2.41	1.58	1.17	1.25	0.4146	3.62	1.95
Sweden	3.13	2.00	2.29	0.9761	1.05	0.8876	1.50	1.69
United Kingdom	39.99	10.98	25.55	4.31	13.64	4.83	2.91	14.60
Panel B. Eastern European countries (*East*)								
Czech Republic	2.04	2.52	0.6688	0.8062	2.01	1.40	2.41	1.69
Hungary	1.01	1.81	1.30	0.7536	1.46	0.9054	3.33	1.51
Poland	2.87	1.98	2.84	1.32	1.37	0.7008	1.92	1.86
Romania	2.20	2.05	1.59	1.27	0.3203	0.2096	2.37	1.43
Slovakia	1.06	0.4984	0.7758	0.0242	0.5941	0.3048	1.52	0.6752
Slovenia	0.8142	0.3832	0.5287	0.3062	0.8045	0.3717	3.63	0.9771
Panel C. Southern European countries (*South*)								
Cyprus	1.68	1.97	–	–	1.03	2.00	7.64	2.86
Greece	11.83	3.32	2.78	1.19	13.46	2.88	5.14	5.80
Italy	24.95	–	21.11	14.34	13.62	13.99	2.57	15.10
Portugal	3.85	3.25	1.96	1.32	2.89	3.04	2.53	2.69
Spain	30.98	27.16	17.81	7.64	21.83	12.72	2.79	17.28
Average	9.43	4.29	6.32	2.81	4.91	2.86	3.23	4.89

Note: This table reports the forecast evaluation of seven tourism demand series for 20 European countries, via the Root mean square error (RMSE) criterion

kNN forecasts. Table 9 reports the kNN parametrization with the most accurate forecasts. The Naive parametrization (M0) concerns nearest neighbors with automatic selection without explanatory variables, and estimated as a simple average on a default number of neighbors.

Many research questions can be retrieved from the empirical evidence. The ones answered in this section are the following:

1. Indicate which model provides most accurate forecasts. The models AR, Naive, NNM1, NNM2, and NNM3 are symbolized as AR, M0, M1, M2, and M3 accordingly.[7]

[7]The ARIMA model, for the remaining of the paper, will be notated as AR for simplicity reasons.

Table 4 Directional (DIR) evaluation of *k*-nearest neighbor European tourism demand forecasts

	ALL–ALL	ALL–HOT	RES–ALL	RES–HOT	NRES–ALL	NRES–HOT	NORBP	Average
Panel A. Northern European countries (*North*)								
Austria	0.8889	1.00	0.7778	0.6667	0.8889	0.8889	0.7778	0.8413
Belgium	0.8889	0.4444	0.7778	0.5556	0.4444	0.4444	0.7778	0.6190
Denmark	0.8889	0.7778	0.6667	0.5556	0.7778	0.7778	1.00	0.7778
Finland	0.7778	0.3333	0.7778	0.7778	0.8889	0.5556	0.7778	0.6984
Germany	0.8889	0.8889	0.8889	0.8889	0.8889	0.7778	0.8889	0.8730
Netherlands	1.00	0.8889	0.8889	0.7778	0.6667	0.7778	1.00	0.8571
Norway	0.8889	0.5556	0.8889	0.6667	0.6667	0.5556	0.7778	0.7143
Sweden	0.8889	0.6667	0.8889	0.7778	0.7778	0.5556	1.00	0.7937
United Kingdom	0.7778	1.00	0.8889	1.00	0.7778	1.00	0.8889	0.9048
Panel B. Eastern European countries (*East*)								
Czech Republic	0.8889	0.7778	1.00	0.6667	0.5556	0.6667	0.8889	0.7778
Hungary	0.8889	0.7778	0.7778	0.6667	0.7778	0.7778	0.7778	0.7778
Poland	1.00	0.8889	0.8889	0.8889	0.6667	0.8889	0.8889	0.8730
Romania	1.00	0.8889	0.7778	0.8889	0.4444	0.5556	1.00	0.7937
Slovakia	0.8889	0.6667	0.6667	0.5556	0.4444	0.6667	1.00	0.6984
Slovenia	0.8889	0.6667	0.3333	0.5556	0.5556	1.00	0.8889	0.6984
Panel C. Southern European countries (*South*)								
Cyprus	0.8889	0.7778	–	–	0.8889	0.8889	0.8889	0.8667
Greece	0.8889	0.7778	0.8889	0.6667	1.00	0.6667	0.7778	0.8095
Italy	0.8889	–	0.6667	0.7778	1.00	1.00	0.8889	0.8704
Portugal	1.00	1.00	0.6667	0.6667	0.7778	0.8889	1.00	0.8571
Spain	0.8889	1.00	0.6667	1.00	1.00	1.00	1.00	0.9365
Average	0.9001	0.7778	0.7778	0.7368	0.7445	0.7667	0.8945	0.8019

Note: This table reports the forecast evaluation of seven tourism demand series for 20 European countries, via the directional (DIR) criterion

2. Does any of the four different parameterizations of *k*-nearest neighbor models (Naive, NNM1, NNM2, and NNM3) retrieve more accurate forecasts than the ARIMA (AR) benchmark forecasting model? If this is true, it is symbolized by Y; Worldwide, by N. The benchmark forecasting model is an ARIMA model (in either univariate or multivariate level), for which the number of orders is selected according to the AIC information criterion.
3. Which of the four (Naive, NNM1, NNM2, and NNM3) parameterizations of the *k*-nearest neighbor model returns the most accurate forecasts? The Naive parametrization concerns nearest neighbors with automatic selection without explanatory variables, and estimated as a simple average on a default number of

Table 5 The significance of the groups of stock indices on *k*-nearest neighbor European tourism demand forecasts

	MAE	RMSE	DIR
Panel A. Northern European countries (*North*)			
Austria	World	Americas	Europe
Belgium	–	Europe	–
Denmark	Americas	Americas	World
Finland	Europe	Americas	–
Germany	Europe	World	–
Netherlands	Americas	Americas	–
Norway	Europe	Americas	World
Sweden	–	Americas	World
United Kingdom	–	Americas	World
NORTH	Europe	Americas	World
Panel B. Eastern European countries (*East*)			
Czech Republic	Europe	Europe	Asia
Hungary	Americas	Americas	World
Poland	Americas	World	World
Romania	Americas	Europe	Americas
Slovakia	World	Europe	World
Slovenia	Americas	Americas	World
EAST	Americas	Europe	World
Panel C. Southern European countries (*South*)			
Cyprus	–	World	World
Greece	Europe	Americas	–
Italy	Americas	Americas	–
Portugal	–	Americas	World
Spain	Americas	Americas	World
SOUTH	Americas	Americas	World
Total	Americas	Americas	World

Note: This table reports the groups of international stock indices (if any), for which the most accurate tourism demand forecasts are retrieved for the Norther, Easter, Southern and Total European countries. The groups of international stock indices which are indicated as Europe, Asia, Americas and World, concern the European, Asian, American and World-wide (all) stock indices. – indicates there is no group of explanatory variables (stock indices) producing more accurate forecasts than the others. *NORTH*, *EAST*, *SOUTH* and *TOTAL* symbolize the group of stock indices with the most accurate *k*NN forecasts for Northern, Eastern, Southern and Total European countries, accordingly

neighbors.[8] This is the only parametrization without explanatory variables. The NNM1 parameterization concerns nearest neighbors with explanatory variables and estimated as a simple average of the first *k* nearest neighbors. The NNM2

[8]The default number of neighbors is selected to be capped to 2% of the number of the out-of-sample period, which in its turn is capped to 1/3 the number of total observations.

Table 6 Univariate versus Multivariate *k*NN forecasting

	ALL–ALL	ALL–HOT	RES–ALL	RES–HOT	NRES–ALL	NRES–HOT	NORBP	All
Panel A. Mean absolute error (*MAE*)								
North	U	M	M	U	U	U	M	U
East	M	M	U	M	M	M	U	M
South	U	U	U	U	M	M	U	U
Total	U	M	U	U	M	M	U	U
Panel B. Root mean square error (*RMSE*)								
North	U	U	M	M	U	M	U	U
East	U	U	M	M	U	U	U	U
South	U	U	U	U	M	U	U	U
Total	U	U	M	M	U	U	U	U
Panel C. Direction (*DIR*)								
North	U	U	M	U	U	U	U	U
East	M	U	M	M	M	U	M	M
South	U	M	U	M	M	U	U	U
Total	U	M	M	M	M	U	U	M

Note: This table reports whether the most accurate forecasts come from univariate or multivariate *k*NN forecasts. Results are summarized for Northern, Eastern, Southern and Total European countries. Results are presented for all tourism demand measures. *All* indicates the results across the board of all tourism demand measures. Forecasts are evaluated via mean absolute error, root mean square error and a directional criterion. They are symbolized as MAE, RMSE and DIR in Panel A, B and C accordingly

parametrization concerns nearest neighbors with explanatory variables and estimated as a weighted average (by distance) of k nearest neighbors. The NNM3 parametrization concerns nearest neighbors with explanatory variables and estimated as a simple average on a default number of neighbors.[9] Four groups of the main equity indices of their corresponding countries' stocks exchanges are employed as explanatory variables. The equity indices data are indicated and graphically examined in Sect. 3. The indication of the most accurate k-nearest neighbor forecasts for the Naive, NNM1, NNM2, and NNM3 parameterizations, is 0, 1, 2, and 3 accordingly.

4. The orders for the k-nearest neighbor models (Naive, NNM1, NNM2, NNM3, and NNM4) are estimated by minimizing either the RMSE or the MAE evaluation criteria. Is the Root mean square error (RMSE) or the Mean absolute error (MAE) loss function (e.g. Eq. 7) that provides more accurate forecasts? If none, the RMSE, or the MAE loss function provides more accurate k-nearest neighbor forecasts than the others, the 0, R, and M symbols are indicated accordingly.[10]

[9]The default number of neighbors is selected to be capped to 2% of the number of the out-of-sample period, which in its turn is capped to 1/3 the number of total observations.

[10]When neither of the RMSE and MAE loss functions can return more accurate forecasts than the other, the 0 symbol is used.

Table 7 RMSE- versus MAE-based *k*NN forecasting

	ALL–ALL	ALL–HOT	RES–ALL	RES–HOT	NRES–ALL	NRES–HOT	NORBP	All
Panel A. Mean absolute error (*MAE*)								
North	0	M	0	0	0	R	0	0
East	0	0	0	0	M	0	0	0
South	0	R	0	0	0	0	0	0
Total	0	0	0	0	0	0	0	0
Panel B. Root mean square error (*RMSE*)								
North	M	M	M	M	M	M	M	M
East	M	M	M	M	M	M	M	M
South	M	M	M	M	M	M	M	M
Total	M	M	M	M	M	M	M	M
Panel C. Direction (*DIR*)								
North	M	0	0	0	0	0	M	0
East	0	0	0	0	R	0	0	0
South	0	M	R	0	0	0	0	0
Total	0	0	0	0	0	0	0	0

Note: This table reports whether the most accurate forecasts come from univariate or multivariate *k*NN forecasts. R depicts minimizing the Root mean square error (RMSE) loss function provides the most accurate forecasts. M indicates the Mean absolute error (MAE) loss function most accurately forecast tourism demand. 0 indicates none of RMSE and MAE loss functions provide more accurate forecasts. Results are summarized for Northern, Eastern, Southern and Total European countries. Results are presented for all tourism demand measures. *All* indicates the results across the board of all tourism demand measures. Forecasts are evaluated via mean absolute error, root mean square error and a directional criterion. They are symbolized as MAE, RMSE and DIR in Panel A, B and C accordingly

For example, a result of M1-Y-1-R in any table indicates that a European country's specific demand measure is more accurately forecast by the NNM1 (or M1) model. This result concerns a specific group of explanatory variables (*Asia*), when the evaluation takes place from the RMSE criterion as well as in either a univariate or multivariate level. Y indicates there is a *k*-nearest neighbor model that provides more accurate forecasts than the AR benchmark model. 1 signifies the NNM1 parametrization is the most accurate parametrization. R depicts minimizing the Root mean square error (RMSE) loss function provides more accurate forecasts than the Mean absolute error (MAE) loss function.

5.2 Empirical Evidence

In most of European countries, demand is better forecast by the back-propagation neural network (or {*k*}NN) model compared to the ARIMA benchmark. This result is also evident in Chen, Lai, and Yeh (2012). Athanasopoulos and Hyndman (2008) found that the inclusion of economic variables improves the forecasting performance

Table 8 *k*NN versus ARIMA forecasting

	ALL–ALL	ALL–HOT	RES–ALL	RES–HOT	NRES–ALL	NRES–HOT	NORBP	All
Panel A. Mean absolute error (*MAE*)								
North	Y	Y	Y	Y	Y	Y	Y	Y
East	Y	Y	Y	N	N	Y	Y	Y
South	Y	Y	Y	Y	Y	N	Y	Y
Total	Y	Y	Y	Y	Y	Y	Y	Y
Panel B. Root mean square error (*RMSE*)								
North	Y	Y	Y	Y	Y	Y	Y	Y
East	Y	Y	Y	Y	N	N	Y	Y
South	Y	Y	Y	Y	Y	Y	Y	Y
Total	Y	Y	Y	Y	Y	Y	Y	Y
Panel C. Direction (*DIR*)								
North	Y	Y	Y	Y	Y	Y	Y	Y
East	Y	Y	Y	Y	N	Y	Y	Y
South	Y	N	Y	Y	Y	Y	Y	Y
Total	Y	Y	Y	Y	Y	Y	Y	Y

Note: This table reports whether the most accurate forecasts come from univariate or multivariate *k*NN forecasts. Y indicates there is a *k*-nearest neighbor model that provides more accurate forecasts than the ARIMA benchmark model. N indicates the opposite. Results are summarized for Northern, Eastern, Southern and Total European countries. Results are presented for all tourism demand measures. *All* indicates the results across the board of all tourism demand measures. Forecasts are evaluated via mean absolute error, root mean square error and a directional criterion. They are symbolized as MAE, RMSE and DIR in Panel A, B and C accordingly

of state space models in forecasting tourism demand. The inclusion of international stock indices as explanatory variables in {*k*}NN models improves the forecasting performance of this class of the models. This result is evident in an either univariate or multivariate forecasting framework. Most of the explanatory variables are positively related to tourism demand. This means as the growth in the international stock markets increase the tourism demand in Europe.

Table 1 reports the classifications of dependent and explanatory variables in order to measure tourism demand via the seven measures: the number of nights spent by both residents and non-residents at any tourist demand establishments (*ALL–ALL*), the number of nights spent by both residents and non-residents at any hotel demand (*ALL–HOT*), the number of nights spent by residents at any tourist demand establishments (*RES–ALL*), the number of nights spent by residents at any hotel demand (*RES–HOT*), the number of nights spent by non-residents at any tourist demand establishments (*NRES–ALL*), the number of nights spent by non-residents at any hotel demand (*NRES–HOT*), and the Net occupancy rate of bed-places and bedrooms in hotels and similar accommodation (*NORBP*).

Table 2 reports the forecast evaluation of seven tourism demand series for 20 European countries, via the Mean absolute error (MAE) criterion. For the Northern European countries in Panel A, the most accurately forecast tourism demand series

Table 9 *k*NN forecasting

	ALL–ALL	*ALL–HOT*	*RES–ALL*	*RES–HOT*	*NRES–ALL*	*NRES–HOT*	*NORBP*	*All*
Panel A. Mean absolute error (*MAE*)								
North	M0	M2	M2	M0	M0	M0	M1	M0
East	M2	M2	M2	M3	M0	M1	M0	M0
South	M2	M2	M2	M3	M0	M1	M0	M2
Total	M2	M2	M2	M0	M0	M0	M0	M0
Panel B. Root mean square error (*RMSE*)								
North	M2	M2	M0	M2	M3	M0	M0	M2
East	M0	M2	M0	M2	M2	M0	M0	M0
South	M2	M2	M2	M2	M2	M0	M0	M2
Total	M2	M2	M0	M2	M2	M0	M0	M2
Panel C. Direction (*DIR*)								
North	M0	M0	M0	M0	M2	M2	M0	M0
East	M2	M2	M2	M2	M0	M0	M1	M2
South	M2	M1	M2	M2	M0	M2	M1	M2
Total	M2	M0	M2	M2	M0	M2	M1	M2

Note: This table reports the *k*NN parameterization with the most accurate forecasts. The Naive parameterization (M0) concerns nearest neighbors with automatic selection without explanatory variables, and estimated as a simple average on a default number of neighbors. The NNM1 parameterization (M1) concerns nearest neighbors with explanatory variables and estimated as a simple average of the first *k* nearest neighbors. The NNM2 parameterization (M2) concerns nearest neighbors with explanatory variables and estimated as a weighted average (by distance) of *k* nearest neighbors. The NNM3 parameterization (M3) concerns nearest neighbors with explanatory variables and estimated as a simple average on a default number of neighbors. Results are summarized for Northern, Eastern, Southern and Total European countries. Results are presented for all tourism demand measures. *All* indicates the results across the board of all tourism demand measures. Forecasts are evaluated via mean absolute error, root mean square error and a directional criterion. They are symbolized as MAE, RMSE and DIR in Panel A, B and C accordingly

with the lowest MAE values in average are observed in Finland, Norway, Belgium and Denmark with values 0.0288, 0.0297, 0.0315 and 0.0732, respectively. In Panel B (Eastern European countries), Slovakia, Poland and Romania, report the best forecasting evaluation in average. As far as the Southern European countries are concerned, Cyprus and Greece obtain better forecasting results in average for all forecast tourism demand series.

When the Root mean square error (RMSE) criterion is used to evaluate forecasts (Table 3) of tourism demand series for the 20 European countries, we observe again that for the Northern European countries in Panel A, the most accurately forecast tourism demand series with the lowest MAE values in average are observed in Finland, Norway, Belgium and Denmark. Slovakia, Romania and Slovenia report the best forecasting evaluation in average as shown in Panel B. For the Southern European countries, Portugal and again Cyprus in average for the seven tourism demand series, report best forecasts when evaluation is implemented via RMSE criterion. Moreover, when we evaluate forecasts via the directional criterion

(Table 4) the results are similar for the Northern European countries in average across the seven tourism demand series, with Finland, Norway, Belgium and Denmark obtaining the lowest DIR values. The same applies for the Eastern European countries in Panel B with the lowest values in average, observed in Slovakia and Slovenia. In Greece and Portugal (Panel C) tourism demand is better forecast in average over all tourism demand series.

Table 5 reports the groups of international stock indices for which the most accurate tourism demand forecasts are retrieved. More specifically, we observe in Panel A for the Northern European countries that the most significant group of stock indices on k-nearest neighbor tourism demand forecasts considering the MAE evaluation criterion Europe equity indices provide additional explanatory power to the specific group of European countries. For the same group, regarding the RMSE evaluation criterion, Americas stock indices produce more accurate forecasts than others, while for the directional criterion all international equity markets give more accurate forecasts. In Panel B, concerning the East group of European countries, the Americas add explanatory power to forecasting European tourism demand according to MAE, Europe stock indices for the RMSE. As for the case of directional criterion, all international equity markets produce accurate forecasts, for the South group of European countries as well. As for Total European countries, the group of stock indices with the most accurate kNN forecasts is the Americas for MAE and RMSE.

In Table 6 we observe the results for most accurate forecasts based on the univariate of multivariate kNN forecasts for all tourism demand measures and for each and all European countries, evaluated by each of the three evaluation criteria (MAE, RMSE and DIR). According to Panel A, for All European countries univariate kNN forecasts give most accurate forecasts measured by MAE. For RMSE evaluation criterion (Panel B), also, demand is better forecast with univariate kNN forecasting across the board of all tourism demand measures. However, for the case of direction criterion (DIR), multivariate kNN forecasting provides most accurate forecasts.

As far as which evaluation criterion loss function provides more accurate forecasts is concerned, in Table 7 the corresponding results are reported. More specifically, in Panel B, we observe that MAE loss function most accurately forecasts tourism demand across all tourism demand measures, as seen in each and all European countries. In Panel A and C we notice that none of the two RMSE or the MAE loss function provides more accurate kNN forecasts than the others when forecast evaluation is implemented via RMSE and DIR criterion. There do exist however, cases in Panel A and C, where RMSE-optimized kNNs provide better forecasts for the North and South group of countries and for the NRES-HOT and ALL-HOT tourism demand measures and East and South for the NRES-ALL and RES-ALL, respectively.

According to our findings reported in Table 8, kNNs outperform ARIMA. In Panel A, B and C of the corresponding table, we observe in the last row and column of each Panel, that for all tourism demand measures and for the Total European countries, demand is better forecast by the kNN model compared to the ARIMA benchmark. There do exist cases though were ARIMA performs better in forecasting

tourism demand such as for the RES-HOT and NRES-ALL measures in Eastern European countries and NRES-NOT for Southern European countries as evaluated by MAE (Panel A). In Panel C concerning the DIR evaluation criterion, ARIMA gives better forecasts for the group of Eastern European countries and the NRES-ALL measure and for the group of Southern European countries for the ALL-HOT measure.

In Table 9, we observe that the M2 parametrization (nearest neighbors with explanatory variables) gives the most accurate forecasts according to the DIR and MAE evaluation criterion for Total European countries and All tourism demand measures. However, when the evaluation criterion is the RMSE criterion (Panel A), the M0 parametrization provides the best forecasts of tourism demand.

6 Concluding Remarks

Across all results, the *k*NNs forecast European tourism demand better than ARIMA. There is not clear evidence on which of the three parameterizations is best for *k*NNs forecasting European tourism demand. The results of the present paper contradict to the results of Claveria and Torra (2014), in terms of NNs performance. As indicated in the Empirical findings section, *k*NNs outperform ARIMA.[11] This result is evident regardless the inclusion as well as the selection of stock indices as explanatory variables. However, this result is evident on Cho (2003).

The NNM3 parametrization, which concerns nearest neighbors with explanatory variables and estimated as a simple average on a default number of neighbors, provides the best forecasts when European stock indices are explanatory variables. When either American or Asian stock indices are used as explanatory variables, the NNM2 parametrization (which concerns nearest neighbors with explanatory variables and estimated as a weighted average—by distance—of k nearest neighbors) provide the best forecasts. When all international stock indices are explanatory variables, the NNM1 parametrization (which concerns nearest neighbors with explanatory variables and estimated as a simple average of the first k nearest neighbors) forecasts best the European tourism demand. When forecast evaluation is implemented via the RMSE criterion, the best forecasts come from the RMSE-optimized *k*NNs.

The tourism demand in the Northern and Eastern Europe is better forecast via the NNM2 parametrization. In Southern Europe, there is not a clear-cut winner between NNM0, NNM2 and NNM3. The common outcome is that only kNNs can forecast accurately European tourism demand, regardless the geographical location of each European country.

[11]The ARIMA model more accurately forecast tourism demand only in Northern Europe in a univariate forecasting framework.

References

Ahmed, N., Atiya, A., Gayar, N., & El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. *Econometric Reviews, 29*(5–6), 594–621.

Akal, M. (2004). Forecasting Turkey's tourism revenues by ARMAX model. *Tourism Management, 26*, 359–365.

Alleyne, D. (2006). Can seasonal unit root testing improve the forecasting accuracy of tourist arrivals? *Tourism Economics, 12*, 45–64.

Arroyo, J., & Mate, C. (2009). Forecasting histogram time series with k-nearest neighbors methods. *International Journal of Forecasting, 25*(1), 192–207.

Athanasopoulos, G., & Hyndman, R. J. (2008). Modelling and forecasting Australian domestic tourism. *Tourism Management, 29*, 19–31.

Basseville, M. (1989). Distance measure for signal processing and pattern recognition. *Signal Process, 18*(4), 349–369.

Bicak, H. A., Altinay, M., & Jenkins, H. (2005). Forecasting tourism demand of North Cyprus. *Journal of Hospitality and Leisure Marketing, 12*, 87–99.

Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis, forecasting and control*. San Francisco: Holden Day.

Bulla, J., & Bulla, I. (2006). Stylized facts of financial time series and hidden semi-Markov models. *Computational Statistics and Data Analysis, 51*(4), 2192–2209.

Casdagli, M., & Weigend, A. S. (1994). Exploring the continuum between deterministic and stochastic modelling. In A. S. Weigend & N. A. Gereshenfeld (Eds.), *Time series prediction: Forecasting the future and understanding the past*. Reading, MA: Addison Wesley.

Chan, F., Lim, C., & McAleer, M. (2005). Modelling multivariate international tourism demand and volatility. *Tourism Management, 26*, 459–471.

Chen, C.-F., Lai, M.-C., & Yeh, C.-C. (2012). Forecasting tourism demand on empirical mode decomposition and neural network. *Knowledge-Based Systems, 26*, 281–287.

Cho, V. (2003). A comparison of three different approaches to tourist arrival forecasting. *Tourism Management, 24*, 323–330.

Choudhary, A., & Haider, A. (2012). Neural network models for inflation forecasting: An appraisal. *Applied Economics, 44*, 2631–2635.

Claveria, O., & Datzira, J. (2010). Forecasting tourism demand using consumer expectations. *Tourism Review, 65*, 18–36.

Claveria, O., & Torra, S. (2014). Forecasting tourism demand to Catalonia: Neural networks vs. time series models. *Economic Modelling, 36*, 220–228.

Coshall, J. T. (2005). A selection strategy for modelling UK tourism flows by air to European destinations. *Tourism Economics, 11*, 141–158.

Cover, T. M., & Hart, P. E. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory, IT-13*(1), 21–27.

Dritsakis, N. (2004). Cointegration analysis of German and British tourism demand for Greece. *Tourism Management, 25*, 111–119.

Fernandez-Rodriguez, F., Sosvilla-Rivero, S., & Andrada-Felix, J. (1999). Exchange-rate forecasts with simultaneous nearest-neighbor methods: Evidence from the EMS. *International Journal of Forecasting, 15*(4), 383–392.

Gencay, R. (1999). Linear, non-linear and essential foreign exchange rate prediction with simple technical trading rules. *Journal of International Economics, 47*(1), 91–107.

Gil-Alana, L. A., Gracia, F. P. D., & Cunado, J. (2004). Seasonal fractional integration in the Spanish tourism quarterly time-series. *Journal of Travel Research, 42*, 408–414.

Gouveia, P. M. D. C. B., & Rodrigues, P. M. M. (2005). Dating and synchronizing tourism growth cycles. *Tourism Economics, 11*, 501–515.

Gustavsson, P., & Nordstrom, J. (2001). The impact of seasonal unit roots and vector ARMA modelling on forecasting monthly tourism flows. *Tourism Economics, 7*, 117–133.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The essentials of statistical learning*. New York: Springer.

Kon, S. C., & Turner, W. L. (2005). Neural network forecasting of tourism demand. *Tourism Economics, 11*, 301–328.

Kulendran, N., & Witt, S. F. (2003). Leading indicator tourism forecasts. *Tourism Management, 24*, 503–510.

Kulesh, M., Holschneider, M., & Kurennaya, K. (2008). Adaptive metrics in the nearest neighbors method. *Physics D, 237*(3), 283–291.

Kulkarni, S. R., Lugosi, G., & Venkatesh, S. S. (1998). Learning pattern classification survey. *IEEE Transactions on Information Theory, 44*(6), 2178–2206.

Law, R. (2000). Back-propagation learning in improving the accuracy of neural network-based demand forecasting. *Tourism Management, 21*, 331–340.

Li, S., Shen, Z., & Xiong, G. (2012). A k-nearest neighbor locally weighted regression method for short-term traffic flow forecasting. In *2012 15th International IEEE Conference on Intelligent Transportation Systems (ITSC)* (pp. 1596–1601).

Li, G., Song, H., & Witt, S. F. (2005). Recent developments in econometric modeling and forecasting. *Journal of Travel Research, 44*, 82–99.

Li, H., & Sun, J. (2012). Forecasting business failure: The use of nearest-neighbor support vectors and correcting imbalanced samples – Evidence from the Chinese hotel industry. *Tourism Management, 33*(3), 622–634.

Li, G., Wong, K. F., Song, H., & Witt, S. F. (2006). Tourism demand forecasting: A time varying parameter error correction model. *Journal of Travel Research, 45*, 175–185.

Lux, T., & Kaizoji, T. (2007). Forecasting volatility and volume in the Tokyo stock market: Long memory, fractality and regime switching. *Journal of Economic Dynamics and Control, 31*(6), 1808–1843.

Maggini, M., Giles, C., & Horne, B. (1997). Financial time series forecasting using k-nearest neighbors classification. In R. B. Cadwell (Ed.), *Nonlinear Financial Forecasting: Proceedings of the First INFFC* (pp. 169–181).

Min, J. C. H. (2005). The effect of the SARS illness on tourism in Taiwan: An empirical study. *International Journal of Management, 22*, 497–506.

Palmer, A., Montano, J. J., & Sese, A. (2006). Designing an artificial neural network for forecasting tourism time-series. *Tourism Management, 27*(781–790), 2006.

Park, B., & Samworth, R. (2008). Choice of neighbor order in nearest-neighbor classification. *Annals of Statistics, 36*(5), 2135–2152.

Patsouratis, V., Frangouli, Z., & Anastasopoulos, G. (2005). Competition in tourism among the Mediterranean countries. *Applied Economics, 37*, 1865–1870.

Pattie, D. C., & Snyder, J. (1996). Using a neural network to forecast visitor behavior. *Annals of Tourism Research, 23*(1), 151–164.

Petropoulos, C., Nikolopoulos, K., Patelis, A., & Assimakopoulos, V. (2005). A technical analysis approach to tourism demand forecasting. *Applied Economics Letters, 12*, 327–333.

Petropoulos, C., Patelis, A., Metaxiotis, K., Nikolopoulos, K., & Assimakopoulos, V. (2003). SFTIS: A decision support system for tourism demand analysis and forecasting. *Journal of Computer Information Systems, 44*, 21–32.

Phillips, P. C. B. (2001a). Trending time series and macroeconomic activity: Some present and future challenges. *Journal of Econometrics, 100*(1), 21–27.

Phillips, P. C. B. (2001b). Descriptive econometrics for non-stationary time series with empirical illustrations. *Journal of Applied Econometrics, 16*(3), 389–413.

Phillips, P. C. B. (2005). Econometric analysis of Fisher's equation. *American Journal of Economics and Sociology, 64*(1), 125–168.

Rodrigues, P. M. M., & Gouveia, P. M. D. C. B. (2004). An application of PAR models for tourism forecasting. *Tourism Economics, 10*, 281–303.

Roget, F. M., & Gonzalez, X. A. R. (2006). Rural tourism demand in Galicia, Spain. *Tourism Economics, 12*, 21–31.

Rust, J. (2005). Comments on "Econometric analysis of Fisher's Equation" by Peter C. B. Phillips. *American Journal of Economics and Sociology, 64*(1), 169–184.

Salman, A. K. (2003). Estimating tourist demand through cointegration analysis: Swedish data. *Current Issues in Tourism, 6*, 323–339.

Shen, S., Li, G., & Song, H. (2011). Combination forecasts of international tourism demand. *Annals of Tourism Research, 38*(1), 72–89.

Smeral, E., & Wuger, M. (2005). Does complexity matter? Methods for improving forecasting accuracy in tourism: The case of Australia. *Journal of Travel Research, 44*, 100–110.

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting – A review of recent research. *Tourism Management, 29*, 203–220.

Song, H., & Witt, S. F. (2006). Forecasting international tourist flows to Macau. *Tourism Management, 27*, 214–224.

Song, H., Witt, S. F., & Jensen, T. C. (2003). Tourism forecasting: Accuracy of alternative econometric models. *International Journal of Forecasting, 19*, 123–141.

Song, H., Wong, K. F., & Chon, K. K. S. (2003). Modelling and forecasting the demand for Hong Kong tourism. *International Journal of Hospitality Management, 22*, 435–451.

Stone, C. J. (1977). Consistent nonparametric regression. *Annals of Statistics, 5*(4), 595–645.

Takens, F. (1981). Detecting strange attractors in turbulence. *Lecture Notes in Mathematics, 898*, 365–381.

Teixeira, J. P., & Fernandes, P. O. (2012). Tourism time series forecast – Different ANN architectures with time index input. *Procedia Technology, 5*, 445–454.

Thawornwong, S., & Enke, D. (2004). The adaptive selection of financial and economic variables for use with artificial networks. *Neurocomputing, 6*, 205–232.

Wichard, J. D. (2011). Forecasting the NN5 time series with hybrid models. *International Journal of Forecasting, 27*(3), 700–707.

Witt, S. F., & Song, H. (2000). Forecasting future tourism flows. In S. Medik & A. Lockwood (Eds.), *Tourism and hospitality in the 21st century* (pp. 106–118). Oxford: Butterworth-Heinemann.

Witt, S. F., Song, H., & Wanhill, S. P. (2004). Forecasting tourism-generated employment: The case of Denmark. *Tourism Economics, 10*, 167–176.

Wong, W. K., Xia, M., & Chu, W. C. (2010). Adaptive neural network model for time-series forecasting. *European Journal of Operational Research, 207*(2), 807–816.

Yang, K., & Shahabi, C. (2007). An efficient k nearest neighbor search for multivariate time series. *Information and Computation, 205*(1), 65–98.

Zhang, S., Jank, W., & Shmueli, G. (2010). Real-time forecasting of online auctions via functional K-nearest neighbors. *International Journal of Forecasting, 26*(4), 666–683.

The Role of Modern Ports and Marine Protected Areas in Eastmed's Environmental Sustainability: The Case of Greece

Basil Tselentis, Theodoros Pelagidis, and Sotirios Manologlou

Abstract Coasts are geographical areas that house many diverse activities, with direct impacts on the environment and the sustainable development of these regions. The concentration of these activities in a small spatial coastal zone creates significant pressures on marine ecosystems threatening further damage to habitats followed by potential loss of species and genetic diversity. Marine reserves and Marine Protected Areas (MPAs) are considered important tools to address these threats to marine resources. Greece is one of the richest countries in marine biodiversity. It is a country with the longest coastline in the Mediterranean (about 15,000 km) and its coastal areas are of great economic, cultural and ecological interest. It is therefore obvious that marine/coastal and island habitats are, or should be, priority areas for Greece, regarding the focus of governmental support and protection, as well as strengthening stakeholder involvement procedures. This realization further enhances the need for developing and applying managerial and financial tools for lasting and sustainable development of coastal areas and Greek seas and thus the study of marine protected areas cannot ignore the ports operating in the surrounding region. A "traditional port" is a coastal marine area with a land zone artificially formed to provide the vessel with not only safe anchorage and mooring for loading and unloading operations, but also a point that can provide the best possible service to ships. Today, ports operate in an extremely competitive framework, continuously trying to reduce cost, increase productivity with simultaneous improvement of the services provided to users. The crucial question concerning the port lies in the fact that by trying to reduce cost, safety and environmental protection maybe partly sacrificed. Port authorities must therefore take all necessary measures to prevent the emergence of pollution (accident-related) or stemming from normal operation, which involves monitoring the frequency and consequences of accidents and other hazards, leading to the estimation of the risk of environmental pollution.

B. Tselentis (✉) · T. Pelagidis · S. Manologlou
Department of Maritime Studies, University of Piraeus, Piraeus, Greece
e-mail: tselenti@unipi.gr; pelagidi@unipi.gr

E. Krassadaki et al. (eds.), *Operational Research in Agriculture and Tourism*, Cooperative Management, https://doi.org/10.1007/978-3-030-38766-2_7

Keywords Marine protected area · Ecotourism · Marine ecosystem · Green port · ESPO

1 Introduction

Over the past few decades, the environment has been subjected to considerable pressure and has been aggravated by human activities. It is now widely acknowledged in the literature, that its limits and endurance have alarmingly deteriorated (Brady & Staffan, 2009; Butler, 1991; Food and Agriculture Organization of the Fisheries Management of the Fisheries Management (FAO), 2011; The Greek Tourism Confederation (SETE), 2019).

The on-going and long-term environmental degradation has led environmentalists to propose a variety of measures and options aimed at its protection and conservation for future generations. Given that water, represents 3/4 of the planet, emphasis has been laid upon the protection of the sea environment and coastal areas, since the survival of local communities and livelihoods largely depends on the sea and its natural habitats (Frederigo, 2009). This chapter focuses on the importance of establishing Marine Protected Areas (MPAs) in coastal waters, since many studies have shown that this approach has many advantages regarding the protection of the marine environment, especially coastal habitats and species, which depend on the quality of the sea and its natural resources (Coad, Campbell, Miles, & Humphries, 2008; Naidoo, Balmord, & Ferraro, 2006). Environmental quality and health are heavily influenced by human activities and their high concentration along the coast. It is clear that an important contradiction lies in the fact that, on the one hand, coastal areas rely on tourism and recreational activities to increase the welfare of local communities and on the other, these activities undermine the basis on which tourism depends on, i.e. the quality of the environment (Coad et al., 2008; Cole, 1999). Many tools have been developed in order to calculate the carrying capacity of coastal areas, supporting policies (Contogiorgi, 2014) to develop coastal areas taking into account the sustainability and non-exhaustion of natural resources many of which are complex and important ecosystems offering several ecosystem services. Greece has the longest Mediterranean coastline as well as many islands, home to unique species of birds and mammals, some of which are already on the endangered and vulnerable lists (Shiffman, 2018). In recent years, Greece shows commitment, but we doubt that the biodiversity is actually protected and the sustainability goals are served. Commitment to the implementation of the 2030 Agenda for Sustainable Development and its 17 Sustainable Development Goals (SDGs), which include fostering a competitive, innovative and sustainable economic growth and strengthening the protection and sustainable management of natural capital as a base for social prosperity and transition to a low-carbon economy (Palumbi, 2009).

In several fora, the latest in 2017 (Vitousek, Mooney, Lubchenco, & Melillo, 1997), Greece reaffirmed its commitment to expand by 2020 the marine protected areas in its territorial waters from 6 to 20%. These sites would aim towards the

conservation of important habitats, such as the *Posidonia oceanica* (sea grass meadows), as well as important bird species.

Such an endeavor will obviously have to consider the many activities (recreation, hoteling, entertainment, food services, etc.), which are concentrated in the coastal zone and have been developing with limited control and definitely without spatial planning policies and norms. In addition, visiting coastal areas and islands requires sea transportation and operation of many ports, which undoubtedly affect marine species and ecosystems, as most visitors are transported by passenger shipping (Higgins-Desbiolles, 2010).

Ports, as service providers, are located and operate on the coasts, at the interface of land and sea, an especially sensitive area, facing a multitude of challenges on both the physical as well as the man made environment. It is obvious that ports must operate in such way as to promote and support the protection of the marine environment, through the implementation of best practices that seek to protect the marine environment in every process. This chapter will raise issues related to this.

Section 2 presents some key facts about ecosystems and their benefits, while stressing the imperative obligation of Greece to adapt sustainable development goals, especially due to its geomorphological and habitat characteristics. Additionally, the environmental management objectives and options are presented at European and International level. The benefits and costs of creating marine protected areas are also discussed.

Section 3 deals with the value of environmental protection and presents ways and practices through which marine protected areas can achieve sustainable development benefitting local communities and the environment. In particular, the benefits of the main sectors (fisheries, tourism) are presented, while reference is made to the environmental challenges ports face and have to deal with. Environmental management objectives and options are presented for the port sector, and their efficiency is discussed.

Finally, Section 4 tries to make a synthesis of all the above, summarizing the most effective options available.

2 Characteristics and Importance of Marine Protected Areas

Nowadays it is generally accepted that oceans and their means of livelihood are threatened. The increasing use of the coastal zone by humans incurs considerable damage to the marine ecosystems now producing tangible unfavorable consequences, such as the exhaustion of a great number of sea fish reserves. The concentration of all activities on a small coastal area creates pressure on the marine ecosystems, further enhancing the deterioration of habitats and possible loss of species and their genetic diversity (Klein et al., 2008).

In this framework, marine refuges and Marine Protected Areas (MPAs) are proposed as important tools for the tackling of such threats against sea resources and ecosystems (Leone, 2017). The term Marine Protected Area stands for an area within the boundaries of which a portion of sea or ocean waters is included. The term is used to describe sea areas protected against human activities, which usually include living, non-living and also historic and cultural resources. Marine protected areas include, according to their importance, either development, fishing and access restrictions or a total ban on the exploitation of the area (Helpern & Warner, 2002). Greece is one of the richest countries in biodiversity as 447 out of the 579 fish species living in the Mediterranean are found in the Greek seas (De Marttino & Morrillo, 2008). Finally, in addition, over 400 species of birds have been recorded in Greece, which hosts a considerable number of endangered bird species at European and international level. It thus becomes obvious that sea/inshore and island habitats are priority fields for Greece, since the country has the longest coastline in the Mediterranean (app. 15,000 km) and the coastal areas are of great economic, cultural and ecological interest (Leone, 2017).

2.1 Special Characteristics of Marine Protected Areas and International State of Affairs

Marine Protected areas often include areas of high ecological value, the marine biodiversity of which is protected and conserved for the benefit of current and future generations. Marine protected areas are considered with the aim of improving both fish management and the protection of the marine environment. Except for the protection of the sea environment, the coastal area and the Mediterranean as a whole is very important for the maintenance and reproduction of fish reserves (World Wildlife Fund, 2010). Their importance and the benefits to local communities, fish stocks, climate regulation, tourism and other areas of human activity have forced European and international governments to legislate for the protection of the marine environment and biodiversity. So, The Convention on Biological Diversity (CBD) of Barcelona to which Greece is a contracting party sets as an aim 10% of the Mediterranean surface being in protection regime until the end of 2014. Specifically, the rate of the Mediterranean surface (2,510,000 km^2) in protection regime (95,660 km^2) is only 3.8%, while the percentage of coastal areas protected (9910 km^2) is just 0.4% of the Mediterranean surface (Notteboom & Rodrigue, 2005). This percentage is considered to be particularly low since the coastal area and the Mediterranean as a whole is very important for the maintenance and reproduction of fish reserves (Helpern & Warner, 2002; Notteboom & Rodrigue, 2005).

Unfortunately, the rate of protection of the marine environment worldwide is judged to be particularly small, with the situation at international level being particularly alarming as only 0.65% of oceans and 1.6% of the coastal area is protected internationally. Unfortunately, however, although MPA's are worldwide

acknowledged as a measure of protection of the sea environment, only 2% of oceans is protected (Cole, 1999). The measures required include increasingly urgent international action to establish a world system of representative Marine Protected Areas' networks by 2015. This goal was set at the United Nations World Summit on Sustainable Development (WSSD) in Johannesburg by the world community in 2002 (Weeks, Russ, Bucol, & Alcala, 2010). The World Commission on Marine Protected Areas promotes this goal although it needs considerable reinforcement and the participation of more areas. In Greece, the creation and the establishment of MPA's has not followed the European standard due to minimal support from government bodies and reactions of stakeholders especially fishermen. This realization further enhances the need for finding managerial and financial tools for lasting and sustainable development of coastal areas and Greek seas (Kelleher & Rechie, 1998).

2.2 Economic Dimension of Marine Protected Areas Benefits

The benefits arising from the creation of Marine Protected Areas are present in all sectors of human activity. They particularly focus on tourism, fishing and infrastructures that are active in the coastal areas. Tourism in areas where MPAs will be created will be considerably strengthened thanks to the interest of scientists, who will want to study the marine environment and its habitats (Institute European Environmental Policy, 2017). At the same time sea tourism will be reinforced to a great extent since a great number of tourists but also local people will be keen on visiting the area thus reinforcing the local community. Furthermore, other sectors of the so-called alternative tourism, such as conference, education, touring, religious, medical tourism as well as other kinds of tourism, will also be developed; these types of tourism are unfortunately not widespread, because the mass tourism model is still dominant (Wood, Fish, Laughren, & Pauly, 2008).

The development of tourism will thus favour the enterprises of the coastal area, since, thanks to the big turnout of visitors, hotels and any type of accommodation facilities, restaurants and tavernas, as well as the local market of souvenirs and traditional products, will be conserved. Thus, benefits are numerous for the environment and the local community provided, however, the model of sustainable tourism is followed and the capacity of each area is respected (Naeem, Thompson, Lawler, Lawton, & Woodfin, 1994). Finally, the benefits for fishing offered by MPA's are also worth mentioning. The protection of the area and the ban of extensive fishing will help protect habitats and conserve them for the future (Greek Ministry of Tourism, 2017). In addition, oceans and their biotic features are used by pharmaceutical companies to produce medicines, as well as by the textile industry to obtain raw materials (Wood et al., 2008).

2.3 Cost: Liabilities

Although the creation of Marine Protected Areas offers benefits to local communities and fishermen, it also has a cost. Thus, the incurring cost-charges must be estimated within the framework of the cost benefit analysis (CBA). Specifically, the creation of protected areas within the marine area will give rise to an area where fishing will be totally banned and some other areas where fishing will be allowed on certain conditions (time, nets, and craft). This will entail an income decrease for local fishermen as a result of the ban. This situation will however be provisional and will last as long as the reproduction of the species. In the long term, following the lift of the ban, the quantity of the fishermen's catch of fish and consequently their income will increase (Maragou & Legakis, 2009). Another category of professionals that will be affected is large hotel units, which will limit their uncontrolled activities due to the protection and renewal of the coastal area. Of course, we can talk about sustainable tourism, provided the over-concentration phenomena of hotels and tourist accommodation facilities, which often do not respect any urban planning regulation and negatively affect the coastal ecosystems, progressively give in towards the protection of the natural land and sea environment (The Greek Tourism Confederation (SETE), 2019). Finally, benefits and costs are hard to quantify and calculate with accuracy because reference is made to qualitative data and the exact income or catch increase or decrease cannot be estimated. However, surveys conducted in many countries have shown that benefits arising from visitors consuming products and services on islands where MPAs were created are considerably higher than the temporary income decrease of fishermen and hotels stemming from the protection of coastal areas (ESPO: European Sea Port Organization, 2013).

2.4 Mediterranean and Greek Marine Protected Areas

The following Diagram 1 shows that marginal benefit (MB) and marginal external cost (MEC) meet at P3 when it comes to a country that has no limitations in terms of admissible pollution, while when it comes to a tourist country such as Greece, France or Australia, MB and MEC meet at P2 further on the left. Thus, we observe that the more on the left they meet (from P3 to P2) the less pollution is allowed in the tourism-oriented country because that country uses environmental protection as a comparative advantage. It enjoys a significant comparative advantage in the field of tourism through MPAs, attracting holidaymakers and visitors, who prefer visiting Greece for its clean, thoroughly protected environment and good weather conditions for the most months of the year. It is interesting to note that countries with limited pollution control (such as China) are beginning to introduce policies aiming at reducing the pollution level, leading to higher marginal environmental benefit, not necessarily driven by tourism development, but by the realization that such a course is unsustainable.

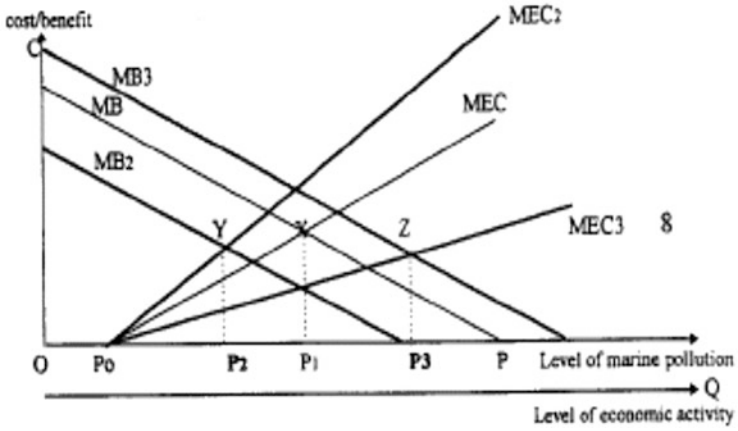


Diagram 1 Level of marine pollution: costs and benefit (Source: Elaborated by the authors)

3 Valorisation of Environmental Quality and Importance of Protection Measures and Tools

Coastal areas are among the most productive and of great biological diversity on the planet. They are also the most densely populated areas where, according to the United Nations 60% of the world's population is at a distance of 60 m from the coast (Galil & Zenetos, 2002). In addition, 80% of tourist activities take place on the coast, with beaches and coral reefs being at the top of the list of tourism preferences (Higgins-Desbiolles, 2010). It is also worth mentioning that, according to the latest data from the Ministry of Mercantile Marine and the Association of Hellenic Ports of Greece, passengers traveling through Piraeus and Rafina to all island destinations, during 2018 in the country increased by 10–20% compared to the corresponding period of 2017. In addition, according to Hellenic Civil Aviation Authority, Greek airports reached their peak in passenger traffic in 2018 (January to October,) with an increase of 9.6% in visitor arrivals compared to 2017. The main airports where most passengers traveled were Athens, Crete (Sitia), Rhodes, Naxos and Milos. It is obvious from the above data that a large portion of visitors visited coastal and Island tourist destinations. Taking into account that the number of tourists visiting Greece in 2018 reached 33 million compared to 30.1 million in 2017, a 4-year high record, according to the Greek Ministry of Tourism, highlights the actual numbers of visitors impacting these areas. For Greece, still facing an ongoing 8-year economic crisis, tourism is a driving force for the economy, accounting for 27.3% of country's GDP for 2017 according to the Greek Tourism Confederation (SETE) (Turley & Gattuso, 2012). It is also worth noting that Greece ranked 2017–2018 in the sixth place in the EU in number of beds. From the data above, it is clear that while tourism enhances the income of coastal-island regions, it also exerts great pressure on the coastal zone and the marine environment to the extent that coastal residents do not really benefit from the development of tourism. Mass tourism often destroys local culture without

greatly contributing to local income as foreign tour operators and investors capitalize most from these earnings-profits.

3.1 Challenges and Weaknesses in the Implementation of the MPAs in Greece

According to KNOEMA Corporation, Greece has been active in increasing the number of MPAs during the last years, since from 1.5% of territorial waters in 2016, marine protected areas rose to 4.5% in 2017, an increase of 210% from the previous year (Robert, 2012).

The protection of the environment and particularly of the marine environment through the establishment of MPAs has been a major issue in recent decades. However, in several countries among them and in Greece, the protected areas and the creation of such areas lag significantly because of some basic weaknesses that prevent the initial planning schemes from being realized. These weaknesses have to do with both the original protection objectives as well as internal weaknesses of each state.

On the global level, the 2006 Convention on Biological Diversity committed States to promote the conservation and protection of 10% of marine and coastal areas by 2012, a goal that has not been achieved. Researchers (Bianchi & Morri, 2000) stated in 2008, 2 years after the adoption of the Convention, that these objectives were unrealistic and unfeasible, as in 2008 only 1.6% of global marine areas under the national jurisdiction were considered effectively protected. Similarly, in 2005 a report by Kelleher et al. (Becker & Chorosh, 2010), indicated that the long time frames required to establish and operate a marine protected area are often due to objections from interested stakeholders such as fishermen, hoteliers and other stakeholders who, erroneously, feel economically threatened from such a development (Latchman, 2008). A further challenge arises from the lack of adequate funding and resources for the creation, operation and maintenance of these areas. Management and control bodies of the already created protected areas, face serious obstacles due to limited finances, thus failing to operate efficiently and achieve the purpose of their creation, which include the protection and conservation of biodiversity, the promotion of scientific research and the management of fish stocks in a sustainable way (Notteboom & Rodrigue, 2005; Prokopiou, Tselentis, Mavridoglou, & Lagos, 2014). However, there are cases where these MPAs help to “rebuild” the local communities by improving the capacity for fisheries management and protection (De Marttino & Morrillo, 2008). Studies have shown that protected areas act as nursing grounds for fish, which reproduce and finally replenish diminished stocks in adjacent areas, thus increasing fish catches (Wolansky, 2017).

3.2 Connection Between Ports and Marine Protected Areas

The study of marine protected areas cannot be disconnected from the existence of the ports of a country. What is known as "traditional port" is a coastal marine area with a land zone artificially formed to provide safe anchorage for vessels, but also secure mooring making sure that loading and unloading operations are carried out in safety, providing the necessary services to ship owners and traders (Contogiorgi, 2014). Different factors have affected the evolution and development of ports (Hillborn, Micheli, & De Leo, 2006):

- Social factors
- Economic factors
- Geographical factors
- Climate-related factors.

However, since ports operate in an extremely competitive framework, they continuously try to reduce cost, increase productivity with simultaneous improvement of the services provided to users. The crucial question concerning ports lies in the fact that the attempt to reduce cost, may lead to sacrifices in the safety and environmental protection. The same holds true for a port, which negatively affects the coastal area in which it operates, by chemical, and noise pollution (Institute European Environmental Policy, 2008). Port authorities must therefore take all necessary measures to prevent the emergence of pollution (accident-related or stemming from normal operation). If the frequency and consequences of accidents are known, the level of risk can be estimated for any type of environmental pollution. A fundamental advantage of shipping and the European port industry over land transport lies in the fact that environmental impacts on the marine environment are considerably lower. Over the past few years, a growing number of European ports have joined ESPO environmental programs (Millazo, Chemello, Badalamenti, Camarda, & Riggio, 2002):

3.3 Environmental Issues Stemming from Port Operation

According to a survey conducted on European ports, the main environmental problems concern the following categories: Vitousek et al. (1997).

- Development of ports towards the sea (45.9%)
- Quality of water (39.5%)
- Refuse and port waste (43.1%)
- Noise pollution (40%)
- Quality of air (51.4%)
- Disposal of materials stemming from dredging (49.8%)
- Dust (57.3%)
- Dangerous loads, traffic volume, ballast (35.9%)

3.4 European Sea Port Organization (ESPO) Objectives

The ESPO code of Practice and the most recent Green Guide, provide a list of important environmental issues and recommendations to ports with a view of complying with the environmental legislation and implementing policies to meet port obligations. In 2003 ESPO reviewed its environmental standards with the publication of a new environmental code of practice. In 2005 the ECOPORTS secretariat was set up to further strengthen voluntary implementation of environmental standards in the port industry. ECOPORTS objectives include the need for environmental management, cost reduction with environmentally friendly solutions and cooperation of ports with projects dealing with common environmental issues (Institute European Environmental Policy, 2008).

3.5 "GREEN" Ports and Marine Protected Areas

Over the last decade, green (sustainable) development has repeatedly been at the forefront and has been promoted by the European Union. This term generally concerns the taking of initiatives in all sectors (including tourism and shipping) to achieve environmental protection and rational use. Marine protected areas mainly foster green growth by protecting not only waters, habitats and the coastal area but also the cultural features of the area: they reinforce the local community by increasing the number of visitors in the area (Ducrotoy Jean-Paul, 2019). A discussion on environmental protection and sustainable development cannot but include ports (Bulling et al., 2010).

The "green" port pursues the protection of the coastal area and generally the mainland from harmful activities for the natural environment. It also promotes the port as a leading player in environmental protection by taking awareness-raising action that can unite the local community and encourage it to help in this direction (Myers, 1993). A green port, also, advances sustainability procedures and practices through the education and awareness of its executives concerning environmental protection, safety and modern technology in order to reduce environmental impacts produced by a diverse range of operations and activities (Law and Environment in Greece, 2017). Thus, the port and the mainland are closely linked and the operation of the former does not hamper the smooth operation of the latter. That means they complete one another; they do not exclude one another (European Commission, 2010).

4 Conclusions

From the above discussion, one can conclude that Marine Protected Areas (MPAs) support the benefit of their respective regions and local communities. They offer a double benefit on both the conservation and protection of the marine environment as well as, at local level, by offering services in all areas of human activity, most notably in the field of tourism. As stated, Greece is lagging behind in the implementation of such a successful and effective tool. The reasons lie in many areas of public governance. These involve access to environmental information, leading to transparency and public participation in the administrative decision-making procedures. Another issue concerns the environmental assessment of projects and the strategic environmental assessment of plans and programs, both of which are continuously hindered especially during the economic crisis. The environmental inspectorate, a very promising establishment, has never reached its true potential and has suffered serious cuts during the last few years, thus disallowing the government of a valuable asset arising from the fines imposed for environmental law violations. Connected to this, most of the Green Fund deposits are ineffective while only 2.5% are available annually for environmental protection and rehabilitation. Serious drawbacks in connection to the institutional protection of national protected areas is observed, since the Council of State issues decisions annulling official designation of national protected areas through joint ministerial decisions, rather than through presidential decrees, as required by Greek law. In reference to coastal waters and areas, problems in the development of the coastal and shoreline legislation and the legal framework for sustainable fishing, are noted, since amongst others, the integrated mapping of Greece's coastal zone and the ratification of the Integrated Coastal Zone Management Protocol of the Barcelona Convention, have not proceeded as planned.

In the port sector, progress has been steady since many ports and marinas have initiated environmental management systems and have acquired certifications such as ISO 14001, Port Environmental Review System (PERS), Blue Flag and Golden Anchor Award. It is important to recognize the major role a port can play in the protection of the environment. The port interacts with the natural and social environment on issues concerning air, soil, noise, quality of water and consumption of natural resources. The green port must thus continuously apply environmental protection policies and activate the responsibility of its executives and the community vis-à-vis the balance between the quality of the environment and socio-economic well-being.

For Greece with such a long coastline and many islands habitable throughout the year, the challenges are numerous and complex. Objectives set in Environmental Management Systems, such as Marine Protected Area establishment and effective management practices, with, no doubt lead to a series of benefits, economic and other, arising in all sectors of human activity. In doing so, Greece will enhance its international position and combined with the already high tourist profile, will become a global point of interest.

References

Becker, N., & Chorosh, Y. (2010). *Economic aspects of marine protected areas* (*MPA's*). Edited by: UNEP-MAP RAC-SPA, Tunis.

Bianchi, C. N., & Morri, C. (2000). Marine biodiversity of the Mediterranean Sea. Situations, problems and prospects for the future research. *Marine Pollution Bulletin, 40*(5), 367–376.

Brady, M., & Staffan, W. (2009). Fixing problems in fisheries – Integrated ITQs CBM MPAs in management. *Marine Policy, 33*(2), 258–263.

Bulling, M. T., Hicks, N., Murray, L., Paterson, D. M., Raffaelli, D., White, P. C. L., et al. (2010). Marine biodiversity–ecosystem functions under uncertain environmental futures. *Philosophical Transactions of the Royal Society B, 365*, 2107–2116.

Butler, R. W. (1991). Tourism, environment, and sustainable development. *Environmental Conservation, 18*(3), 201–209.

Coad, L., Campbell, A., Miles, L., & Humphries, K. (2008). *The costs and benefits of protected areas for local livelihoods: A review of the current literature* (Working Paper). Cambridge, UK: UNEP World Conservation Monitoring Centre

Cole, M. A. (1999). *Limits to growth, sustainable development and environmental Kuznets curves: An examination of the environmental impact of economic development, sustainable development*. Wiley Online Library.

Contogiorgi, C. (2014) *1st Ecoports workshop on waste facilities*. Viewed February 25, 2015, from http://ecoports.com/templates/workshop/PiraeusPortAuthority.pdf

De Marttino, M., & Morrillo, A. (2008). Activities, resources and inter-organizational relationships: Key factors in port competiveness. *Maritime Policy and Management, 35*(6), 571–589. https://doi.org/10.19044/esj.2014.v10n10p%25p.

Ducrotoy Jean-Paul. (2019). *Threats to the coastal zone*. Retrieved from http://www.coastalwiki.org/wiki/Threats_to_the_coastal_zone

ESPO: European Sea Port Organization. (2013). *The environmental priorities of European ports for 2013. An analysis taking part size and geography consideration*. Retrieved from http://ecoports.com/publications

European Commission. (2010). *Maritime affairs and fisheries: Study on the economic effects of maritime spatial planning*. Brussels: European Commission Marine Affairs and Fisheries.

Food and Agriculture Organization of the Fisheries Management of the Fisheries Management (FAO). (2011). *Marine protected areas and fisheries*. Rome: FAO.

Frederigo, C. (2009). *Marine protected areas: An essential tool for the future European maritime policy*. UK: University Press.

Galil, B. S., & Zenetos, A. (2002). A sea change – Exotics in the Eastern Mediterranean Sea. In E. Leppäkoski, S. Gollasch, & S. Olenin (Eds.), *Invasive aquatic species of Europe: Distribution, impacts and management*. Dordrecht: Springer.

Greek Ministry of Tourism. (2017). *Annual report about tourism in Greece*. Retrieved from http://ec.europa.eu/enterprise/sectors/tourism/documents/annual-reports/index_en.html

Helpern, B. S., & Warner, R. R. (2002). Marine reserves have rapid and lasting effects. *Ecology Letters, 5*(3), 361–366.

Higgins-Desbiolles, F. (2010). The elusiveness of sustainability in tourism: The culture-ideology of consumerism and its implications. *Tourism and Hospitality Research, 10*(2), 79–115. https://doi.org/10.1057/thr.2009.31.

Hillborn, R., Micheli, E., & De Leo, G. A. (2006). Integrating marine protected areas with catch regulation. *Canadian Journal of Fisheries and Aquatic Sciences, 63*, 642–649.

Institute European Environmental Policy. (2008). *Marine protected areas in Europe and the United States*. Edited by European Commission. Belgium: IEEP.

Institute European Environmental Policy. (2017). *Socio economic benefits of the EU marine protected areas*. London: Institute European Environmental Policy.

Kelleher, G., & Rechie, C. (1998, June). Lessons from marine protected areas around the world. *Parks, 8*(2).

Klein, C. J., Chan, A., Kircher, L., Cundiff, A. J., Gardner, N., Hrovat, Y., et al. (2008). Striking a balance between biodiversity conversation and socioeconomic viability in the design of marine protected areas. *Conservation Biology, 22*, 691–700.

Latchman, I. (2008). *Marine protected areas, benefits and costs for islands*. Zeist: WWF.

Law and Environment in Greece. (2017, October). *WWF*. Retrieved from http://www.wwf.gr/images/pdfs/WWF-NOMO-2017-%20Synopsis-EN.pdf

Leone, G. (2017). *Mediterranean quality status report. U.N Environment programme*. Mediterranean Action Plan (MAP), Barcelona Convection.

Maragou, P., & Legakis, A. (2009). *The update of the Greek red data book of threatened fauna: Summary results and trends*. Athens: Hellenic Zoology Society.

Millazo, M., Chemello, R., Badalamenti, F., Camarda, R., & Riggio, S. (2002). The impact of human recreational activities in marine protected areas: What lessons should be learnt in the Mediterranean Sea? *Marine Ecology, 23*(1), 280–290.

Myers, N. (1993). Biodiversity and the precautionary principle. *Ambio, 22*, 74–79.

Naeem, S., Thompson, L. J., Lawler, S. P., Lawton, J. H., & Woodfin, R. M. (1994). Declining biodiversity can alter the performance of ecosystems. *Nature, 368*, 734–737.

Naidoo, R., Balmord, A., Ferraro, P. J., Poalsky, S., Rickets, T. H., & Rouget, M. (2006). Integrating economic costs into conservation planning. *Trends in Ecology & Evolution, 21*, 681–687.

Notteboom, T., & Rodrigue, J.-P. (2005). Port regionalization: Towards a new phase in port environment. *Marine Policy & Management, 32*(3), 297–313.

Palumbi, S. R. (2009). *The ecology of marine protected areas*. Vancouver, BC: University of British Columbia.

Prokopiou, D. G., Tselentis, B. S., Mavridoglou, G., & Lagos, D. (2014). Prototype model of carrying capacity in tourism: The implementations for the Island of Rhodes. *European Scientific Journal, 2*, 434–442.

Robert, P. (2012, April, 2). *Tourism management in MPA's as an economical way* (p. 23). Medaset workshop, Milos.

Shiffman, D. (2018). *We've only really protected 2 percent of the ocean*. Retrieved from www.earther.gizmodo.com

The Greek Tourism Confederation (SETE). (2019). Retrieved from https://sete.gr/

Turley, C., & Gattuso, J.-P. (2012, July). WWF – Threats to oceans and coasts, Future biological and ecosystem impacts of ocean acidification and their socioeconomic policy implications. *Current Opinion in Environmental Sustainability, 4*(3), 278–286.

Vitousek, P. M., Mooney, H. A., Lubchenco, J., & Melillo, J. M. (1997). Human domination of Earth's ecosystems. *Science, 277*, 494–499.

Weeks, R., Russ, G. R., Bucol, A. A., & Alcala, A. C. (2010). Shortcuts for marine conservation planning: The effectiveness of socio-economic data surrogates. *Biological Conservation, 143*, 1236–1244.

Wolansky, E. (2017). *Protective functions of coastal forests and trees against natural hazards*. Townsville: Australian Institute of Marine Science.

Wood, L. J., Fish, L., Laughren, J., & Pauly, D. (2008). Assessing progress towards global marine protection targets: Shortfalls in information and action. *Oryx, 42*(3), 340–351.

World Wildlife Fund. (2010). *Conservation science program*. Washington, DC: Society of Conservation Biology.

A Hybrid Firefly Algorithm Based on Coordinates for the Prize-Collecting Vehicle Routing Problem

Manousos Rigakis, Dimitra Trachanatzi, Magdalene Marinaki, and Yannis Marinakis

Abstract This paper investigates the Prize-Collecting Vehicle Routing Problem (PCVRP), to simulate a tourist trip design problem, and the solution of it via a hybrid Firefly Algorithm (FA), namely the Firefly Algorithm based on Coordinates (FAC). To the best of our knowledge, there is no publication found in the literature, focusing on the solution of the PCVR via FA. The hybridization that we propose is founded on the position, in the 2D-space, of each node included in a solution. Thus, the update mechanism of the original FA can be applied on non-probabilistic, continuous, problem-rated values. In order to demonstrate the effectiveness of the proposed algorithm, computational experiments were conducted over benchmark instances found in the literature. The results obtained by the FAC were compared to the corresponding solutions of another hybrid metaheuristic algorithm, the Distance Related Differential Evolution (DRDE) Algorithm and the CPLEX solver.

Keywords Prize-collecting vehicle routing problem · Firefly algorithm based on coordinates · Solution encoding

1 Introduction

The ever increasing importance of the tourism industry to the global economy drives the scientific community close to the simulation, modelling and optimization of problems that promote and aid this sector. For instance, a number of publications are focused on algorithmic frameworks that optimize itineraries for tourists when visiting a city, a country or a museum, while they are bound by several constraints, e.g. time restrictions, opening hours, transport connection/time schedule, ticket availability, budget limits and others. Currently, an area of special interest is the

M. Rigakis · D. Trachanatzi (✉) · M. Marinaki · Y. Marinakis
School of Production Engineering and Management, Technical University of Crete, Chania, Greece
e-mail: mrigakis@isc.tuc.gr; dtrachanatzi@isc.tuc.gr; magda@dssl.tuc.gr; marinakis@ergasya.tuc.gr

E. Krassadaki et al. (eds.), *Operational Research in Agriculture and Tourism*, Cooperative Management, https://doi.org/10.1007/978-3-030-38766-2_8

tourist recommendation systems, as through these individuals or even tourist groups, are able to customise their trip based on their preferences and needs (Gavalas, Konstantopoulos, Mastakas, & Pantziou, 2014). Thus, route-planning problems emerge to help tourists determine the most suitable tour for them. Such problems are: the Orienteering Problem (OP), the Team Orienteering Problem (TOP), the Tourist Trip Design Problems (TTDP), the Prize-Collecting Vehicle Routing Problem (PCVRP) and others, as different variants of the Vehicle Routing Problem with Profits (VRPPs) (Stavropoulou, Repoussis, & Tarantilis, 2019). The common principle among the aforementioned problems is that each Point of Interest (POI), that is considered as an available node in a tourist trip, is associated with a profit/prize value.

This paper focuses on the PCVRP, a problem originally inspired by the hot rolling production of the iron and steel industry, that has been studied in a number of publications relevant to its practical applications. Nevertheless, we consider that the formulation of the problem is able to successfully simulate a tourist trip and that it could be integrated in a tourist recommendation system. In order to describe the PCVRP, we identify the following characteristics: a homogeneous fleet of vehicles is available to conduct a specified number of routes; each vehicle has a specific capacity, that is not to be exceeded; all vehicles must initiate from, and end their route at the depot, while each node (POI) can only be visited at most once. Moreover, the PCVRP includes profit values associated to each available node, while, based in its definition, it is not compulsory to visit every node. The objective of the problem is to determine a set of feasible routes that minimize the total occurred cost and simultaneously maximize the total prize collected from the visited nodes. Calculating the total cost, both the travelled distance among the nodes in the optimal route and a fixed cost for every constructed route, are considered. One important feature of the PCVRP is the task completion ratio, as it restricts the minimization process to a lower bound of the total demand aggregated from the visited nodes. Expressing the PCVRP as a tourist route-planning problem, we can relate: the total distance travelled as cost (travel expenses, fuel price, or even time); the fixed cost of each vehicle/route as accommodation expenses (based on the number of days); and the prize collected as the satisfaction that individuals get following the proposed trip. As for the demand that typically is associated with each node, one could perceive it as the time spend at a POI when it is visited, and that, consequently, the task completion constraint facilitates the quality of the trip, in terms of sufficient time devoted to the satisfactory POIs.

As a VRP variant, the PCVRP is considered to be an NP-hard problem, and such, large instances of it could not be solved by exact approaches. In order to address that, we propose the Firefly Algorithm (FA), a meta-heuristic algorithm inspired by the natural phenomenon of fireflies' flashing light, to solve the PCVRP. The main occurred obstacle is that like other swarm based algorithms, the FA was originally introduced for the solution of continuous optimization problems. According to the literature, researchers proposed the hybridization of the FA to discrete versions along with other heuristic techniques or even, other algorithms. In this paper, we consider that the original framework of the FA should remain intact and we propose a new

encoding of the discrete representation of the route solution, in a way that the movement equation could be applied directly, along with a reversion mechanism. The presented approach is referred as the Firefly Algorithm based on Coordinates (FAC). The quality of the solutions produced by the FAC are compared to a hybrid Differential Evolution Algorithm, since there is a lack of relative published results. The results indicate that the FAC produces high quality solutions and it could be effectively incorporated into a tourist recommendation system.

The rest of the paper is organized as follows: Sect. 2 contains a brief literature review; in Sect. 3 the mathematical model of the PCVRP is described; in Sect. 4 the original FA is presented; in Sect. 5 the FAC algorithm is introduced and described in detail; Sect. 6 contains the experimental results; and Sect. 7 gives the conclusion.

2 Brief Literature Review

In 2006, Tang and Wang introduced the Prize-Collecting Vehicle Routing Problem (PCVRP) to solve the hot rolling production scheduling problem (Tang & Wang, 2006). The proposed mathematical formulation of the PCVRP incorporated a linear combination of three objectives: the total distance minimization; the number of utilized vehicles minimization and the maximization of the collected prizes. Thus, the model can satisfy different requirements by altering the value of the three coefficients, respectively. To optimize the PCVRP, they utilized an Iterated Local Search (ILS) algorithm based on Very Large-Scale Neighbourhood (VLSN). In the same field, Zhang et al. proposed a multi-objective PCVRP-based model for the hot-rolling batch scheduling problem, that included penalty for the non-visited customers and a fixed number of vehicles, and solved it using a Particle Swarm Optimization (PSO) variant (Zhang, Chaovalitwongse, Zhang, & Pardalos, 2009). The aforementioned problem with similar mathematical formulation was also solved by Jia, Yi, Yang, Du, and Zhu (2013) using a Pareto Max-Min Ant System (P-MMAS).

Hence, two variants of the PCVRP formulation have been derived from the literature: when the number of vehicles to be utilized is predetermined (PCVRP-P) and when is not predetermined (PCVRP-NP) (Long et al., 2019). In 2015, Tiwari et al. proposed a hybrid edge recombination approach for the PCVRP-P (Tiwari, Chang, Elangovan, & Annadurai, 2015). Subsequently, Li and Tian developed a two-level self-adaptive variable neighbourhood search for the PCVRP-NP, aggregating the multiple objectives in one function through the weighted sum method (Li & Tian, 2016). Lately, in 2018, Long et al. studied both version of the PCVRP as a multi-objective problem, using a Pareto-based evolutionary algorithm, combined with a local search strategy. In order to handle the PCVRP-NP, they proposed a decomposition strategy that divides the problem into multiple PCVRP-P problems (Long et al., 2019).

Firefly Algorithm (FA) is a swarm based meta-heuristic algorithm that was developed to solve continuous optimization problems and due to its success, researchers utilized it to solve different optimization problems and proposed various

modifications and improvements, as shown in surveys (Fister, Yang, & Brest, 2013; Tilahun & Ngnotchouye, 2017; Yang, 2014). The FA has been rarely adopted in solving vehicle routing problems, due to the discrete nature of the problem and the corresponding encoding of the solutions-fireflies. Initially, Jati (2011) addressed the Traveling Salesman Problem (TSP), enhancing the classic FA with an evolutionary and discrete behavior. In 2013, Pan, Ye, Wang, and Cao (2013) proposed the FA for the solution of the Vehicle Routing Problem with Time Windows (VRPTW), including coding and disturbance mechanisms. In addition, a practical application on the solution of the VRP with heterogenous fleet was presented by Simić, Kovačević, Svirčević, and Simić (2015), utilizing a hybrid genetic–firefly model. Alinaghian and Naderipour (2016) proposed a hybrid variant of the FA, the Gaussian FA, to solve a time-dependent VRP with multi-alternative graph, aiming to reduce the fuel consumption. A discrete FA variant was, also, employed to solve a rich vehicle routing problem, which can be more specifically considered as an asymmetric and clustered vehicle routing problem with simultaneous pickup and deliveries, variable costs and forbidden paths (AC-VRP-SPDVCFP) (Osaba et al., 2017). Recently, a hybrid combination of the FA and the Ant Colony System (ACS) was developed to solve VRPs, proposing a new distance measurement technique (Goel & Maini, 2018).

3 Prize-Collecting Vehicle Routing Problem

The present paper is focused on the solution of the PCVRP-P, thus, a predefined number of vehicles have to be used, i.e. a predefined number of feasible routes M have to be constructed to visit the available points of interest. Following the formulation presented by Li and Tian (2016), the Prize-collecting Vehicle Routing Problem can be described through a complete graph $Z = (V, A)$, where $V = \{0, \ldots, N\}$ is the node set and $A = \{(i,j) | i,j \in V\}$ is the set of corresponding arcs.

Each node i included in the set $\{1, \ldots, N\}$, represents a customer and such, specific values of prize, p_i and demand d_i are associated to it. The depot, as the initial/starting point, is denoted by node 0, and no prize or demand value is attributed to it. In addition, for each pair of nodes i, j, the travelling time between them can be expressed by their Euclidean distance, noted as c_{ij}, the symmetry of the problem defines that $c_{ij} = c_{ji}$. Furthermore, each vehicle has a maximum capacity Q and a large fixed usage cost G associated to it. Additionally, r denotes the task completion parameter (minimum ratio), obtained by dividing the predetermined amount of demand by the total demand of all nodes. Finally, as N_m is considered the set of nodes that are visited by vehicle m, $(m = 1, \ldots, M)$. The following decision variables are used:

- $x_{ij} (i \neq j) = 1$ when node j is visited immediately after node i in the optimal solution, otherwise $x_{ij} = 0$, (i, j ∈ V).
- $x_{ii} = 0$ when customer i is visited in the optimal solution, otherwise $x_{ii} = 1$, $(i \in V)$.

The mathematical formulation of the PCVRP-P is:

$$Minimize: \sum_{i\in N}\sum_{j\in N, i\neq j} c_{ij}x_{ij} + G \times M - \sum_{i=1}^{N} p_i(1 - x_{ii}) \quad (1)$$

s.t.

$$\sum_{i=1}^{N} x_{i0} = \sum_{i=1}^{N} x_{0i} = M \quad (2)$$

$$\sum_{i=1}^{N} x_{ij} \leq 1, j = 1, \ldots, N \quad (3)$$

$$\sum_{j=1}^{N} x_{ij} \leq 1, i = 1, \ldots, N \quad (4)$$

$$\sum_{i\in N_m} d_i(1 - x_{ii}) \leq Q, m = 1, \ldots, M \quad (5)$$

$$\sum_{i\in S}\sum_{j\in S,(j\neq i)} x_{ij} \leq |S| - 1, \quad \forall S \subset V \quad (6)$$

$$\frac{\sum_{i=1}^{N} d_i(1 - x_{ii})}{\sum_{i=1}^{N} d_i} \geq r \quad (7)$$

$$x_{ij} \in \{0, 1\}, i, j \in V \quad (8)$$

The goal of the objective function (Eq. 1) is the minimization of the total cost (travelled distance and fixed vehicle usage) by taking into account the total collected prize from the visited nodes. Constraint (2) requires that each vehicle conducts a route that initiates from the depot and returns to it. Constraints (3) and (4) ensure that each node is visited at most once, while constraints (5) facilitate the capacity restrictions of the vehicle. Additionally, constraints (6) are used to eliminate the sub-tours for each vehicle route. Constraint (7) ensures the minimum ratio of demand to be covered. Finally, constraints (8) specify the integrity conditions on the variables.

4 The Original Firefly Algorithm

The Firefly Algorithm was inspired by the process of bio-luminescence, which, in nature, fireflies utilize to attract both potential partners and prey or as a warning mechanism. The FA was proposed by Yang (2009), using each firefly's *light intensity* as the stepping stone to simulate the quality of a solution, i.e. the objective function, and the movement mechanism of the fireflies in a D-dimensional space. In general, a bright firefly should attract a less-brighter one to move towards it, but light intensity is not the sole factor of this mechanism. As the distance between two fireflies increases, the light intensity decreases according to inverse square law and

such a firefly becomes less attractive. Yang established three main rules to simulate the described behavior:

1. All fireflies are unisex and all are able to attract each-other regardless of their sex;
2. Attractiveness is proportional to their brightness, thus, for any two flashing fireflies the less brighter one will move towards the brighter one. The attractiveness is proportional to the brightness and they both decrease as their distance increases. If there is no brighter one than a particular firefly it will move randomly;
3. The brightness of a firefly is affected or determined by the landscape of the objective function.

Thus, each firefly holds two important traits, the **light intensity** (I) and the **attractiveness** (β). The light intensity of a firefly at position x, regarding an optimization problem, is directly related to the value of the objective function f(x) (in case of maximization) or reciprocal of it (in case of minimization) (Yang, 2010). The attractiveness is proportional to the light intensity seen by adjacent fireflies and will vary with the distance r_{ij} between fireflies i and j, with respect to the 2-D Cartesian distance equation (Eq. 9) in a medium with fixed light absorption coefficient γ (Eq. 10), while β_0 denotes the attractiveness at $r = 0$.

$$r_{ij} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2} \tag{9}$$

$$\beta = \beta_0 e^{-\gamma r^2} \tag{10}$$

The movement of a firefly i that is attracted to a more attractive firefly j, at time t, is determined by the following equation:

$$x_i^{t+1} = x_i^t + \beta_0 e^{-\gamma r_{ij}^2}\left(x_j^t - x_i^t\right) + \alpha_t \epsilon_i^t, \tag{11}$$

where the last term is randomization that includes a randomization parameter (α_t) and a vector of random numbers drawn from a Gaussian or uniform distribution $\left(\epsilon_i^t\right)$, at time t, respectively. The parameter setting is important in order to control the diversity of the solutions, while parametric studies (Yang & He, 2013) suggest that:

$$\alpha_t = \alpha_0 \delta^t, \delta \in [0.95{,}0.97] \tag{12}$$

where for average problem scale $L, \alpha_0 = 0.01L$ and $\gamma = 1/\sqrt{L}$. In addition, $\beta_0 = 1$ can be used for most applications. Algorithm 1 shows the pseudo-code of the FA as originally proposed by Yang (2009).

Algorithm 1 Firefly Algorithm

```
Define the objective function f(x)
Initialise the firefly population X = x_1, ..., x_n
Define the light absorption coefficient γ
for each x_i in the population do
  Initialise light intensity I_i
end for
repeat
  for each x_i do
    for each x_j do
      if I_j > I_i then
        Move firefly x_i towards x_j, (Equation 3.3)
      end if
      Vary the attractiveness, (Equation 3.2)
      Evaluate new solutions and update light intensity
    end for
  end for
  Rank the fireflies and find the current best
until termination criterion reached
Rank the fireflies and return the best one
```

5 Firefly Algorithm Based on Coordinates

As already mentioned, the FA, in the original formulation, could not be directly applied to solve a Vehicle Routing Problem, as the solution should be expressed in a discrete way to represent a node sequence. Other studies focusing on optimization problems with non-continuous variables, adopt techniques that update either the continuous space or the discrete space. Specifically, in the former technique the movement equation (Eq. 11) is applied to the solution and, then, a mechanism is used converting it to discrete values. For instance, such mechanisms entail the use of: the sigmoid function (Sayadi, Ramezanian, & Ghaffari-Nasab, 2010); the hyperbolic function (Chandrasekaran, Simon, & Padhy, 2013) and rounding after movement in the continuous space (Baghlani, Makiabadi, & Sarcheshmehpour, 2014). In the later technique, the update solution will remain discrete after the movement mechanism. An example of this technique can be found in the solution of a VRP when Osaba et al. (2017) proposed the use of Hamming distance to calculate the distance between the fireflies, and, based on that, to compute the number of nodes to relocate via a relocation technique. A similar edge-based sequential movement is proposed by Singh (Singh, Thapar, Bhatia, Singh, & Goyal, 2015). A recent, detailed review on the aforementioned methods is presented by Tilahun and Ngnotchouye (2017).

5.1 Description of the Coordinates Related Method

Our method is related to "updating the continuous space" case, while in the same time all the probabilistic conditions that a function or rounding method will entail, are eliminated. We denote our method as "Coordinates Related" (CR) and is based on: (a) the coordinates of each node (the position in the 2D-space) and (b) the Euclidean distance between the nodes of a given VRP. To be exact, a VRP solution is typically represented by a vector of discrete values, as a sequence of nodes, as visited by a specific number of vehicles. In the proposed method, the solution vector is related to two new vectors (of the same length) with continuous values, the respective x and y coordinates of the each node in the 2D-space. An example of the transformation is depicted in Table 1, of a solution with two routes, where the depot (node: 1) is located at [0,0].

Following the algorithmic framework of the FA, a firefly is attracted to a brighter one, namely, a solution will move towards a better one. The movement is controlled by the update equation (Eq. 11), that includes the distance between the fireflies (solutions). We consider that the distance of two given solutions, can be computed as the minimum Euclidean distance of each pair of nodes, in the same position, in the two different vectors. In case of solution vectors that differ in their length, only the distance of each possible pair (of nodes) is included in the computation. Figure 1 depicts the distance of the solution vectors presented in Table 2. The solution vectors are depicted in Table 2 and 3 with bold values.

At this point, we can calculate the attractiveness, following Eq. (10). Assuming that the light intensity of Solution 1 is bigger than the light intensity of Solution 2, Solution 2 should be updated, following Eq. (11). In practice, the later equation corresponds to two different computations, one for each dimension (x and y), and one new set of coordinates is produced as the updated version of Solution 2. The main principle is that the CR attempts to relate the new values of [x,y] to those of nodes from the originally given set. For example, as seen in Table 3, correlating the new coordinates regarding the positions 2 and 4, i.e. [61.020,61.875] and [3.214,39.035], the new updated solution emerges incorporating the nodes 15 and 8, with coordinates [61,59] and [84,39], respectively.

5.2 Description of the Proposed Algorithm

The proposed Firefly Algorithm based on Coordinates (FAC), is based on the original FA framework and includes heuristic components to: (1) create the initial solution population; (2) ensure the feasibility of the updated solutions; and (3) to improve the updated solution. Thus, the framework of FAC incorporates the following processes: the *Initial − Swarm*, the *CR*, the *Enforce − Feasibility*, the *Augment − Route*, the *Remove − node* and the *Exchange − node*, as shown in Algorithm 2.

Table 1 Example of the transformed vector according to DR

Position	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Solution vector	1	4	8	9	5	2	1	12	3	6	7	10	11	1
Transformed vector-x	0	13	14	2	29	50	0	3	49	58	84	14	5	0
Transformed vector-y	0	7	24	39	89	5	0	82	8	30	39	24	10	0

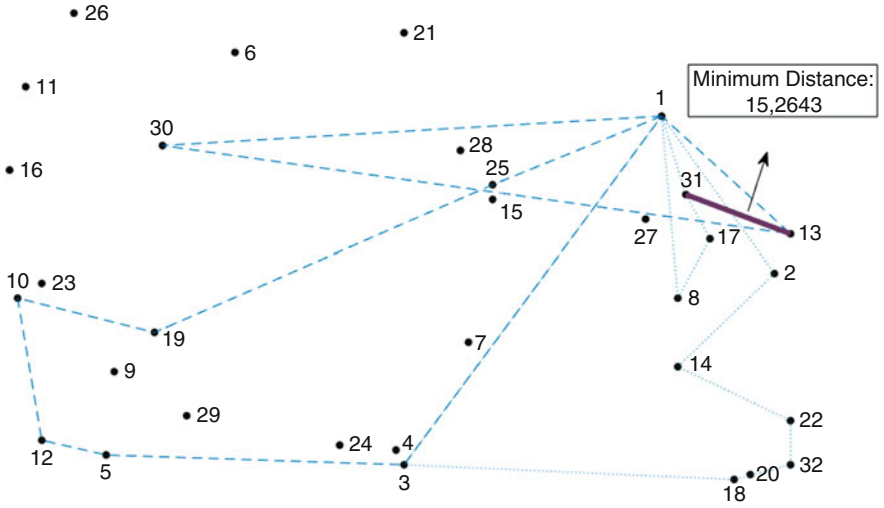


Fig. 1 Distance example

Table 2 Example of Solution vectors

Position	1	2	3	4	5	6	7	8	9	10	11	12
Solution vector 1	**1**	**2**	**14**	**22**	**32**	**18**	**3**	**1**	**31**	**17**	**8**	**1**
Coordinate x	82	96	84	98	98	91	50	82	85	88	84	82
Coordinate y	76	44	25	14	5	2	5	76	60	51	39	76
Solution vector 2	**1**	**25**	**19**	**10**	**12**	**5**	**3**	**1**	**13**	**30**	**1**	
Coordinate x	82	61	19	2	5	13	50	82	98	20	82	
Coordinate y	73	62	32	39	10	7	5	76	52	70	76	

Table 3 Example of updated Solution 2

Position	1	2	3	4	5	6	7	8	9	10	11
Solution vector 2	**1**	**25**	**19**	**10**	**12**	**5**	**3**	**1**	**13**	**30**	**1**
Coordinate x-NEW	82.012	61.020	18.998	3.214	5.122	13.425	49.897	82.350	98.314	20.056	82.100
Coordinate y-NEW	73.101	61.875	32.042	39.035	10.010	7.865	5.087	75.999	52.610	70.804	76.125
Updated Solution 2	**1**	**15**	**19**	**8**	**12**	**5**	**3**	**1**	**13**	**30**	**1**

Algorithm 2 Firefly Algorithm based on Coordinates

```
Define: size population W, number of iterations L
Initialise the firefly population via Initial − Swarm (Algorithm 3)
Define the light absorption coefficient γ
for each x_i in the population do
  Compute the objective function value f_i
  Initialise light intensity I_i < −1/f_i
end for
repeat
  for each x_i do
    for each x_j do
      if I_j > I_i then
        Calculate distance of x_i and x_j, (Equation 3.1)
        Move firefly x_i towards x_j, using the node coordinates, (Equation
        3.3)
        Construct the updated x_i: x'_i based on the new coordinates
        Make the x'_i feasible via Enforce − Feasibility
        Improve x'_i via Augment − Route
        Improve x'_i via Remove − node (Algorithm 4)
        Improve x'_i via Exchange − node (Algorithm 5)
      end if
      Vary the attractiveness (Equation 3.2)
      Evaluate new solutions and update light intensity
    end for
  end for
  Rank the fireflies and find the current best
  Update the randomisation parameter (Equation 3.4)
until L iterations are reached
Rank the fireflies and return the best one
```

To create the initial population the *Initial* − *Swarm* heuristic method is employed, that is based on the savings algorithm published by Clarke and Wright in 1964 (Clarke & Wright, 1964). The first step is the formulation of M initial routes that include only the depot and one other node. Afterwards, these routes are combined into an initial solution vector and the non-included nodes are sorted based on their prize value. One-by-one these nodes are positioned into the solution vector according to the savings method. That implies that the most efficient position (between nodes i, j) of a new node k to be inserted is where the following expression is minimized: $c_{i,k} + c_{k,j} - c_{i,j}$. Each route of the solution is augmented by nodes up to the predefined limit of capacity Q, while the capacity reached from the complete solution should exceed the capacity threshold Q_{low} (where $Q_{low} = rQ$).

Algorithm 3 *Initial − Swarm*

```
repeat
  Create M initial routes: [1 node 1]
  Combine the initial routes to vector
  while Total capacity < Q_low do
    Create stack: sorted non-included nodes
    for Each node k in stuck do
      Calculate the efficient position of node k
      Correlate position to route m
      if including node k in route m does not violate the capacity con-
      straint Q then
        Include k and update the solution
      end if
    end for
  end while
  Calculate the solutions' value in the objective function
until Population W is constructed
```

The most important heuristic technique is the one that enforces the feasibility of the solution. When the estimation of nodes into the updated solution is completed, after the movement equation, the new solution vector could include some error, e.g. more or fewer routes, repeated node usage and out of bounds achieved capacity (per route or total). Thus, the *Enforce − Feasibility* is completely problem oriented and should cover all the possible scenarios in a robust way. Another used technique is the *Augment − Route*, where the savings method is, once more, employed to enlarge each route with respect to all the aforementioned constraints. Due to the contribution of the total collected prize to the objective function, is highly probable, that by adding more nodes, into cost efficient positions, increasing the total prize, a better solution could be produced.

Subsequently, two local search techniques are employed for a predefined number of iterations, the *Remove − node* and the *Exchange − node*. First, under the *Remove − node*, the algorithm attempts to remove nodes from the solution, reconstructing it and achieves a better quality solution without violating the capacity constraints. The adopted principle is that, when a node with low prize value is removed from the solution, the remaining solution is connected with a new arc, and if that new connection is efficient in terms of cost (distance), then, a better solution (with lower objective function value) could emerge, see Algorithm 4. Subsequently, the *Remove − node* is used, when for a number of iterations, nodes are exchanged between two random routes of the solution. The first node to be determined is the one that has the greater distance from its predecessor, while the second one is selected randomly. If the new solution is changed to a better one, based on the value of the objective function, it replaces the initial one in the population.

Algorithm 4 *Remove − node*

repeat
  Find route m with the highest value of total distance travelled
  Find the node k with the smaller prize in route m
  if Reducing the capacity of route m by the prize of node k does not violate the constraints **then**
    Calculate new distance travelled, connecting the nodes immediate before and after k
    Calculate the objective function of the new formation
    if New objective function value is smaller than the initial one **then**
      Update the solution accordingly
    end if
  end if
until Max iterations reached

Afterwards, the local search technique *Exchange − node* takes place, that is a repetitive "1-1 exchange" procedure among random routes in a solution, for a specified number of iterations. Every pair of routes is randomly selected, as is one of the positions (a node) from one route.

Algorithm 5 *Exchange − node*

repeat
  Randomly select routes: m_1 and m_2
  Randomly select position in m_1, node: i
  Compute distance vector of route m_2
  Select position with the greatest distance value, node j
  Compute total demand of both routes after the exchange
  if New total demand values do not exceed the capacity limit **then**
    Calculate new distance travelled, exchanging the position of nodes i, j
    Calculate the objective function of the new formation
    if New objective function value is smaller than the initial one **then**
      Update the solution accordingly
    end if
  end if
until Max iterations reached

The second position, i.e. the other node from the second route, is calculated using a distance vector. The end of the most expensive (greatest length) connection indicates the position of interest. The described node exchange aims to create a solution with less distance travelled (lower in cost), by omitting inefficient sequences of nodes and creating more efficient connections, as seen in Algorithm 5.

6 Experimental Study

This section presents the experimental results of this study, solving the PCVRP-P, in order to prove the solution quality of the proposed FAC. The benchmark instances are drawn from the recent literature (Long et al., 2019), while there is a lack of comparison with other published algorithms, due to different benchmark sets, unpublished results and multi-objective solutions (Pareto-fronts). Thus, the obtained results of FAC are compared with those of the Distance Related Differential Evolution (DRDE) algorithm. The DRDE, is part of the authors previous work, while it follows the same principle and utilizes the Euclidean distance between nodes to apply the update mechanism of the original Differential Evolution (DE) (Trachanatzi et al., 2019).

6.1 Benchmark Instances

As the PCVRP has not been studied extensively, previous publications used the benchmark instances of the Capacitated Vehicle Routing Problem (CVRP), that can be downloaded from: http://www.coin-or.org/SYMPHO/NY/branchandcut/VRP/data/index.htm.old, incorporating a prize value to each customer and a minimum demand served ratio r, both generated uniformly from [1,80] or [1,100] and [0.6,0.8], respectively. However, recently, Long et al. published a set of 120 instances for the PCVRP-P, accessible at https://github.com/longjianyuGH/PCVRP.git (Long et al., 2019). In this study, we consider a set that consists in total of 115 benchmark problems, as for each of the 24 CVRP instances (group: A, B, E, M), 5 versions were generated, by changing the ratio r: {0.60,0.65,0.70,0.75,0.80}. The different variants include problems with number of nodes from 32 to 200, and number of vehicles from 4 to 17.

6.2 Computational Results

The FAC advantage is that it requires the same parametrization as the original FA. Following the study of Yang and He (2013) and considering that the problem scale is related to the number of the available nodes N in the solutions space, we compute the parameter values, as seen in Table 4.

Moreover, an exploration study was conducted that resulted in the efficient values of: the population size $W = 80$, the number of iterations $L = 500$, and the number of iterations of both local search techniques, i.e. the $Remove - node$ and the $Exchange - node$. In Fig. 2 we present the value of the objective function on average, over five executions, for a specific instance ($A - n32 - k5 - 1$), as the number of local search iterations increases. Accordingly, the percentage shows the

Table 4 Setting the parameters

Parameter	Description	Expression
a_0	Initial randomness scaling factor	0.01 N
δ	Cooling factor	*rand* [0.95,0.97]
γ	Light absorption coefficient	$1/\sqrt{N}$

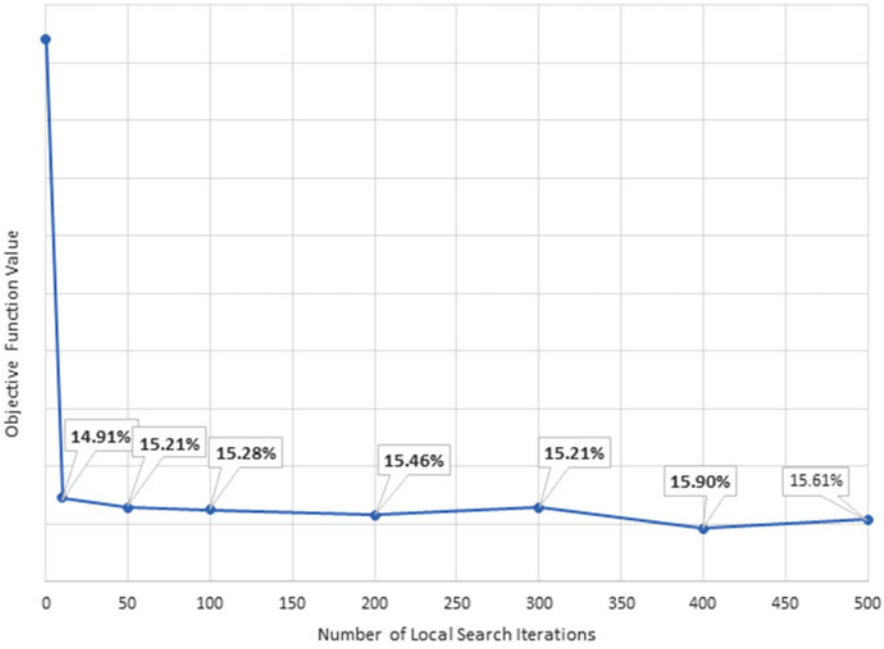


Fig. 2 Number of maximum local search iterations exploration

solution improvement for a specific number of local search iterations over the initial obtained objective function value without incorporating the local search procedures. As an instance, the enhancement ability of the local search can be seen at 400 iterations, that improved the solution by 15.90%.

From that, we conclude that as we increase the local search iterations the solution quality improves up to a certain threshold; after the 500 iterations the solution is not improving and even more, it deteriorates. The later phenomenon can be explained by the fact, that high exploitation from the first algorithmic iterations, results in local optimum, which is difficult for the FAC to overcome. Consequently, the maximum number of local search iterations was set to 400. In the following, the conducted experimental results on the benchmark instances categories: A, B, E, M, are presented in Tables 5, 6, 7, and 8, respectively. For each of the 115 considered instances, the minimum (w_{min}) and average (w_{avg}) achieved values of the objective function, over five algorithmic executions can be seen in the aforementioned tables.

Table 5 Computational results of group: A

	DRDE		FAC		
Instance	w_{min}	w_{avg}	w_{min}	w_{avg}	Deviation (%)
A-n32-k5-1	4.443E+03	4.459E+03	4.402E+03	4.436E+03	−0.94
A-n32-k5-2	4.521E+03	4.533E+03	4.497E+03	4.503E+03	−0.55
A-n32-k5-3	4.431E+03	4.437E+03	4.411E+03	4.421E+03	−0.44
A-n32-k5-4	4.441E+03	4.470E+03	4.428E+03	4.440E+03	−0.30
A-n32-k5-5	4.772E+03	4.813E+03	4.735E+03	4.764E+03	−0.79
A-n37-k6-1	5.604E+03	5.619E+03	5.586E+03	5.600E+03	−0.33
A-n37-k6-2	5.618E+03	5.625E+03	5.595E+03	5.604E+03	−0.40
A-n37-k6-3	5.554E+03	5.565E+03	5.481E+03	5.505E+03	−1.34
A-n37-k6-4	5.615E+03	5.616E+03	5.554E+03	5.579E+03	−1.10
A-n37-k6-5	5.334E+03	5.363E+03	5.291E+03	5.305E+03	−0.82
A-n44-k6-1	5.211E+03	5.253E+03	5.101E+03	5.161E+03	−2.15
A-n44-k6-2	5.072E+03	5.082E+03	4.965E+03	4.972E+03	−2.16
A-n44-k6-3	5.475E+03	5.480E+03	5.370E+03	5.396E+03	−1.95
A-n44-k6-4	5.024E+03	5.041E+03	4.911E+03	4.954E+03	−2.30
A-n44-k6-5	5.249E+03	5.266E+03	5.168E+03	5.183E+03	−1.57
A-n48-k7-1	6.194E+03	6.238E+03	6.109E+03	6.117E+03	−1.40
A-n48-k7-2	6.366E+03	6.461E+03	6.388E+03	6.411E+03	0.33
A-n48-k7-3	6.266E+03	6.341E+03	6.224E+03	6.264E+03	−0.68
A-n48-k7-4	6.377E+03	6.387E+03	6.209E+03	6.244E+03	−2.70
A-n48-k7-5	6.210E+03	6.246E+03	6.154E+03	6.174E+03	−0.92
A-n53-k7-1	6.247E+03	6.275E+03	6.057E+03	6.120E+03	−3.14
A-n53-k7-2	5.935E+03	5.980E+03	5.832E+03	5.857E+03	−1.76
A-n53-k7-3	5.966E+03	6.026E+03	5.904E+03	5.936E+03	−1.06
A-n53-k7-4	5.744E+03	5.787E+03	5.758E+03	5.782E+03	0.23
A-n53-k7-5	6.126E+03	6.168E+03	6.089E+03	6.120E+03	−0.61
A-n60-k9-1	8.025E+03	8.111E+03	8.032E+03	8.055E+03	0.10
A-n60-k9-2	8.133E+03	8.196E+03	8.062E+03	8.128E+03	−0.89
A-n60-k9-3	8.387E+03	8.440E+03	8.355E+03	8.383E+03	−0.38
A-n60-k9-4	8.104E+03	8.126E+03	7.982E+03	8.008E+03	−1.53
A-n60-k9-5	8.368E+03	8.408E+03	8.252E+03	8.305E+03	−1.41
A-n65-k9-1	7.680E+03	7.690E+03	7.671E+03	7.750E+03	−0.12
A-n65-k9-2	7.767E+03	7.832E+03	7.712E+03	7.813E+03	−0.71
A-n65-k9-3	7.999E+03	8.022E+03	7.749E+03	7.877E+03	−3.23
A-n65-k9-4	7.848E+03	7.891E+03	7.710E+03	7.784E+03	−1.80
A-n65-k9-5	7.586E+03	7.673E+03	7.524E+03	7.559E+03	−0.82
A-n69-k9-1	7.740E+03	7.776E+03	7.600E+03	7.662E+03	−1.84
A-n69-k9-2	7.625E+03	7.758E+03	7.556E+03	7.627E+03	−0.92
A-n69-k9-3	7.525E+03	7.571E+03	7.422E+03	7.470E+03	−1.39
A-n69-k9-4	7.601E+03	7.658E+03	7.385E+03	7.432E+03	−2.94
A-n69-k9-5	7.618E+03	7.672E+03	7.532E+03	7.644E+03	−1.15
A-n80-k10-1	9.193E+03	9.265E+03	8.989E+03	9.027E+03	−2.27

(continued)

Table 5 (continued)

	DRDE		FAC		
Instance	w_{min}	w_{avg}	w_{min}	w_{avg}	Deviation (%)
A-n80-k10-2	8.842E+03	8.955E+03	8.753E+03	8.783E+03	−1.02
A-n80-k10-3	9.048E+03	9.154E+03	8.815E+03	8.898E+03	−2.65
A-n80-k10-4	9.031E+03	9.098E+03	8.931E+03	8.965E+03	−1.12
A-n80-k10-5	8.668E+03	8.773E+03	8.533E+03	8.572E+03	−1.58

Table 6 Computational results of group: B

	DRDE		FAC		
Instance	w_{min}	w_{avg}	w_{min}	w_{avg}	Deviation (%)
B-n39-k5-1	4.095E+03	4.108E+03	4.061E+03	4.073E+03	−0.83
B-n39-k5-2	4.205E+03	4.215E+03	4.188E+03	4.198E+03	−0.41
B-n39-k5-3	4.012E+03	4.037E+03	4.015E+03	4.024E+03	0.06
B-n39-k5-4	4.137E+03	4.157E+03	4.166E+03	4.180E+03	0.70
B-n39-k5-5	4.070E+03	4.088E+03	4.039E+03	4.053E+03	−0.75
B-n41-k6-1	5.279E+03	5.308E+03	5.226E+03	5.239E+03	−1.01
B-n41-k6-2	5.177E+03	5.196E+03	5.147E+03	5.164E+03	−0.59
B-n41-k6-3	5.089E+03	1.710E+07	5.053E+03	5.078E+03	−0.71
B-n41-k6-4	4.990E+03	5.000E+03	4.947E+03	4.957E+03	−0.87
B-n41-k6-5	5.179E+03	5.197E+03	5.072E+03	5.105E+03	−2.11
B-n50-k7-1	6.210E+03	6.217E+03	6.090E+03	6.125E+03	−1.97
B-n50-k7-2	5.716E+03	5.733E+03	5.650E+03	5.666E+03	−1.17
B-n50-k7-3	5.805E+03	5.877E+03	5.818E+03	5.827E+03	0.23
B-n50-k7-4	5.927E+03	5.938E+03	5.862E+03	5.884E+03	−1.10
B-n50-k7-5	5.896E+03	5.912E+03	5.905E+03	5.915E+03	0.15
B-n56-k7-1	5.390E+03	5.432E+03	5.358E+03	5.399E+03	−0.60
B-n56-k7-2	5.335E+03	5.416E+03	5.371E+03	5.405E+03	0.67
B-n56-k7-3	5.389E+03	5.442E+03	5.336E+03	5.356E+03	−1.00
B-n56-k7-4	5.565E+03	5.627E+03	5.625E+03	5.631E+03	1.07
B-n56-k7-5	5.619E+03	5.636E+03	5.525E+03	5.583E+03	−1.71
B-n63-k10-1	9.136E+03	9.172E+03	8.991E+03	9.058E+03	−1.62
B-n63-k10-2	8.858E+03	8.986E+03	8.949E+03	8.980E+03	1.02
B-n63-k10-3	9.444E+03	9.522E+03	9.231E+03	9.327E+03	−2.31
B-n63-k10-4	9.444E+03	9.522E+03	9.088E+03	9.135E+03	−3.92
B-n63-k10-5	9.440E+03	9.455E+03	9.393E+03	9.441E+03	−0.51
B-n78-k10-1	8.049E+03	8.089E+03	7.892E+03	7.929E+03	−1.98
B-n78-k10-2	8.112E+03	8.227E+03	7.954E+03	8.035E+03	−1.99
B-n78-k10-3	8.458E+03	8.498E+03	8.366E+03	8.402E+03	−1.10
B-n78-k10-4	8.151E+03	8.236E+03	8.021E+03	8.072E+03	−1.62
B-n78-k10-5	8.304E+03	8.542E+03	8.355E+03	8.361E+03	0.60

Table 7 Computational results of group: E

Instance	DRDE		FAC		Deviation (%)
	w_{min}	w_{avg}	w_{min}	w_{avg}	
E-n33-k4-1	3.834E+03	3.853E+03	3.822E+03	3.851E+03	−0.31
E-n33-k4-2	3.576E+03	3.626E+03	3.616E+03	3.640E+03	1.11
E-n33-k4-3	3.474E+03	3.478E+03	3.445E+03	3.461E+03	−0.84
E-n33-k4-4	3.437E+03	3.505E+03	3.487E+03	3.511E+03	1.42
E-n33-k4-5	3.451E+03	3.476E+03	3.435E+03	3.447E+03	−0.47
E-n51-k5-1	3.455E+03	3.489E+03	3.438E+03	3.443E+03	−0.47
E-n51-k5-2	3.592E+03	3.642E+03	3.553E+03	3.563E+03	−1.12
E-n51-k5-3	3.503E+03	3.528E+03	3.463E+03	3.491E+03	−1.16
E-n51-k5-4	3.493E+03	3.550E+03	3.491E+03	3.500E+03	−0.08
E-n51-k5-5	3.594E+03	3.620E+03	3.586E+03	3.622E+03	−0.22
E-n76-k10-1	8.180E+03	3.264E+03	8.205E+03	8.269E+03	0.30
E-n76-k10-2	8.023E+03	8.054E+03	7.971E+03	7.979E+03	−0.65
E-n76-k10-3	8.006E+03	8.042E+03	7.829E+03	7.846E+03	−2.27
E-n76-k10-4	8.024E+03	8.081E+03	7.946E+03	7.992E+03	−0.97
E-n76-k10-5	8.125E+03	8.158E+03	8.048E+03	8.087E+03	−0.96
E-n101-k14-1	1.142E+04	1.145E+04	1.105E+04	1.117E+04	−3.38
E-n101-k14-2	1.182E+04	1.183E+04	1.152E+04	1.154E+04	−2.59
E-n101-k14-3	1.135E+04	1.143E+04	1.122E+04	1.127E+04	−1.15
E-n101-k14-4	1.140E+04	1.141E-04	1.104E+04	1.105E+04	−3.24
E-n101-k14-5	1.180E+04	1.183E+04	1.146E+334	1.154E+04	−3.04

The presented values correspond to the solution of PCVRP-P, via both the FAC and DRDE. The performance of the two algorithms can be found in the last column of every table as the deviation among their minimum obtained values. As an example, considering the instance $A - n32 - k5 - 1$, the FAC achieved a better value than DRDE by 0.94% (−0.94% in terms of minimization), while for the $A - n48 - k7 - 2$, DRDE surpassed FAC by 0.33%. The FAC obtained better solutions than DRDE, on average, by 1.22%, 0.83%, 1.00% and 3.19%, respectively, in each group. In total, only in 16 out of 115 instances the DRDE achieved a lower objective function value, and with respect to the average obtained value, only in 8.

Additionally, to prove the quality of the FAC results and its effectiveness over the DRDE, we solved the PCVRP-P using CPLEX 12.8. Table 9 presents the computational results for a subset of the original instances, i.e. an indicative set. In more detail, for each instance in the indicative set, the minimum objective function value (w_{min}) and the average solution time (t_{avg}), over five algorithmic executions, for both algorithms, are reported in the table above. Moreover, Table 9 contains the corresponding, to each instance, best solution value obtained by CPLEX, while the available solution time was limited to 1800 s. The solution time upper-bound was included in terms of fair comparison, considering real-life application in the future,

Table 8 Computational results of group: M

Instance	DRDE		FAC		Deviation (%)
	w_{min}	w_{avg}	w_{min}	w_{avg}	
M-n101-k10-1	7.217E+03	7.320E+03	7.197E+03	7.245E+03	−0.28
M-n101-k10-2	7.460E+03	7.507E+03	7.357E+03	7.377E+03	−1.40
M-n101-k10-3	7.108E+03	7.258E+03	7.156E+03	7.191E+03	0.67
M-n101-k10-4	7.187E+03	7.265E+03	7.152E+03	7.181E+03	−0.49
M-n101-k10-5	7.145E+03	7.224E+03	7.137E+03	7.160E+03	−0.12
M-n121-k7-1	3.337E+03	3.418E+03	3.146E+03	3.249E+03	−6.06
M-n121-k7-2	3.683E+03	3.904E+03	3.599E+03	3.612E+03	−2.31
M-n121-k7-3	3.497E+03	3.774E+03	3.427E+03	3.537E+03	−2.04
M-n121-k7-4	3.396E+03	3.507E+03	3.228E+03	3.288E+03	−5.18
M-n121-k7-5	3.373E+03	3.380E+03	3.215E+03	3.290E+03	−4.04
M-n151-k12-1	7.988E+03	8.013E+03	7.630E+03	7.650E+03	−4.69
M-n151-k12-2	7.573E+03	7.700E+03	7.182E+03	7.228E+03	−5.44
M-n151-k12-3	7.759E+03	7.807E+03	7.466E+03	7.481E+03	−3.93
M-n151-k12-4	7.054E+03	7.307E+03	7.099E+03	7.149E+03	0.63
M-n151-k12-5	7.205E+03	7.326E+03	6.876E+03	7.025E+03	−4.78
M-n200-k17-1	1.152E+04	1.162E+04	1.110E+04	1.120E+04	−3.76
M-n200-k17-2	1.164E+04	1.181E+04	1.126E+04	1.152E+04	−3.43
M-n200-k17-3	1.115E+04	1.145E+04	1.082E+04	1.101E+04	−3.03
M-n200-k17-4	1.147E+04	1.149E+04	1.071E+04	1.078E+04	−7.05
M-n200-k17-5	1.150E+04	1.153E+04	1.084E+04	1.115E+04	−6.11

since on average the DRDE execution time (in seconds) is in the range [19.311,68.967] and the FAC is in the [25.794,70.254].

Figure 3 depicts the percentage deviation of the w_{min} from the corresponding best solution obtained by CPLEX, for the indicative set. In the aforementioned figure, the negative values demonstrate the percentage deviation of the inferior algorithmic solution, while the positive values, correspond to superior solutions that DRDE or FAC achieved over the CPLEX solver. For example, regarding the instance $A-n32-k5-1$, DRDE obtained a worst solution, with 1.67% gap from the CPLEX best, and FAC with 0.73% gap (negative values in the figure), while regarding the instance $E-n76-k10-1$, both algorithms obtained a better solution with 0.71% and 0.41% gap (positive values in the figure), respectively. These values can be explained as within the 1800 s, CPLEX achieved a solution of better quality for most of the instances with small number of customers and routes, while with respect to large instances, e.g. $M-n151-k12-1$, the solver could not achieve a good quality solution within the time limit. Nevertheless, the computational results presented above, demonstrate the effectiveness of both algorithms, and as showed, the FAC provides solutions of better quality than the DRDE. Even though the latter algorithm exhibits smaller execution times, their execution time difference could not counterbalance the FAC's solution quality.

Table 9 The comparison of DRDE, FAC and CPLEX 12.8, based on an indicative set of instances

	DRDE		FAC		
Instance	w_{min}	t_{avg} (s)	w_{min}	t_{avg} (s)	CPLEX 12.8 (Time limit: 1800 s)
A-n32-k5-1	4.443E+03	25.506	4.402E+03	27.856	4.370E+03
A-n37-k6-1	5.604E+03	19.463	5.502E+03	26.894	5.467E+03
A-n44-k6-1	5.211E+03	22.007	5.101E+03	37.529	5.084E+03
A-n48-k7-1	6.194E+03	22.490	6.109E+03	37.555	6.085E+03
A-n53-k7-1	6.247E+03	24.638	6.057E+03	40.872	6.012E+03
A-n60-k9-1	8.025E+03	27.244	8.032E+03	48.995	8.016E+03
A-n65-k9-1	7.680E+03	57.934	7.671E+03	53.312	7.657E+03
A-n69-k9-1	7.740E+03	63.497	7.600E+03	54.159	7.637E+03
A-n80-k10-1	9.193E+03	68.967	8.989E+03	50.532	8.955E+03
B-n39-k5-1	4.095E+03	37.641	4.061E+03	28.025	4.028E+03
B-n41-k6-1	5.279E+03	21.186	5.226E+03	28.415	5.223E+03
B-n50-k7-1	6.210E+03	22.685	6.090E+03	31.183	6.110E+03
B-n56-k7-1	5.390E+03	30.923	5.358E+03	34.574	5.375E+03
B-n63-k10-1	9.136E+03	31.765	8.991E+03	39.708	9.029E+03
B-n78-k10-1	8.049E+03	37.515	7.892E+03	49.819	8.147E+03
E-n33-k4-1	3.834E+03	19.311	3.822E+03	25.794	3.814E+03
E-n51-k5-1	3.455E+03	23.741	3.438E+03	54.470	3.419E+03
E-n76-k10-1	8.180E+03	35.135	8.205E+03	55.354	8.239E+03
E-n101-k14-1	1.142E+04	55.354	1.105E+04	60.123	1.134E+04
M-n151-k12-1	7.988E+03	68.525	7.630E+03	70.254	1.102E+04

7 Conclusions

In this paper we consider a variant of the VRP, namely PCVRP-P in the literature, that could be applied in tourist trip design problems, and we propose the Firefly Algorithm based on Coordinates (FAC) to solve it. In the PCVRP-P a pre-specified number of routes have to be constructed from a set of nodes, bounded by capacity restrictions. Each node is characterized by a prize value and the objective of the problem is to visit nodes to maximize the aggregated collected prize, while minimizing the cost of travelling between the nodes. The FAC is a hybrid Firefly Algorithm, which incorporates a new technique to encode the solutions, according to the coordinates of the included nodes. Through this representation, the movement equation of the FA can be applied without transformation, as originally that implies continuous valued encoding of solutions. A comparison of solving the PCVRP-P, is made between the proposed FAC and the DRDE, based on suitable benchmark instances found in the literature, while both algorithms were compared to the best solution obtained by the CPLEX solver over an indicative set of instances, to prove

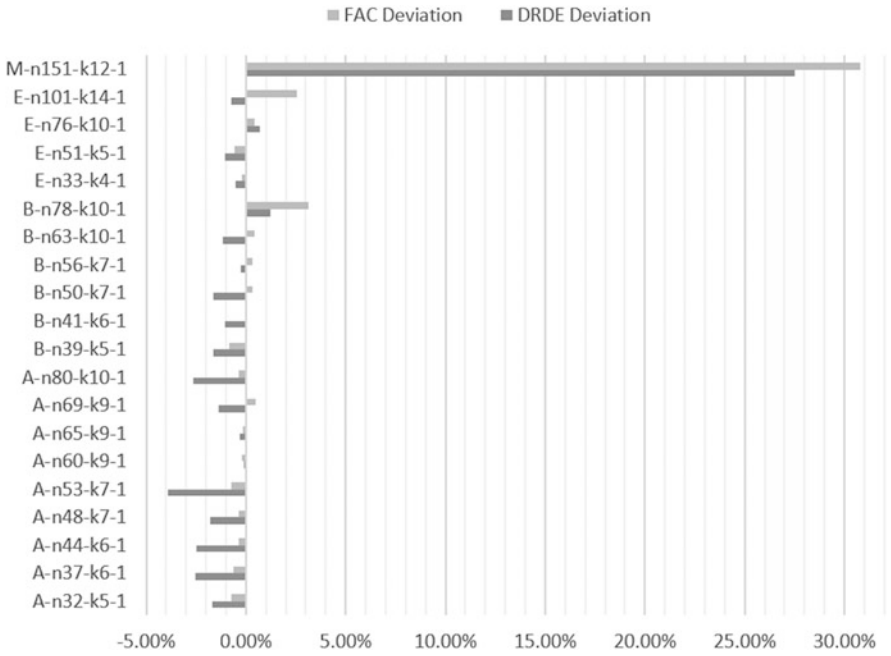


Fig. 3 Deviation of DRDE and FAC w_{min} from the obtained CPLEX value

the quality of their solutions. Based on the conducted experiments, the FAC is more effective in solving the PCVRP-P, compared to DRDE.

To further explore the capabilities of the FAC, as future research, we propose the exploration of the PCVRP-NP, where the number of constructed routes is not a pre-specified value. In addition, in order to incorporate the PCVRP into a tourist recommendation system, the weighted sum method could be introduced, by taking into account a user's preferences.

Acknowledgments This research is co-financed by Greece and the European Union (European Social Fund—ESF) through the Operational Programme "Human Resources Development, Education and Lifelong Learning" in the context of the project "Strengthening Human Resources Research Potential via Doctorate Research" (MIS-5000432), implemented by the State Scholarships Foundation (IKY).

References

Alinaghian, M., & Naderipour, M. (2016). A novel comprehensive macroscopic model for time-dependent vehicle routing problem with multi-alternative graph to reduce fuel consumption: A case study. *Computers & Industrial Engineering, 99*, 210–222.

Baghlani, A., Makiabadi, M. H., & Sarcheshmehpour, M. (2014). Discrete optimum design of truss structures by an improved firefly algorithm. *Advances in Structural Engineering, 17*(10), 1517–1530.

Chandrasekaran, K., Simon, S. P., & Padhy, N. P. (2013). Binary real coded firefly algorithm for solving unit commitment problem. *Information Sciences, 249*, 67–84.

Clarke, G., & Wright, J. W. (1964). Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research, 12*(4), 568–581.

Fister, I., Yang, X. S., & Brest, J. (2013). A comprehensive review of firefly algorithms. *Swarm and Evolutionary Computation, 13*, 34–46.

Gavalas, D., Konstantopoulos, C., Mastakas, K., & Pantziou, G. (2014). A survey on algorithmic approaches for solving tourist trip design problems. *Journal of Heuristics, 20*(3), 291–328.

Goel, R., & Maini, R. (2018). A hybrid of ant colony and firefly algorithms (HAFA) for solving vehicle routing problems. *Journal of Computational Science, 25*, 28–37.

Jati, G. K. (2011). Evolutionary discrete firefly algorithm for travelling salesman problem. In *Adaptive and intelligent systems* (pp. 393–403). Berlin: Springer.

Jia, S. J., Yi, J., Yang, G. K., Du, B., & Zhu, J. (2013). A multi-objective optimisation algorithm for the hot rolling batch scheduling problem. *International Journal of Production Research, 51*(3), 667–681.

Li, K., & Tian, H. (2016). A two-level self-adaptive variable neighborhood search algorithm for the prize-collecting vehicle routing problem. *Applied Soft Computing, 43*, 469–479.

Long, J., Sun, Z., Pardalos, P. M., Hong, Y., Zhang, S., & Li, C. (2019). A hybrid multi-objective genetic local search algorithm for the prize-collecting vehicle routing problem. *Information Sciences, 478*, 40–61.

Osaba, E., Yang, X. S., Diaz, F., Onieva, E., Masegosa, A. D., & Perallos, A. (2017). A discrete firefly algorithm to solve a rich vehicle routing problem modelling a newspaper distribution system with recycling policy. *Soft Computing, 21*(18), 5295–5308.

Pan, F., Ye, C., Wang, K., & Cao, J. (2013). Research on the vehicle routing problem with time windows using firefly algorithm. *JCP, 8*(9), 2256–2261.

Sayadi, M., Ramezanian, R., & Ghaffari-Nasab, N. (2010). A discrete firefly meta-heuristic with local search for makespan minimization in permutation flow shop scheduling problems. *International Journal of Industrial Engineering Computations, 1*(1), 1–10.

Simić, D., Kovačević, I., Svirčević, V., & Simić, S. (2015). Hybrid firefly model in routing heterogeneous fleet of vehicles in logistics distribution. *Logic Journal of the IGPL, 23*(3), 521–532.

Singh, A., Thapar, S., Bhatia, A., Singh, S., & Goyal, R. (2015). Disk scheduling using a customized discrete firefly algorithm. *Cogent Engineering, 2*(1), 1011929.

Stavropoulou, F., Repoussis, P. P., & Tarantilis, C. D. (2019). The vehicle routing problem with profits and consistency constraints. *European Journal of Operational Research, 274*(1), 340–356.

Tang, L., & Wang, X. (2006). Iterated local search algorithm based on very large-scale neighborhood for prize-collecting vehicle routing problem. *The International Journal of Advanced Manufacturing Technology, 29*(11–12), 1246–1258.

Tilahun, S. L., & Ngnotchouye, J. M. T. (2017). Firefly algorithm for discrete optimization problems: A survey. *KSCE Journal of Civil Engineering, 21*(2), 535–545.

Tiwari, A., Chang, P. C., Elangovan, G., & Annadurai, S. P. (2015, May). A hybrid edge recombination approach to solve price collecting vehicle routing problem. In *2015 International Conference on Control, Automation and Robotics* (pp. 200–203). IEEE.

Trachanatzi, D., Rigakis, M., Taxidou, A., Marinaki, M., Marinakis, Y., & Matsatsinis, N. (2019). A novel solution encoding in the differential evolution algorithm for optimizing tourist trip design problems. In *International conference on learning and intelligent optimization* (pp. 253–267). Cham: Springer.

Yang, X. S. (2009, October). Firefly algorithms for multimodal optimization. In *International symposium on stochastic algorithms* (pp. 169–178). Berlin: Springer.

Yang, X. S. (2010). *Nature-inspired metaheuristic algorithms*. Bristol: Luniver Press.
Yang, X. S. (2014). Cuckoo search and firefly algorithm: Overview and analysis. In X. S. Yang (Ed.), *Cuckoo search and firefly algorithm* (Studies in Computational Intelligence) (Vol. 516). Cham: Springer.
Yang, X. S., & He, X. (2013). Firefly algorithm: Recent advances and applications. *International Journal of Swarm Intelligence, 1*(1), 36–50.
Zhang, T., Chaovalitwongse, W. A., Zhang, Y. J., & Pardalos, P. M. (2009). The hot-rolling batch scheduling method based on the prize collecting vehicle routing problem. *Journal of Industrial and Management Optimization, 5*(4), 749–765.

Green Supplier Evaluation and Selection: An Updated Literature Review

Syrine Jemaa, Ahmed Alayidi, Athanasios Migdalas, George Baourakis, and Periklis Drakos

Abstract Green supply chain management is concerned with the integration of environmental criteria and sustainability issues in the management of the supply chain. Within this framework, firms applying internally green strategy have, naturally, the interest to pressure purchases of goods and services from suppliers that are themselves green, at least to a certain extent. Thus, Supplier Evaluation and Selection is crucial due to its big impact on business function. The prospect of applying green principles has consequently become an important feature of a supplier's overall performance. Hence, green supplier selection and evaluation, although a relatively new research subject, has grown quite rapidly. It develops and studies the decision and evaluation models based on environmental criteria. Its main tools are based on multi-criteria decision making approaches. Our objective, in this paper, is to review journal articles published in the period 2012–2019 on this topic, in order to

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-3-030-38766-2_9) contains supplementary material, which is available to authorized users.

S. Jemaa (✉)
Department of Business, Economics and Management, Mediterranean Agronomic Institute of Chania (CIHEAM-MAICh), Chania, Greece

A. Alayidi
Department of Sustainable Agriculture, Mediterranean Agronomic Institute of Chania (CIHEAM-MAICh), Chania, Greece

A. Migdalas
Department of Business Administration, Technology and Social Sciences (ETS), Luleå University of Technology, Luleå, Sweden
e-mail: athanasios.migdalas@ltu.se

G. Baourakis
Mediterranean Agronomic Institute of Chania (CIHEAM-MAICh), Chania, Greece
e-mail: baouraki@maich.gr

P. Drakos
Department of Economics, School of Social Sciences, University of Crete, Rethymno, Greece
e-mail: drakosp@uoc.gr

E. Krassadaki et al. (eds.), *Operational Research in Agriculture and Tourism*, Cooperative Management, https://doi.org/10.1007/978-3-030-38766-2_9

identify the most widely applied approaches for green supplier evaluation and selection and the most cited green criteria.

Our research methodology in inducing the present review has been based on several criteria for the selection of the appropriate articles relevant to the subject. The analysis of those papers has led to an extended classification with respect to the multi-criteria decision making approaches, as well as with respect to the adopted green criteria.

We conclude that the majority of applied approaches are based on integrated fuzzy models and that the most popular criteria are "Environmental management system" and "Green design".

Keywords Green supplier evaluation and selection · Multi-criteria decision-making approach · Green criteria · Green supply chain management

List of Abbreviations of Identified Approaches

AD	Axiomatic design
AHP	Analytic hierarchy process
ANP	Analytic network process
AQM	Alternative queuing method
BOM	Bi-objective model
BWM	Best-worst method
CIO	Choquet integral operator
CM	C-means
COPRAS	COmplex PRoportional ASsessment
COPRAS-G	modified COmplex Proportional ASsessment of alternatives with Grey relations
DANP	DEMATEL-based analytical network process
DEA	Data envelopment analysis
DEMATEL	DEcision-MAking Trial and Evaluation Laboratory
ELECTRE	ELimination and Choice Expressing the REality
ERM	Enhanced Russell measure
FA	Factor analysis
GCM	Grey cognitive map
GDEA	Green DEA
GP	Genetic programming
GRA	Grey relational analysis
GREY	Grey system theory
IGRA	Interval grey analysis approach
IS	Inference system
KAM	Kourosh and Arash method
MOLP	Multi-objective linear programming
MOMM	Multi-objective mathematical model
MOOM	Multi-objective optimization model

MOP	Multi-objective programming
PP	Preference programming
QFD	Quality function deployment
SD	System dynamics
SE	Shannon entropy
SWARA	Step-wise weight assessment ratio analysis
TLF	Taguchi loss functions
TODIM	An acronym in Portuguese of interactive and multi-criteria decision making
TOPSIS	Technique for order preference by similarity to ideal solution
VAHP	Voting analytic hierarchy process
VIKOR	An acronym in Serbian: "VlseKriterijumska Optimizacija I Kompromisno Resenje" of multicriteria optimization and compromise solution
WASPAS	Weighted aggregated sum product assessment

1 Introduction

The rise of ecological and sustainability issues as a matter of public concern is increasingly influencing buyers' procurement behavior, an influence which expands to a firm's strategy and competitiveness. Although an early study demonstrated the effect of ecological concern on buyers' brand perception (Kinnear, 1973), research effort dedicated to the conceptualization of incorporating sustainability in firms' strategic management prior to the 1990s was lacking (Fowler, 2007). Business strategy literature was based on two main conceptual frameworks: the industrial-organizational perspective which is based on the idea that industry-sector-borne factors are detrimental to the profitability of companies which are active in it (Fowler, 2007), and the resource-based view which considers a company's resources as determinant factors, i.e., defining resources of strategic priority such as brand names, employee recruitment strategy, commercial contacts, efficient managerial procedures, etc., and utilizing them with efficiency to provide competitive advantage for the company (Amit, 1993; Barney, 1986). Both views were in accordance with the neoclassical theory doctrine, which defines the ceaseless increase of shareholders' profits as the fundamental role of the management body (Friedman, 1984), with environmental considerations being important only in the context of legal compliance and the provision of shareholder value (Walley, 1994).

Hart (1995) presented the Natural Resource-Based View (NRBV) as a conceptual framework for incorporating sustainability in strategic management. NRBV is based on three interconnected principles of pollution prevention, product stewardship, and sustainable development. The author's main argument was that firms will increasingly be ecosystem-dependent. Therefore, the efficacy of firms' development of resources and means to handle the ceaselessly changing environment is a determinant factor when it comes to competitive advantage (Hart, 1995).

On the other hand, recent years have witnessed an increase in outsourcing activities conducted by companies (Wu, 2013), which consequently increased the importance of procurement management. Thus, for incorporating sustainability in a company's strategy, the evaluation and selection of suppliers based on sustainability-related criteria have become a necessity. Criteria for sustainability include three main aspects which are environmental, social and economic responsibilities (Purvis, 2018), any supplier who meets these three criteria is called a "Green Supplier" (Gurel, 2015). Research on criteria for evaluation and selection have been conducted since the 1960s. Dickson (1966) defined 23 criteria based on importance to practitioners. Although price, delivery, and quality turned out to be the most widely used filters through which a supplier is evaluated and selected (Rezaei, 2016), Dickson's work represents the foundation of most of the criteria in use today (Weber, 1991). However, the "green supplier" concept leaves us with a very important question: In the context of the environment, what are the specific practical criteria and methods by which a company evaluates and selects green suppliers to date?

This work reviews green supplier evaluation and selection literature published in the period 2012–present. The main objective of this review is to identify the methods and criteria for supplier selection and evaluation documented in the aforementioned period in an effort to facilitate further research.

2 Methodology

2.1 Search Strategy and Criteria

Literature reviews constitute documented investigative studies which shed light on research effort in the current fields of interest which assists in orienting further research efforts and tendencies (Cronin, 2008).

This document is based on the following research methodology:

1. General literature research was conducted.
2. A criteria-based classification framework was developed.
3. The collected literature was filtered through the framework and organized in a table.
4. The literature review was presented based on the classification framework.
5. Analytical discussion of the literature review was done in addition to suggesting recommendations for future research work.

The literature review is a commonly used procedure to conduct investigative studies of the diverse approaches of the subject under study (Lage Junior, 2010). The literature was collected from different online databases using "Google Scholar" as the main search engine. The main keywords used in the search engine are "Supplier selection and evaluation: environment". Articles in which the keywords were synonyms of the ones we targeted, for example, "green" and "sustainable" instead of

"environment", similarly, "vendor", "contractor" and "partner" instead of "supplier" were also accepted. The literature selection was based on the following criteria:

1. Papers which were published between 2012 and 2019.
2. Papers which are journal articles.
3. Papers which are not literature reviews themselves.
4. Papers which are relevant to the topic of this review paper.
5. Papers which include multi-criteria decision making approaches.
6. Papers which include operating with green and/or sustainable and/or environmental criteria.
7. Papers which include real case studies.

55 papers met the criteria of selection including 35 papers from Elsevier's Science direct (63.64%), 7 papers from Springer (12.73%), 4 papers from Taylor & Francis (7.27%), 4 papers from MDPI (7.27%), 1 paper from Growing Science (1.82%) in addition to 4 papers from different sources which are included in details in the reference list (7.27%). Figure 1 depicts schematically the literature gathering and selection method adopted in this review paper.

2.2 The Classification Framework

As this research is mainly focused on the green supplier selection and evaluation process in which environmental criteria are included and of great importance, the conventional criteria which are termed "Traditional criteria", being mainly centered on economic considerations, are neglected in the classification we developed. Concerning the decision-making approach, all of the selected articles adopted a multi-criteria decision-making approach. However, the articles differed when it comes to the implementation of the fuzzy set theory. Non-fuzzy sets are composed of elements or objects that can be finite, countable or over-countable. The component is either part of or not part of a certain set. Thus, they are characterized by being dichotomous, i.e., true-or-false type answers which, implicitly, assumes that the developed model based on it contains no uncertainties which are, in some cases, not in accordance with reality. On the other hand, the fuzzy set theory is an exclusively mathematical framework which facilitates the study of elements presenting specific types of uncertainty problems resulting from the lack of well-defined criteria of class membership, i.e., it is more based on the degree of truth of the element rather than the binary true/false perception (Zimmermann, 2001).

The criteria used in all of the articles were isolated and examined for categorization purposes. We chose to classify them according to their area of interest and we noticed that there were three main areas of interest, being green governance and capabilities, green manufacturing and product quality, and innovation and direct ecological impact. Regarding green governance and capabilities, the authors focused on the supplier's policies, management practices, and green capabilities. For the green manufacturing and product quality, we divided the criteria based on their relation to production phases which lead to pre-production, production process, post-production, and product

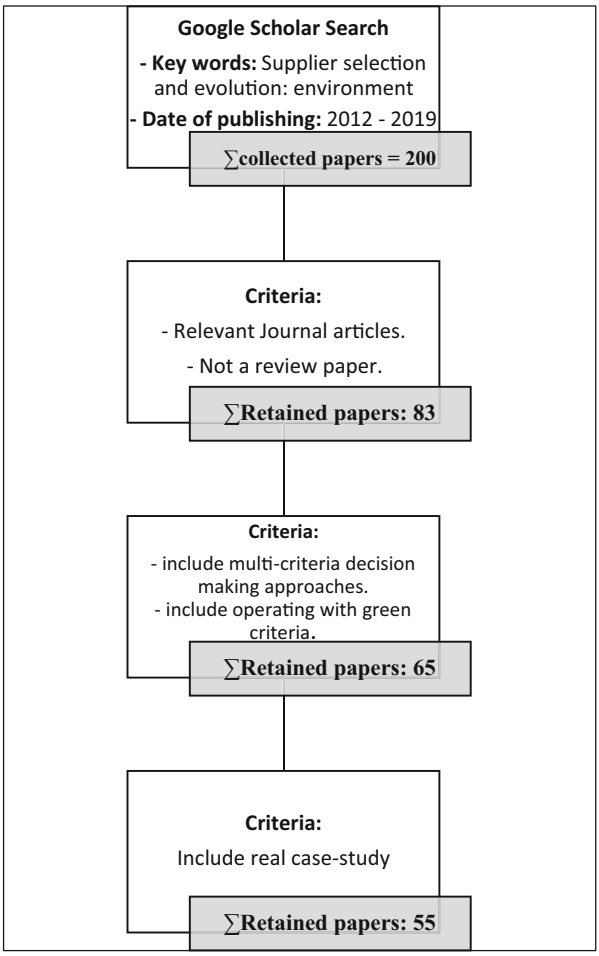


Fig. 1 Literature collection and selection method flow

sub-classes. As for the innovation and ecological impact class, we classified it based on a criteria focus on the measurement of the ecological impact of the supplier's operations or the supplier's Research & Development (R&D) and innovation potential. Figure 2 provides a general schematic description of the classification framework we developed.

3 Results

The examination of the 55 selected articles permit the extraction of the required information like the applied approach, the used criteria, the objective of the study and the industry. These information are summarized in Appendix 3 (Supplementary Material). Also, a list of abbreviations of the identified approaches is presented in "List of Abbreviations".

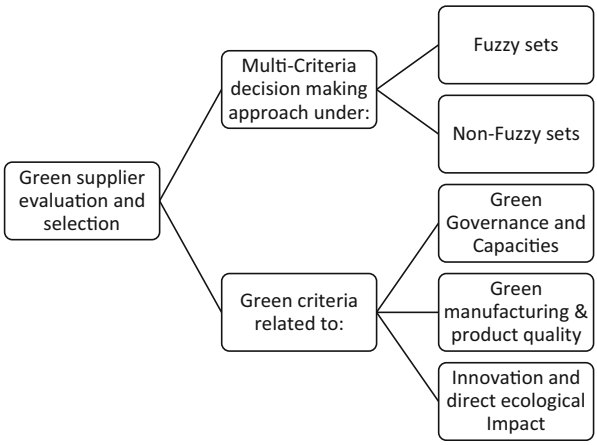


Fig. 2 General description of the classification framework

The analysis of the collected information led to the classification of Multi-Criteria Decision Making (MCDM) approaches and criteria of evaluation related to green supplier selection and evaluation, as presented in Fig. 2.

A summary of our findings is presented according to the MCDM approaches and the specification of the main used class of evaluation criteria for each article.

3.1 Multi-criteria Decision Making Approach Under Fuzzy Sets

3.1.1 Single Approach

Qin (2017) used TODIM under Interval type-2 fuzzy sets in an automobile industry, to solve the problem of selecting the best environmental supplier based mainly on criteria related to Green Governance and Capacities.

Tuzkaya (2013) took a study case from the electronics company and applied the CIO under intuitionistic fuzzy to solve green supplier selection and evaluation problem; most of the green criteria utilized were from the class of green governance and capacities.

Rostamzadeh (2015) utilized the FVIKOR method in the manufacturing industry based on green quantitative multi-criteria that are mainly part of the class of green governance and capacities to select and evaluate the prominent ecological supplier.

Kannan (2014) applied the FTOPSIS method for the management of the green supply chain and the selection of the best eco-friendly supplier in an electronics company based a mix of green criteria related to the class of green governance and capacities and the class of innovation and green manufacturing and product quality.

3.1.2 Hybrid Approach

Shaw (2012) proposed the approaches of FAHP and FMOLP, able to determine the weight of green criteria related to the class of innovation and direct ecological impact in a garment factory, and select the leading supplier to reduce negative environmental issues.

Banaeian (2018) based his study on the food industry, integrating FTOPSIS, FVIKOR and FGRA to evaluate and select the best green supplier on environmental criteria related to green governance and capacities.

Kannan (2013) presented a methodology composed of FAHP, FTOPSIS and MOLP in the automotive industry, using basically environmental criteria in relation with green governance and capacities in order to pick the most ecological supplier.

Fallahpour (2016) proposed an approach combining FKAM, FDEA and FGP to help decision-makers take strategic decisions by ranking and selecting the best green suppliers in the garment company by taking into consideration green criteria from the class of innovation and direct ecological impact.

Azadi (2015) proposed an integrated DEA called the Enhanced Russell Measure (ERM) under fuzzy context in the manufacturing industry, basing his choice of green criteria on the class of green manufacturing and product quality for the selection and evaluation of the best sustainable supplier.

Yazdani (2014) proposed an approach using AHP and FTOPSIS to achieve the selection of the best environmental supplier by determining the weight of green criteria, basically from the class of green governance and capabilities, obtaining then the optimal solution in an automotive company.

Azadnia (2015) illustrated an approach using FAHP and MOLP to select the appropriate environmental supplier for a food company using a set of green criteria mainly from the class of innovation and direct ecological impact.

Kannan (2015) applied FAD and FMOOM in the manufacturing industry. It is a structure to examine and select the best alternative sustainable supplier based on green criteria mainly from the class of green governance and capacities.

Shaw (2013) identified the supreme sustainable supplier by applying AHP and FMOLP in a global model monitoring carbon foot print and green criteria mainly from the class of green governance and capacities and that innovation and direct impact in a garment factory.

Orji (2015) mixed FSD and FTOPSIS to form a model applied for the manufacturing industry to evaluate the total sustainable performance of vendors and to select the most sustainable supplier selection and long-term relationship, focusing principally on green criteria from the class of green governance and capacities.

Orji (2014) integrated FDEMATEL and FTOPSIS in order to reach the selection of sustainable supplier in the manufacturing industry, using for that a mix of green criteria, mainly those from the class of green governance and capacities.

Banaeian (2015) implemented AHP, the Delphi method, FGRA and MOLP in a model for a food company, starting the study with the determination of appropriate green factors which were included mostly in the class of green governance and capacities, in order to make a preferential order of suppliers and pick the most environmental one.

Govindan (2016) proposed the FTOPSIS and FMOLP methods in the manufacturing industry based on a bench of criteria belonging mainly to the class

of green governance and capacities to select the suitable eco-friendly supplier and assign a descending order for potential ones.

Awasthi (2018) presented an approach using FAHP and FVIKOR in an electronics company taking into consideration sustainable threats of sub-suppliers to reach better sustainable supplier selection and working with green criteria mainly from the list of innovation and direct ecological impact.

Gupta (2017) integrated BWM and FTOPSIS to pick the appropriate green supplier in the automotive industry by focusing on environmental attributes related to the class of green governance and capacities.

Akman (2015) presented FCM and VIKOR, for the purpose of assembling green suppliers into groups to guide managers of an automotive company to select the most suitable suppliers based on green criteria which belong to the classes of green governance and capacities and green manufacturing product quality.

Fallahpour (2017) implemented two methods FPP and FTOPSIS to pick and evaluate the best sustainable supplier under identified green criteria and sub-criteria related principally to the class of green governance and capacities adapted to a garment factory.

Zhao (2014) presented an integrated approach by applying FEntropy and FTOPSIS to develop the environmental performance of an electronics company, and evaluate and select the best green supplier based on environmental consciousness as a green attribute.

Büyüközkan (2012) integrated FANP, FDEMATEL and FTOPSIS to deal with green supplier selection and evaluation in the automotive industry based on green factors related principally to green governance and capacities.

Haeri (2019) presented a hybrid approach using BWM, FGCM and IGRA to take the weight of environmental criteria that belong mainly to the class of innovation and direct ecological impact and allocate green suppliers in an automotive firm.

Mohammed (2019) implemented a mix of methods using FAHP, FTOPSIS and FMOOM based on green criteria related mostly to green governance and capacities in a metal company to evaluate the performance of sustainable suppliers, make an allocative plan for them and pick the optimal one.

Liu (2019) presented an approach using AQM and BWM under IVIUL to treat the problem of environmental supplier selection and evaluation in the manufacturing industry based on green attributes related to green governance and capacities.

Dos Santos (2019) worked with FEntropy and FTOPSIS for the purpose of evaluating the ecological performance of suppliers in a manufacturing company and select the most efficient green supplier focusing relatively on green criteria from the class of green manufacturing and direct ecological impact.

Çalık (2019) illustrated FAHP and MOLP in order to solve sustainable supplier selection and evaluation and applied it in a case study of an electronics company based on appropriate environmental criteria primarily related to the class of green governance and capacities.

Abdel-baset (2019) implemented an approach composed of ANP and VIKOR under Neutrosophic numbers in the trade industry for sustainable selection and evaluation of suppliers focusing mainly on green attributes from the class of green governance and capacities.

Van (2018) presented a model with QFD and TOPSIS under an interval neutrosophic set in the manufacturing industry to select and evaluate sustainable suppliers by focusing on environmental criteria from the class of innovation and direct ecological impact.

Babbar (2018) proposed a framework composed of FQFD and FMOMM, applied to a company from the food industry to select and allocate the best green suppliers by the means of quantitative and qualitative green criteria from the class of green governance and capacities.

Khan (2018) illustrated the integration of FSE and FIS and its application on the automotive industry to identify the best sustainable supplier out of a list of potential suppliers based on a mix of green criteria from the three classes.

Gören (2018) used an integrated approach of FDEMATEL, TLF and a BOM for the selection of the prominent sustainable supplier in the trade industry by utilizing an equilibriated mix of green criteria from the three classes.

Abdel-Basset (2018) implemented a hybrid approach of ANP and TOPSIS under an interval neutrosophic set to evaluate and select the most sustainable supplier in the food industry by taking in consideration principally green criteria from the class of green governance and capacities.

Lo (2018) illustrated an approach composed of BWM, modified FTOPSIS and FMOLP methods to manage the purchase process in an electronics company by evaluating and selecting the optimal ecological supplier, focusing on the class of green governance and capacities as green factors.

3.2 Multi-criteria Decision Making Approach Under Non-fuzzy Sets

3.2.1 Single Approach

Hsu (2013) presented the DEMATEL method in the electronics industry to evaluate the capacity of green suppliers and select the leader among them by using green governance and capacities as the main criteria.

Baskaran (2012) presented the method of GRA in a framework of helping managers of a garment company to evaluate and select the best sustainable supplier based on green criteria from the list of innovation and direct ecological impact.

Falatoonitoosi (2013) implemented DEMATEL in an automotive company for the purpose of supporting managers to choose suitable green suppliers and helping suppliers to ameliorate their environmental performance by considering green criteria mostly from the class of green governance and capacities.

Chen (2012) used the ANP method to clear up the problem of green supplier selection and evaluation in a company of electronics given the importance of this process in the business function and by taking into consideration mainly green factors from the class of green governance and capacities.

Shi (2015) applied the DEA method for the evaluation and selection of the prominent green supplier in a context of environmental economy by taking green criteria adapted to the manufacturing industry, mainly related to green governance and the capacities class.

Pishchulov (2019) derived the VAHP method to rank and select objectively sustainable suppliers in the manufacturing industry based on a mixture of green criteria from the three classes to meet the objectives of the company.

Mathiyazhagan (2018) applied the AHP method on the context of the automotive industry to adopt a green procurement process based mainly on green criteria from the class of innovation and direct ecological impact to analyse and select the best green supplier.

Erbiyik (2016) implemented the AHP method in the metal industry to determine the importance of green logistic criteria belonging to the class of green governance and capacities to select the best environmental supplier.

3.2.2 Hybrid Approach

Luthra (2017) integrated AHP and VIKOR in an automotive company to help decision makers to select the most sustainable supplier and maintain the company position in the market by taking into consideration mainly criteria in relation with green governance and capacities.

Kumar (2014) established an integrated method Green DEA with carbon footprint criteria belonging to the innovation and direct ecological impact class in an automotive company to select the most eco-friendly supplier.

Hashemi (2015) implemented an integrated ANP and an Improved GRA for selecting the prominent environmental suppliers in the automotive industry on a mix of green criteria related to green governance and capacities, innovation and direct ecological impact.

Freeman (2015) mixed AHP, Entropy and TOPSIS to order potential suppliers based mostly on green criteria from the class of green governance and capacities in a company of electronics and to select the best eco-friendly supplier.

Sivakumar (2015) presented the integration of AHP and TLF in the mining industry for the selection of the suitable green supplier using environmental criteria from the class of green governance and capacities.

Chithambaranathan (2015) applied GREY, Electre and VIKOR in two study cases of the manufacturing industry and catering industry. This tool aimed to evaluate the environmental capacity of firms with a mix of green criteria, focusing on the ones from green governance and capacities, to end up with the selection of the most efficient green supplier.

Kuo (2015) integrated DANP and VIKOR to evaluate the environmental performance of suppliers in the electronics industry and to select the appropriate green supplier by using mainly green criteria from the class of green manufacturing and product quality.

Yazdani (2017) proposed an approach of DEMATEL, QFD and COPRAS to select and rank eco-friendly suppliers in the food industry by focusing on green governance and capacities as environmental criteria.

Yazdani (2016) proposed a mix of methods: SWARA, QFD and WASPAS, for solving the problem of selecting the best environmental supplier and using principally green criteria from the class of green governance and capacities.

Wu (2016) implemented ANP and MOP for selecting environmental partners in an electronics firm to make a balance between reducing impacts on the environment and increasing the business function by using a mix of green criteria from the three classes.

Liou (2016) illustrated a methodology composed of DEMATEL, DANP and modified COPRAS-G to ameliorate and select the appropriate green suppliers in an electronics company by choosing as environmental criteria the ones from the green governance and capacities class.

He (2018) worked with FA, DEA and AHP in order to reach a supply chain with low-carbon levels in the metal industry. This approach permitted green supplier selection and evaluation by taking into consideration appropriate green criteria related principally to the class of green governance and capacities.

4 Discussion

In this paper, 55 journal articles related to green supplier selection and evaluation were retained and examined in order to reach the implemented objective. Our observations are presented as follows:

1. The Most Cited Multi-criteria Decision Making Approach

After the analysis of the collected literature, the integration of the fuzzy sets in solving green supplier selection and evaluation is common in many articles. In fact, 35 out of 55 (64%) papers worked under fuzzy sets. This integration of fuzzy sets in multi-criteria decision making approaches by Bellman and Zadeh (1970) made a revolution in this field because of its capacity to make the models more close to reality by decreasing the subjectivity, unreliability and ambiguity behind the decision-maker's opinions (Shen, 2013). In contrast, 20 papers out of 55 (36%) presented studies under non-fuzzy sets.

Four out of fifty-five (7%) journal articles are identified as single approaches and thirty-one out of fifty-five (57%) articles are implementing hybrid Multi-Criteria Decision Making approaches under Fuzzy sets (summarized in Appendix 1). Eight out of fifty-five (15%) papers presented single methods while twelve out of fifty-five (21%) articles applied hybrid Multi-Criteria Decision Making approaches under non-Fuzzy sets (illustrated in Appendix 1).

So, this reveals the predominance of hybrid Multi-Criteria Decision Making approaches over individual approaches. It leads the amelioration of the quality of researches related to this subject and its evolution. In fact, Govindan (2015) had

recognized in his review paper applied for literature collected from 1997 to 2011 that the majority of the researchers were implementing individual approaches due to their easy application while hybrid approaches are considered closer in modeling the complexity of real-world decision making problems (Govindan, 2015). They are, also, able to solve conflicting objectives in the context of multi-criteria decision making problems that aims, in the same time, to preserve the environment and increase the business function of companies (Zhou, 2018).

Concerning the widely used method, TOPSIS (integrated or single) is number one, followed by AHP under fuzzy sets, AHP (integrated or single) is still the most utilized under non-fuzzy sets. Actually, when a comparative analysis is run between FAHP and FTOPSIS, each method has shown its strength and limits but has the same degree of adequacy when dealing with imprecision and the relative weight of multi-attributes, even though FAHP is more adequate when the study's objective is to change a current supplier (Junior, 2014).

2. The Most Cited Green Evaluation Criteria

For the purpose to establish sustainability, studies made on supplier selection and evaluation are mostly combining traditional, ecological and social criteria. In our paper, the objective is to identify the widely used environmental attributes. This is why we isolated green criteria in each article and we determined the hundreds that were summarized in Appendix 3 (Supplementary Material). The next step was to classify them, and identify the number of use of these environmental factors that are illustrated in Appendix 2. The individual description of articles in the previous section (results) contains the principal utilized class of green criteria in each study.

The statistical analysis of green evaluation criteria shows that the class of green governance and capacities is the most widely considered with a percentage of 54% which includes 13% concerning the sub-class of policies, 12% management and 29% green capabilities. In each sub-class we determined the most cited criteria that were respectively; "Staff environmental training", "Environmental management systems" and "Re-cycle rate". The focus on this class and its dominance is explained by the fact that "Very often, the preference structure is based on political rather than only technical criteria" (Opricovic, 2004).

The class of green manufacturing and product quality is used at a rate of 25% and consists of 6% for pre-production, 1% for production, 8% for product and 10% for post-production. The criteria in each sub-class are, respectively, "Green material selection", "Green production", "Green design", "Green packing and labelling". And the class of innovation and direct ecological impact is considered by a percentage of 21% divided by 18% for the sub-class of direct ecological impact and 3% for the sub-class of development potential. The most pointed criteria in the sub-classes are, respectively, "Gas emission" and "Green product innovation".

The top of the list of environmental criteria for green supplier and evaluation are "Environmental Management Systems" (EMS) and "Green Design", followed by "Re-cycle rate" then "Gas emission", "Environmental competencies", "Green technology capability", "Environmental, legislative management", "Green waste management", "Green packing and labelling", and "Natural resources consumption".

5 Conclusion

The main objective of this paper is the review of the literature related to Green Supplier Evaluation and Selection published between 2012 and 2019. A structured search methodology was implemented to filter the collected articles (200 articles) based on some criteria to obtain the most relevant journal articles. The retained papers were examined to extract the needed information and statistical analysis was done to reach our objective. After the observation of the information, a classification framework was presented for this purpose. So, our findings related to Multi-Criteria Decision Making approaches are that the majority of studies run under fuzzy sets, are integrated approaches and the most widely used Multi-Criteria Decision Making approach is TOPSIS under fuzzy sets and AHP under non-fuzzy sets. Concerning green evaluation criteria, the dominant class was Green Governance and capacities and the most widely used green criteria for green supplier and evaluation were "Environmental Management Systems" and "Green Design" with the same importance.

There are several future research directions, among others, the investigation of the environmental criteria for selection and accreditation of suppliers, the development of models that better balance the intermix of economic, environmental and social criteria, and, furthermore, the development and evaluation of integrated approaches based on fuzzy sets in order to deal with problems in practice.

Acknowledgments We would like to thank all the members of the Business Economics and Management department in the CIHEAM-Mediterranean Agronomic Institute of Chania for their help and support.

Appendix 1 Table of Classification of Multi-criteria Decision Making Approaches

Class	Approaches	No. of articles	Authors
Single approaches under fuzzy sets	TODIM under IT2FSs	1	Qin (2017)
	CIO under IFSs	1	Tuzkaya (2013)
	FTOPSIS	1	Kannan (2014)
	FVIKOR	1	Rostamzadeh (2015)
	Total	4	
Hybrid approaches under fuzzy sets	FAHP-FMOLP	1	Shaw (2012)
	FAHP-FTOPSIS-MOLP	1	Kannan (2013)
	AHP-FTOPSIS	1	Yazdani (2014)
	FAHP-MOLP	2	Azadnia (2015), Çalık (2019)
	FAD-FMOOM	1	Kannan (2015)

(continued)

Class	Approaches	No. of articles	Authors
	FERM	1	Azadi (2015)
	AHP-FMOLP	1	Shaw (2013)
	FSD-FTOPSIS	1	Orji (2015)
	FDEMATEL-FTOPSIS	1	Orji (2014)
	FAHP-FVIKOR	1	Awasthi (2018)
	BWM-FTOPSIS	1	Gupta (2017)
	FCM-VIKOR	1	Akman (2015)
	FPP-FTOPSIS	1	Fallahpour (2017)
	FEntropy-FTOPSIS	2	Zhao (2014), dos Santos (2019)
	FANP-FDEMATEL-FTOPSIS	1	Büyüközkan (2012)
	BWM-FGCM-IGRA	1	Haeri (2019)
	FAHP-FTOPSIS-FMOOM	1	Mohammed (2019)
	AQM-BWM under IVIUL	1	Liu (2019)
	ANP-VIKOR under INSs	1	Abdel-Baset (2019)
	QFD-TOPSIS under INSs	1	Van (2018)
	FQFD-FMOMM	1	Babbar (2018)
	FSE-FIS	1	Khan (2018)
	FDEMATEL-TLF-BOM	1	Gören (2018)
	ANP-TOPSIS under INSs	1	Abdel-Basset (2018)
	BWM-modified FTOPSIS-FMOLP	1	Lo (2018)
	FTOPSIS-FVIKOR-FGRA	1	Banaeian (2018)
	AHP-Delphi-FGRA-MOLP	1	Banaeian (2015)
	KAM-DEA-GP	1	Fallahpour (2016)
	FTOPSIS-FMOLP	1	Govindan (2016)
	Total	31	
Single approaches under non-fuzzy sets	DEMATEL	2	Hsu (2013), Falatoonitoosi (2013)
	GRA	1	Baskaran (2012)
	ANP	1	Chen (2012)
	DEA	1	Shi (2015)
	VAHP	1	Pishchulov (2019)
	AHP	2	Mathiyazhagan (2018), Erbiyik (2016)
	Total	8	
Hybrid approaches under non-fuzzy sets	AHP-VIKOR	1	Luthra (2017)
	GDEA	1	Kumar (2014)
	Integrated ANP-improved GRA	1	Hashemi (2015)
	AHP-entropy-TOPSIS	1	Freeman (2015)
	AHP-TLF	1	Sivakumar (2015)

(continued)

Class	Approaches	No. of articles	Authors
	Grey-Electre-VIKOR	1	Chithambaranathan (2015)
	DANP-VIKOR	1	Kuo (2015)
	DEMATEL-QFD-COPRAS	1	Yazdani (2017
	SWARA-QFD-WASPAS	1	Yazdani (2016)
	ANP-MOP	1	Wu (2016)
	DEMATEL-DANP-modified COPRAS-G	1	Liou (2016)
	FA-DEA-AHP	1	He (2018)
	Total	12	

Appendix 2 Table of Classification of Evaluation Criteria

Class	Sub-class	Criteria	Authors	No. of articles
Green governance and capacities	Policies	Staff environmental training	Sivakumar (2015), Liou (2016), Qin (2017), Banaeian (2015), Chithambaranathan (2015), Mathiyazhagan (2018), Gupta (2017), Pishchulov (2019), Hsu (2013), Shaw (2013)	10
		Adherence to environmental policies	Chithambaranathan (2015), Kannan (2015), Gupta (2017), Orji (2014, 2015), Abdel-Basset (2018)	6
		Environmental protection policy	Mathiyazhagan (2018)	1
		Information disclosure	Shaw (2013), Liou (2016), Orji (2014, 2015), Kannan (2015), Chen (2012)	6
		Green supplier evaluation	Babbar (2018), Kannan (2014), Büyüközkan (2012), Rostamzadeh (2015), Awasthi (2018), Falatoonitoosi (2013), Gupta (2017), He (2018)	8
		Pressuring suppliers for green initiatives on their end	Rostamzadeh (2015), Gupta (2017)	2

(continued)

Class	Sub-class	Criteria	Authors	No. of articles
		Green certification	Chithambaranathan (2015), Fallahpour (2017), Liou (2016), Pishchulov (2019), Çalık (2019), Kannan (2015), Mathiyazhagan (2018)	7
		Adoption of environmental code of conduct	Pishchulov (2019)	1
		Commitment to green investment	Gupta (2017), Pishchulov (2019), Shaw (2013)	3
		Carbon policy	Hsu (2013)	1
		Environmental planning	Banaeian (2015), Pishchulov (2019), Chithambaranathan (2015), Gupta (2017)	4
		Cooperation for environmental improvement	Kannan (2014), Rostamzadeh (2015), Pishchulov (2019), Chithambaranathan (2015), Büyüközkan (2012)	5
		Sub-supplier evaluation and training	Pishchulov (2019), Kannan (2014)	2
	Management	Environmental management systems	Luthra (2017), Banaeian (2018), Kannan (2013), Freeman (2015), Yazdani (2014, 2016, 2017), Azadnia (2015), Orji (2014), Banaeian (2015), Akman (2015), Mohammed (2019), dos Santos (2019), Van (2018), Babbar (2018), Khan (2018), Gören (2018), Fallahpour (2017), Rostamzadeh (2015), Liou (2016)	20
		Audit of carbon management	Hsu (2013)	1
		Management system of carbon information	Hsu (2013)	1
		Commitment to environmental management	dos Santos (2019), Liou (2016), Pishchulov (2019), Chen (2012), Hashemi (2015), Chithambaranathan (2015), Yazdani (2017), Haeri (2019), Gupta (2017), Kannan (2014)	10
		Green process management	Tuzkaya (2013), Gupta (2017), Kannan (2015)	3

(continued)

Class	Sub-class	Criteria	Authors	No. of articles
		Environmental, legislative management	Qin (2017), Kannan (2014, 2015), Falatoonitoosi (2013), Gupta (2017), Mathiyazhagan (2018), Tuzkaya (2013), Luthra (2017), Chen (2012), Pishchulov (2019), Rostamzadeh (2015)	11
		Process auditing	Pishchulov (2019), Kannan (2015)	2
		Ensuring supplier EMS adoption	Gupta (2017), Rostamzadeh (2015)	2
		Green supply chain management and the level of commitment to it	Wu (2016)	1
	Green capabilities	Environmental competencies	Luthra (2017), Qin (2017) Kannan (2013), Freeman (2015), Yazdani (2014), Orji (2014, 2015), Akman (2015), Chen (2012), Büyüközkan (2012), Pishchulov (2019), Liu (2019)	12
		Environmental activity control	Sivakumar (2015)	1
		Carbon governance	Hsu (2013)	1
		Carbon accounting and inventory	Kuo (2015), Kannan (2015), Wu (2016), Fallahpour (2017)	4
		ISO-14001 certification	Freeman (2015), Shaw (2013), Fallahpour (2017), Rostamzadeh (2015)	4
		Green market share	Kannan (2015), Chithambaranathan (2015)	2
		Ratio of green costumers to total costumers	Kannan (2015)	1
		Green organizational activities	Falatoonitoosi (2013), Büyüközkan (2012)	2
		Applying lifecycle analysis to execute an eco-report	Liou (2016)	1
		Green transportation		3

(continued)

Class	Sub-class	Criteria	Authors	No. of articles
			Gupta (2017), Fallahpour (2017), Rostamzadeh (2015)	
		Green warehousing	Kannan (2015), Fallahpour (2017), Gupta (2017), Rostamzadeh (2015)	4
		Inventory of non-hazardous substances	Mathiyazhagan (2018), Fallahpour (2017)	2
		Inventory of hazardous substances	Kuo (2015), Kannan (2015), Abdel-Basset (2018), Pishchulov (2019)	4
		Using modern eco-efficient transportation fleet	Rostamzadeh (2015), Fallahpour (2017)	2
		Using green fuels	Rostamzadeh (2015), Fallahpour (2017)	2
		Encouraging eco-driving to decrease fuel consumption	Rostamzadeh (2015)	1
		Avoidance and treatment of hazardous materials	Kannan (2014), Fallahpour (2017), Çalık (2019), Liou (2016), Chen (2012), Kannan (2015)	6
		Energy consumption and saving energy	Yazdani (2016), Govindan (2016), Wu (2016), Shi (2015), Pishchulov (2019), Chen (2012), Gupta (2017), Çalık (2019), Shaw (2013)	10
		Green manufacturing	Luthra (2017), Abdel-Baset (2019), Lo (2018), Gupta (2017), Chen (2012)	5
		Level of use of renewable energy	Pishchulov (2019), Rostamzadeh (2015)	2
		Green waste management	Freeman (2015), Yazdani (2014, 2016, 2017), Shaw (2013), Mohammed (2019), Luthra (2017), Pishchulov (2019), Awasthi (2018), Abdel-Baset (2019), Büyüközkan (2012)	11
		Re-cycle rate	He (2018), Rostamzadeh (2015), Büyüközkan (2012), Govindan (2016), Kannan (2015), Mathiyazhagan (2018),	16

(continued)

Class	Sub-class	Criteria	Authors	No. of articles
			Pishchulov (2019), Yazdani (2014, 2016, 2017), Fallahpour (2017), Khan (2018), Erbiyik (2016), Shi (2015), Gupta (2017), Babbar (2018)	
		Re-use rate	Kannan (2015), Fallahpour (2017), Büyüközkan (2012), Mathiyazhagan (2018), Erbiyik (2016), Yazdani (2014), Shi (2015), Yazdani (2016, 2017), Babbar (2018)	10
		Remanufacturing	Rostamzadeh (2015), Freeman (2015), Fallahpour (2017), Büyüközkan (2012), Mathiyazhagan (2018), Pishchulov (2019), Kannan (2015)	7
		Recycled product-treatment-potential	Mathiyazhagan (2018)	1
		Cleaner and/or green technology capability	Freeman (2015), Khan (2018), Liou (2016), Qin (2017), Sivakumar (2015), Chithambaranathan (2015), Kannan (2014, 2015), Banaeian (2015), Fallahpour (2017), Büyüközkan (2012), Mathiyazhagan (2018)	12
		Environmental amelioration cost	He (2018), Luthra (2017), Tuzkaya (2013), Erbiyik (2016), Gupta (2017)	4
Green manufacturing and product quality	Pre-production	Good use of ICT tools	Chen (2012), Gupta (2017)	2
		Green purchasing	Gupta (2017), Chen (2012), Rostamzadeh (2015), Büyüközkan (2012), Mathiyazhagan (2018)	5
		Green material selection	Freeman (2015), Chithambaranathan (2015), Yazdani (2014), Banaeian (2015), Awasthi (2018), dos Santos (2019), Khan (2018),	10

(continued)

Class	Sub-class	Criteria	Authors	No. of articles
			Fallahpour (2017), Rostamzadeh (2015), Gupta (2017)	
		Green material coding and recoding	Kannan (2015)	1
		Prevention of mixed material	Kannan (2015)	1
		Use of waste of other companies	Rostamzadeh (2015), Pishchulov (2019)	2
		Commitment to green logistics	Falatoonitoosi (2013), Lo (2018), Büyüközkan (2012), Pishchulov (2019)	4
	Production process	Green production	Kannan (2015), Rostamzadeh (2015), Büyüközkan (2012), Pishchulov (2019), Kannan (2014)	5
		Level of use of toxic/ restricted substances	Freeman (2015)	1
		Materials consumption	Pishchulov (2019)	1
	Product	Green design	Kannan (2013), Chithambaranathan (2015), Yazdani (2014), Yazdani (2016, 2017), Orji (2014, 2015), Kannan (2015), Akman (2015), dos Santos (2019), Van (2018), Babbar (2018), Gören (2018), Chen (2012), Fallahpour (2017), Rostamzadeh (2015), Kannan (2014), Gupta (2017), Luthra (2017), Pishchulov (2019)	20
		Green products	Rostamzadeh (2015), Liou (2016), Pishchulov (2019), Kuo (2015), Kannan (2015), Gupta (2017)	6
		Eco-design cost	Azadi (2015)	1
		Products reduce the impact on natural resources	Kannan (2015), Fallahpour (2017)	2
		Ability to alter the product and process in order to reduce impact on natural resources	Kannan (2015), Fallahpour (2017), Liou (2016), Rostamzadeh (2015), Gupta (2017)	5

(continued)

Class	Sub-class	Criteria	Authors	No. of articles
		Renewable product design	Kannan (2015)	1
		Biodegradability	Pishchulov (2019)	1
	Post production	Carbon disclosure and report	Hsu (2013), Shaw (2013)	2
		Commitment to green marketing	Chen (2012), Pishchulov (2019) Qin (2017), Freeman (2015), Tuzkaya (2013), Kannan (2015), Akman (2015), Haeri (2019), dos Santos (2019)	9
		Sale of excess stock or raw material	Rostamzadeh (2015), Kannan (2014)	2
		Sale of scrap and used materials	Rostamzadeh (2015), Kannan (2014)	2
		Sale of used equipment	Rostamzadeh (2015), Kannan (2014), Pishchulov (2019)	3
		Green packing and labelling	Fallahpour (2017), Rostamzadeh (2015), Kannan (2014), Büyüközkan (2012), Mathiyazhagan (2018), Pishchulov (2019), Freeman (2015), Kannan (2015), Abdel-Basset (2018), Fallahpour (2017), Luthra (2017), Abdel-Baset (2019)	11
		Wastewater management	Kannan (2015), Awasthi (2018), Wu (2016), Fallahpour (2017), Shi (2015), Kuo (2015), Pishchulov (2019)	7
		Solid waste management	Shaw (2013), Wu (2016), Shi (2015), Kuo (2015)	4
		Cost of component disposal	Kannan (2015)	1
		Waste electronical equipment management	Orji (2014)	1
		Green distribution	Rostamzadeh (2015), Büyüközkan (2012)	2
Innovation and direct ecological impact	Ecological impact	Carbon reduction targets	Gupta (2017), Hsu (2013), Shaw (2013)	3
		Natural resources consumption	Qin (2017), Hashemi (2015), Kannan (2013),	11

(continued)

Class	Sub-class	Criteria	Authors	No. of articles
			Fallahpour (2016), Yazdani (2017), Wu (2016), Haeri (2019), dos Santos (2019), Gören (2018), Chen (2012), Gupta (2017)	
		Sustainable land use	Pishchulov (2019), Awasthi (2018), Khan (2018)	3
		Volume of waste water produced	Awasthi (2018), Khan (2018), Pishchulov (2019)	3
		Volume of solid waste	Pishchulov (2019)	1
		Volume of liquid waste	Pishchulov (2019), Shaw (2013)	2
		Volume of hazardous waste	Fallahpour (2017), Pishchulov (2019), Kannan (2015), Chen (2012)	4
		Pollution production	Qin (2017), Azadnia (2015) Baskaran (2012), Haeri (2019), Mohammed (2019), dos Santos (2019)	6
		Pollution prevention	Akman (2015), Kuo (2015)	2
		Pollution control	Fallahpour (2016), Chithambaranathan (2015), Tuzkaya (2013), Kannan (2015), Orji (2014), Wu (2016), Haeri (2019), Van (2018), Liou (2016), Mathiyazhagan (2018)	10
		Environmental footprint	Shi (2015), Shaw (2013)	2
		Gas emission	Kuo (2015), Kannan (2015), Wu (2016), Fallahpour (2017), Van (2018, Çalık (2019), He (2018), Pishchulov (2019), Kumar (2014), Shaw (2012), Azadnia (2015), Govindan (2016), Shi (2015), Pishchulov (2019), Awasthi (2018), Khan (2018)	15
		Environmental efficiency	Pishchulov (2019), Lo (2018), Fallahpour (2017), Kannan (2015), Chen (2012),	7

(continued)

Class	Sub-class	Criteria	Authors	No. of articles
			Rostamzadeh (2015), Gupta (2017)	
		Environmental risk	Shaw (2013), Chen (2012)	2
		Carbon risk assessment	Hsu (2013), Shaw (2013)	2
		Biodiversity	Awasthi (2018)	1
		Degree of water pollution	Pishchulov (2019)	1
		Environmental adaptability	Abdel-Basset (2018)	1
		Carbon verification	Hsu (2013)	1
		Level of use of non-renewable resources	Wu (2016)	1
		Gaseous residue of manufacturing	Shaw (2013)	1
	Development potential	Green R&D investments	Gupta (2017), Fallahpour (2017), Pishchulov (2019)	3
		Green product innovation	Tuzkaya (2013), Kannan (2015), Fallahpour (2017), Qin (2017), Luthra (2017), Chithambaranathan (2015), Gupta (2017)	7
		Technical assistance for technology up gradation	Gupta (2017)	1
		Increasing innovation capabilities	Chen (2012), Lo (2018), Yazdani (2017)	3

References

Abdel-Baset, M. C. (2019). An integrated neutrosophic ANP and VIKOR method for achieving sustainable supplier selection: A case study in importing field. *Computers in Industry, 106*, 94–110.

Abdel-Basset, M. M. (2018). A hybrid neutrosophic group ANP-TOPSIS framework for supplier selection problems. *Symmetry, 10*(6), 226.

Akman, G. (2015). Evaluating suppliers to include green supplier development programs via fuzzy c-means and VIKOR methods. *Computers & Industrial Engineering, 86*, 69–82.

Amit, R. S. (1993). Strategic assets and organizational rent. *Strategic Management Journal, 14*, 33–46.

Awasthi, A. G. (2018). Multi-tier sustainable global supplier selection using a fuzzy AHP-VIKOR based approach. *International Journal of Production Economics, 195*, 106–117.

Azadi, M. J. (2015). A new fuzzy DEA model for evaluation of efficiency and effectiveness of suppliers in sustainable supply chain management context. *Computers & Operations Research, 54*, 274–285.

Azadnia, A. H. (2015). Sustainable supplier selection and order lot-sizing: An integrated multi-objective decision-making process. *International Journal of Production Research, 53*(2), 383–408.

Babbar, C. (2018). A multi-objective mathematical model integrating environmental concerns for supplier selection and order allocation based on fuzzy QFD in beverages industry. *Expert Systems with Applications, 92*, 27–38.

Banaeian, N. M. (2015). Criteria definition and approaches in green supplier selection–a case study for raw material and packaging of food industry. *Production & Manufacturing Research, 3*(1), 149–168.

Banaeian, N. M. (2018). Green supplier selection using fuzzy group decision making methods: A case study from the agri-food industry. *Computers & Operations Research, 89*, 337–347.

Barney, J. (1986). Organizational culture: Can it be a source of sustained competitive advantage? *Academy of Management Review, 2*(3), 656–665.

Baskaran, V. N. (2012). Indian textile suppliers' sustainability evaluation using the grey approach. *International Journal of Production Economics, 135*(2), 647–658.

Bellman, R. E., & Zadeh, L. A. (1970). Decision-making in a fuzzy environment. *Management Science, 17*(4), B-141.

Büyüközkan, G. (2012). A novel hybrid MCDM approach based on fuzzy DEMATEL, fuzzy ANP and fuzzy TOPSIS to evaluate green suppliers. *Expert Systems with Applications, 39*(3), 3000–3011.

Çalık, A. P. (2019). Lean and green supplier selection problem: A novel multi objective linear programming model for an electronics board manufacturing company in Turkey. In *Multiple criteria decision making and aiding* (pp. 281–309). Cham: Springer.

Chen, C. C. (2012). A business strategy selection of green supply chain management via an analytic network process. *Computers & Mathematics with Applications, 64*(8), 2544–2557.

Chithambaranathan, P. S. (2015). Service supply chain environmental performance evaluation using grey based hybrid MCDM approach. *International Journal of Production Economics, 166*, 163–176.

Cronin, P. R. (2008). Undertaking a literature review: A step-by-step approach. *British Journal of Nursing, 17*(1), 38–43.

Dickson, G. W. (1966). An analysis of vendor selection systems and decisions. *Journal of Purchasing, 2*(1), 12.

dos Santos, B. M. (2019). Performance evaluation of green suppliers using entropy-TOPSIS-F. *Journal of Cleaner Production, 207*, 498–509.

Erbiyik, H. B. (2016). Analytic hierarchy process method for supplier selection considering green logistics: Case study of aluminum production sector. *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering, 10*(8), 2867–2871.

Falatoonitoosi, E. L. (2013). Modeling for green supply chain evaluation. *Mathematical Problems in Engineering, 2013*, 1–9.

Fallahpour, A. O. (2016). An integrated model for green supplier selection under fuzzy environment: Application of data envelopment analysis and genetic programming approach. *Neural Computing and Applications, 27*(3), 707–725.

Fallahpour, A. O. (2017). A decision support model for sustainable supplier selection in sustainable supply chain management. *Computers & Industrial Engineering, 105*, 391–410.

Fowler, S. J. (2007). Incorporating sustainable business practices into company strategy. *Business Strategy and the Environment, 16*, 26–38.

Freeman, J. (2015). Green supplier selection using an AHP-entropy-TOPSIS framework. *Supply Chain Management. An International Journal, 20*(3), 327–340.

Friedman, R. (1984). *Strategic management: A stakeholder approach*. Marshfield, MA: Pitman.

Gören, H. G. (2018). A decision framework for sustainable supplier selection and order allocation with lost sales. *Journal of Cleaner Production, 183*, 1156–1169.
Govindan, K. R. (2015). Multi criteria decision making approaches for green supplier evaluation and selection: A literature review. *Journal of Cleaner Production, 98*, 66–83.
Govindan, K. S. (2016). Green supplier selection and order allocation in a low-carbon paper industry: Integrated multi-criteria heterogeneous decision-making and multi-objective linear programming approaches. *Annals of Operations Research, 238*(1–2), 243–276.
Gupta, H. (2017). Supplier selection among SMEs on the basis of their green innovation ability using BWM and fuzzy TOPSIS. *Journal of Cleaner Production, 152*, 242–258.
Gurel, O. A. (2015). Determinants of the green supplier selection. *Procedia-Social and Behavioral Sciences, 181*, 131–139.
Haeri, S. A. (2019). A grey-based green supplier selection model for uncertain environments. *Journal of Cleaner Production, 221*, 768–784.
Hart, S. (1995). The natural resource-based view of the firm. *Academy of Management Review, 20* (4), 986–1014.
Hashemi, S. H. (2015). An integrated green supplier selection approach with analytic network process and improved Grey relational analysis. *International Journal of Production Economics, 159*, 178–191.
He, X. (2018). Supplier selection study under the respective of low-carbon supply chain: A hybrid evaluation model based on FA-DEA-AHP. *Sustainability, 10*(2), 564.
Hsu, C. W. (2013). Using DEMATEL to develop a carbon management model of supplier selection in green supply chain management. *Journal of Cleaner Production, 56*, 164–172.
Junior, F. R. (2014). A comparison between fuzzy AHP and fuzzy TOPSIS methods to supplier selection. *Applied Soft Computing, 21*, 194–209.
Kannan, D. K. (2013). Integrated fuzzy multi criteria decision making method and multi-objective programming approach for supplier selection and order allocation in a green supply chain. *Journal of Cleaner Production, 47*, 355–367.
Kannan, D. d. (2014). Selecting green suppliers based on GSCM practices: Using fuzzy TOPSIS applied to a Brazilian electronics company. *European Journal of Operational Research, 233*(2), 432–447.
Kannan, D. G. (2015). Fuzzy axiomatic design approach based green supplier selection: A case study from Singapore. *Journal of Cleaner Production, 96*, 194–208.
Khan, S. A.-S.-S. (2018). Supplier sustainability performance evaluation and selection: A framework and methodology. *Journal of Cleaner Production, 205*, 964–979.
Kinnear, T. C. (1973). The effect of ecological concern. *Journal of Marketing Research, 10*(2), 191–197.
Kumar, A. J. (2014). A comprehensive environment friendly approach for supplier selection. *Omega, 42*(1), 109–123.
Kuo, T. H. (2015). Developing a green supplier selection model by using the DANP with VIKOR. *Sustainability, 7*(2), 1661–1689.
Lage Junior, M. (2010). Variations of the kanban system: Literature review and classification. *International Journal of Production Economics, 125*(1), 13–21.
Liou, J. J. (2016). New hybrid COPRAS-G MADM model for improving and selecting suppliers in green supply chain management. *International Journal of Production Research, 54*(1), 114–134.
Liu, H. C. (2019). A new integrated MCDM model for sustainable supplier selection under interval-valued intuitionistic uncertain linguistic environment. *Information Sciences, 486*, 254–270.
Lo, H. W. (2018). An integrated model for solving problems in green supplier selection and order allocation. *Journal of Cleaner Production, 190*, 339–352.
Luthra, S. G. (2017). An integrated framework for sustainable supplier selection and evaluation in supply chains. *Journal of Cleaner Production, 140*, 1686–1698.
Mathiyazhagan, K. S. (2018). Modeling the criteria for selection of suppliers towards green aspect: A case in Indian automobile industry. *Opsearch, 55*(1), 65–84.

Mohammed, A. H. (2019). A hybrid MCDM-FMOO approach for sustainable supplier selection and order allocation. *International Journal of Production Economics., 217*, 171–184.

Opricovic, S. (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. *European Journal of Operational Research, 156*(2), 445–455.

Orji, I. J. (2014). A decision support tool for sustainable supplier selection in manufacturing firms. *Journal of Industrial Engineering and Management (JIEM), 7*(5), 1293–1315.

Orji, I. J. (2015). An innovative integration of fuzzy-logic and systems dynamics in sustainable supplier selection: A case on manufacturing industry. *Computers & Industrial Engineering, 88*, 1–12.

Pishchulov, G. T. (2019). The voting analytic hierarchy process revisited: A revised method with application to sustainable supplier selection. *International Journal of Production Economics, 211*, 166–179.

Purvis, B. M. (2018). Three pillars of sustainability: In search of conceptual origins. *Sustainability Science, 14*, 1–15.

Qin, J. L. (2017). An extended TODIM multi-criteria group decision making method for green supplier selection in interval type-2 fuzzy environment. *European Journal of Operational Research, 258*(2), 626–638.

Rezaei, J. N. (2016). A supplier selection life cycle approach integrating traditional and environmental criteria using the best worst method. *Journal of Cleaner Production, 135*, 577–588.

Rostamzadeh, R. G. (2015). Application of fuzzy VIKOR for evaluation of green supply chain management practices. *Ecological Indicators, 49*, 188–203.

Shaw, K. S. (2012). Supplier selection using fuzzy AHP and fuzzy multi-objective linear programming for developing low carbon supply chain. *Expert Systems with Applications, 39*(9), 8182–8192.

Shaw, K. S. (2013). Global supplier selection considering sustainability and carbon footprint issue: AHP multi-objective fuzzy linear programming approach. *International Journal of Operational Research, 17*(2), 215–247.

Shen, L. O. (2013). A fuzzy multi criteria approach for evaluating green supplier's performance in green supply chain with linguistic preferences. *Resources, Conservation and Recycling, 74*, 170–179.

Shi, P. Y. (2015). A decision support system to select suppliers for a sustainable supply chain based on a systematic DEA approach. *Information Technology and Management, 16*(1), 39–49.

Sivakumar, R. K. (2015). Green vendor evaluation and selection using AHP and Taguchi loss functions in production outsourcing in mining industry. *Resources Policy, 46*, 64–75.

Tuzkaya, G. (2013). An intuitionistic fuzzy Choquet integral operator based methodology for environmental criteria integrated supplier evaluation process. *International journal of Environmental Science and Technology, 10*(3), 423–432.

Van, L. Y. (2018). New integrated quality function deployment approach based on interval neutrosophic set for green supplier evaluation and selection. *Sustainability, 10*(3), 1–13.

Walley, N. B. (1994). It's not easy being green. *Harvard Business Review, 72*, 46–52.

Weber, C. A. (1991). Vendor selection criteria and methods. *European Journal of Operation Research, 50*(2), 2–18.

Wu, F. L. (2013). Supplier selection for outsourcing from the perspective of protecting crucial product knowledge. *International Journal of Production Research, 51*(5), 1508–1519.

Wu, C. (2016). An integrated model for green partner selection and supply chain construction. *Journal of Cleaner Production, 112*, 2114–2132.

Yazdani, M. (2014). An integrated MCDM approach to green supplier selection. *International Journal of Industrial Engineering Computations, 5*(3), 443–458.

Yazdani, M. H. (2016). New integration of MCDM methods and QFD in the selection of green suppliers. *Journal of Business Economics and Management, 17*(6), 1097–1113.

Yazdani, M. C. (2017). Integrated QFD-MCDM framework for green supplier selection. *Journal of Cleaner Production, 142*, 3728–3740.

Zhao, H. (2014). Selecting green supplier of thermal power equipment by using a hybrid MCDM method for sustainability. *Sustainability, 6*(1), 217–235.
Zhou, X. (2018). An integrated sustainable supplier selection approach based on hybrid information aggregation. *Sustainability, 10*(7), 2543.
Zimmermann, H.-J. (2001). *Fuzzy set theory and its applications* (4th ed.). New York: Springer Science & Business Media. Recuperado el 26 de 05 de 2019.

A Similarity Hybrid Harmony Search Algorithm for the Orienteering Problem

Eleftherios Tsakirakis, Magdalene Marinaki, and Yannis Marinakis

Abstract In the last few years there is a continuous growth in modeling and solving of problems of different fields (logistics, tourism, games) as Orienteering Problems (OPs). The Orienteering Problem is a combinatorial optimization problem where a standard amount of nodes are given, each with a specific score. The main goal is to find a path, limited in length, from the start point to the end point through a subset of locations in order to maximize the total path score. In this paper, we present a variant of the classic Harmony Search (HS) algorithm, the Similarity Hybrid Harmony Search (SHHS) algorithm, for the solution of the Orienteering Problem. The SHHS follows the basic steps of the standard HS with some minor changes and includes a new idea considering the similarity of the feasible routes such as the musical notes of a suitable frequency for the Harmony Memory. The algorithm was tested in a number of benchmark instances from the literature and in most of them the best known solutions were found.

Keywords Harmony search · Orienteering problem · Metaheuristic algorithms

1 Introduction

In this paper, we present an evolutionary algorithm based on Harmony Search (Geem, Kim, & Loganathan, 2001) for the solution of the Orienteering Problem. The Orienteering Problem was introduced by Golden, Levy, and Vohra (1987). The initial application of the problem was a game where a number of players starting from a specified control point and using a map that informs them about a number of checkpoints and the scores associated with each one of them, are trying to pass from as many checkpoints as possible adding to their total score the score of the specific point and to return to the control point (or to go to a different control point) within a

E. Tsakirakis (✉) · M. Marinaki · Y. Marinakis
School of Production Engineering and Management, Technical University of Crete, Chania, Greece
e-mail: etsakirakis@isc.tuc.gr; magda@dssl.tuc.gr; marinakis@ergasya.tuc.gr

E. Krassadaki et al. (eds.), *Operational Research in Agriculture and Tourism*, Cooperative Management, https://doi.org/10.1007/978-3-030-38766-2_10

specific time period (Chao, Golden, & Wasil, 1996a). The winner is the one that maximizes its total collected score (Vansteenwegen, Souffriau, & Van Oudheusden, 2011).

This problem is one of the first problems that belongs in a category that is called Vehicle Routing Problems with Profit (Archetti, Speranza, & Vigo, 2014). In these problems, a profit and a demand are associated in each node (customer) and, thus, the main goal is instead of (or in addition to) the minimization of the total distance or the total travel time, the maximization of the additive profit of visiting of the most profitable customers (Marinakis, Politis, Marinaki, & Matsatsinis, 2015). In all these problems, as there is, usually, a time limit in order to perform the service of the customers (Archetti et al., 2014; Marinakis et al., 2015), it is possible to visit only some customers and not all the customers. The most known variant of the Orienteering Problem is the Team Orienteering Problem where instead of one path, a number of P paths should be determined where the total collected score should be maximized (Chao, Golden, & Wasil, 1996b; Tang & Miller-Hooks, 2005). Other variants of the Orienteering Problem are the Orienteering and the Team Orienteering Problems with Time Windows (Montemanni & Gambardella, 2009; Righini & Salani, 2009) where, also, time windows are assigned in each arc of the graph.

The Harmony Search (HS) algorithm is a very successful algorithm (Geem et al., 2001) that can be used for solving effectively both continuous and combinatorial optimization problems. The inspiration of the HS came from the music composition process. When the musicians play their musical instruments, they try to match the right acoustic frequencies in order to compose the perfect harmony. Thus, the musicians try to find the right acoustic frequency just like HS seeks the best optimal solution for the objective function. A number of variants of the basic Harmony Search algorithm have been proposed, some of them are the improved harmony search (IHS) (Mahdavi, Fesanghary, & Damangir, 2007), the global-best harmony search (GHS) (Omran & Mahdavi, 2008), the self-adaptive global best harmony search (SGHS) (Pan, Suganthan, Tasgetiren, & Liang, 2010) and the local-best harmony search algorithm with dynamic sub-harmony memories (DLHS) (Pan, Suganthan, Liang, & Tasgetiren, 2011). Also, the Harmony Search algorithm has been applied in different applications such as engineering optimization problems (Mahdavi et al., 2007), the flowshop scheduling problem (Pan et al., 2011), the school bus routing problem (Geem, 2005), the generalized orienteering problem (GOP) (Geem, Tseng, & Park, 2005), the capacitated vehicle routing problem (Pichpibul & Kawtummachai, 2013), inventory problems (Taleizadeh, Niaki, & Barzinpour, 2011), wireless communications networks (Ser, Matinmikko, Gil-Lopez, & Mustonen, 2012), the analysis and processing of image (Fourie, Mills, & Green, 2010) and for structure design optimization (Erdal, Dogan, & Saka, 2011; Kaveh & Talataha, 2009; Miguel & Miguel, 2012).

Our approach, the Similarity Hybrid Harmony Search (SHHS), is a variant of the classic HS. The Orienteering Problem is a discrete optimization problem. Thus, we kept the basic concept of the method but we modified some of the main operations of the HS. Moreover, we added a method called "similarity process" as a new feature. Our algorithm took its name from this method. The "similarity process" is applied

during the main procedure of the SHHS algorithm. The aim of this technique is to amplify the score of each solution. The "similarity process" focuses in the exclusion of the problematic nodes from the solutions and to insert better ones with the Nearest Insertion algorithm.

The rest of the paper is organized as follows. In the next section, the Orienteering Problem is formulated and described. Afterwards, the basic Harmony Search algorithm is presented in brief and, then, the proposed algorithm is given and analyzed in detail. Then, the computational results in the five different benchmark sets are given. Finally, the conclusions and the future perspectives of the algorithm are given.

2 Orienteering Problem

The Orienteering Problem (OP) can be described using a graph $G = (V, A)$, where V denotes the set of nodes, each one having a score s_i and A is the set of arcs between points in V. There are two fixed points, the starting point (usually node 1) and the ending point (usually different from the starting point and is denoted as node N), with score of these two nodes equal to zero. There is a symmetric nonnegative cost c_{ij} associated with each arc, where c_{ij} denotes the time between point i and point j. Each node can be visited at most once and the total time taken to visit all nodes cannot exceed the specified limit T_{max} (Vansteenwegen et al., 2011). The main target of the solution of the OP is to find a path, limited by T_{max}, that visits some of the nodes in order to maximize the total collected score. The scores are assumed to be entirely additive and each node can be visited at most once (Vansteenwegen et al., 2011).

The OP can be formulated as an integer problem. The following decision variables are used:

$$x_{ij} = \begin{cases} 1, & \text{if a visit to node } i \text{ is followed by a visit to node } j \\ 0, & \text{otherwise} \end{cases}$$

and

$$y_{ij} = \begin{cases} 1, & \text{if node } i \text{ is in path} \\ 0, & \text{otherwise} \end{cases}$$

Then, the formulation of the problem is the following:

$$z = max \sum_{j \epsilon V} s_j \, y_j$$

$$\sum_{j \epsilon V} x_{1j} = 1,$$

$$\sum_{j \epsilon V} x_{jN} = 1,$$

$$\sum_{(i,j) \epsilon A} x_{ij} + \sum_{(j,k) \epsilon A} x_{jk} = 2y_i \forall j \in V,$$

$$\sum_{(i,j) \epsilon A} c_{ij} x_{ij} \leq T_{max}$$

$$\sum_{(i,j) \epsilon S} x_{ij} \leq |S| - 1 \forall S \; in \; V$$

3 The Harmony Search Algorithm

The Harmony Search (HS) algorithm was presented by Geem et al. (2001) as a metaheuristic algorithm to solve the traveling salesman problem (TSP) back in 2001. An improvement and application of the algorithm in continuous optimization problems was given by Lee and Geem (2004). The concept of this metaheuristic is based on the music composition procedure that is performed by the musicians. The musicians improvise with the acoustic frequencies in order to create the perfect harmony. The meta-heuristic HS has five steps. These are:

1. Initialize the optimization procedure and algorithm parameters
2. Initialize the harmony memory
3. Improvise a new harmony from the harmony memory
4. Update the harmony memory
5. Repeat the previous two steps while the stopping criterion is not satisfied

A brief presentation of the previous five steps is the following:

Step 1 An optimization (minimization or maximization) problem is formulated where $f(\mathbf{x})$ denotes the objective function and $x_i \in X_i$, with $i = 1, \ldots, N$ are the decision variables where N defines the number of the decision variables.

Additionally, for the HS algorithm a set of parameters are used with most important of them the harmony memory (HM), a table where the solution vectors are included. The size of the harmony memory is denoted as HMS and indicates the

number of the stored solutions in memory. Other parameters are the harmony memory consideration rate (HMCR), the pitch adjustment rate (PAR) and the number of improvisations (NI). The HMCR and the PAR are used to improve the solution vectors and NI defines the stopping criteria of the algorithm.

Step 2 Afterwards, the HM is initialized with randomly generated solution vectors as the HMS designates. Below an example of an HM is shown:

$$HM = \begin{bmatrix} & & x_1^1 & x_2^1 & \dots & x_{N-1}^1 & x_N^1 \\ & & & & & & \dots \\ & & x_1^2 & x_2^2 & \dots & x_{N-1}^2 & x_N^2 \\ & & & & & & \dots \\ x_1^{HMS-1} & x_2^{HMS-1} & & \dots & x_{N-1}^{HMS-1} & x_N^{HMS-1} & \\ & x_1^{HMS} & x_2^{HMS} & & \dots & x_{N-1}^{HMS} & x_N^{HMS} \end{bmatrix}$$

Step 3 In this step, the improvisation of the harmonies begins. A new harmony vector $x' = (x'_1, x'_2, \dots, x'_N)$, is generated through the harmony memory consideration rate (HMCR), the pitch adjustment rate (PAR) or through randomization. These options are applied to every variable in the new harmony vector in every iteration. Generating a new harmony is referred as improvisation.

During the improvisation procedure, the value of every decision variable (x') is selected from any value of the specified HM range. Taking into account the option of memory consideration, the HS algorithm picks the new value via the HMCR parameter, which varies between 0 and 1 where p defines the probability:

$$x'_i = \begin{cases} x'_i \in x_1, x_2, \dots, x_i^{HMS}, \text{where } p \leq HMCR \\ x'_i \in X_i, \quad \text{otherwise} \end{cases}$$

If $p \leq HMCR$, then, a decision value will be selected from the stored values in the HM. Otherwise, a decision value will be randomly chosen from the possible range of values. For instance, an HMCR of 0.9 declares that the HS algorithm will pick the decision variable from the stored values in the HM with a 90% probability and from the entire value range with 10% probability.

The chosen decision variables are examined one by one to determine whether it should be pitch-adjusted. To achieve this, the HS algorithm uses the PAR parameter in order to define the rate of pitch-adjustment. The concept is as follows:

$$\text{Pitch adjustment decision for } x_i' \leftarrow \begin{cases} \text{Yes}, & \text{where } p \leq PAR \\ \text{No}, & otherwise \end{cases}$$

The modification of x_i' is given from:

$$x_i' \leftarrow x_i' \pm rand() * bw$$

The *rand*() varies between 0 and 1. The *bw* value defines an arbitrary distance value for the decision variables.

Step 4 In this step, HS algorithm finds the worst harmony vector in the HM and compares it with the new one. If the new harmony vector is better, then, it is inserted in the HM and the worst one is excluded from the HM.

Step 5 The HS algorithm repeats Steps 3 and 4 until the number of improvisation *NI* is satisfied.

4 The Proposed Similarity Hybrid Harmony Search

In this part of the paper, the proposed Similarity Hybrid Harmony Search (SHHS) algorithm is presented. As we mentioned earlier, our method is a mix of the classic HS with the new technique "similarity process". The SHHS algorithm is fully explained step-by-step and analyzed in the below subsections. This section is organized with the illustration of the SHHS parameters and the main phases of the algorithm as follows:

- Phase 1: Definition of the parameters
- Phase 2: Composition of the initial harmonies and Harmony Memory construction
- Phase 3: The Improvisation stage
- Phase 4: New technique called "**similarity process**"
- Phase 5: Finding best harmonies and update Harmony Memory

4.1 Definition of the Parameters

The SHHS algorithm operates using variables from the classic HS, the operators for the "similarity process" and of the parameters needed for solving the Orienteering Problem. Initially, we have a set of Nodes, $V = 1, 2, \ldots, N$. Then, we add the rest of variables for the OP. These are the score of each node, s_i, the distance from node i to node j which is denoted as t_{ij}, the vehicle routes, M and the total time limit, T_{max}. Moreover, in the beginning of the SHHS, the harmony memory *HM* is initialized,

which will contain all the harmonies (the solutions). Also, the number of improvisations *NI* of the algorithm (total executions) is defined. We use the variables of harmony memory consideration rate *HMCR* and the pitch adjustment rate *PAR*, which both are included in the classic HS and are explained with more details in the next subsections. In the following, an example of a harmony memory with five different solutions is given. As it can be seen each solution does not include all the nodes as in the OP only a subset of nodes will be used in each solution. The starting node is the node 1 and the ending node is the node 15. In general, the harmonies consist of *M* frequencies. The frequencies are important for the SHHS to solve OP. For this reason, we use a harmony memory list *HML* as a list variable. The *HML* includes all the frequencies from all the created harmonies. Last but not least, in the subsection of the "similarity process" there are the similarity parameter *SP* and the similarity matrix *SM*. The *SP* is vital for that procedure, because it unifies the frequencies, which are picked from the *HML*, in order to produce new harmonies. The usages of the parameters *HML*, *SP* and *SM* are explained in the next subsections with more details.

4.2 Composition of the Initial Harmonies and Harmony Memory Construction

Before the solution process of the OP starts, we need to initialize some of the above parameters. Initially, it is essential to determine the size of the harmony memory *HM*. In practice, we choose the number of the harmonies, which will be created during the operations of the SHHS algorithm and will be stored in the *HM*. Once the size of the *HM* is fixed, it remains the same until the end of the SHHS. Next, the algorithm has to be launched for a specific number of iterations. For this reason, the number of improvisations *NI* is settled. The amount of the *NI* cannot be neither quite few nor too much. The choice of the right *NI* depends on the sum of the music notes (nodes), that each instance of the OP has. Before the composition of the initial harmonies begins, we initialize the possibilities of *HMCR* and *PAR*, which will be used in the next phases of the SHHS.

The first phase of the SHHS referred to the creation of the harmonies (solutions). For their production we use the heuristic Nearest Insertion algorithm. The number of the produced harmonies is equal with the range of the *HM*. During the construction procedure all of them will be inserted one by one in the *HM*. To begin with, every harmony consists of *M* frequencies. Structurally each frequency includes two early nodes, the first 1 and the last *N*. Then, the SHHS picks a random node, which is different for every of the *M* frequencies and inserts it into them. So, they enter the Nearest Insertion process including three initial nodes. The remaining nodes are possible candidates for insertion into the frequencies. For each one of the remaining nodes, the algorithm searches and selects the best frequency to put into. In order to achieve this, the duration time of the frequency is calculated. If the duration time is

above the total time limit T_{max}, the candidate node cannot be inserted. Otherwise, the node is acceptable and enters to the right frequency. With the insertion of the candidate node, it changes the duration time and increases the total score of the selected frequency. The Nearest Insertion continues until all the harmonies are constructed. The first phase of the SHHS ends with the filling up of the *HM*.

4.3 The Improvisation Stage

The Improvisation stage is the second phase of the SHHS. In this one, the algorithm follows the logic and the basic steps of the Harmony Search. In particular, we refer to the usage of the harmony memory *HM* and to the variables of *HMCR* and *PAR*. Moreover, the parameter harmony memory list *HML* will be added and will, also, be used in the next phase of the algorithm. As we mentioned before, during this phase, the SHHS produces the 75% of the new harmonies. The production of them is achieved considering the already constructed harmonies from the *HM*. The new ones are created via some procedures according to the possibilities of *HMCR* and *PAR*. All the procedures are analyzed in the next paragraph.

This phase begins with the construction of the *HML*. This list will contain all the frequencies from the stored harmonies in the *HM*. The size of the *HML* will vary according to *M*. For instance, if the *HM* contains 100 harmonies and the *M* is 3, the *HML* will have 300 frequencies. Of course, the size of the list will be shortened as all the same between them (except of one) frequencies will be removed and only the different frequencies are stored. When the preparation of the *HML* is finished, the algorithm enters the Improvisation stage.

The Improvisation stage is based on the philosophy of the HS algorithm. It constructs new harmonies considering the *HML*. The frequencies in the list may be or may not be taken into account by the SHHS. That depends on the possibility of the memory consideration *HMCR*. A new harmony is fully created when the *M* frequencies are appropriately modified via three operations:

- Memory consideration list *HMCR*
- Pitch adjustment *PAR*—Local search
- Random frequency

If the first operation is chosen, the algorithm chooses one of the frequencies from the Memory consideration list. However, the candidate frequency does not immediately be inserted into the harmony. This happens as the insertion of this frequency may violate the restrictions of OP. And that is why the selected frequency contains nodes, which have already been used. Thus, the algorithm removes the not suitable nodes from the candidate frequency and calculates its new score and distance time in order to continue to the pitch adjustment process.

In this point, the SHHS checks the other possibility of *PAR* and takes into consideration the modified frequency from the previous operation. In pitch adjustment, a Local Search (LS) algorithm is applied. If the *PAR* is activated, the algorithm

uses the LS to improve the score of the frequency. Before the procedure of LS, it is essential to remove all the used nodes from the Memory consideration list operation. Excluding the unwanted nodes, the LS alters the selected frequency without violating the rules of OP. When the procedure is finished, the new frequency joins successfully the harmony. On the other hand, if the *PAR* is not activated, the selected frequency embodies immediately the harmony without any modification. The third and last operation of Random frequency does not take into account the *HML*. If this operation is activated, the SHHS produces randomly a frequency. This operation happens, when the first operation is not activated. Du-ring the creation of the frequency, the algorithm allows only the nodes, which have not been used in the harmony. Then, with the node insertion it increases the score and the total distance time of the random frequency. This operation ends when all of the remaining nodes are checked.

4.4 Similarity Process

In this section, we introduce a new technique called "Similarity Process". This method contributes to the increase of the total profit of the harmonies. Similarity process allows the SHHS algorithm to expand the area of the solutions seeking and creating more efficient harmonies than the improvisation stage. In order to achieve that, similarity process consists of two steps. The first is the searching part, in which the algorithm uses the *SP* parameter and the *SM* matrix to find suitable harmonies in order to compare them with the most profitable ones. In the second part, we apply the Nearest Insertion algorithm to fill up with valid nodes the harmonies from the first part. Before the start of the first part of similarity process, the SHHS constructs the similarity matrix (*SM*). The rows and columns of the *SM* represent the corresponding harmonies from the harmony memory list (*HML*). Moreover, each row or column contains the numbers of the similar nodes for all the candidate harmonies. The searching part begins by finding the most profitable harmonies. Then, we seek harmonies according to the *SP* by taking value from the *SM*. The *SP* designates the number of similar nodes, which must include a selected harmony compared to the best one. Considering the *SP*, the algorithm extracts the similar nodes from the selected harmony, reduces the distance time and calculates its new score in order to proceed to the next step. In the second part, the algorithm uses the Nearest Insertion method to enter appropriate nodes into the harmonies. With every node insertion we check the feasibility of the harmony through the total distance time. If the T_{max} is violated, then, the selected node is not suitable. Otherwise, the algorithm enters the node into the harmony and calculates its new score. This process continues until the set of the nodes is emptied.

4.5 *Finding the Best Harmonies and Update Harmony Memory*

At the end of the "similarity process", the SHHS reaches the final phase. Here, the algorithm seeks the harmonies with best score in order to be stored in the harmony memory *HM*. When the *HM* is updated, the algorithm proceeds to the next iteration. This happens until the total *NI* is reached. At the end of *NI*, the algorithm produces the final *HM*.

During the two previous phases, the SHHS algorithm constructs new harmonies considering the *HML*, which includes the frequencies from all harmonies in the *HM*. In order to proceed to the next iteration, the algorithm updates the *HM* only with the best harmonies. The aim in each of the *NI* iterations is to fill the *HM* with the old and new harmonies with the best score. For this reason, it keeps in the *HM* the 50% of the already stored harmonies and the other 50% is consisted of the best created ones. That is why the new harmonies are temporarily saved. Thus, the algorithm selects the 50% of the best and transfer them to the *HM*.

5 Computational Results

In this section, the computational results of the algorithm are presented and discussed in detail. As we would like to test the effectiveness of the algorithm, we have a number of sets of benchmark instances in the literature to be used for the comparisons. Thus, in the webpage http://www.mech.kuleuven.be/en/cib/op/ there is a number of benchmark instances for different variants of Orienteering Problem. We selected five sets of benchmark instances, three of them proposed in Tsiligirides (1984) and two of them proposed in Chao (1993). The three sets of Tsiligirides (1984) have 18, 11 and 20 instances, respectively, with number of nodes 32, 21 and 33, respectively, each one having different value in T_{max} in the limit constraint of the problem. For example, for the first set of instances of Tsiligirides the value of T_{max} varies between 5 and 85. The increase of the value allows more nodes to be visited in the best route. For each set of benchmark instances the value of T_{max} is presented in the corresponding Table. Finally, the last two sets of benchmark instances have 26 and 14 instances with 66 and 64 nodes, respectively.

In Tables 1, 2, and 3, the results of the proposed algorithm in the three set of instances of Tsiligirides are presented. More analytically, in the first column of each one of the Tables, the number of the instance and the value of the T_{max} are presented. In the second column, the best known solution (BKS) from the literature for each one of the instances are given, while in the fourth column the results of the proposed algorithm are presented. Finally, in the last column, the quality (deviation from the best known solution) is given. The quality of a solution is calculated using the following equation $RD = \frac{c_{BKS} - c_{SHHS}}{c_{BKS}}$ where c_{SHHS} denotes the score of the solution found by SHHS and c_{BKS} is the score of the best known solution (the smaller this

Table 1 Tsiligiride's Set 1

Instances	SHHS	BKS	RD (%)
t1-05	10	10	0
t1-10	15	15	0
t1-15	45	45	0
t1-20	65	65	0
t1-25	90	90	0
t1-30	110	110	0
t1-35	130	135	3.70
t1-40	150	155	3.22
t1-45	175	175	0
t1-50	190	190	0
t1-55	200	205	2.43
t1-60	225	225	0
t1-65	240	240	0
t1-70	260	260	0
t1-73	265	265	0
t1-75	270	270	0
t1-80	280	280	0
t1-85	285	285	0

Table 2 Tsiligiride's Set 2

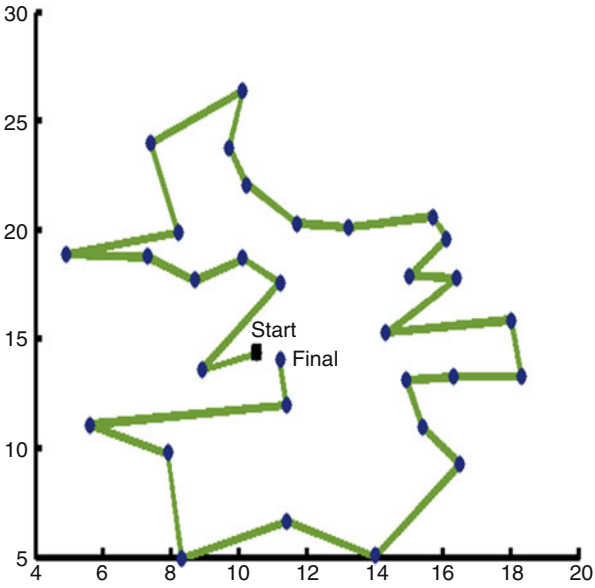
Instances	SHHS	BKS	RD (%)
t2-15	120	120	0
t2-20	200	200	0
t2-23	210	210	0
t2-25	230	230	0
t2-27	265	265	0
t2-30	300	300	0
t2-32	320	320	0
t2-38	360	360	0
t2-40	395	395	0
t2-45	450	450	0

value, the better the performance of the algorithm). As it can be seen in the first set of Tsiligirides instances the proposed algorithm finds the best known solution in 12 out of 18 instances and the average quality is equal to 1.087%. In the second set, the proposed algorithm finds the best known solution in 9 out of 10 instances and the average quality is equal to 0.138%. Finally, in the third set of instances, the proposed algorithm finds the best known solution in 10 out of 20 instances and the average quality is equal to 1.403%. In total in the three set of instances proposed by Tsiligirides the proposed algorithm finds the best known solution in 31 out of 48 instances (64.58%) with average quality equal to 1.021%. The results are very satisfactory as they prove the effectiveness of the algorithm in the randomly distributed in the spaces instances.

Table 3 Tsiligiride's Set 3

Instances	SHHS	BKS	RD (%)
t3-15	170	170	0
t3-20	200	200	0
t3-25	260	260	0
t3-30	320	320	0
t3-35	390	390	0
t3-40	420	430	2.32
t3-45	470	490	4.08
t3-50	500	520	3.84
t3-55	540	550	1.81
t3-60	580	580	0
t3-65	610	610	0
t3-70	630	640	1.56
t3-75	660	670	1.49
t3-80	700	710	1.40
t3-85	730	740	1.35
t3-90	760	770	1.29
t3-95	790	790	0
t3-100	800	800	0
t3-105	800	800	0
t3-110	800	800	0

Fig. 1 Total score = 285



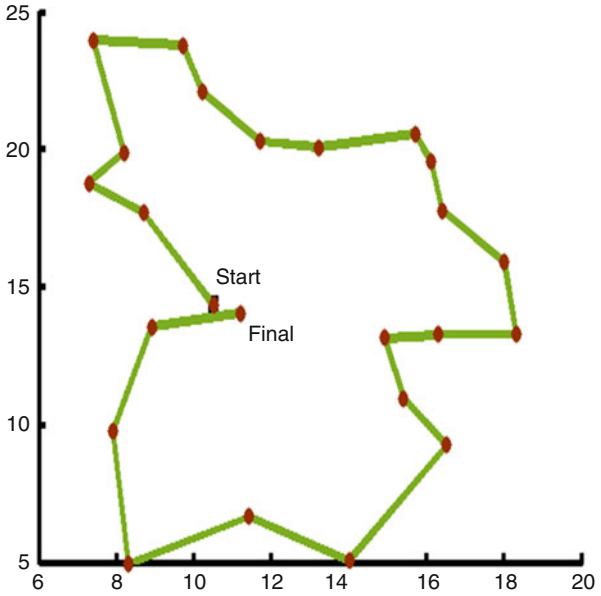


Fig. 2 Total score = 225

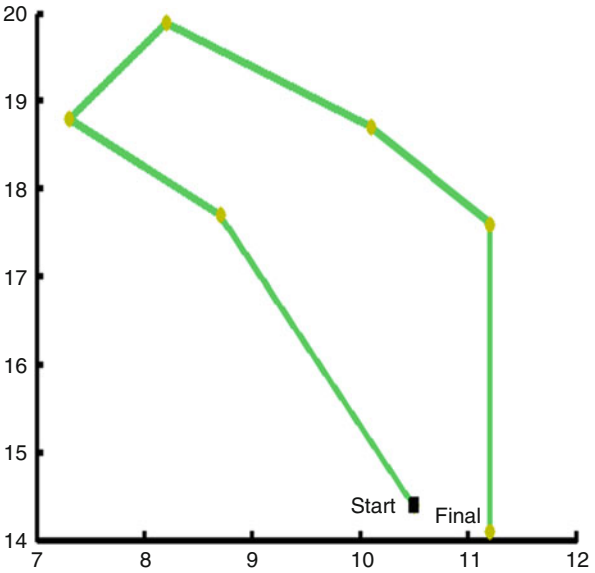


Fig. 3 Total score = 45

In Figs. 1, 2, 3, 4, 5, 6, 7, 8, and 9, three different instances from each one of the three sets of benchmark instances proposed by Tsiligirides are given. More precisely, we select to present one instance for each one of the data sets with the whole set of nodes included in the solution (Figs. 1, 4, and 7) and other instances with less number of nodes for each set of benchmark instances (Figs. 2, 3, 5, 6, 8, and 9). In all

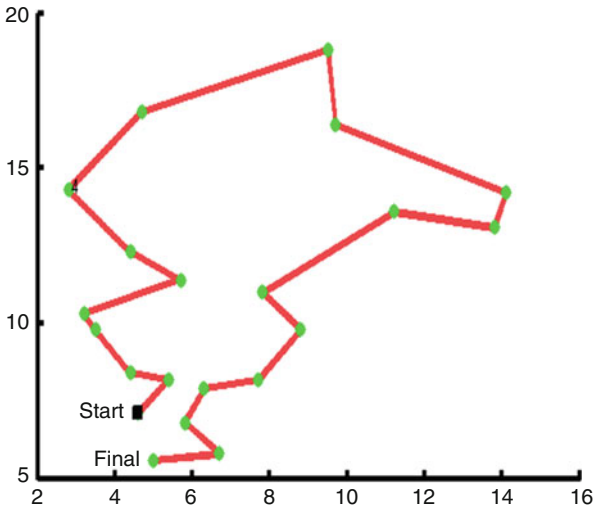


Fig. 4 Total score = 450

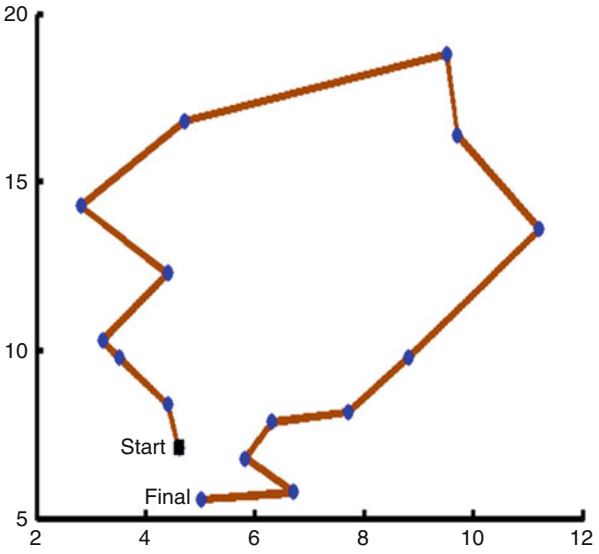


Fig. 5 Total score = 320

nine figures there are not any crossings between the nodes and, thus, the solutions (even when they are not equal to the best known solutions) are very effective.

In Tables 4 and 5, the results of the proposed algorithm for the two last set of instances are presented. The difference of these two sets of instances from the other three is that the nodes are in a clustered form, in a diamond shape or in a square shape. The columns of the Tables present the number of instance and the T_{max}, the best values from the literature, the results of the proposed algorithm and the quality of the instances as in the previous three Tables. As it can be seen in the Diamond shaped set of instances, the proposed algorithm finds the best known solution in

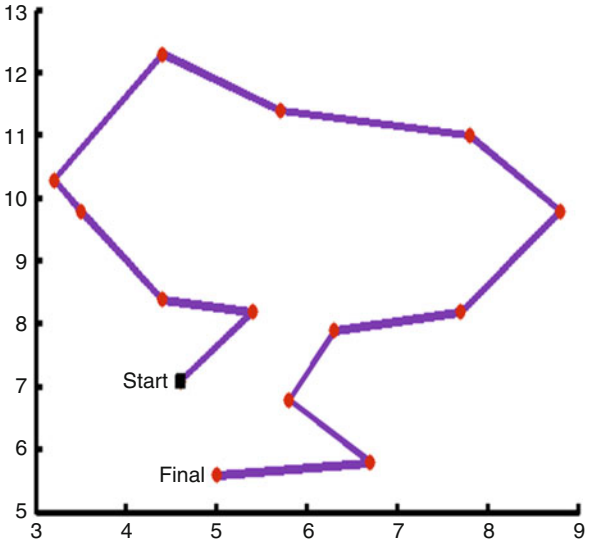


Fig. 6 Total score = 200

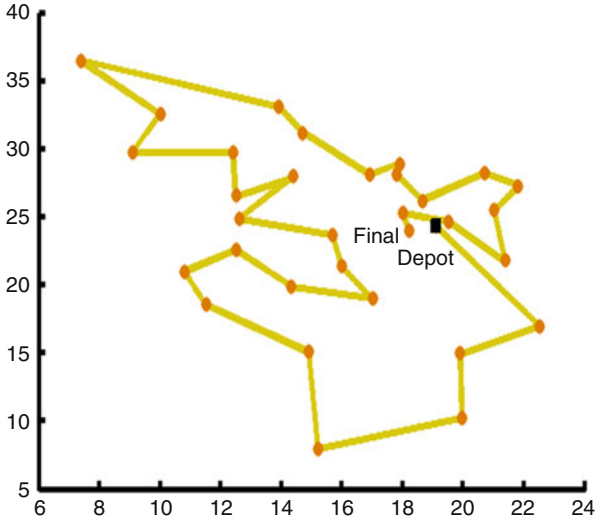


Fig. 7 Total score = 800

4 out of 13 instances and the average quality is equal to 1.919%. In the Square shaped set of instances, the proposed algorithm finds the best known solution in 6 out of 24 instances and the average quality is equal to 2.842%. In total in the two sets of instances, the proposed algorithm finds the best known solution in 10 out of 37 instances (27.02%) with average quality equal to 2.518%. The results are worst than the results of the first three instances, however, they are still very effective, which proves that the proposed algorithm performs better when the nodes are randomly distributed in the space. Finally, the proposed algorithm in 85 instances

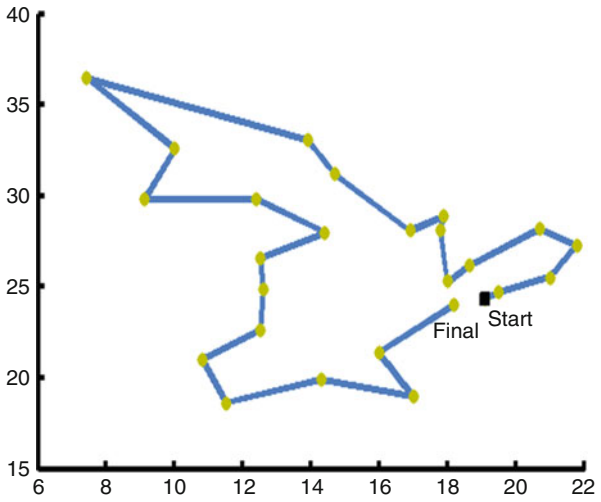


Fig. 8 Total score = 610

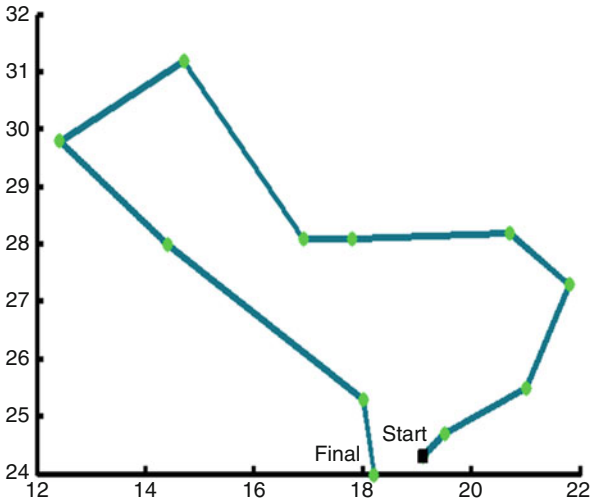


Fig. 9 Total score = 260

found the best known solution in 41 of them (48.23%) and the average quality of the instances is equal to 1.672% which ranks the proposed algorithm as one of the most effective algorithms for solving the Orienteering Problem.

In Figs. 10, 11, 12, 13, 14, and 15 three different instances for each of the two clustered sets of benchmark instances are given. As in the previous figures, we select to present one instance for each one of the data sets with the whole set of nodes included in the solution (Figs. 10 and 13) and other instances with less number of nodes for each set of benchmark instances (Figs. 11, 12, 14, and 15). In all six figures there are not any crossings between the nodes and, thus, the solutions (even when they are not equal to the best known solutions) are very effective.

Table 4 Diamond-shaped set tests

Instances	SHHS	BKS	RD (%)
64-1-15	96	96	0
64-1-20	294	294	0
64-1-25	390	390	0
64-1-30	468	474	1.26
64-1-35	570	576	1.04
64-1-40	714	714	0
64-1-45	798	816	2.20
64-1-50	876	894	2.01
64-1-55	954	978	2.45
64-1-60	1020	1062	3.95
64-1-65	1086	1116	2.68
64-1-70	1158	1188	2.52
64-1-75	1200	1230	2.43
64-1-80	1260	1278	1.40

Table 5 Square-shaped test sets

Instances	SHHS	BKS	RD (%)
66-1-10	40	40	0
66-1-15	120	120	0
66-1-20	205	205	0
66-1-25	290	290	0
66-1-30	400	400	0
66-1-35	465	465	0
66-1-40	545	575	5.21
66-1-45	645	650	0.76
66-1-50	710	730	2.73
66-1-55	825	825	0
66-1-60	885	915	3.27
66-1-65	965	980	1.53
66-1-70	1040	1070	2.80
66-1-75	1120	1140	1.75
66-1-80	1155	1215	4.93
66-1-85	1235	1260	2.77
66-1-90	1300	1340	2.98
66-1-95	1325	1385	4.33
66-1-100	1380	1445	4.49
66-1-105	1425	1515	5.94
66-1-110	1500	1545	2.91
66-1-115	1555	1590	2.20
66-1-120	1615	1635	1.22
66-1-125	1655	1665	0.60
66-1-130	1680	1680	0

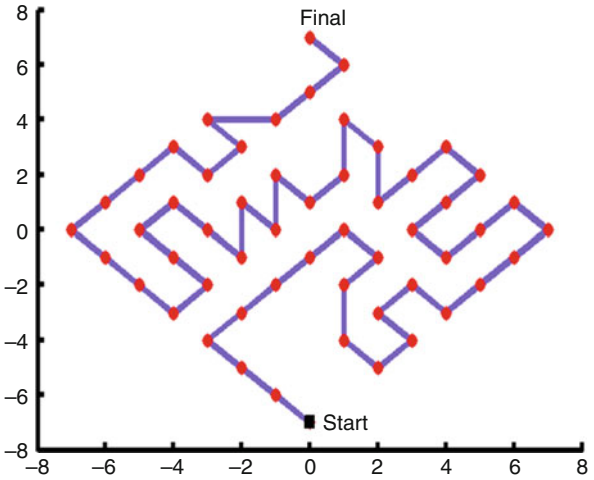


Fig. 10 Total score = 1248

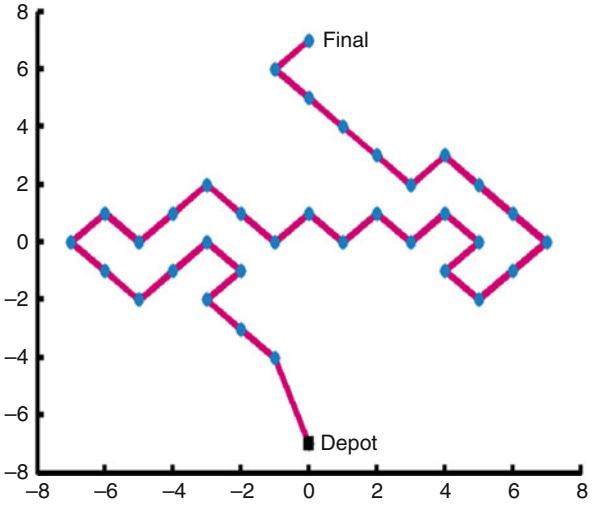


Fig. 11 Total score = 87

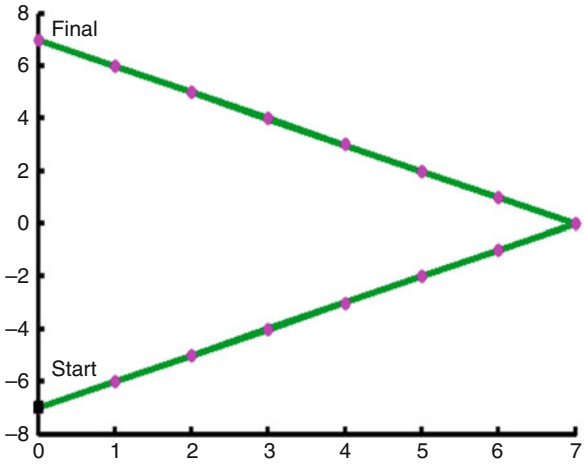


Fig. 12 Total score = 294

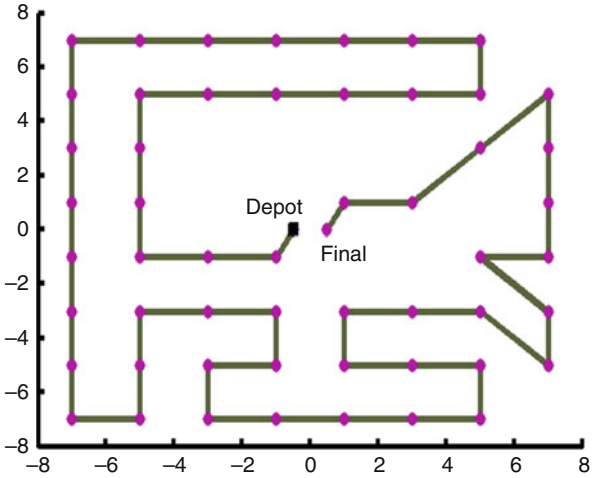


Fig. 13 Total Score = 1680

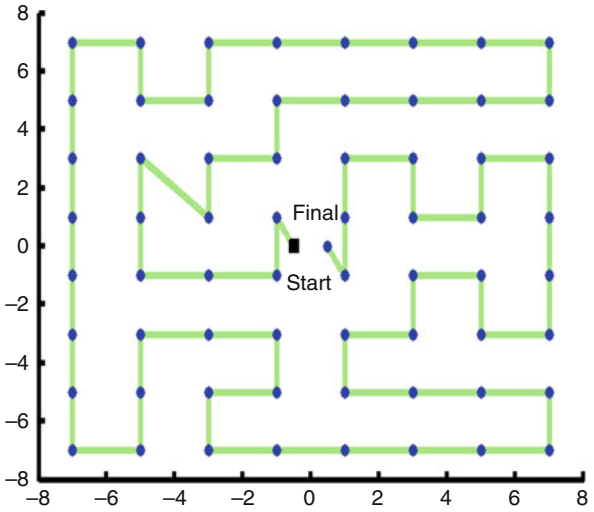


Fig. 14 Total score = 1485

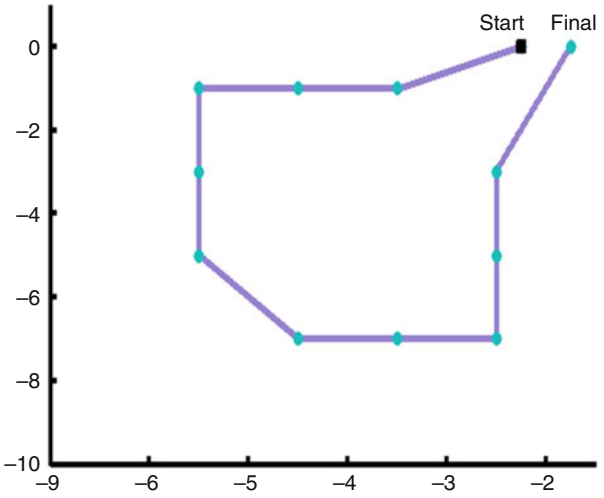


Fig. 15 Total score = 290

6 Conclusions and Future Research

In this paper, an algorithm for the solution of the Orienteering Problem is presented. The algorithm, denoted as Similarity Hybrid Harmony Search algorithm, is a variant suitable for routing problems, of the well known nature inspired algorithm Harmony Search. The algorithm was tested in classic sets of benchmark instances for the Orienteering Problem and in half of the cases it found the best known solutions. The future steps of our research will be to test the algorithm in more difficult Vehicle Routing Problems with Profits, like the Team Orienteering Problem or the Capacitated Team Orienteering Problem.

References

Archetti, C., Speranza, M. G., & Vigo, D. (2014). Vehicle routing problems with profits. In P. Toth & D. Vigo (Eds.), *Vehicle routing: Problems, methods, and applications* (MOS-SIAM Series on Optimization) (pp. 273–298). Philadelphia, PA: SIAM.

Chao, I. (1993). *Algorithms and solutions to multi-level vehicle routing problems*. Ph.D. Dissertation, Applied Mathematics Program, University of Maryland, College Park.

Chao, I. M., Golden, B. L., & Wasil, E. (1996a). A fast and effective heuristic for the orienteering problem. *European Journal of Operational Research, 88*(3), 475–489.

Chao, I. M., Golden, B. L., & Wasil, E. (1996b). The team orienteering problem. *European Journal of Operational Research, 88*(3), 464–474.

Erdal, F., Dogan, E., & Saka, M. P. (2011). Optimum design of cellular beams using harmony search and particle swarm optimizers. *Journal of Constructional Steel Research, 67*(2), 237–247.

Fourie, J., Mills, S., & Green, R. (2010). Harmony filter: A robust visual tracking system using the improved harmony search algorithm. *Image and Vision Computing, 28*(12), 1702–1716.

Geem, Z. W. (2005). School bus routing using harmony search. In *GECCO 2005*. Washington, DC: ACM.

Geem, Z. W., Kim, J. H., & Loganathan, G. V. (2001). A new heuristic optimization algorithm: Harmony search. *Simulation, 76*(2), 60–68.

Geem, Z. W., Tseng, C.-L., & Park, Y. (2005). Harmony search for generalized orienteering problem: Best touring in China. In L. Wang, K. Chen, & Y. Ong (Eds.), *ICNC 2005*, LNCS 3612 (pp. 741–750). Berlin: Springer.

Golden, B., Levy, L., & Vohra, R. (1987). The orienteering problem. *Naval Research Logistics, 34*, 307–318.

Kaveh, A., & Talataha, S. (2009). Particle swarm optimizer, ant colony strategy and harmony search scheme hybridized for optimization of truss structures. *Computers and Structures, 87* (5–6), 267–283.

Lee, K. S., & Geem, Z. W. (2004). A new meta-heuristic algorithm for continuous engineering optimization: Harmony search theory and practice. *Computer Methods in Applied Mechanics and Engineering, 194*(36–38), 3902–3933.

Mahdavi, M., Fesanghary, M., & Damangir, E. (2007). An improved harmony search algorithm for solving optimization problems. *Applied Mathematics and Computation, 188*(2), 1567–1579.

Marinakis, Y., Politis, M., Marinaki, M., & Matsatsinis, N. (2015). A memetic-GRASP algorithm for the solution of the orienteering problem. In H. Le Thi, T. Pham Dinh, & N. Nguyen (Eds.), *Modelling, computation and optimization in information systems and management sciences* (pp. 105–116). Cham: Springer.

Miguel, L. F. F., & Miguel, L. F. F. (2012). Shape and size optimization of truss structures considering dynamic constraints through modern metaheuristic algorithms. *Expert Systems with Applications, 39*(10), 9458–9467.

Montemanni, R., & Gambardella, L. (2009). Ant colony system for team orienteering problems with time windows. *Foundations of Computing and Decision Sciences, 34*(4), 287–306.

Omran, M. G. H., & Mahdavi, M. (2008). Global-best harmony search. *Applied Mathematics and Computation, 198*(2), 643–656.

Pan, Q.-K., Suganthan, P. N., Liang, J. J., & Tasgetiren, M. F. (2011). A local-best harmony search algorithm with dynamic sub-harmony memories for lot-streaming flow shop scheduling problem. *Expert Systems with Applications, 38*(4), 3252–3259.

Pan, Q.-K., Suganthan, P. N., Tasgetiren, M. F., & Liang, J. J. (2010). A self-adaptive global best harmony search algorithm for continuous optimization problems. *Applied Mathematics and Computation, 216*(3), 830–848.

Pichpibul, T., & Kawtummachai, R. (2013). Modified harmony search algorithm for the capacitated vehicle routing problem. In *Proceedings of the International Multi Conference of Engineers and Computer Scientists* (Vol. II), IMECS 2013, Hong Kong.

Righini, G., & Salani, M. (2009). Decremental state space relaxation strategies and initialization heuristics for solving the orienteering problem with time windows with dynamic programming. *Computers and Operations Research, 36*(4), 1191–1203.

Ser, J. D., Matinmikko, M., Gil-Lopez, S., & Mustonen, M. (2012). Centralized and distributed spectrum channel assignment in cognitive wireless networks: A harmony search approach. *Applied Soft Computing, 12*(2), 921–930.

Taleizadeh, A. A., Niaki, S. T. A., & Barzinpour, F. (2011). Multiple-buyer multiple-vendor multi-product multi-constraint supply chain problem with stochastic demand and variable lead-time: A harmony search algorithm. *Applied Mathematics and Computation, 217*(22), 9234–9253.

Tang, H., & Miller-Hooks, E. (2005). A TABU heuristic for the team orienteering problem. *Computers and Industrial Engineering, 32*(6), 1379–1407.

Tsiligirides, T. (1984). Heuristic methods applied to orienteering. *Journal of Operational Research Society, 35*(9), 797–809.

Vansteenwegen, P., Souffriau, W., & Van Oudheusden, D. (2011). The orienteering problem: A survey. *European Journal of Operational Research, 209*(1), 1–10.