Advances in Intelligent Systems and Computing 572

Jacek Żak
Yuval Hadas
Riccardo Rossi *Editors*

Advanced Concepts, Methodologies and Technologies for Transportation and Logistics

Springer

Advances in Intelligent Systems and Computing

Volume 572

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

About this Series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within "Advances in Intelligent Systems and Computing" are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Advisory Board

Chairman

Nikhil R. Pal, Indian Statistical Institute, Kolkata, India
e-mail: nikhil@isical.ac.in

Members

Rafael Bello Perez, Universidad Central "Marta Abreu" de Las Villas, Santa Clara, Cuba
e-mail: rbellop@uclv.edu.cu

Emilio S. Corchado, University of Salamanca, Salamanca, Spain
e-mail: escorchado@usal.es

Hani Hagras, University of Essex, Colchester, UK
e-mail: hani@essex.ac.uk

László T. Kóczy, Széchenyi István University, Győr, Hungary
e-mail: koczy@sze.hu

Vladik Kreinovich, University of Texas at El Paso, El Paso, USA
e-mail: vladik@utep.edu

Chin-Teng Lin, National Chiao Tung University, Hsinchu, Taiwan
e-mail: ctlin@mail.nctu.edu.tw

Jie Lu, University of Technology, Sydney, Australia
e-mail: Jie.Lu@uts.edu.au

Patricia Melin, Tijuana Institute of Technology, Tijuana, Mexico
e-mail: epmelin@hafsamx.org

Nadia Nedjah, State University of Rio de Janeiro, Rio de Janeiro, Brazil
e-mail: nadia@eng.uerj.br

Ngoc Thanh Nguyen, Wroclaw University of Technology, Wroclaw, Poland
e-mail: Ngoc-Thanh.Nguyen@pwr.edu.pl

Jun Wang, The Chinese University of Hong Kong, Shatin, Hong Kong
e-mail: jwang@mae.cuhk.edu.hk

More information about this series at http://www.springer.com/series/11156

Jacek Żak · Yuval Hadas
Riccardo Rossi
Editors

Advanced Concepts, Methodologies and Technologies for Transportation and Logistics

Springer

Editors
Jacek Żak
Faculty of Engineering Management
Poznań University of Technology
Poznań
Poland

Yuval Hadas
Department of Management, Head of Supply Chain Management and Logistics Studies;
Faculty of Social Sciences
Bar-Ilan University
Ramat Gan
Israel

Riccardo Rossi
Department of Civil, Environmental and Architectural Engineering
University of Padova
Padova
Italy

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-57104-1 ISBN 978-3-319-57105-8 (eBook)
DOI 10.1007/978-3-319-57105-8

Library of Congress Control Number: 2017937910

© Springer International Publishing AG 2018
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

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

Preface

The presented book titled: "Advanced concepts, methodologies and technologies for transportation and logistics" is a collection of original works/articles produced by the members of the Euro Working Group on Transportation (EWGT) in the last several years (2015–2017). The scope of this book falls within the generic theme of the Springer Series "*Advances in Intelligent Systems and Computing*". The chapters included in this book present the results of various research projects carried out by the members of the EWGT, the extended versions of their conference presentations (usually organized under the patronage of the Association of European Operational Research Societies), and original findings produced by the authors. This book is a sample of EWGT research activities and covers the state of the art in quantitative-oriented transportation/logistics research.

The Euro Working Group on Transportation (EWGT) is a specialized chapter of EURO—the Association of European Operational Research Societies. The EWGT activities are focused on the development and application of operations research techniques, methods, and algorithms in transportation. In recent years, the scope of the group has extended toward logistics. The EWGT gathers transportation and logistics professionals: researchers, analysts, consultants, and practitioners from all the World (primarily from Europe), who are interested in mathematical modeling, quantitative analysis, and optimization of transportation/logistics processes and systems. The EWGT members are involved in the analysis of transportation – logistics phenomena and behavior, developing solution procedures for complex transportation and logistics decision situations and assisting decision makers (DM-s) who face difficult and unstructured transportation/logistics – oriented decision problems. The interests of EWGT cover both methodological and technological advances at the boundary of disciplines such as transportation and logistics (T&L), computer science (CS), information technology (IT), operations research (OR), decision aiding (DA).

The EWGT was founded in July 1991 in Cetraro (Italy), and the 1st Meeting of the Group was held in Landshut (Germany) in October 1992. Many activities

have been performed since then, including organization of international seminars and workshops, publishing books and conference proceedings (e.g., Springer books and Elsevier series Transportation Research Proceedia), launching common research projects at the European level, organization of Summer and Winter Euro Institutes for young researchers, arranging transportation and logistics streams at the Euro Conferences, organizing thematic Mini-Euro Conferences, and, last but not least, holding 20 International Meetings of the Group in different European cities, such as Paris (twice), Barcelona, Newcastle upon Tyne, Goteborg, Helsinki, Budva, Rome, Bari (twice), Poznan (twice), Ischia, Padova, Porto, Seville, Delft, Istanbul, and Budapest (next meeting).

The Guest Editors of this book are strongly involved in the activities of EWGT. Prof. Riccardo Rossi is the Group Coordinator (since 2011) and the Chair of one of the previous EWGT Meetings, Prof. Jacek Żak is the member of the EWGT Scientific Committee and the Chair of two previous EWGT Meetings, and Prof. Yuval Hadas is an active participant of the group activities for the last 5 years.

The presented book is a follow-up of a previously published (2014) Springer Monograph on "Computer based modeling and optimization in transportation" edited by Prof. Jorge Freire de Sousa and Prof. Riccardo Rossi. The latter covered the research output of EWGT for the years 2012–2014. Its success measured by the number of the downloads (more than 33 000) has indicated the need for the presentation of new works and findings of EWGT community. Thus, the Guest Editors have launched this project to demonstrate the current level of research carried out by the members of EWGT.

This book is composed of 23 chapters split into four major thematic streams/sections. Due to the complexity and diversity of the topic considered, the Guest Editors assigned specific persons to coordinate each of the above-mentioned streams/sections. Each stream covers a specific, important area of transportation and logistics research, including

- Section 1: Multiple Criteria Analysis in Transportation and Logistics—7 chapters, coordinated by Prof. Jacek Żak (Poznań University of Technology, Poland).
- Section 2: Urban Transportation and City Logistics—7 chapters, coordinated by Prof. Yuval Hadas (Bar-Ilan University, Israel).
- Section 3: Road Safety—4 chapters, coordinated by Prof. Massimiliano Gastaldi (University of Padova, Italy)
- Section 4: Artificial Intelligence and Soft Computing in Transportation and Logistics—5 chapters, coordinated by Prof. Michele Ottomanelli (Technical University of Bari, Italy).

Each stream/section is initiated by an introductory chapter prepared by its coordinator. It presents an overview of the considered theme and shortly describes all its components, i.e., the contributions (chapters) included in the stream. It is followed

by 4–7 chapters, each of which presents an original output generated by the authors (contributors), falling within the thematic scope of the stream/section.

Roughly, 50 articles have been considered for publication in this book. All of them have gone through a stringent, peer-reviewed procedure and have been finally recommended for publication by the referees and Guest Editors. The finally accepted 23 chapters present a variety of research topics and problems, transportation modes, and proposed approaches:

- Network design problems; traffic control strategies; transportation and logistics projects analysis; assessment of supply chain management models; selection of suppliers; analysis of drivers' behavior; terminal design problems.
- Road, railway, air, and water transportation considered both in passenger and freight environment.
- Exact (e.g., integer programming) and approximate algorithms (e.g., metaheuristics: genetic algorithms; evolutionary approaches); single and multiple criteria methods (e.g., Electre, AHP); deterministic and non-deterministic methods (e.g., fuzzy algorithms).

The Guest Editors are very thankful to all 63 contributors, representing 11 countries and 3 continents (Europe, Asia, and North America), who prepared particular chapters of this book. Their interesting research has produced an original outcome of this book and has created an added value for transportation and logistics research. The Guest Editors realize that the picture presented in each stream is not complete. At the same time, they would like to indicate that the works collected in this book cover a representative selection of subjects significant for transportation and logistics research. This book is focused on advanced methodologies and technologies applied in transportation and logistics. In the Guest Editors' opinion, the presented selection of topics should generate an interest of a wide spectrum of readers and at the same time give them a good theoretical background and inspiration for further in-depth investigation in selected topics.

The Guest Editors would like to express their sincere gratitude to all four coordinators of thematic streams for their assistance in shaping this book and to all 40 anonymous referees (2–3 per chapter) representing 13 countries, who responded willingly to our requests for reviewing the chapters. The refereeing process required from the referees comprehensive and interdisciplinary knowledge, diverse expertise and fair justification of the submitted contributions. Their efforts are highly appreciated. Special thanks are addressed to our partner—editors at Springer who are responsible for book typesetting and production, in particular to Dr. Thomas Ditzinger representing in Springer the area of Applied Sciences and Engineering, who contributed substantially to the final success of this book.

This book is dedicated to Professor Shinya Kikuchi (Virginia Polytechnic, USA) and Professor Matthew G. Karlaftis (National Technical University of Athens, Greece) who passed away unexpectedly in the recent years and left emptiness behind. Shinya and Mathew were great friends for the EWGT community and

contributed enormously to our development. They still inspire many of us. We are proud that Shinya and Matthew carried out common research with us and participated in our activities. The EWGT community will always miss them.

Jacek Żak
Yuval Hadas
Riccardo Rossi

Contents

Multiple Criteria Analysis in Transportation and Logistics

The Methodology of Multiple Criteria Decision Making/Aiding in Transportation 9
Jacek Żak

The Multi-Actor Multi-Criteria Analysis (MAMCA) Tool: Methodological Adaptations and Visualizations 39
Sheida Hadavi, Cathy Macharis, and Koen Van Raemdonck

Dealing with the Complexity of Stakeholder Interaction in Participatory Transport Planning 54
Michela Le Pira, Giuseppe Inturri, Matteo Ignaccolo, and Alessandro Pluchino

A Methodology of Redesigning and Evaluating Medium-Sized Public Transportation Systems 73
Jacek Żak and Maja Kiba-Janiak

A Framework for Solving Real-Time Multi-objective VRP 103
Oren E. Nahum and Yuval Hadas

Multiple Criteria Evaluation of Suppliers in Different Industries - Comparative Analysis of Three Case Studies 121
Jacek Żak and Barbara Galińska

Aircraft Type Selection Problem: Application of Different MCDM Methods 156
Slavica Dožić and Milica Kalić

Urban Transportation and City Logistics

Transit Network Design Problem: An Expansion of the Route Generation Algorithm 183
Iran Khanzad, Seyedehsan Seyedabrishami, Mohsen Nazemi, and Amirali Zarrinmehr

The Railway Network Design, Line Planning and Capacity Problem: An Adaptive Large Neighborhood Search Metaheuristic 198
David Canca, Alicia De-Los-Santos, Gilbert Laporte, and Juan A. Mesa

Diagnostic of the Balance and Equity of Public Transport for Tourists and Inhabitants 220
Maurici Ruiz and Joana Maria Seguí-Pons

On the Design of Leisure Devoted Cycling Networks 237
Alessandro Giovannini, Federico Malucelli, and Maddalena Nonato

The Influence of Transportation Service Level on a Municipal Service Center's Costs: A Numerical Study Based on Supply Chain Management Models 256
Matan Shnaiderman

A Stochastic Version of the Strategy-Based Congested Transit Assignment Model and a Technique by Smoothing Approximations 272
Esteve Codina and Francisca Rosell

Evaluation of CO_2 Emission Reduction from Vehicles by Information Provision Using Driving Simulator 290
Yukimasa Matsumoto and Shogo Ishiguro

Road Safety

Power and Exponential Functions Relating Accidents to Traffic and Rain. Calibration on a French Network 309
Maurice Aron, Romain Billot, Neïla Bhouri, Nour-Eddin El Faouzi, and Régine Seidowsky

Development of Coordinated Ramp-Metering Based on Multi-objective Nonlinear Optimization Functions: Traffic and Safety 334
Habib Haj-Salem, Nadir Farhi, Jean Patrique Lebacque, and Neïla Bhouri

Drivers' Behavior at Bicycle Crossroads 355
Francesco Bella and Manuel Silvestri

Automatic Incident Detection on Freeway Ramp Junctions. A Fuzzy Logic-Based System Using Loop Detector Data 370
Riccardo Rossi, Massimiliano Gastaldi, and Gregorio Gecchele

Artificial Intelligence and Soft Computing in Transportation and Logistics

Optimizing Airport Gate Assignments Through a Hybrid Metaheuristic Approach 389
Mario Marinelli, Gianvito Palmisano, Mauro Dell'Orco, and Michele Ottomanelli

Shipping Liner Company Stowage Plans: An Optimization Approach 405
Daniela Ambrosino, Massimo Paolucci, and Anna Sciomachen

A Multi-depot Periodic Vehicle Routing Model for Petrol Station Replenishment 421
Pasquale Carotenuto, Stefano Giordani, Simone Massari, and Fabrizio Vagaggini

A Multi-criteria Intelligent Control for Traffic Lights Using Reinforcement Learning 438
Junchen Jin and Xiaoliang Ma

Analysis of Imprecise Perception in Route Choice Considering Fuzzy Costs 452
Mario Binetti, Marco De Mitri, and Michele Ottomanelli

Author Index 469

Multiple Criteria Analysis in Transportation and Logistics

Multiple Criteria Analysis in Transportation and Logistics

Jacek Żak

Poznań University of Technology, ul. Piotrowo 3, Poznań, Poland
jacek.zak@put.poznan.pl

The first thematic stream of the Academic Monograph "Advanced concepts, methodologies and technologies for transportation and logistics" edited by Jacek Żak, Yuval Hadas and Riccardo Rossi in the Springer Series "Advances in Intelligent Systems and Computing" is the one focused on "Multiple Criteria Analysis in Transportation and Logistics". This section shows methodological principles and theoretical background of Multiple Criteria Analysis (MCA), often called Multiple Criteria Decision Making/Aiding (MCDM/A) and its application in the fields of Transportation and Logistics. It covers 7 chapters that deal with various multiple criteria, transportation – logistics oriented decision problems and present a variety of computational methods and approaches of solving these problems. The contributions presented in this stream range in terms of the problematics and transportation modes considered.

MCA or MCDM/A is a field of study which aims at giving the Decision Maker (DM) different computational methods and computer-based tools that enable him/her to advance in solving, the so called, multiple criteria decision problems, i.e.: such complex decision situations in which several – often contradictory – points of view (criteria) must be taken into account [1, 4]. These complex decision situations can be classified as multiple criteria [1, 4, 5]: choice (optimization) problems; classification problems and ranking problems. In the areas of transportation and logistics the examples of these three categories of decisions problems are as follows:

- Choice problems: design of the structure of the transportation - distribution network (finding the optimal locations of warehouses and depots, definition of optimal routes); finding the optimal size of the fleet or crew for selected transportation processes;
- Classification problems: allocation of drivers into predefined classes based on their qualifications; categorization of distribution centers (warehouses) based on their characteristics;
- Ranking problems: ordering/ranking from the best to the worst the trams for the public transportation system or logistics service providers delivering transportation/ logistics services to a customer.

The origins of MCA or MCDM/A can be found in the theoretical works of a mathematician Cantor and economists Edgeworth and Pareto in the second half of the 19th century. These initial works were followed by landmark research carried out in the first half of the 20th century by: Ramsey, van Neumann and Morgenstern, Nash, Samuelson and Simon [1]. The dynamic development of MCA/MCDM(A) and its split into two major schools: the European one (based on the outranking approach) [1, 4] and the

American one (based on Multiattribute Utility Theory) [1, 4] dates back to the 1950s and 1960s of the 20th century.

From a methodological point of view MCA or MCDM/A originates in Operations Research (OR) and its methodology is similar to the profile and overall approach applied in Operations Research [1]. For these reasons it is so important for the research and academic activities of the Euro Working Group on Transportation. MCA/MCDM(A), similarly to OR, has a quantitative and interdisciplinary character, uses scientific methods and rigid, systematic way of thinking and provides a consistent set of rules that support the DM in the process of solving complex decision problems. As opposed to OR, MCA/MCDM(A) uses several characteristics (criteria) while evaluating and optimizing different solutions (variants), applies multiple criteria methods/algorithms and does not yield "objectively best" solutions. Instead of the single objective optimal solution the MCA/MCDM(A) methodology accepts the "trade – offs" between criteria, the resulting idea of a family of Pareto optimal (non-dominated) solutions and the concept of a finally accepted satisfactory or compromise solution.

The methodology of MCA or MCDM(A) has developed a variety of multiple criteria methods that are used to solve the above defined multiple criteria decision problems. In general, they can be divided into 3 major families [1, 4, 5]:

- The methods of American inspiration based on the utility function;
- The outranking methods of the European origin;
- Interactive methods.

The methodology of MCA or MCDM/A has a universal character and can be applied in different fields. In transportation and logistics it helps the DM in analyzing, modeling, structuring and solving complex, unstructured transportation/logistics decision situations, in assessing innovative transportation/logistics projects, concepts and solutions, in balancing trade-offs and conflicting interests associated with the operations of certain transportation and logistics processes and systems and in searching for the most desired, compromise transportation/logistics decisions. While applying MCA/MCDM(A) for solving transportation/logistics oriented decision problems two major attributes/components of these problem must be defined, i.e. [1, 4, 5]: a set of variants (actions, solutions) and a consistent family of criteria. A set of actions can be defined in a different way, as: constant, evolving or post-experimental. It can be formulated directly (as a list) or indirectly (in the form of rules and formulas). The examples of different sets of transportation/logistics – oriented variants are as follows: a list of trams considered for implementation in a public transportation system or a set of constraints defining the space of feasible solutions for a vehicle routing problem. A consistent family of criteria should satisfy the following conditions: exhaustiveness (completeness) of evaluation; cohesiveness with the global preferences of the DM; non-redundancy (criteria should not overlap). Due to psychological reasons and limited human perception, in the majority of transportation/logistics – oriented decision problems, it is advised to use 7 +/− 2 criteria for the evaluation of variants. In many cases when transportation/logistics systems are redesigned and/or transportation and logistics projects are implemented a consistent family of criteria can be composed of the following characteristics: financial efficiency or profitability; assets (infrastructure) utilization; safety; reliability;

comfort (of travel/delivery); environmental friendliness; performance characteristics (speed; travel time; delays). This issue is thoroughly discussed in this stream.

The MCA/MCDM(A) methodology clearly identifies major participants of the decision making/aiding process, such as: the decision maker (DM), the analyst and the interveners (stakeholders) and describes their roles in this process. It also emphasizes the critical role of computerized Decision Support Systems that assist the DM while solving multiple criteria decision problems.

As revealed in the chapters presented in this stream the application of MCA or MCDM/A in transportation and logistics is natural and obvious. It results from the following facts:

- The majority of transportation/logistics decision problems has a multiple criteria character.
- There are many stakeholders in transportation/logistics processes and systems that express various interests and a compromise must be found to balance them.
- The complexity of transportation and logistics activities and phenomena requires that in their analysis a variety of measures and characteristics must be used and many aspects must be considered.

For the above mentioned reasons MCA or MCDM/A methodology has gained more and more appreciation and popularity among transportation researchers, recently. In the last 20 years several successful applications of this methodology in transportation planning, analysis and control were reported. These include, among others, the works of: Satty [3], Macharis et al. [2], Żak [5]. A comprehensive review of literature concerning the application of MCA or MCDM/A methodology in transportation and logistics can be found in several chapters of this book/stream. It is worth mentioning that J. Żak, member of the EWGT Scientific Committee, is very active in this field.

As mentioned above the stream on MCA in Transportation and Logistics includes 7 chapters. Altogether they present the state of the art in MCA/MCDM(A) in Transportation and Logistics. They cover such issues as:

- The methodological background of MCA/MCDM(A).
- Comprehensive review of the literature concerning MCA/MCDM(A) and its application in transportation and logistics.
- The classification, presentation and comparison of different MCDM/A methods, including: AHP, FAHP, ESM, ELECTRE, MAMCA (PROMETHEE and AHP) and Multiple Criteria Evolutionary Heuristic Algorithm.
- The handling of complex multiple – criteria transportation and logistics decision problems, including: fleet composition/selection problem, transportation projects evaluation problem, public transportation redesign problem, vehicle routing problem and suppliers selection problem.
- The analysis of hybrid methodologies integrating MCDM/A with Group Decision Making (GDM), multi-actor involvement and stakeholders' active participation in the decision making process.

J. Żak gives in his chapter, titled: The methodology of Multiple Criteria Decision Making/Aiding (MCDM/A) in Transportation a methodological background of

MCDM/A and its application in the field of Transportation. The Methodology of MCDM/A is presented as a set of universal rules, procedures and paradigms that support Decision Makers (DMs) in analyzing and solving complex transportation decision problems of a multiple criteria/objective character. The process of Multiple Criteria Decision Making/Aiding with its all phases is characterized and major actors participating in this process are described. The nature and specific characteristics of MCDM/A are explained and transportation – oriented interpretation of discussed terms and ideas is provided. The classification of the MCDM/A methods is given. The theoretical background is supported by the real world transportation case study focused on the analysis, mathematical modeling and a solution procedure of a selected multiple criteria, transportation decision problem, i.e.: the fleet composition problem. A multiple criteria evaluation and selection of low-floor trams is presented. The application of selected MCDM/A methods, including: ELECTRE III/IV and AHP is demonstrated. The paper has an instructional character and it is addressed to the academic community (researchers, students), professionals, politicians and managers dealing with complex transportation decision problems and facing the problematic of transportation – oriented decision making. The paper is prepared for EWGT community, in particular, as a guide book for their research and consulting projects.

S. Hadavi, C. Macharis and K. Van Raemdonck present further advancement of the MAMCA method in their chapter titled: "The Multi-Actor Multi-Criteria Analysis (MAMCA) Tool: Methodological Adaptations and Visualizations". Multi Actor Multi-Criteria Analysis (MAMCA) method, being a combination of AHP and PROMETHEE methods, had been developed to support groups of decision makers (DM-s) and stakeholders in complex decision-making processes. In the authors' opinion the MAMCA methodology differs from the classical approaches of Multi-Criteria Decision-Analysis (MCDA) in the fact that the different actors (stakeholders) who are engaged in a decision making process are explicitly involved throughout all its steps. In the presented chapter, after a brief summary of the MAMCA methodology, the authors present the MAMCA software, an interactive web tool which has been developed based on the aforementioned methodology. In the software, several visualizations are provided to aid decision makers with analysis of the problem at hand. Furthermore the authors discuss how the PROMETHEE method has been utilized to provide a comprehensive overview on the performance of different alternatives/solutions to the problem. The software and the visualizations are demonstrated within a case study focused on the selection of an Urban Consolidation Center (UCC) in Brussels.

In the chapter titled: "Dealing with the Complexity of Stakeholder Interaction in Participatory Transport Planning" M. Le Pira, G. Inturri, M. Ignaccolo and A. Pluchino raise an important question of citizens' and stakeholders' active participation in transportation – oriented decision making processes. The authors claim that stakeholders involvement in decision – making processes is a critical condition of reaching a reasonable consensus (compromise) in the selection and implementation of certain transportation projects, policies or plans. In the presented chapter they use a combination of multiple criteria quantitative methods, stakeholder interaction mechanisms and simulation models to guide and reproduce a participatory experiment aimed at building a consensus concerning mobility management strategies. The authors apply

Analytic Hierarchy Process (AHP) method to elicit stakeholders' preferences and different voting methods to aggregate the individual preferences into a common group-oriented model of preferences. The authors perform the group interaction experiment through a facilitated dialogue to reach a consensus among stakeholders and they apply an agent-based model (ABM) to simulate the same consensus building process. In addition, the social network of stakeholders has been analyzed to gain insights on its influence on the consensus formation. The results of this integrated procedure, applied in a pilot experiment with University students as stakeholders, provide useful suggestions on how to use different methods and guide effective and efficient participation processes aimed at consensus (compromise) building.

J. Żak and M. Kiba-Janiak co-author the chapter titled: "The methodology of redesigning and evaluating the medium-sized public transportation system". They propose a universal and original, MCDM/A – based paradigm of redesigning a public transportation system. The developed procedure is composed of two major stages: 1. A heuristic construction of the public transportation system redesign variants/scenarios; 2. Multiple criteria evaluation of these variants. In the first stage 6 redesign concepts of the public transportation system are developed. The variants are constructed heuristically, based on "common sense", best practices and authors' expert knowledge in the field. In the second stage the evaluation of redesign scenarios is formulated as a multiple criteria ranking problem. A consistent family of criteria is defined, preferences of the decision maker (DM) and stakeholders are identified and modeled, and finally computational experiments are performed. In the computational phase selected Multiple Criteria Decision Making/Aiding (MCDM/A) methods, such as: AHP and ELECTRE III/IV are applied. The authors carry out a thorough discussion of generated results. The analysis is carried out based on a real world case study in a medium sized city in Poland. The proposed approach can be applied as a decision making/aiding tool for municipal authorities (local governments) and management of urban transportation systems facing a problem of upgrading local transportation.

O. Nahum and Y. Hadas present "A framework for solving Real-Time Multi-Objective VRP", which is one of the most popular transportation – logistics problems. As opposed to a traditional, single-criterion approach the authors formulate the Vehicle Routing Problem (VRP) as a multiple objective decision problem. They depict its multi-objective character and point out the necessity of balancing different interests, including: customers' satisfaction and transportation costs while solving the problem. Their formulation has a dynamic character and it is based on real-time data generated by modern, advanced technologies. Having in mind the fact that VRP is an NP-hard problem the authors propose a generic, heuristic framework for solving the real-time VRP. They focus their efforts on automatic selection of one solution from a set of solutions (being a good approximation of a Pareto front). The authors claim that in real-time dynamic problems, a solution is given based on known data. As time progresses, new data are added to the problem, and the initial solution has to be re-evaluated in order to suit the new data. This is usually done at pre-defined time intervals. If the time intervals are small enough the amount of information added in each step/iteration is limited. Therefore, the new solution is similar to the previous one. Based on this reasoning the automatic selection of a single solution from a set of similar solutions makes sense and satisfies the requirements of quick selection that must be

done within a short time window. The proposed approach is based on traditional and evolutionary algorithms, which incorporate multi-criteria decision making methods, for solving real-time multi-objective VRP.

Next chapter co-authored by J. Żak and B. Galińska is focused on the analysis of another classical logistics problem, i.e. selection of suppliers. The authors present in their work a "Multiple Criteria Evaluation of suppliers in different industries – comparative analysis of three case studies". They carry out a multiple criteria evaluation and ranking of three distinctive categories of suppliers, i.e.: meat suppliers in a food industry, logistics service providers (LSPs) in a household chemistry industry and suppliers of packaging and supplementary materials in a printing industry. The definition of variants (suppliers), family of evaluation criteria and modeling of the DM's preferences are presented. The results of computational experiments performed with the application of ELECTRE III/IV and AHP methods are demonstrated. The processes of selecting the most desired suppliers in the above mentioned cases are compared and different aspects of the proposed multiple criteria - based procedures are thoroughly discussed.

S. Dožić and M. Kalić in their chapter, titled: "Aircraft type selection problem: application of different MCDM methods", compare three alternative MCDM/A methods, including: Analytic Hierarchy Process (AHP), Fuzzy Analytic Hierarchy Process (FAHP) and Even Swaps Method (ESM) and analyse their suitability for solving the multiple criteria fleet selection problem. The latter is a dominance-based method, while the two former belong to popular hierarchical methods. All methods are applied for solving the aircraft type selection problem in a hypothetical airline, operating under the same conditions (as described in a case study). All three methods are compared from a methodological perspective and based on the solutions they arrive at. Considering the differences between particular methods the authors perform different forms of sensitivity analysis. While considering AHP and FAHP methods, the sensitivity of alternative ratings in respect to different pairwise comparisons of the variants is analysed. It shows that the methods are sensitive to this kind of alterations. In ESM, the objective ranking across alternatives varies, showing that the ESM is not sensitive at all.

References

1. Figueira, J., Greco, S., Ehrgott, M.: Multiple Criteria Decision Analysis: State of The Art Surveys. Springer Verlag, London (2005)
2. Macharis, C., DeWitte, A., Ampe, J.: The multi-actor, multi-criteria analysis methodology (MAMCA) for the evaluation of transport projects: theory and practice. J. Adv. Transp. **43**(2), 183–202 (2009)
3. Saaty, T.: Transport planning with multiple criteria: the analytic hierarchy process applications and progress review. J. Adv. Transp. **29**(1), 81–126 (1995)
4. Vincke, P.: Multicriteria Decision – Aid. John Wiley & Sons, New York (1992)
5. Żak, J.: The methodology of multiple criteria decision making/aiding in public transportation. J. Adv. Transp. **45**(1), 1–20 (2011)

The Methodology of Multiple Criteria Decision Making/Aiding in Transportation

Jacek Żak(✉)

Poznań University of Technology, 3 Piotrowo Street, 60-965 Poznań, Poland
jacek.zak@put.poznan.pl

Abstract. The chapter presents the methodological background of **Multiple Criteria Decision Making/Aiding (MCDM/A)** and its application in the field of **Transportation.** The Methodology of MCDM/A is presented as a set of universal rules, procedures and paradigms that support Decision Makers (DMs) in analyzing and solving complex transportation decision problems of a multiple criteria/objective character. **The process of Multiple Criteria Decision Making/ Aiding** with its all phases is characterized and major actors participating in this process are described. The nature and specific characteristics of MCDM/A are explained and transportation – oriented interpretation of discussed terms and ideas is provided. The classification of the MCDM/A methods is given. The theoretical background is supported by the **real world transportation case study** focused on the analysis, mathematical modeling and a solution procedure of a selected multiple criteria, transportation decision problem, i.e.: the fleet composition problem. A multiple criteria evaluation and selection of low-floor trams is presented. The application of selected MCDM/A methods, including: ELECTRE III/IV and AHP is demonstrated. The chapter has an instructional character and it is addressed to the academic community (researchers, students), professionals, politicians and managers dealing with complex transportation decision problems and facing the problematic of transportation – oriented decision making. The paper is prepared for EWGT community, in particular, as a guide book for their research and consulting projects.

Keywords: Multiple Criteria Decision Making/Aiding · Transportation · Methodological guidelines · Real world case studies

1 Introduction

There are many definitions of **transportation,** depending on the meaning/understanding and content of this term. Transportation can be considered as a field of knowledge, study or science. It can be also perceived as a (business) process composed of certain stages or operations. Transportation is also defined as a professional activity requiring specialized knowledge and resources to be carried out. In addition, it can be viewed as an industry comprising all business organizations delivering transportation services.

In all these meanings **transportation** always refers to transferring/moving people and goods from their origins to destinations by certain means of transport (vehicles) and

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_1

with the application of a certain infrastructure (roads, railways, terminals) [9, 29, 32, 34]. Based on the character of the environment in which transportation is carried out it can be divided into the following modes: air and space; rail, road, cable and pipeline (often denominated by land/surface transportation modes), and finally water (including: sea and inland water). As far as the carried load is concerned transportation can be generally classified as passenger and freight transportation [9, 32, 34].

There are **many decision problems in transportation** that require a solution and/or searching for the most desirable (rational) decision/choice/option. The typical transportation decision problems include: finding a location of a piece of infrastructure (terminals, hubs, depots), selecting routes and designing transportation corridors, fleet management (composition, routing and scheduling, assignment, replacement), planning and designing of transportation solutions (segments of roads/railways, elements of a transportation network), traffic control/management, design and adjustment of the transportation portfolio, management of transportation processes, assessment and implementation of transportation projects.

Based on his research Żak [32] proves that the vast majority of transportation decision problems has a multiple criteria character. In the survey research carried out in roughly 120 companies [31] he recognizes 21 major transportation decision problems. These decision problems are perceived by 80% of the questionnaire respondents as multiple criteria problems. Almost 90% of respondents recognizes the **trade-offs** and **contradictory interests** associated with these problems. In their opinion the investigated decision problems require handling of the above mentioned **trade-offs** between various interests (aspects) and searching for the balance between **quality** of the solution and its **costs.** The respondents claim that there are many groups of stakeholders in transportation, including: managers and owners (shareholders) of transportation units, their employees and customers. In some instances transportation solutions may also affect such groups as: local communities (residents), roads/railways users, local authorities. All these bodies express their own expectations and aspirations and are interested in finding a good, balanced and compromise solution to major transportation decision problems.

As revealed in several reports [16, 23, 32, 35] the existence of many stakeholders and the complexity of transportation activities, processes and systems requires that in their analysis **a variety of measures and characteristics must be used. Many authors apply different criteria and attributes while** evaluating transportation solutions, designing and assessing original transportation concepts, developing and implementing transportation projects or assessing the operations of a transportation system/network [5, 16, 23, 39]. In their opinion transportation needs to be assessed from different perspectives, including, among others: economic, technical, social, safety-oriented and environmental aspects. Due to the existence of the above mentioned trade-offs, conflicts of interests and different perspectives of transportation evaluation the search for balanced, sustainable and compromise solutions is necessary.

For all the above mentioned reasons the necessity to apply **Multiple Criteria Decision Making/Aiding (MCDM/A)** in transportation becomes transparent. **MCDM/A** is a field of study that attempts to equip the decision maker (DM) in a set of tools and methods that help him or her **to solve complex decision problems** in which **several** –

often contradictory – **points of view (criteria) must be taken into account** [7, 27, 32]. The methodology of MCDM/A has a universal character and can be applied in different fields, from engineering, through medicine, economics, finance, agriculture up to humanities and arts. It is well suited for the analysis of transportation. The methodology of Multiple Criteria Decision Making/Aiding (MCDM/A) can help the Decision Makers (DM-s) in analyzing and solving complex transportation decision situations/problems, in assessing innovative transportation projects, concepts and solutions, in analyzing trade-offs and balancing conflicting interests associated with the operations of certain transportation processes and systems, in searching for the most desired, compromise transportation decisions.

MCDM/A supports the DM and enables him/her to **advance in solving** the so called **multiple criteria decision problems** (defined in Sect. 2). It leads the DM through the whole decision making process, including the recognition of variants/solutions; definition of a consistent family of criteria; modeling of preferences, analysis and selection of appropriate decision making/aiding tools, performance of computational experiments, analysis of results and final selection of the most desired solution. **The MCDM/A methodology helps the DM to structure** and solve the unstructured, complex decision problem and **reach the final compromise solution**, based on the analysis of the existing trade-offs and the DM's and stakeholders' preferences. In recent years several successful applications of multiple criteria analysis in transportation have been reported [13, 16, 26, 30, 33, 35].

The major objective of this chapter is **to demonstrate the advantages of Multiple Criteria Decision Making/Aiding (MCDM/A) and its usefulness and applicability in transportation.** The basic concepts of MCDM/A are defined and its transportation interpretation is given. The chapter refers to a traditional MCDM/A-based paradigm of solving multiple criteria decision problems. It shows how real life, complex, multiple criteria transportation decision problems can be solved with the application of MCDM/A methodology.

The chapter is composed of 4 sections. The introduction includes the background of the presented analysis, the definition of basic terms and the research objectives of the chapter. In Sect. 2 the methodology of MCDM/A is presented and its major notions and concepts are defined. Section 3 is focused on the presentation of the real world transportation case study in which the application of MCDM/A methods is demonstrated. The case study is focused on the analysis of a multiple criteria tram selection problem. Section 4 includes conclusions and final remarks. The chapter is completed by a list of references.

2 The Methodology of Multiple Criteria Decision Making/Aiding (MCDM/A)

2.1 The Formal Definition of the Field and Its Major Features

Multiple Criteria Decision Making/Aiding (MCDM/A) is a field which originates in **Operations Research** [11], the area that conducts comprehensive "research on operations" and provides a variety of quantitative tools and methods that help the Decision

Maker (DM) to make rational decisions. Operations Research techniques have been applied extensively by researchers, managers, analysts and engineers to solve complex decision problems that arise in different organizations, business units and systems. Operations Research, as a field of knowledge, focuses its efforts on conducting a comprehensive analysis of a certain decision situation, constructing its mathematical description (mathematical model) and finally finding an optimal solution to the decision problem faced, using appropriate computer-based, quantitative, analytical methods and tools. The classical, state-of-the-art Operations Research methods include: linear and non-linear programming procedures, transportation and assignment algorithms, network optimization methods, integer and dynamic programming techniques, metaheuristic algorithms, game theory and decision analysis tools, simulation and Markov Chains methods, queueing and inventory theory - based procedures. All these techniques are applied in such situations in which planning, control and coordination of various operations (activities) is required. Operations Research attempts to allocate scarce resources (labor, cash, machines and equipment, facilities, ground and space, etc.) to competing operations/activities in an optimal manner. Thus, it searches for certain solutions that optimize a single objective function. The area of application of Operations Research is widespread and includes: manufacturing, transportation and logistics, construction, telecommunication, financial planning and banking, health care and public services, the military and utilities, to mention just a few.

The nature of MCDM/A, as a field of study, is similar to the profile and overall approach applied in Operations Research. Similarly to OR, MCDM/A attempts to equip the DM in a set of tools and methods that help him or her to solve complex decision problems. At the same time, MCDM/A, as opposed to OR, focuses its efforts on solving **multiple criteria decision problems**, that is such complex decision situations in which several – often contradictory – points of view must be taken into account [7, 27]. By definition MCDM/A is a field which aims at giving the DM different tools, methods and algorithms that enable him/her to advance in solving the above defined multiple criteria decision problems. MCDM/A and OR have the following major features in common [7, 11]:

- Both MCDM/A and OR have a quantitative character and are focused on providing tools and methods that solve certain decision situations.
- They both use scientific methods and rigid, systematic way of thinking to investigate the problem of concern.
- Both fields look at the problem at stake from a broad perspective, applying system approach to problem analysis and solution.
- Both MCDM/A and OR search for solutions that attempt to resolve conflicts between different components of the organization or system, satisfy interests of various stakeholders and finally accept and implement a solution that guarantees the fulfillment of the DM's overall objectives.
- Both areas have an interdisciplinary character, require a variety of skills and competences and thus a team approach to carry out a complete study of a certain decision problem.
- Both MCDM/A and OR can be considered as consistent and comprehensive methodologies that provide a set of rules that support the DM in the process of solving

complex decision problems. The rules elaborated/developed within each of the fields are to lead to the generation of rational solutions of the considered decision problems.

At the same time certain features of both fields are in contrast, including [7, 11]:

- OR evaluates and optimizes different solutions (actions, variants) using a single measure of merit, i.e. single objective function or single criterion, while the nature of MCDM/A implies that several magnitudes - criteria or, in other words, a multiple criterion objective function is applied when the solutions' evaluation or optimization process is carried out.
- As a result of the above mentioned characteristic the problem solving methods and algorithms proposed within OR have a single objective character, while all MCDM/A methods are focused on multiple criteria analysis and allow for taking into account several criteria to solve a decision problem.
- OR attempts to find the best/optimal solution for the problem under consideration, which means that instead of improving the status quo OR is focused on identifying the best possible course of action. Thus, OR techniques generate single and the only best solution of the considered decision problem. At the same time MCDM/A accepts the "trade – offs" between criteria and the resulting concept of a satisfactory or compromise solution [25, 27]. The MCDM/A methods do not yield "objectively best" solutions, because such solutions that are the best from all points of view, simultaneously, do not exist in real world situations. While solving a multiple criteria decision problem a family or sample of "optimal" in a multiple objective sense, or Pareto optimal (non-dominated) solutions [25, 27] is generated. Finally accepted compromise solution is selected from the family of Pareto optimal (non-dominated) solutions, based on the DM's expressed preferences, expectations and aspirations.
- As opposed to OR computational procedures, which are usually single – phase methods that result in generating a final result in one step computational experiment, MCDM/A methods require more steps and computer – human being interaction to generate and finally select the desired result/solution.

2.2 Basic Notions, Principles and Concepts

As mentioned before MCDM/A is focused on solving, so called, multiple criteria decision problems, i.e. situations in which, having defined a set of actions/variants/solutions $\boldsymbol{A}$ and a consistent family of criteria $\boldsymbol{F}$ the DM tends to [7, 27, 32]:

- determine the best subset of actions/variants/solutions in $\boldsymbol{A}$ according to $\boldsymbol{F}$ (choice problem),
- divide $\boldsymbol{A}$ into subsets representing specific classes of actions/variants/solutions, according to concrete classification rules (sorting problem),
- rank actions/variants/solutions in $\boldsymbol{A}$ from the best to the worst, according to $\boldsymbol{F}$ (ranking problem).

The proposed definition lets us distinguish the following major categories of MCDM/A problems [27, 32]:

– choice problems (multiple criteria optimization problems),
– sorting problems,
– ranking problems.

Transportation specialists frequently face multiple criteria choice (optimization) problems in their professional life. The following examples belong to this category:

- Designing an optimal structure of a Park & Ride System in a certain metropolitan area. Defining the most suitable locations and sizes of the parking lots, taking into consideration contradictory interests of travelers (drivers and passengers), local residents, local authorities, public transport operators, and other road users. Analyzing: comfort of travel, safety and security, travel costs and times, investment costs, accessibility, reliability, environmental friendliness, comfort of life.
- Searching for an optimal allocation of transportation jobs (duties) to the drivers and trucks (vehicles), taking into account the characteristics of the jobs (distances to cover, category and size of the load), capacities and technical features of the trucks, interests and preferences of drivers, satisfaction of customers and fulfilment of management goals.

Sorting problems are also quite common in transportation. Classical examples are as follows:

- Classification of national transportation projects and their allocation to the predefined classes based on the assessment of such measures of merit as: profitability, market potential, social potential, technical feasibility, environmental friendliness. Constructing the annual governmental budget based of projects evaluation. The projects can be categorized as follows: very promising and fully supported financially, well evaluated and partly financed, moderate and waiting for additional financial resources, poorly assessed and not receiving financial support.
- Technical diagnostics of vehicles (fleet) and their assignment to various categories, based on multi-dimensional analysis (evaluation) of their technical condition (e.g.: vehicles in an excellent technical condition, vehicles in an acceptable condition, requiring tune-ups and periodical inspections, vehicles in a poor condition that must be repaired or replaced and vehicles requiring braking).

Transportation management also involves solving the multiple criteria ranking problems, such as:

- Comprehensive evaluation of different transportation solutions (e.g. redesign of the critical section of the railway transportation system) based on their technical, economic, social, safety and environmental characteristics resulting in their prioritization and ranking.
- Evaluation and selection of freight carriers (transportation companies) delivering transportation services to a certain customer. Ordering transportation units from the best to the worst, taking into account such aspects as: transportation/delivery costs,

reliability, accessibility, timeliness, market position and experience, safety and security, environmental friendliness, modernity, productivity/efficiency.
- Fleet selection problem focused on the analysis and evaluation of the considered transportation fleet (trucks, trams, vans, buses) and their ordering from the best to the worst.

As indicated earlier a transportation decision problem considered as a case study in this chapter, Sect. 3, is a multiple criteria ranking problem. It is focused on the fleet assessment.

Based on the above quoted definition one can easily determine major components that exist in each formulation of the multiple objective decision problem, i.e. a set of objects/actions/alternatives/variants/solutions $\boldsymbol{A}$ and a consistent family of criteria $\boldsymbol{F}$. The set $\boldsymbol{A}$ can be defined [7, 20, 27]:

- directly, in the form of a complete list of all components of $\boldsymbol{A}$ that contains their names or symbolic denominations, or
- indirectly in the form of certain rules and formulas that determine feasible actions/variants/solutions, e.g. in the form of constraints.

It can be:

- constant, defined beforehand (*a' priori*), prior to the start of computational experiments and not being subject to changes and modifications during the decision process;
- evolving or alternating, i.e. being influenced by the course of computational experiments and subject to changes and modifications during the decision process;
- post-experimental (*a posteriori*), i.e. being the result of the computational experiments/decision process.

In addition, one can distinguish between:

- global set $\boldsymbol{A}$, in which all its components are independent and are not interrelated with each other; in the global set $\boldsymbol{A}$ the selection of one element, as a solution of the problem, excludes the possibility of choosing another alternative at the same time;
- fragmentary set $\boldsymbol{A}$, in which some components may interact with others; in fragmentary set $\boldsymbol{A}$ there is a possibility of selecting a combination of different objects of $\boldsymbol{A}$ that constitute the final solution of the problem.

Several examples of set $\boldsymbol{A}$ in different transportation problems have been given below:

- when the problem at stake consists in evaluating different types/makes of trams for a certain public transportation system and selecting the most suitable one the set $\boldsymbol{A}$ is composed of a complete list of legally and technically acceptable offers submitted by various manufacturers of trams;
- if the considered decision problem refers to the selection of the most rational composition (portfolio) of transportation services (jobs) offered by a certain carrier at the market the set $\boldsymbol{A}$ is defined indirectly in the form of constraints that shape the space of feasible solutions for the service portfolio optimization problem; each combination of products that satisfy these constraints belongs to $\boldsymbol{A}$;

- when the structure of the freight transportation system is designed and the locations of the depots/distribution centers must be determined the set $\boldsymbol{A}$ consists of a table of all available locations that satisfy predefined critical conditions of technical (space, infrastructure), legal and environmental nature.
- in case a manager of a transportation company/dispatcher recruits and selects drivers to be employed in a considered company the set $\boldsymbol{A}$ encompasses several alternative candidates who submitted their applications for this position.

Table 1. The examples of three different CONSISTENT FAMILIES OF CRITERIA for three transportation decision problems: fleet selection problem, design of a park and ride system and redesign of a section of a transportation system

Criterion number and symbol	**CONSISTENT FAMILIES OF CRITERIA**		
	Fleet selection problem	**Design of a park & ride system**	**Redesign of a section of a transportation system**
C1	Comfort of travel in a tram	Investment costs	Waiting time
C2	Traction – operational characteristics of a tram	Accessibility (by individual and public transportation)	Riding time
C3	Tram reliability	Travel time (from major origins; to the city center)	Reliability
C4	Tram durability/life expectancy	Generated traffic flow (passengers)	Safety
C5	Operational costs	Safety and security	Directness of connections
C6	Environmental friendliness of a tram	Environmental friendliness	Comfort of travel
C7	Utility & functionality of a tram (incl. safety)	Capacity utilization (cars)	Financial efficiency
C8	Price of a tram		Investment profitability (IRR)
C9	Experience of a tram manufacturer/supplier		Environmental friendliness

The consistent/coherent family of criteria $\boldsymbol{F}$ is a set of criterion functions $\boldsymbol{f}$, defined on $\boldsymbol{A}$ and designed to evaluate all components of the set $\boldsymbol{A}$ in a comprehensive and consistent manner. Each criterion $\boldsymbol{f}$ in $\boldsymbol{F}$ represents subjective preferences of the DM with respect to a concrete aspect/dimension of the decision problem, such as: cost, quality, reliability, efficiency, timeliness, etc. Thus, all the criteria in $\boldsymbol{F}$ compose a set of characteristics that exhaust the spectrum of issues associated with the considered decision problem and corresponding to the DM's expressed interests and expectations. The consistent family of criteria should be characterized by the following features [20, 27]:

- **exhaustiveness of evaluation**, which essentially means that it should provide a complete (exhaustive) evaluation of $\boldsymbol{A}$, and guarantee that no aspect of the considered decision problem has been omitted;
- **cohesiveness,** which corresponds to the rule that each criterion in $\boldsymbol{A}$, having a specific direction of preferences (minimized – min or maximized – max) contributes to the overall/global model of the DM's aspirations and expectations;
- **non redundancy,** meaning that any criterion in $\boldsymbol{F}$ should not be co - related with other criteria being the elements of $\boldsymbol{F}$ and should not have a common/joint domain with the domains of other criteria.

In Table 1 one can see three different **CONSISTENT FAMILIES OF CRITERIA** applied in the following cases:

- Fleet Selection Problem, when different types/makes of trams for a certain public transportation system are evaluated by a family of criteria;
- Design of a Park and Ride System, when the most suitable locations and sizes of the parking lots are determined (considering different aspects and interests);
- Redesign of a Section of a Transportation System, when several, alternative transportation solutions are assessed by various characteristics.

It is worth pointing out that from a practical point of view the number of criteria used to solve a multiple criteria transportation decision problem should range from 2–9. While handling the multiple criteria decision problem one must apply at least two criteria as the minimal number of characteristics used. Otherwise, the decision problem has a single-criterion character and cannot be classified as a multiple criteria decision problem. The upper boundary of the criteria number is defined by a magic number of 7+/- 2 measures [32]. This number corresponds to psychological capabilities of a human being. It has been proved by many researchers [7, 25] that an average human individual cannot handle more than 7+/- 2 aspects while considering a complex situation. Thus, it is recommended that when a DM handles a multiple criteria transportation-oriented decision problem he or she should define the maximum number of considered aspects/characteristics not exceeding 9. In some very comprehensive situations this number may not cover all the required aspects. As a result the assessment may not be complete. In such specific cases the family of criteria may be slightly extended and include 10-11 measures. This is, however, not recommended. If the number of criteria is large it is advised to aggregate certain measures, redefine their scopes or use alternative formulations.

2.3 Solution Procedures of the Multiple Criteria Decision Problems. MCDM/A Methods, Tools and Computational Procedures

Multiple criteria decision problems require specific procedures, methods and decision-making/aiding tools to be solved. Precise rules must be applied to solve a multiple criteria decision problem.

Roy [19, 20, 27] distinguishes four major stages of the solution procedure of the multiple criteria decision problems, including: definition of variants, definition of criteria, modeling and aggregation of the DM preferences, solving the decision problem.

These indications, further extended by Figueira et al. [7] and Żak [32] lead to a universal procedure of solving a multiple criteria decision problem, composed of the following 5 stages:

- Investigation of the decision problem considered and its verbal description. Recognition of the category of the decision problem, definition of the major objective of the analysis, the DM and stakeholders.
- Problem structuring and mathematical modeling. Analysis of the parameters characterizing the decision problem, collecting data, definition of the set of variants/feasible solutions $\boldsymbol{A}$ and construction of the consistent family of criteria $\boldsymbol{F}$. Modeling and aggregating the DM's and stakeholders' preferences (synthesis of the global model of preferences).
- Review, evaluation and selection of the appropriate methods and algorithms to solve the considered decision problem. Matching the specific characteristics of the decision problem with appropriate multiple criteria methods.
- Carrying out a series of computational experiments. Solving the decision problem with an application of the global model of preferences and selected MCDM/A method. Review and analysis of the generated results. Sensitivity analysis – investigation of the stability of generated results.
- Selection of the most desired solution/variant/action considered to be a compromise solution. Possible implementation of this solution in the real world.

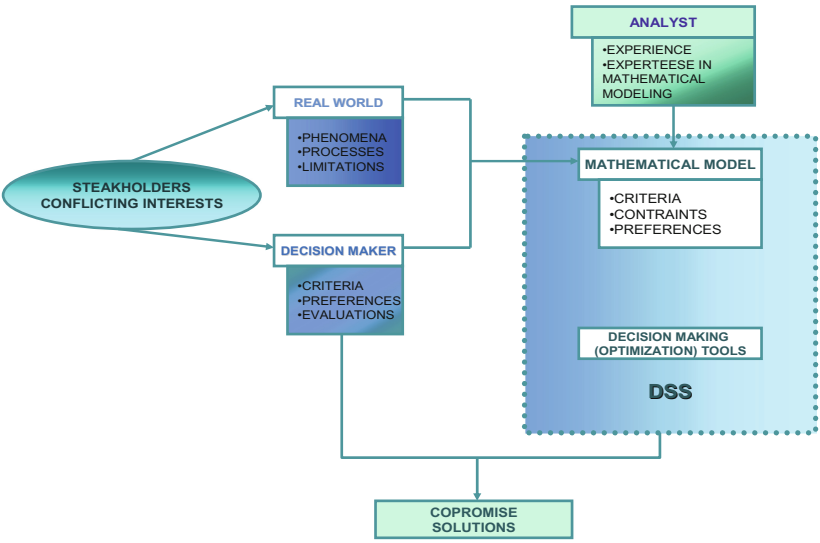


Fig. 1. The generic paradigm of the MCDM/A-based decision making/aiding process

These stages are carried out within the decision process based on the application of the MCDM/A methodology (see Fig. 1). This methodology places the decision problem at stake in the real world and helps to solve it. For transportation – oriented situations the real world has a transportation character and is limited to transportation companies,

processes and/or systems. It involves such phenomena as: travelers' behavior, traffic flows in certain transportation corridors, road and fleet capacity availability and utilization, freight movement and storage, volatile demand for transportation services (seasonal effects) and many others. Transportation – oriented real world is the object of investigations, with a special emphasis on specific decision problems that arise in a transportation environment. These decision problems are complex tasks or questions/issues that require solution or settlement [32, 36]. The decision problem arises when the decision maker – DM (see the definition below) searches for the most desired action (variant, solution) among all feasible actions (variants, solutions). The decision problem is the result of the observation of reality by the DM and depicting problem-oriented situations that require decisions to be made [32, 36].

Żak [31, 32] distinguishes the following major categories of road (primarily freight) transportation decision problems: crew sizing, selection (recruiting) and assignment; definition of the transportation services' portfolio; transportation jobs selection, pricing and assignment; fleet composition (selection and sizing) combined with fleet assignment; fleet replacement; vehicle routing and scheduling combined with crew assignment and scheduling.

The MCDM/A methodology [1, 7, 15] clearly identifies major participants of the decision making/aiding process, such as: the decision maker (DM), the analyst and the interveners (stakeholders) and describes their roles in this process. The DM (an individual or a group of individuals) defines the objectives of the decision process, expresses preferences and finally evaluates the generated results. He or she, being a subject of the decision making process and an internal body to the decision problem, is responsible for making the final decision. The most frequent decision makers in the transportation environment are: transportation policy makers; local and governmental officials (authorities); top and middle level managers in the transportation companies (carriers); traffic engineers; managers of the public transportation systems operators and dispatchers. The analyst is an expert/consultant in transportation, mathematical modeling and decision making/aiding. He/she shares with the DM his/her knowledge, experience and expertise. The analyst, who is external to the decision problem, handles the decision making/aiding process and supports the DM during its course. His/her role is to construct a decision model and suggest the most appropriate tools and methods in order to advance the DM in solving the decision problem. The analyst explains to the DM the consequences of certain actions/solutions and finally recommends the most desired one. Both parties (DM and Analyst) observe and analyze together the "real transportation world" and recognize together all the features of the considered decision problem. The interveners, frequently called stakeholders, represent different individuals, organizations and groups that are interested in generating a rational solution to a certain decision problem. In many decision situations they express their own preferences and present concrete stand points that should be taken into account while analyzing the decision problem. Final decision usually has a strong impact on satisfaction of the interveners' interests. The most common interveners/stakeholders in transportation are: travelers/road users (passengers, drivers); crews/employees of the transportation companies (carriers/operators); local residents (communities); customers and suppliers of transportation companies (carriers/operators).

As presented in Fig. 1 the MCDM/A - based decision process is strongly supported by computer-based systems. These decision support systems (DSS) usually have an interactive character and help the DM to generate final, compromise solutions in an iterative manner. They never replace the DM and do not make the final decision instead of him, her or them. Their role is to assist the DM in the whole decision – making process, present a variety of solutions and help him/her/them to investigate those solutions. The computer – based decision support systems help the DM to express and model his/her or their preferences and better understand the considered decision problem.

For each specific transportation decision problem the above described generic scheme of the MCDM/A - based decision making process is customized to the concrete transportation environment and specific character of the considered decision situation. The proposed approach has been applied in solving a concrete transportation decision problem (case study) analyzed in this chapter.

The multiple criteria decision making/aiding methods, used in this process, can be classified according to several criteria, including: the overall objective of the decision method correlated with the category of the decision problem [32], the moment of the definition of the DM's preferences [7, 27] and the manner of the preference aggregation [20, 21, 27].

Based on the first classification criterion one can distinguish the following categories of MCDM/A methods [27, 32]:

- multiple criteria choice (optimization) methods (e.g. LBS [14]; Steuer Procedure [25]),
- multiple criteria sorting (classification) methods (e.g. Electre Tri, 4eMka [7]),
- multiple criteria ranking methods (e.g. Electre III/IV [7, 21]; AHP [22]).

With respect to the second division criterion three categories of methods are identified:

- methods with an a'priori defined preferences (e.g. Electre methods [20, 27], Promethee I and II [3, 7], UTA [12],
- methods with an a'posteriori defined preferences (e.g. PSA method and other multiple criteria metaheuristics [17]),
- interactive methods (e.g. Steuer Procedure [25], STEM [2], [LBS [14]).

According to the third classification criterion one can distinguish:

- the methods of American inspiration, based on the utility function [15] (e.g. AHP [22], UTA [12]),
- the methods of the European/French origin, based on the outranking relation (e.g. Electre methods [20, 21], Promethee I and II [3, 7]).

3 The Application of MCDM/A Methodology in Transportation. Selected Case Study

3.1 Overall Characteristics of the Presented Decision Problems

Chapter 3 presents a real life case study that refers to solving a complex transportation decision problem of a multiple criteria character. The case study demonstrates the applicability of MCDM/A concepts and rules. It shows how the universal paradigm of solving multiple criteria decision problems can be applied in a transportation environment. The case study is focused on the analysis of the suitability of different models/makes of trams for a public transportation system. Based on the methodological principles presented in Sect. 2.3 the standardized stages of the solution procedure of a multiple criteria decision problem are applied in the analysis of the case study. They are characterized below.

3.2 Multiple Criteria Fleet Selection Problem – Evaluation of Alternative Tram Makes/Models for a Public Transportation System

3.2.1 The Background of the Fleet Selection Problem

The fleet selection problem belongs to traditional transportation decision problems and constitutes one of strategic questions within a broader area of fleet management [4]. It is associated with a fleet composition problem that involves two major sub-problems [4, 18, 32, 40]: fleet sizing problem which is focused on defining the most appropriate number of vehicles in the fleet and fleet selection problem featured by determining the most suitable category/type of vehicles in the fleet. Both problems are rooted in matching transportation demand and supply. They require the definition of the optimal number and kind/type of vehicles in a fleet to guarantee, on the one hand, a complete fulfillment of transportation jobs and elimination of excessive fixed costs associated with the fleet underutilization on the other. Both questions are widely discussed in literature [4, 32].

As mentioned above the fleet composition problem, including the fleet selection component, for a public transportation system in particular, has a strategic character and its solution may result in critical, long – term consequences, such as: the level of passengers' service (reliability, comfort of travel, travel costs); the fleet operations, maintenance and replacement policies and the overall performance and efficiency of the public transportation system (profitability, level of subsidies). One can distinguish two levels in the structure of the fleet composition problem for a public transportation system [5, 28, 35]:

- Selection of the transportation mode (e.g. subway, light city train, bus, tram), resulting in a certain modal split of passengers' trips;
- Definition of the vehicles' types (makes) and their numbers in each transportation mode.

The second of the above mentioned problems is discussed in this case study. It is associated with a comprehensive analysis, evaluation and selection of specific vehicles, i.e. bus makes, tram models, train types.

By nature the fleet selection problem has a multiple objective character. This statement is supported by the works of many authors [28, 32, 35] who agree that the evaluation of vehicles should have a multidimensional character and should involve, among others, such aspects as: technical – operations parameters, economic factors, comfort and safety of travel, environmental aspects [32]. The following criteria are frequently used when the evaluation of vehicles is carried out [32]: purchasing cost (price), operations and maintenance costs, safety and comfort of transportation, durability and reliability of vehicles, technological advancement, esthetics, environmental friendliness. Several authors [28, 32] claim that while selecting vehicles for the public transportation system the interests of different stakeholders should be taken into account. In particular the following entities are mentioned as the stakeholders in the vehicles' purchasing process [28, 32]: passengers using the fleet operating in the public transportation system, operator of the public transportation system that utilizes the purchased fleet and municipal authorities that substantially support the purchase financially. In addition, some authors report that the selection of the transportation fleet for a public transportation system is supported by the application of multiple criteria methods [16, 35].

3.2.2 Verbal Description of the Tram Selection Problem

The decision problem considered in the case study consists in evaluating and selecting the most suitable model of a low-floor tram car for the public transportation system in a medium sized metropolitan area of Poznan city in Poland. It is strictly linked with an open bid competition organized in the city for tram manufacturers to deliver 40 tram - cars that should operate in the system and provide passengers' transportation services.

The decision maker (DM) who makes final decision regarding the selection of a concrete model/type of a tram is an Open Bid Commission (OBC) composed of 5 persons representing the operator, city transport organizer and municipal authorities. Thus, the decision is made by a group DM. The stakeholders in this process are: municipal authorities (City Board), city transport organizer, public transport operator and local community (residents), in particular passengers using the public transportation system. The interests of these stakeholders should be taken into consideration in the tram selection procedure. The DM and stakeholders want to evaluate the considered variants – low – floor trams with a set of criteria of technical, economic, social and environmental character. They want to take into account expectations and requirements of various groups of interest, and cover all possible aspects of the analysis. The DM wants to carry out a quantitative evaluation of all the variants and generate its final, aggregated ranking (from the best to the worst). Thus, the decision problem has been defined as a multiple criteria ranking problem and for solving it certain multiple criteria ranking methods, including AHP, ELECTRE, ANP, ORESTE or MAPPAC should be applied.

3.2.3 Definition of Variants and a Consistent Family of Criteria

The analysis has been carried out for 7 models/types of trams designed and manufactured by different manufacturers, legally allowed to participate in the open bid. These models/makes of the trams, denominated by abbreviations: A0001, A0002, ….. A0007 constitute the set of the considered variants. All of them represent low-floor trams featured by

100% of the floor surface being placed not higher than 35–40 cm from the level of the ground (tram stop platform). Their bodies are 5-module structures made of aluminum and steel. Each tram car is equipped in 4 pairs of doors of different width. Two middle doors are wider (130–150 cm) while front and back doors are narrower (70–90 cm). The capacity of tram cars ranges between 209 and 229 passengers, including 48 to 69 seats. Each of the tram cars is equipped in 2 to 8 electric engines, whose overall power falls into the 200 kW to 560 kW interval. All tram models are designed for a wider gauge of 143.5 cm. They are constructed for a maximum vehicle gauge of 210 cm and a minimum turning radius of 18 m.

The car trams are delivered by different European manufacturers whose headquarters are located in such countries as: France, Germany, Italy, Spain and Poland. Tram suppliers have different market experience ranging between 1 to 26 years of successful operations at the market. Some of the manufacturers delivered more than 1000 tram cars worldwide while others have not produced trams yet. Depending of the manufacturers profile their trams are utilized in Europe, North America, Asia and Australia. The considered variants/tram cars are as follows: Alstom Citadis (France) – A0001, Ansaldo Breda Sirio (Italy) – A0002, CAF Bilbao (Spain) – A0003, PESA 120 N (Poland) – A0004, Siemens Combino (Germany) – A0005, Solaris Tramino (Poland) – A0006 and Stadler Variotram (Germany) – A0007.

The evaluation of these tram – cars has had a comprehensive and a multiple criteria character. The OBC has assumed that several aspects of technical, economic, market-oriented, safety and environmental nature/character should be taken into consideration. The members of the OBC agreed that the interests of the major stakeholders must be taken into account in the evaluation process. They also accepted the suggestion of the external consultant (analyst) that the evaluation of tram – cars should be formulated as a multiple – criteria ranking problem and a selected MCDM/A method must be applied to rank the variants (trams) from the best to the worst.

The evaluation of tram cars has been carried out with application of the following criteria:

- **Comfort of travelling in a tram - car (K1)**, called **COMFORT** in short. This criterion evaluates the quality standard and safety of traveling in a tram. It includes such components as: overall tram capacity, number and quality of seats in the tram, ergonomic features of the tram, including dimensions of the interior (e.g.: minimum width of a central aisle, average width of the doors; height of the floor).
- **Traction – exploitation characteristics of the tram - car (K2)**, called **TRACTION**. This criterion measures traction and safety features of the tram as well as its ability to operate in different external conditions. The construction of the criterion is based on the analysis of such parameters as: braking distance, gradeability, maximum speed, acceleration/smoothness of riding.
- **Tram - car reliability (K3); RELIABILITY.** This criterion characterizes tram - car ability to failure-free operations (without any brake-downs) in the required time horizon and its susceptibility to restoring (recovery) in case the failure/brake-down occurs. The construction of this criterion is based on several parameters that strongly influence on the probability of tram failure – free operations, such as: general

warranty period and conditions, technical availability ratio (including: statistical failure rate).

- **Tram – car durability (K4); DURABILITY.** This criterion evaluates the overall life expectancy of the tram - car, which corresponds to the period of its technical fitness and availability. In the formulation of this criterion two elements have been taken into consideration: average age of a tram taken out of service and mileage (vehicle-km) to the major repair.
- **Operating and maintenance costs (K5); OPERATING COSTS.** This criterion evaluates all cost components of tram - car utilization, including: maintenance and repair costs as well as energy consumption costs.
- **Environmental friendliness of the tram - car (K6); ENVIRONMENTAL FRIENDLINESS.** This criterion evaluates the influence of the tram - car on the environment, including such factors as: noise and vibration generated by the tram, share of renewable materials used in the tram.
- **Tram – car utility and functionality (K7)**; **UTILITY & FUNCTIONALITY.** This criterion evaluates the utility values and functionality of the tram - car. In particular, it assesses tram's friendliness for passengers and provision of equipment and technology enhancing its utility. The criterion includes such aspects as: availability and quality of the audio – visual information systems, provision of automatic platforms and designated areas for disabled/handicapped people, availability and quality of the air-conditioning equipment and other devices.
- **Price of the tram - car (K8); PRICE.** This criterion defines an estimated cost of purchasing a single tram-car, assuming a scale effect of the contract (purchasing of 40 trams).
- **Suppliers' experience (K9); EXPERIENCE.** This criterion evaluates market - competitive position and credibility of the tram – car manufacturer and its experience in manufacturing and delivering to the market low-floor tram – cars. The criterion includes the analysis of such parameters as: overall market experience, number of low-floor tram – cars manufactured by a certain manufacturer and number of cities (metropolitan areas) served by a specific supplier.

As presented in the description above and in Table 2 some of the criteria, including: K5 and K8 are composed of a single characteristics (measure), others (K1, K2, K3, K4, K6, K7 and K9) include several sub-criteria. The units of criteria and sub-criteria belonging to: K1, K2, K3, K7, K8 and K9 are presented in Table 2, while the units of the remaining criteria and sub-criteria are as follows:

– K4: average age of a tram taken out of service [years] and mileage (vehicle-km) to the major repair [vkm];
– K5 – [zl/vkm];
– K6: noise and vibration generated by the tram [dB] and share of renewable materials used in the tram [%].

Due to the fact that the evaluation/selection of trams has been carried out with the application of two alternative multiple criteria ranking methods, i.e.: ELECTRE III/IV and AHP (see the description in Sect. 3.2.5) the structure of the proposed family of criteria has been adjusted accordingly (along with the requirements of the applied

Table 2. A segment of the evaluation matrix of the considered variants – low – floor trams

CRITERIA	SUB -CRITERIA	UNITS	DIRECTION OF PREFERENCE	ALSTOM	ANSALDO BREDA	CAF	PESA	SIEMENS	SOLARIS	STADLER
				A0001	A0002	A0003	A0004	A0005	A0006	A0007
K1	No. of seats	[pcs.]	Max	48	46	48	63	57	61	57
	Quality of seats	[points]	Max	8	8	7	8	9	8	6
	Overall capacity	[pcs.]	Max	212	209	223	222	210	229	212
	Ergonomic Features									
	– Min. width of an aisle	[cm]	Max	59	55	55	74	72	75	52
	– Avg. width of doors	[cm]	Max	113	117	113	117	110	125	110
	– Floor height	[cm]	Min	35	35	35	35	36	35	35
K2	Braking distance	[m]	Min	11,4	11,5	11,5	11,5	11,6	11,5	11,5
	Gradeability	[%]	Max	7	6	6	7	7	6	7
	Maximum speed	[km/h]	Max	70	65	70	70	70	65	60
	Acceleration/smoothness of riding	[points]	Max	9	8	6	8	7	5	7
K3	General warranty period	[months]	Max	36	36	36	60	36	60	36
	Technical availability	[%]	Max	98	97	96	99	95	98	96
…	…	…	…	…	…	…	…	…	…	…
K7	Audio-visual information systems	[points]	Max	9	4	7	8	6	8	5
	Automatic platforms & designated areas for disabled	[0/1]	Max	0	1	0	1	0	0	1
	Air-conditioning	[points]	Max	9	5	5	8	5	8	7
K8	-	[mln zl]	Min	15,4	9,0	12,3	7,1	8,5	8,3	10,4
K9	Market experience	[years]	Max	84	159	120	161	165	18	70
	No. of low-floor tram – cars manufactured	[pcs.]	Max	1058	322	209	33	780	0	255
	No. of cities served	[years]	Max	26	10	7	4	16	0	14

methods). In all the computational experiments based on the application of AHP method the evaluations of variants/trams on all criteria and sub-criteria has been based on the raw data presented in Table 2 and described above. At the same time the raw data for the computational experiments carried out with the application of ELECTRE III/IV method required certain transformation. For all instances in which single, separate criteria have been applied to evaluate trams (criteria K5 and K8) all the evaluations remained unchanged. For all the remaining criteria (K1, K2, K3, K4, K6, K7, K9), structured as quantities composed of sub-criteria, the sub-criterion evaluations have been normalized [11], i.e. transformed into 0–1 intervals and then aggregated (arithmetically or weighted averaged) within each criterion. As a result for all criteria composed of sub-criteria standardized and normalized evaluations have been computed. Selected examples of these transformed, normalized evaluations used in the computational experiment based on the application of ELECTRE III/IV method are presented in Table 3.

Table 3. A segment of the evaluation matrix of the considered variants – low – floor trams. The data in a normalized form used in the computational experiment carried out with the application of ELECTRE III/IV method

CRITERIA	VARIANTS – LOW FLOOR TRAMS						
	A0001	A0002	A0003	A0004	A0005	A0006	A0007
COMFORT - K1 [0-1]	0.36	0.30	0.40	0.77	0.57	0.85	0.27
TRACTION - K2 [0-1]	1.00	0.56	0.56	0.94	0.88	0.50	0.63
RELIABILITY - K3 [0-1]	0.38	0.25	0.13	1.00	0.00	0.88	0.13
DURABILITY - K4 [0-1]	.....	.....	.....	.....	.....	.....	.....
OPERATING COSTS - K5 [zl/vkm]							
ENVIRONMENTAL FRIENDLINESS -K6 [0-1]							
UTILITY & FUNCTIONALITY - K7 [0-1]	0.67	0.33	0.20	0.85	0.13	0.52	0.57
PRICE - K8 [mln zl]	**15.4**	**9.0**	**12.3**	**7.1**	**8.5**	**8.3**	**10.4**
EXPERIENCE - K9 [0-1]	0.82	0.55	0.39	0.38	0.79	0.00	0.38

As presented in Table 3 all data for criteria K1, K2, K3, K4, K6, K7 and K9 is normalized and transformed into coefficients falling into a [0–1] interval. Only the raw data characterizing criteria K5 and K8 remains unchanged (red, bolded text).

3.2.4 Modeling of DM's and Stakeholders' Preferences

After having defined variants and criteria the model of preferences for the DM and considered stakeholders has been defined. As described in Sect. 3.2.2. the DM - an Open Bid Commission (OBC) has taken into consideration the interests of such bodies as: municipal authorities (City Board), city transport organizer, public transportation system operator and local community (residents), in particular passengers using the public transportation system. Each of the analyzed groups, composed of several to several dozen persons has been surveyed. In this process individual expectations and preferences of particular persons – representatives of different groups have been defined. Afterwards the aggregated preferences (arithmetic averages) of particular groups have been computed. As a result the weights of criteria have been defined and sensitivity of the group DM and stakeholders with respect to the changes of the criteria values has been scaled. In the next phase the representatives of all groups of interests have been invited for a meeting and the initial results of preference modeling have been demonstrated. During the panel discussion adjusted group models of preferences for municipal authorities (City Board), city transport organizer, public transportation system operator and passengers (residents) have been defined and ultimately, aggregated into one, final model of preferences. In the aggregation process the arithmetic averages of criteria weights and sensitivity coefficients have been computed.

Table 4. The final, aggregated model of preferences defined for the computational experiment carried out with the application of (a) ELECTRE III/IV Method and (b) AHP Method.

(a) ELECTRE III/IV Method (importance of criteria – WEIGHTS; sensitivity of the DM and stakeholders – THRESHOLDS)

CRITERIA	WEIGHTS	THRESHOLDS		
	w	*q*	*p*	*v*
COMFORT (K1)	8.00	0.03	0.10	0.30
TRACTION (K2)	2.00	0.02	0.09	0.30
RELIABILITY (K3)	8.00	0.02	0.12	0.50
DURABILITY (K4)	7.50	…	…	…
OPERATING COSTS (K5) [zl/ vkm]	**7.00**	…	…	…
ENVIRONMENTAL FRIENDLINESS (K6)	5.00	…	…	…
UTILITY & FUNCTIONALITY (K7)	6.00	0.03	0.10	0.40
PRICE (K8) [mln zl]	**9.00**	**0.35**	**1.00**	**3.00**
EXPERIENCE (K9)	2.50	0.02	0.10	0.40

(b) AHP Method (importance of criteria – RELATIVE WEIGHTS and ABSOLUTE WEIGHTS

	RELATIVE WEIGHTS – w_r									ABSOLUTE WEIGHTS w_a
	K1	**K2**	**K3**	**K4**	**K5**	**K6**	**K7**	**K8**	**K9**	
K1	*1*	7	1	1	2	3	2	1/2	6	0,157
K2	1/7	*1*	1/7	1/5	1/5	1/3	1/4	1/7	1	0,033
K3	1	7	*1*	1	2	3	2	1/2	6	0,156
K4	1	5	1	*1*	1	3	2	1/2	6	0,140
K5	1/2	5	1	1	*1*	3	2	1/3	4	0,139
K6	1/3	3	1/3	1/3	1/3	*1*	1/2	1/5	3	0,079
K7	1/2	4	1/2	1/2	1/2	2	*1*	1/3	3	0,100
K8	2	7	2	2	3	5	3	*1*	5	0,160
K9	1/6	1	1/6	1/6	1/4	1/3	1/3	1/5	*1*	0,036

The preference modeling phase was different for the data used in the experiment based on the application of ELECTRE III/IV and AHP methods, respectively (see the description of both methods in Sect. 3.2.5). As presented in Table 4 – part (a) the weights of criteria, for the computational experiment carried out with application of ELECTRE III/IV method, are defined on a cardinal/linear scale [7] from 0 to 10. For the considered DM and stakeholders the most important criterion is PRICE – K8 with a weight of 9, while the least important criteria are: TRACTION – K2 and EXPERIENCE – K9, with the weights of 2 and 2.5, respectively. The DM and stakeholders sensitivity intervals of indifference (I), weak preference (Q), strong preference (P) and incomparability (R), defined by the thresholds of indifference – q, preference – p and veto – v, are expressed in the standardized, dimensionless form for the majority of criteria (K1, K2, K3, K4, K6, K7, K9), i.e. those criteria that are composed of sub-criteria and require raw data transformation and adjustment. For all these criterion evaluations that are expressed in the [0-1] scale the values of q, p and v thresholds are also standardized and normalized. They take the following values, depending on the range of the criterion values for all the variants and the minimal difference/dispersion of evaluations between them: $q = 0.02 – 0.03$; $p = 0.08 – 0.12$ and $v = 0.30 – 0.50$. For the two remaining criteria: K5 and K8 the thresholds are expressed as the values defined in the criteria natural units. They are developed in the following way: q – falls into the interval equivalent to 5% of the minimal value of the criterion range and 3% of the arithmetic average of the criterion value for all considered variants; p – corresponds to roughly 10% of the arithmetic average or median of the criterion values; v – is defined as roughly 30% of the arithmetic average or median of the criterion values and not more than 40% of the minimal value of the criterion range.

Table 4 – part (b) demonstrates the preference modeling technique characteristic for the AHP method. The importance of criteria is defined through pair-wise comparisons between particular characteristics (K1, K2, …, K9), which constitute the matrix of relative weights $\boldsymbol{w_r}$. As one can see the relative weights $\boldsymbol{w_r}$ express the strength of one element (criterion) against another on a standardized scale (1–9 points). The compensatory character of the data is demonstrated, i.e. the values characterizing the less important element (e.g. K6 against K5 - 1/3 or K7 against K3 - 1/2) is the inverse of the value characterizing the more important element in the compared pair (K5 against K6 – 3 or K3 against K7 – 2). As one can see the rows of the matrix featured by substantial number of relative weights $\boldsymbol{w_r}$, being natural numbers larger than 1 reveal criteria of the greatest importance (e.g. criteria K8, K1 and K3). At the same time the least important criteria are described by such rows of the matrix that include fraction values, mainly (e.g. criteria K2 and K9). The absolute weights $\boldsymbol{w_a}$, computed according to the AHP algorithm, represent the overall measures of criteria importance expressed in the form of the vector. As indicated in the $\boldsymbol{w_a}$ column the most important criteria, characterized by the largest values of absolute weights, are K8 ($\boldsymbol{w_a} = 0.160$), K1 ($\boldsymbol{w_a} = 0.157$) and K3 ($\boldsymbol{w_a} = 0.156$). This confirms the previously described status of these criteria and proves quantitatively that these three criteria are substantially more important than the others and at the same time quite close one to another in terms of importance. Based on the same computations one may conclude that criteria K2 and K9 are the least important characteristics, with $\boldsymbol{w_a} = 0.033$ and $\boldsymbol{w_a} = 0.036$, respectively. Based on the methodological principles of the

AHP method further modeling of the DM's and stakeholders preferences, including analysis of their sensitivity, corresponds to the construction of similar matrices of relative weights $\boldsymbol{w_r}$ at all levels of the pre-defined hierarchy, i.e. for sub-criteria and variants. In the analyzed case, for the proposed structure of 7 variants, 9 criteria and 2–4 sub-criteria within particular criteria, this has resulted in the construction of 23 matrices of relative weights $\boldsymbol{w_r}$ and corresponding 23 vectors of absolute weights $\boldsymbol{w_a}$. In each of the matrices at the lower levels of hierarchy the variants have been pair-wised compared for all criteria and/or sub-criteria.

3.2.5 Evaluation and Selection of Appropriate MCDM/A Methods. Major Features of the AHP and ELECTRE III/IV Methods Used in the Computational Experiments

Analysis, evaluation and selection of the MCDM/A methods. Based on different publications [6, 8, 10] and similarly to the approach presented in other chapters of this book [37] the author of this paper have carried out a comprehensive analysis of different MCDM/A methods. He has examined three major aspects characterizing MCDM/A methods, such as [6]:

(1) How the features of the method correspond to the type, scope and the specific character of the considered decision problem (DP)?
(2) How easy and accurate can the method model the verbally expressed preferences of the DM?
(3) What is the form and reliability of generated results (final rankings)? Are they consistent with the DM's expectations?

Taking into account that the considered decision problem (DP) belongs to the category of multiple criteria ranking problems the author has reviewed a set of several multiple criteria ranking methods, capable of solving the tram selection problem. He has considered the following computational procedures: Electre III/IV, Oreste, AHP and UTA methods.

The results of his considerations have been presented in Table 5 in the abbreviated form. For each of the above mentioned major evaluation aspects certain detailed issues/questions (3–5) have been considered. As a result 11 sub-aspects have been used in the comparison of 4 MCDM/A ranking methods. Positive match between the method characteristics and the features of the decision situation as well as the expectations of the DM is marked by ✔. As presented in Table 5 Electre III/IV method is featured by the highest compatibility with the considered decision situation and DM's expectations (9 positive matches), followed by equally good AHP and Oreste methods (8 positive matches). UTA method is characterized by the lowest number of strengths (5 positive points).

Table 5. The analysis of the suitability of the MCDM/A methods for solving the considered decision problem

Fitness to the	Features	Electre III/IV	Oreste	AHP	UTA
Type, scope & specific character of the DP	The category of DP - ranking problem	✔	✔	✔	✔
	Small size of the set of variants	✔	✔	✔	
	Deterministic character of input information	✔	✔	✔	✔
	Cardinal character of input information	✔	✔	✔	✔
	Structure of criteria (subcriteria)			✔	
Way of modeling and aggregating DMs' preferences	High precision of the preference model	✔		✔	
	Low labor intensity of modeling preferences	✔	✔		✔
	User friendliness of the decision support process	✔	✔	✔	
DMs' expectations regarding the final ranking	Ordinal, graphical form	✔	✔		
	Cardinal, graphical form			✔	✔
	Incomparability of variants	✔	✔		

Due to the fact that Electre III/IV and Oreste methods belong to the same methodological school of MCDM/A, based on the outranking relation, it was not recommended to use both of them in the same computational experiment. For sake of objectivity the application of a computational method based on a different methodological background sounded more appropriate. Thus, Electre III/IV and AHP methods have been selected and used in the computational phase. An additional argument for the AHP use, was its ability to reflect the hierarchical structure of evaluation characteristics with criteria and sub-criteria in its original form, with evaluations based on the raw data. This, has allowed the author to compare the computational results generated by two methods (Electre III/IV and AHP) based on different methodological principles, alternative ways of defining and structuring criteria and at the same time featured by the highest compatibility with the considered decision situation (tram selection problem) and the DM's expectations.

Major characteristics of the AHP Method [22, 23]. The AHP (Analytic Hierarchy Process) method is a multiple objective ranking procedure, proposed by Saaty [22], focused on the hierarchical analysis of the decision problem. The method is based on the multiattribute utility theory [15] and allows to rank a finite set of variants $\boldsymbol{A}$. Through the definition of the overall objective, evaluation criteria, subcriteria and variants the method

constructs the hierarchy of the decision problem. On each level of the hierarchy, based on the pair-wise comparisons of criteria, subcriteria and variants, the DM's preferential information is defined in the form of relative weights $\boldsymbol{w_r}$ [22]. Each weight represents relative strength of the compared element against another and it is expressed as a number from 1 to 9. All weights have a compensatory character, i.e.: the value characterizing the less important element (1/2, 1/5, 1/9) is the inverse of the value characterizing the more important element in the compared pair (2, 5, 9).

The algorithm of the AHP method focuses on finding a solution for a, so called, eigenvalue problem [22] on each level of the hierarchy. As a result a set of vectors containing normalized, absolute values of weights $\boldsymbol{w_a}$ for criteria, subcriteria and variants is generated. The sum of the elements of the vector is 1 (100%). The absolute weights $\boldsymbol{w_a}$ are aggregated by an additive utility function. The utility of each variant $\boldsymbol{i - U_i}$ is calculated as a sum of products of absolute weights $\boldsymbol{w_a}$ on the path in the hierarchy tree (from the overall goal, through criteria and subcritearia) the variant is associated with. The utility $\boldsymbol{U_i}$ represents the contribution of variant $\boldsymbol{i}$ in reaching an overall goal and constitutes its aggregated evaluation that defines its position in the final ranking.

The important element of the AHP algorithm is the investigation of the consistency level of matrices of relative weights $\boldsymbol{w_r}$ on each level of hierarchy. Through the calculation of a, so called, consistency index ***CI*** one can measure how consistent is the preferential information given by the DM. If the value of ***CI*** is close to 0 the preferential information given by the DM is considered to be almost perfect. The acceptable level of ***CI*** is below 0.1

Major characteristics of the ELECTRE III/IV Method [7, 20, 21]. The Electre III/IV method belongs to a family of multiple criteria ranking procedures based on the outranking relation [20, 21]. It generates final rankings of a finite set of variants and orders them from the best to the worst, taking into account the following relationships between variants: indifference (I), preference (P) or (>), non-preference (P ~) or (<) and incomparability (R).

To start the computational procedure of the Electre III/IV method the following input data is required: matrix of performances, which comprises the evaluation of each variant by each criterion and the DM's preference model constructed with the application of indifference q_j, preference p_j and veto v_j thresholds as well as the weight w_j, defined for each criterion. The thresholds define the sensitivity of the DM to the changes of the criteria values and the weight w_j, expresses the importance of each criterion.

The computational procedure starts with [20, 21] the calculation of concordance indicators $\boldsymbol{C(a, b)}$ for each pair of variants $\boldsymbol{a}$ and $\boldsymbol{b}$. They are presented as a concordance matrix. $\boldsymbol{C(a, b)}$ expresses the extent to which the scores of $\boldsymbol{a}$ and $\boldsymbol{b}$ on all criteria are in concordance with the proposition $\boldsymbol{a}$ outranks $\boldsymbol{b}$. Next, the discordance index $\boldsymbol{D_j(a, b)}$ is calculated for each criterion $\boldsymbol{j}$, taking into account arguments represented by the values of criteria for which one cannot accept the statement that $\boldsymbol{a}$ outranks $\boldsymbol{b}$. Finally the outranking relation $\boldsymbol{S}$ is constructed. $\boldsymbol{S}$ indicates the extent to which "a outranks b" overall. It is expressed by the degree of credibility $\boldsymbol{d(a, b)}$.

Based on $\boldsymbol{d(a, b)}$ the method establishes two preliminary rankings (complete preorders) using a classification algorithm (distillation procedure). During this procedure one can obtain a descending and an ascending preorder. In the descending distillation the

variants are ranked from the best to the worst, while in the ascending distillation they are ranked in the inverse order. The intersection of two preorders gives the final ranking. Further details of the computational algorithm applied in Electre III/IV method can be found in the works of Figueira et al. [7], Roy [20, 21], and Vincke [27].

3.2.6 Computational Experiments and Analysis of Generated Results. Selection of the Most Desired Solution

After having structured the decision problem, modeled preferences and selected the appropriate decision methods the computational experiments have been performed. They have been carried out according to the computational algorithms of ELECTRE III/IV and AHP methods.

In the first case (application of ELECTRE III/IV method) concordance indicators $C(a, b)$ and discordance indexes $D_j(a, b)$ for all criteria j have been calculated. Based on these results the outranking relation $S(a, b)$ has been constructed and expressed in the form of the degree of credibility $d(a, b)$ that indicates the extent to which "a outranks b" overall. The values of $d(a, b)$ from the 0-1 interval have been placed in the credibility matrix.

Based on $d(a, b)$ two preliminary rankings (complete preorders), called descending and ascending distillations, have been established. Their aggregation (intersection) have resulted in the generation of a final ranking.

In the second case (application of AHP method) the consistency indexes CI for all matrices of relative weights w_r (representing the preferences of DM for criteria, subcriteria and variants) have been computed. Their values have ranged between 0.0 and 0.04, which is a very good indication of the consistency of the preferential information. In all considered cases (matrices) the values of CI-s were below the acceptable threshold of 0.1 ($CI < 0.1$). Thus, the redefinition of the preferential information was not required and further stages of the computational procedure have been performed.

As a result vectors of absolute values of weights w_a for criteria, subcriteria and variants have been generated (see an example for criteria in Table 4 (b) and aggregated by an additive utility function. The utility of each variant $i - U_i$ has been computed and used to define the position of each variant in the final ranking. The variants featured by higher utility have been placed before the variants characterized by lower utility.

Due to space limitation the intermediary steps of the computational procedure have not been presented in this chapter. Only the final results – rankings have been demonstrated. The interested reader can find all steps of the ELECTRE III/IV and AHP procedures in other publications by Żak [6, 37, 39].

Figure 2 demonstrates the final ranking of low-floor trams (variants A0001, A0002, …., A0007) generated by the ELECTRE III/IV method in the graphical and tabular form. The Final Graph and the corresponding Ranking Matrix generated by the ELECTRE III/IV method show the positions and the mutual relationships between variants. The arrows on the graph denominate the preference relation (P) between variants. The variant above its counterpart is preferred against it (e.g. A0004 P A0006 and A0003 P 0001). At the same time non-preferred variant is placed below its competitor. This means that the latter outperforms the former, e.g.: A0003 P* A0005 and A0002 P* A0006. The variants that are indifferent (I) are placed in the same rectangular box in the graph and their relation is marked I in the Ranking Matrix, e.g. A0002 and A0005. As indicated in the final ranking

the ELECTRE III/IV method allows for taking into consideration the incomparability (R) between variants. This is the case of variants A0003 and A0007 and A0001 and A0007. These variants are not connected by any arrow in the graph and their mutual relations are marked R in the ranking matrix. The winning variant A0004 – tram PESA is placed at the top of the graph and its row in the ranking matrix is featured by all P-s (except itself - I), indicating that A0004 is preferred against all the remaining variants. At the same time variant A0001 – tram ALSTOM is considered to be the least desirable option, and as such, placed at the bottom of the ranking. Its scores in the Ranking Matrix are mainly P*, which corresponds to the fact that it is not preferred against the remaining variants. The ELECTRE III/IV method-based ranking does not reveal the distance between variants. The distance measure is not computed in any form.

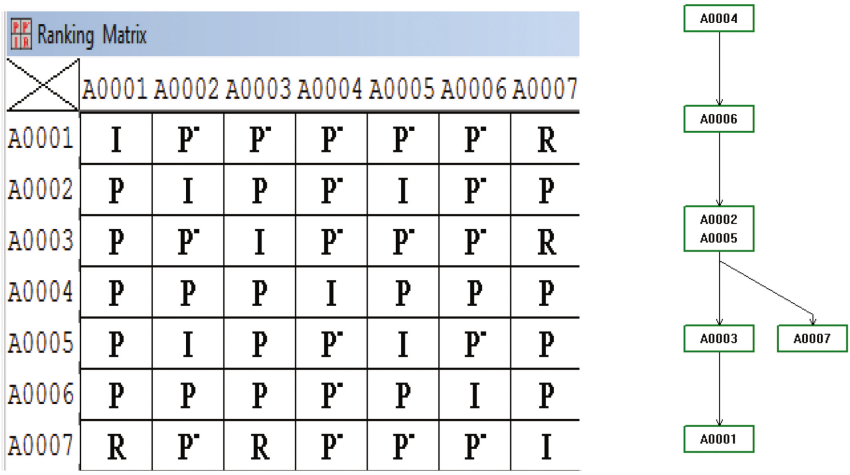


Fig. 2. Final results of the computational experiment carried out with the application of the ELECTRE III/IV method. Final graph and the corresponding ranking matrix

The final results generated by the ELECTRE III/IV method show that the single winner of the ranking is variant A0004 – tram PESA, followed by variant A0006 – tram SOLARIS. At the other end of this hierarchy one can find variant A0001 – tram ALSTOM, which is considered the weakest option in the considered tram competition. In the ranking matrix the row of variant A0001 is filled in by all P* (non-preference) except one (incomparability – R). Variant A0007 – tram STADLER is associated with the similar weaknesses (4 P*-s and 2 R-s).

Figure 3 presents the final ranking of the same trams (variants A0001, A0002, ….., A0007) generated by the AHP method. The ranking is constructed based on the aggregated utility of each variant, which makes the comparison between variants quite transparent. All trams are comparable (indifferent or preferred) and the distance between particular options is clearly shown. As presented in the chart, in the considered case, the overall utilities of the trams range between 0.105 and 0.211. The final scores allow for ranking them from the best (variant A0004 – Tram PESA) to the worst (variant A0002 – Tram ANSALDO BREDA). The colorful bars show the position of each variant –

tram in the ranking and its quantitative overall evaluation. The contribution of each criterion (K1, K2, …, K9) to the final score of each variant is different and indicated by various colors. Based on the ranking generated by the AHP method the winning position of variant A0004 – Tram PESA is undisputable. Variant A0004 substantially outperforms all other variants, including variants A0001 – tram ALSTOM and A0006 – tram SOLARIS, placed at the second and third positions in the ranking, by 30% and 47%, respectively. The distance between the winner of the ranking and its closest competitors is clearly visible.

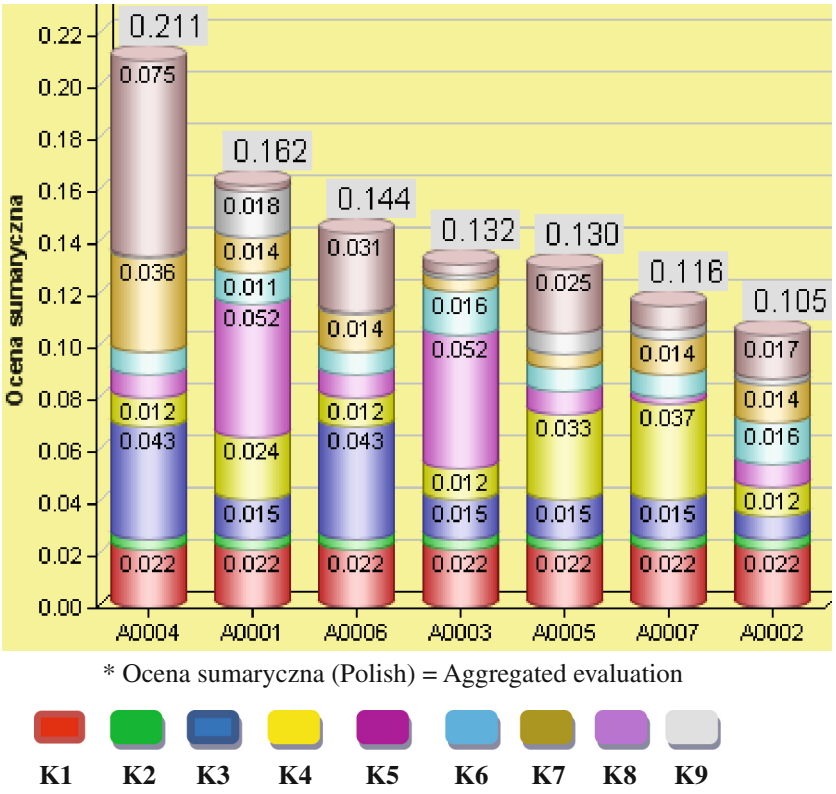


Fig. 3. Final results of the computational experiment carried out with application of the AHP method. Ranking of variants – low-floor trams based on the values of their utilities U_i computed by the AHP method algorithm

Based on both presented rankings, variant A0004 - tram PESA (first positions in both rankings) is recommended as the most desirable solution. The strongest advantages of this variant are as follows: Reliability (criterion K3), Environmental Friendliness (criterion K6), Utility and Functionality (criterion K7) and Price (criterion K8). The winner of the ranking is also excellent on such characteristics as Comfort (criterion K1) and Traction (criterion K2). PESA is an excellent tram by all means. It satisfies the interests of different stakeholders at the highest degree and fulfils various expectations and

requirements. Variant A0004 does not have weak points – serious disadvantages. Its performance on all considered characteristics is very good, placing the tram at least at the third position on each criterion. It is also worth noticing that variant A0004 is featured by excellent performance on three criteria that are the most important characteristics for the DM, such as: K8 – Price, K1 – Comfort and K3 – Reliability. Two of them contribute substantially to its overall utility and the resulting winning position in the ranking (see colors: dark purple – K8 and dark blue – K3 in Fig. 3).

4 Conclusions

The paper presents the theoretical background and practical application of MCDM/A methodology in transportation. Basic notions and concepts of MCDM/A are described. Their transportation – oriented interpretation is given. The chapter demonstrates also the universal procedure of solving multiple criteria decision problems and shows how this procedure should be applied in real world case studies. In the case study a multiple criteria fleet selection problem is analyzed and solved.

From a methodological point of view the major output of this research is the presentation of basic definitions, concepts and rules of MCDM/A and their explanation in a transportation – oriented language. The author clarifies basic MCDM/A terms in a reader friendly manner and provides their transportation interpretation. Thus, a complex and broad problematic is presented to a community of transportation researchers, analysts and consultants in a condensed manner. The author claims that the MCDM/A methodology can be applied in a variety of transportation – oriented decision situations. He demonstrates, further, the applicability of different MCDM/A methods in transportation. He characterizes shortly two multiple criteria ranking methods, i.e.: ELECTRE III/IV and AHP and shows a set of computational experiments in which both methods are applied. They are used to solve a multiple criteria fleet selection problem. The results of multiple criteria rankings of low-floor trams, generated by both methods, are compared.

It is worth emphasizing that a majority of transportation decision problems have a multiple criteria character [31] and thus they should be solved with the application of MCDM/A methodology. The author of this paper has presented the application of MCDM/A methods and rules in such instances as: fleet replacement problem [32], fleet composition problem [18, 40], fleet classification problem [24]; crew sizing problem [36], vehicle routing problem [32], vehicle and network assignment problems [5, 38], transportation projects evaluation problem [39].

From a practical point of view the important value added of the chapter is a demonstration of all phases of the solution procedure of a real world multiple criteria decision problem, i.e. a tram ranking problem. The author shows in detail all phases of this process, including: investigation of the decision problem and its verbal description; problem structuring and mathematical modeling; selection of the appropriate solution methods; computational experiments; final selection of the most desirable solution. The author proves that MCDM/A rules can be successfully applied to solve complex, multiple criteria decision problems, such as fleet selection problem. The generated

results show that the recommended solution – tram PESA substantially outperforms other competitors. It guarantees that the purchase price of the tram is reduced by 15% and 54% in relation to the second cheapest and the most expensive competitor, respectively. For the considered decision situation (purchase of 40 trams) it generates savings of approximately 48 mln zl (11 mln Euro) – 332 mln zł (77 mln Euro), respectively. Using the benchmark of the second (better of the pair: A0001 – ALSTOM and A0006 – SOLARIS) and the last (worse of the pair A0002 – ANSALDO BREDA and A0001 – ALSTOM) tram in the generated rankings, tram PESA provides 14% and 300% improvement reliability-wise and 27% and 158% improvement in terms of utility and functionality. Its overall score (generated by the AHP method) is 30% better than the score of the second tram in the ranking (A0001 – ALSTOM). These results strongly support the statement that MCDM/A has practical applicability in transportation and can generate improvements and enhancement of the comfort of travel, cost levels, fleet utilization, safety and security, fleet reliability and other parameters.

References

1. Belton, V., Stewart, T.: Multiple Criteria Decision Analysis. An Integrated Approach. Kluwer Academic Publishers, London (2002)
2. Benayoun, R., Montgolfier, J., Tergny, J., Laritchev, O.: Linear programming with multiple objective functions: step method (STEM). Math. Program. **1**, 366–375 (1971)
3. Brans, J., Vincke, P., Mareschal, B.: How to select and how to rank projects: the PROMETHEE method. Eur. J. Oper. Res. **24**, 228–238 (1986)
4. Crainic, T., Laporte, G.: Planning models for freight transportation. Eur. J. Oper. Res. **97**(3), 409–438 (1997)
5. Czyżak, P., Żak, J.: A model of an urban transportation system formulated as a multiobjective mathematical programming problem under uncertainty. J. Adv. Transp. **29**(1), 43–62 (1995)
6. Fierek, S., Żak, J., Solecka, K., Kruszyński, M.: Multiple criteria evaluation of the mass transit systems in European cities. Logistyka **2**, 509–522 (2012). CD 2
7. Figueira, J., Greco, S., Ehrgott, M. (eds.): Multiple Criteria Decision Analysis. State of the Art Surveys. Springer, New York (2005)
8. Frederic, A., Patrick, H.: Encyclopedia of Decision Making and Decision Support Technologies. Hershey, New York (2008)
9. Garett, M. (ed.): Encyklopedia of Transportation: Social Science and Policy. Sage Publishing, Los Angeles (2014)
10. Guitouni, A., Martel, J.: Tentative guidelines to help choosing an appropriate MCDA method. EJOR **109**(2), 501–521 (1998)
11. Hillier, F., Lieberman, G.: Introduction to Operations Research. McGraw-Hill, New York (2005)
12. Jacquet-Lagreze, E., Siskos, J.: Assessing a set of additive utility functions for multicriteria decision-making, the UTA method. EJOR **10**, 151–164 (1982)
13. Janic, M.: Multicriteria evaluation of high speed rail, transrapid maglev and air passenger transport in Europe. Transp. Planning and Technol. **26**(6), 491–512 (2003)
14. Jaszkiewicz, A., Słowiński, R.: The light beam search approach – an overview of methodology and applications. Eur. J. Oper. Res. **113**(2), 300–314 (1999)
15. Keeney, R., Raiffa, H.: Decisions with Multiple Objectives. Preferences and Value Tradeoffs. Cambridge University Press, Cambridge (1993)

16. Macharis, C., De Witte, A., Ampe, J.: The multi-actor, multi-criteria analysis methodology (MAMCA) for the evaluation of transport projects: theory and practice. J. Adv. Transp. **43**(2), 183–202 (2009)
17. Michalewicz, Z., Fogel, D.: How to Solve It: Modern Heuristics. Springer, Heindelberg (2000)
18. Redmer, A., Żak, J., Sawicki, P., Maciejewski, M.: Heuristic approach to fleet composition problem. Procedia Soc. Behav. Sci. **54**, 414–427 (2012)
19. Roy, B.: Decision-aid and decision making. Eur. J. Oper. Res. **45**, 324–331 (1990)
20. Roy, B.: Multiple Criteria Decision Aiding. Wydawnictwo Naukowo-Techniczne, Warszawa (1990). (in Polish)
21. Roy, B.: The outranking approach and the foundations of Electre methods. In: Bana e Costa, C.A. (ed.) Readings in Multiple Criteria Decision Aid, pp. 155–183. Springer, Heidelberg (1990)
22. Saaty, T.: The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation. McGraw-Hill, New York (1980)
23. Saaty, T.: Transport planning with multiple criteria: the Analytic Hierarchy Process applications and progress review. J. Adv. Transp. **29**(1), 81–126 (1995)
24. Sawicki, P., Żak, J.: Technical diagnostic of a fleet of vehicles using rough set theory. Eur. J. Oper. Res. **193**(3), 891–903 (2009)
25. Steuer, R.: Multiple Criteria Optimization: Theory, Computation and Application. Wiley, New York (1986)
26. Toloie-Eshlaghy, A., Homayonfor, M.: MCDM methodologies applications: a literature review from 1999 to 2009. Res. J. Int. Stud. **21**, 86–137 (2001)
27. Vincke, P.: Multicriteria Decision-Aid. Wiley, Chichester (1992)
28. Vuchic, V.: Urban Transit: Operations, Planning and Economics. Wiley, New York (2005)
29. Webster's Encyclopedic Dictionary of the English Language. Lexicon Publications, New York (2005)
30. Żak, J.: The MCDA methodology applied to solve complex transportation decision problems. In: Proceedings of the 9th Meeting of the EURO Working Group on Transportation: Intermodality, Sustainability and Intelligent Transportation Systems, Bari, 10–13 June, pp. 685–693 (2002)
31. Żak, J.: Identification of the most important road transportation decision problems. Arch. Transp. **16**(2), 89–109 (2004)
32. Żak, J.: Multiple Criteria Decision Aiding in Road Transportation. Poznań University of Technology Publishers, Poznań (2005). (in Polish)
33. Żak, J.: Application of operations research techniques to the redesign of the distribution systems. In: Dangelmaier, W., Blecken, A., Delius, R., Klöpfer, S. (eds.) IHNS 2010. LNBIP, vol. 46, pp. 57–72. Springer, Heidelberg (2010). doi:10.1007/978-3-642-12494-5_6
34. Żak, J.: Decision support systems in transportation. In: Jain, L., Lim, C. (eds.) Handbook on Decision Making – Vol. 1: Techniques and Applications. Intelligent Systems Reference Library, vol. 4, pp. 249–294. Springer, Heidelberg (2010)
35. Żak, J.: The methodology of multiple criteria decision making/aiding in public transportation. J. Adv. Transp. **45**(1), 1–20 (2011)
36. Żak, J.: The methodology of multiple criteria decision making/aiding as a system-oriented analysis for transportation and logistics. In: Świątek, J., Tomczak, J.M. (eds.) ICSS 2016. AISC, vol. 539, pp. 265–284. Springer, Cham (2017). doi:10.1007/978-3-319-48944-5_25
37. Żak, J., Galińska, B.: Multiple criteria evaluation of suppliers in different industries – comparative analysis of three case studies. In.: Żak, J., Hadas, Y., Rossi, R. (eds.) Advanced Concepts, Methodologies and Technologies for Transportation and Logistics. AISC, vol. 572, pp. 121–155. Springer, Cham (2017)

38. Żak, J., Jaszkiewicz, A., Redmer, A.: Multiple criteria optimization method for the vehicle assignment problem in a bus transportation company. J. Adv. Transp. **43**(2), 203–243 (2009)
39. Żak, J., Kruszyński, M.: Application of AHP and ELECTRE III/IV methods to multiple level, multiple criteria evaluation of urban transportation projects. Transp. Res. Procedia **10**, 820–830 (2015)
40. Żak, J., Redmer, A., Sawicki, P.: Multiple objective optimization of the fleet sizing problem for road freight transportation. J. Adv. Transp. **45**(4), 321–347 (2011)

The Multi-Actor Multi-Criteria Analysis (MAMCA) Tool: Methodological Adaptations and Visualizations

Sheida Hadavi(✉), Cathy Macharis, and Koen Van Raemdonck

Mobi Research Group, Vrije Universiteit Brussel,
Pleinlaan 2, 1050 Brussels, Belgium
{sheida.hadavi,cathy.macharis,koen.van.raemdonck}@vub.ac.be

Abstract. In order to aid groups in complex decision-making processes, the Multi-Actor Multi-Criteria Analysis (MAMCA) methodology has been developed. The MAMCA methodology differs from the classical approaches of Multi-Criteria Decision-Analysis (MCDA) in the fact that the different actors who are involved in a project, the so-called stakeholders, are explicitly involved throughout the steps. In this research, after a brief summary of the MAMCA methodology, we will present the MAMCA software, an interactive web tool which was established based on the aforementioned methodology. In the software, several visualizations are provided to aid decision makers with analysis of the problem at hand. Furthermore, we discuss how we adapted the PROMETHEE method to provide a comprehensive overview on the performance of the different solutions to the problem, so-called alternatives. The software and the visualizations are demonstrated with an example [13], a case study for the choice of an Urban Consolidation Center (UCC) in Brussels.

Keywords: MCDA · MAMCA · PROMETHEE

1 Introduction

In order to achieve more sustainable solutions, in addition to the economical effects, the ecological, spatial and social impacts of a project should be considered [2,17,28]. The MCDA approach makes it possible to evaluate several alternatives on various quantitative and qualitative criteria [22]. A further step in the evolution of appraisal methods is the explicit introduction of the stakeholder notion in the analysis. A stakeholder is by definition any individual or group of individuals that can influence or are influenced by the decision [10].

The MAMCA methodology makes these stakeholders explicit in the appraisal methodology. MAMCA is a methodology to evaluate, amongst others, transport projects, and aid groups in complex decision-making processes. In order to include the stakeholders' opinions into the decision-making process, already in the very beginning, during the problem formulation phase, the stakeholders

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_2

are identified together with the possible alternatives that can be considered. Thereafter, the stakeholders' criteria and priorities will be gathered. Finally, the alternatives will be evaluated on the gathered criteria. Hence, the MAMCA methodology adds an extra layer to the traditional MCDA method, namely the actor layer. In other words, per stakeholder a Multi-Criteria Analysis (MCA) model is built. All these models are aggregated in the final step.

The MAMCA methodology has already proven its effectiveness in evaluating complex sustainable mobility and transport policy decisions [20]. However, in order to create an accessible and global multi-actor multi-criteria analysis application, the MAMCA software has been developed.

The MAMCA software, follows the steps of the MAMCA methodology. First, a project is set up with a goal in mind and the alternatives are defined. Thereafter, the stakeholders are identified. Various stakeholder groups with different criteria per group can be created in the MAMCA software. While several representatives can be added to each group, the concept of participants has been introduced to the MAMCA software, to allow large groups, such as citizens, to have a direct impact on the decision by participating in a survey.

Allocating weights is possible with three methods: pairwise comparison (introduced by Saaty in [23]), equalization and manual allocation. The Preference Ranking Organization Method for Enrichment Evaluations (PROMETHEE) method, developed by Brans in [3] and further extended by Brans and Vincke in [5], is provided for evaluation in the MAMCA software. Alternatively, Analytical Hierarchy Process (AHP) method, introduced by Saaty in [23] is provided as well. Moreover, diverse visualizations have been adopted to assist stakeholders in analyzing different decisions.

During the development of the software, further effort has been put into the adaption of the PROMETHEE method. In this paper, we propose an alternative series of steps to calculate the final net scores of the alternatives. The proposed steps, reach the same net scores as calculated by the original PROMETHEE method. The benefit gained by the proposed steps is that we receive a performance indicator for each alternative on all criteria, providing a comprehensive analysis on the alternatives.

In Sect. 2, some literature in the transport sector, which use MCDA methods in their appraisal are reviewed and some related software packages are discussed. The theoretical foundations of the MAMCA methodology will be reviewed in Sect. 3. Thereupon, the MAMCA software is established and introduced in Sect. 4. Herein, the focus will lie on the provided visualizations and the adaption of the PROMETHEE technique. Finally, in Sect. 5 we will discuss the future necessary steps and we will conclude in Sect. 6.

In this paper, in order to present some features of the tool and discuss the provided visualizations, we will take a case study for the choice of an Urban Consolidation Center (UCC) in Brussels from [13], as an example. In the example, the MAMCA software has been used as a methodology to consider the interest

of the stakeholders while evaluating the different implementation scenarios in the Brussels-Capital Region[1].

2 Related Work

The application of MCDA in the transport sector has a broad scope, ranging from the evaluation of policy measures in passenger transport [2], to strategic decisions [8,26], technologies [19,25], locations [15,24], and finally infrastructure projects [7]. Żak and Thiel in [28] give a comparison of several MCDA methods to solve mass transit systems' decision problems. For an overview of transport applications and applied MCDA methods see [17]. For the transport sector many different stakeholders are often involved [27,28] such as: users of the transportation, owners, managers and employees of the transportation companies, and responsible authorities for the transport operations. Not taking into account the divergent objectives of stakeholders, will make implementation of the selected alternative impractical.

PROMCALC [4] was the first published software by Mareschal and Brans, the authors of the PROMETHEE method and the GAIA visual module [6]. PROMCALC, published in 1990's, was replaced by the Decision Lab [9] developed in collaboration with the Canadian company Visual Decision in 2000s [1]. Thereafter, Visual PROMETHEE [9] has been introduced in 2010, which has spreadsheet interface, complete and partial rankings and sensitivity analysis for evaluation of PROMETHEE rankings' robustness. D-Sight Collaborative Decision-Making (CDM) [11] is a web Group Decision-Support System (GDSS) tool published in February 2010. D-Sight CDM has PROMETHEE I and II rankings, the GAIA visualization tool, sensitivity analysis and a modern user interface. Expert Choice [12] is a GDSS based on AHP, evaluating the alternatives based on pairwise comparisons. The Expert Choice software is equipped with sensitivity analysis and uses three types of comparisons namely verbal, numerical and graphical to evaluate the alternatives.

3 MAMCA Methodology

In a classical MCDA approach the following steps are taken: problem definition, developing the alternatives, developing a set of criteria and an evaluation matrix of their importance, the general evaluation of the alternatives, and finally the implementation [7,21]. Moreover, various stakeholders are often involved in the decision-making process. It is important to explicitly involve stakeholders' objectives in the process and analysis. The MAMCA methodology differs from the classical approach of MCDA in the explicit introduction of stakeholders in a very early stage. These stakeholders will be key to identify and evaluate the criteria, which are here equal to the objectives of the stakeholders.

[1] Understanding of the example is not necessary in this article. It is uniquely used to present the software.

The first step in the MAMCA approach is the definition of the problem and the identification of the possible alternatives submitted for evaluation. The stakeholders are identified in the second step. Stakeholders are people who have an interest in the consequences of any decisions taken. An in-depth understanding of each stakeholder group's objectives is critical in order to appropriately assess the different alternatives. Stakeholder analysis should be viewed as an aid to properly identify the range of stakeholders who need to be consulted and whose views should be taken into account in the evaluation process.

The choice and definition of the evaluation criteria are primarily based on the identified stakeholder objectives and the purposes of the alternatives considered. In the MAMCA methodology, the criteria for the evaluation are the goals and objectives of the stakeholders, and not the effects or impacts of the actions per se as is usually done in a multi-criteria analysis. The given weights represent the importance the stakeholders attach to these objectives.

In the fourth step, for each criterion, one or more indicators are constructed: direct quantitative indicators such as money spent, number of lives saved, reductions in CO2 emissions achieved, etc. or scores on an ordinal indicator such as high/medium/low for criteria with values that are difficult to express in quantitative terms. The measurement method for each indicator is also made explicit; for instance for the "economic activity" criterion, the number of jobs created counted by the number of people is considered. This allows measuring of the performance of each alternative in terms of its contribution to the objectives of specific stakeholder groups. Steps 1 to 4 can be considered as mainly analytical, and they precede the 'overall analysis', which takes into account the objectives of all stakeholder groups simultaneously and is more 'synthetic' in nature.

The fifth step is the construction of an evaluation matrix, aggregating each alternative's contribution to the objectives of all stakeholders. Any MCDA-method can be used to assess the different strategic alternatives. The PROMETHEE method has, for example, been extended in [18], the Analytical hierarchy process (AHP) method in [23] and ELECTRE in [14]. These methods give each stakeholder group the liberty of having their own criteria, weights and preference structure and only at the end of the analysis the different points of view are confronted.

The multi-criteria analysis developed in the previous step eventually (in step six) leads to a classification of the proposed alternatives revealing the strengths and weaknesses of the proposed alternatives. A sensitivity analysis is performed in this stage in order to assess the stability of this ranking. More important than the ranking, the multi-criteria analysis reveals the critical stakeholders and their criteria. The multi-actor, multi-criteria analysis provides a comparison of different strategic alternatives, and supports the decision makers in making their final decision by pointing out for each stakeholder which elements have a clearly positive or a clearly negative impact on the sustainability of the considered alternatives.

The last stage of the methodology (step 7) includes the actual implementation. When the decision is made, steps have to be taken to implement the chosen

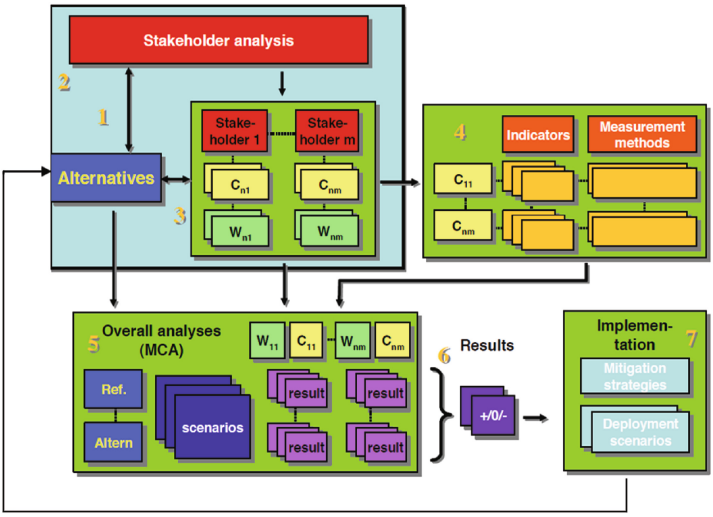


Fig. 1. Methodology for the multi-actor multi-criteria analysis (MAMCA) [16]

alternative by creating deployment schemes. The information on the points of view of each stakeholder, received from the previous steps, tremendously helps to define the implementation paths. The overall methodology of the MAMCA is shown in Fig. 1.

4 MAMCA Software

The MAMCA software is developed in Java and PHP and runs inside the GlassFish and Apache application containers, respectively. It contains three layers: the first layer is the data access layer, the second layer is the logic layer, programmed in Java, and the user interface is the third layer.



Fig. 2. MAMCA software

Figure 2 demonstrates a screen-shot of the MAMCA software. It displays the first page of the workspace after a project manager has logged in to the software and has selected a project. We can see the project title and the evaluation type (PROMETHEE or AHP). Afterwards, we can see the navigation bar which follows the steps defined in the MAMCA methodology.

The MAMCA tool can be accessed after logging in to the platform by a user-name and password. Thereafter, either a new project should be created or we can access an existing project. When creating a new project, the title, goal, description and evaluation type of the project should be identified.

4.1 Alternatives

In the first step, demonstrated in Fig. 2, we can define new alternatives, modify or remove them. The alternatives can be added one by one or imported from a Microsoft Excel sheet. The alternatives list can be exported as a Microsoft Excel sheet. In the presented project to implement an urban consolidation center in Brussels, six alternatives have been defined. Each alternative has a name and description, which is shown if the pointer stays on the alternative name in the table.

4.2 Stakeholder Analysis

In the second step, involved stakeholders are introduced. Four types of actors are introduced in the MAMCA tool: project manager, actors, participants and experts. Project managers are the creators of a project. They will define the project and the alternatives, which are the same for all the stakeholders. Thereafter, the project manager will introduce the involved stakeholders as groups of actors, participants or experts and will invite one or many people to each group.

Actors are stakeholder groups that are engaged in the project and are interested in the outcome of the project. Thus, apart from their defined task they will have access to the result of the project. Each actor group has different privileges. When creating the actor groups, it has to be defined: (a) if the actors in each group will add their own criteria or if they are pre-defined and added by the project manager, (b) if the evaluation table will be filled in by the actors or if it will be filled in by the project manager or experts.

In the MAMCA process, often the representatives of the identified stakeholder groups are invited. In order to find more equitable and sustainable solutions, it is necessary to involve all stakeholders, and not only their representatives, especially for large groups, such as citizens. The increasing digitalization of the world and the knowledge-based society ensure that people will increasingly be able to participate in the decision process. Even stronger, people will ask to be involved in these processes. In order to explicitly integrate a broad set of various and often conflicting points of view, the concept of participants has been introduced in the MAMCA software. Participants are people from large groups such as citizens, which are important to consider but are hard to contact in person. In order to include more stakeholders' opinions, the process has been made

brief, and the participant groups receive a link to fill in a pairwise comparison based on their priorities.

Furthermore, experts have been introduced to the MAMCA methodology to assist on the fourth step of MAMCA methodology, where indicators and measurement methods to evaluate alternatives are defined. For each stakeholder group introduced in the second step, we can define whether the stakeholder group has the expertise to fill in the evaluation table or if an expert will be invited to carry out the evaluation. Similar to the alternatives list, the actor groups list can be imported and exported. Figure 3 shows the actors for the UCC case project.

Group Name	Group Type	Criteria Definition	Evaluations Input	Parameter Lock	Action
Receivers	Actor	Project Manager	Project Manager	No	
Shippers	Actor	Actor	Project Manager	No	
LSP	Actor	Actor	Expert	No	
Citizens	Actor	Project Manager	Expert	No	
Authorities	Actor	Actor	Actor	No	
Citizens_Survey					
Citizens_Expert					
LSP_Expert					

Fig. 3. MAMCA software: stakeholder groups

4.3 Criteria and Weights

In the third step, per stakeholder group the criteria are defined. Each criterion has a name, group and description. It is important to realize that the criteria are different for each stakeholder group. The criteria are added by actors during the project or are predefined by the project manager. The criteria can be imported and exported as an Microsoft Excel sheet.

The fourth step is the weight allocation. In this step, three methods are provided to allocate a weight to each criterion. If we equalize the weights, the criteria groups will each get an equal weight, thereafter those weights will be equally divided to the criteria in that group. We can assign the weight manually. If the weights do not sum up to 1, a ratio of the weights will be saved. Finally, the pairwise comparison method introduced by [23] can be used to weigh the criteria.

4.4 Evaluation

The fifth step is the evaluation step. When creating the project, we must choose between AHP and PROMETHEE. If the chosen method is AHP, alternatives will be compared pairwise for each criterion. Please refer to [23] for further explanation on the AHP technique. If PROMETHEE is the chosen appraisal method, the evaluation table has to be filled in, indicating on a quantitative or qualitative scale how each alternative has scored on each criterion. The parameters for each criterion have to be defined as well. In the following, we will further discuss the PROMETHEE technique.

PROMETHEE Technique. PROMETHEE is an outranking method to rank a finite set of alternatives based on a set of often conflicting criteria. PROMETHEE I and PROMETHEE II has been originally introduced by Brans in [3]. Thereafter, several other versions of PROMETHEE have been introduced, which we will not review in this paper as they are not relevant to our research.

The basic principle of PROMETHEE II is based on the pairwise comparison of alternatives along each recognized criterion. In PROMETHEE II, each alternative is evaluated on different criteria, providing the evaluation table. In order to get the ranking of the alternatives, in addition to the evaluation table, three sets of information are required: the weight allocated to each criterion, the preference function for each criterion and whether the criterion is to be maximized or minimized. Brans and Vincke suppose that all the criteria have the same importance or the stakeholders will introduce normalized weights as their preference index [5]. The preference function takes the difference between the evaluations of each two alternatives, and gives a preference degree ranging from zero to one. Brans and Vincke in [5] proposed six basic types of preference function: usual criterion, U-shape criterion, V-shape criterion, level criterion, V-shape with indifference criterion and Gaussian criterion. Furthermore, for each criterion, the value of an indifference threshold: q, and the value of a preference threshold: p, and the value of an intermediate value between p and q, s has to be defined. The stepwise procedure for calculating the ranking of alternatives has been discussed in [3], and demonstrated, as in Fig. 4 by Behzadian et al. in [1].

4.5 Results

The sixth step which is the multi-actor tab, includes the visualizations and analysis which will be discussed in details in the following.

Multi-Actor Analysis. In Fig. 5 the scores of the alternatives derived from all stakeholder groups are aggregated in one chart providing the multi-actor view. The scores of alternatives for each group is an arithmetic mean of the alternatives' scores for each stakeholder in that group. As Fig. 5 reveals, there is not always a single best alternative available. On the other hand, we can realize Scenario 2b and 3b have performed better compared to the other alternatives. Scenario 2b has scored positive for all the stakeholder groups except for the Citizens. It has been the best alternative for the Shippers and LSP. Furthermore, it has scored positive for the Receivers and Authorities. It has to be investigated how to adapt Scenario 2b to better perform for the Citizens. Scenario 3b has been the best alternative for Citizens and Authorities, and has received positive scores for the LSP and Receivers. Further analysis can be applied to become aware of how to adjust Scenario 3b to better perform for the Shippers.

Multi-Actor Box Chart Analysis. Figure 6 provides a comprehensive overview on the alternatives' scores. It demonstrates the range of the scores each alternative has received and where the scores are more concentrated, while

Step 1. Determination of deviations based on pair-wise comparisons

$$d_j(a,b) = g_j(a) - g_j(b) \qquad (1)$$

Where $d_j(a,b)$ denotes the difference between the evaluations of a and b on each criterion.

Step 2: Application of the preference function

$$P_j(a,b) = F_j[d_j(a,b)] \qquad j=1,\ldots,k \qquad (2)$$

Where $P_j(a,b)$ denotes the preference of alternative a with regard to alternative b on each criterion, as a function of $d_j(a,b)$.

Step 3: Calculation of an overall or global preference index

$$\forall a,b \in A, \qquad \pi(a,b) = \sum_{j=1}^{k} P_j(a,b) w_j \qquad (3)$$

Where $\pi(a,b)$ of a over b (from 0 to 1) is defined as the weighted sum $p(a,b)$ of for each criterion, and w_j is the weight associated with jth criterion.

Step 4: Calculation of outranking flows/ The PROMETHEE I partial ranking

$$\phi^+(a) = \frac{1}{n-1} \sum_{x \in A} \pi(a,x) \qquad (4) \qquad \text{and} \qquad \phi^-(a) = \frac{1}{n-1} \sum_{x \in A} \pi(x,a) \qquad (5)$$

Where $\phi^+(a)$ and $\phi^-(a)$ denote the positive outranking flow and negative outranking flow for each alternative, respectively.

Step 5: Calculation of net outranking flow/ The PROMETHEE II complete ranking

$$\phi(a) = \phi^+(a) - \phi^-(a) \qquad (6)$$

Where $\phi(a)$ denotes the net outranking flow for each alternative.

Fig. 4. Stepwise procedure of PROMETHEE II [1]

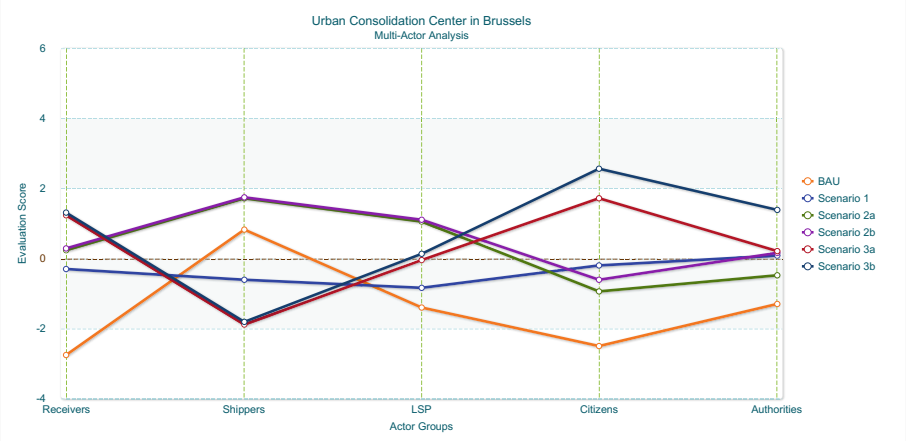


Fig. 5. MAMCA software: multi-actor analysis

not directly showing each stakeholder's opinion. This visualization is especially beneficial when we have a large number of stakeholders and stakeholder groups. The top and bottom quartiles of the alternative's scores are demonstrated by the lines, and the green and blue boxes show the middle top an middle bottom

quartiles respectively. Scenario 1 has a high concentration in the negative scores, representing that it has received negative scores from all the stakeholders. Scenario 2b has a rather high concentration in the positive scores, showing that it has received positive scores from most of the stakeholders. While Scenario 3b has a very high top score, the opinion of stakeholders has not been united about it, as the box has a large area.

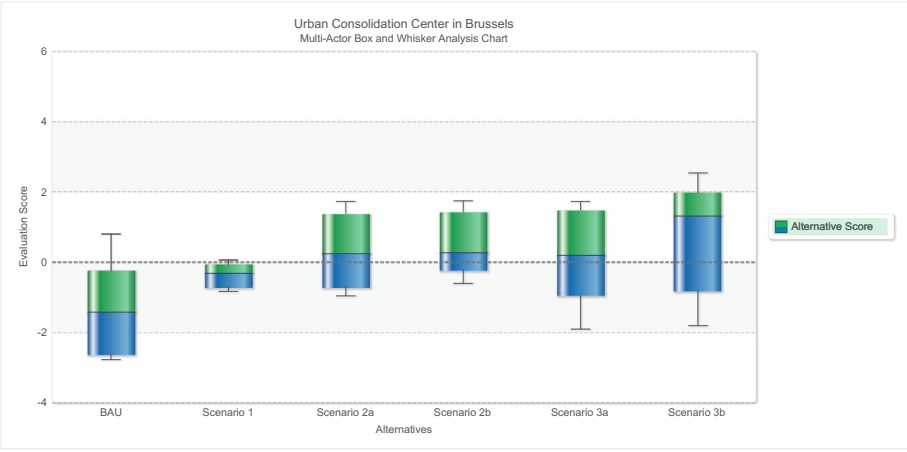


Fig. 6. MAMCA software: multi-actor box chart analysis

Alternatives Evaluation Analysis. In order to calculate alternatives' net scores with PROMETHEE II, a five-step procedure has been introduced by [3], demonstrated in Fig. 4 by Behzadian et al. [1]. In this research, we propose a four-step procedure to calculate the alternatives' net scores, presented in Fig. 7. In the proposed procedure, the first two steps are identical to the original steps.

The benefit of using the proposed procedure is that we evaluate the performance of each alternative on all criteria. In the third step of the procedure, a matrix, shown in Fig. 8, is built representing the performance of alternatives on the criteria set. Thereafter, in step four, the net score of each alternative is calculated summing up their performance on the criteria set, multiplied by the criteria weight. It is important to mention that the final net scores of alternatives calculated by the proposed four-step procedure are the same as the net scores calculated by the original five-step procedure.

Figures 9 and 10 exhibit results of the calculation based on the four-step procedure suggested in Fig. 7. In these figures, we can see how alternatives have performed over the criteria set of each stakeholder group. Figure 9 is from analysis of the Citizens. Scenario 2b has scored positive for all the stakeholder groups except Citizens. Thus, in Fig. 9 we can investigate how we can adapt this alternative in order to perform better for Citizens. While Scenario 2b has scored negative on almost all the criteria of this group, we assume if the noise problem would be addressed, the score of the alternative will change greatly.

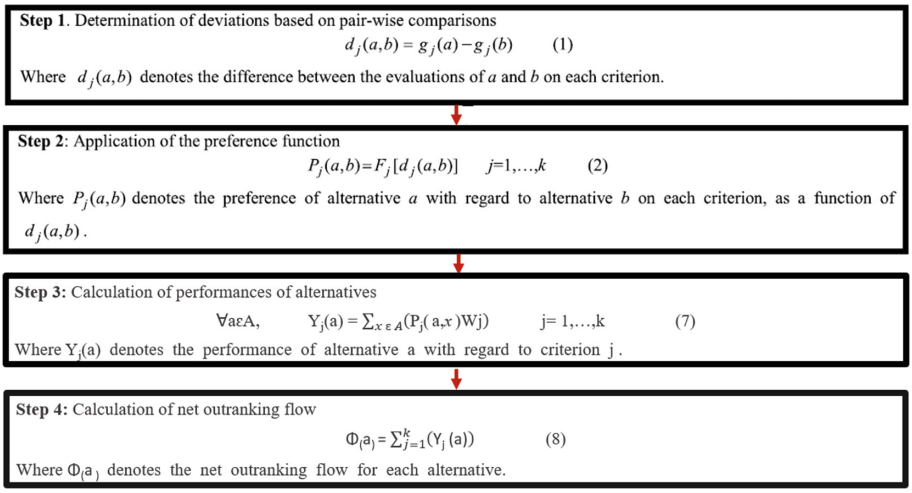


Fig. 7. Stepwise procedure of adapted PROMETHEE calculation (Step 1 and 2 from [1], Step 3 and 4 from own setup)

This hypothesis will be tested using sensitivity analysis in the following. Furthermore, in Fig. 10, we observe that the reason for scenario 3b receiving a negative score for the Shippers, while receiving a positive score for all the other stakeholder groups, is due to the high transport cost of this scenario. Thus, a solution such as a subsidy from the government, in order to lower the cost of this scenario for the Shippers, can provide an alternative that is satisfying to all stakeholders.

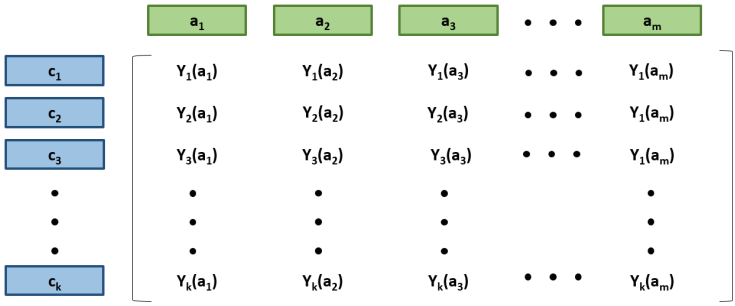


Fig. 8. The matrix representing the performance of alternatives on the criteria (own setup)

Sensitivity Analysis. In the MAMCA software, we provide two sensitivity analysis: weight analysis and evaluation analysis. In the weight analysis, using a flowing bar chart, criteria weights of different stakeholder groups can be changed

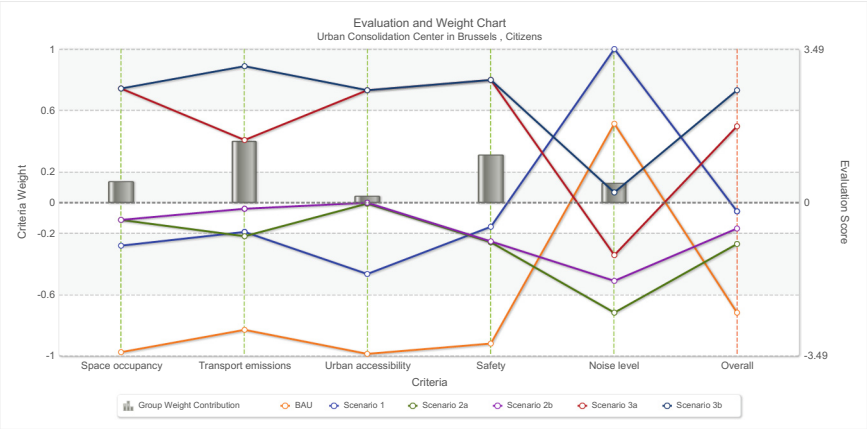


Fig. 9. MAMCA software: citizen's alternatives evaluation analysis

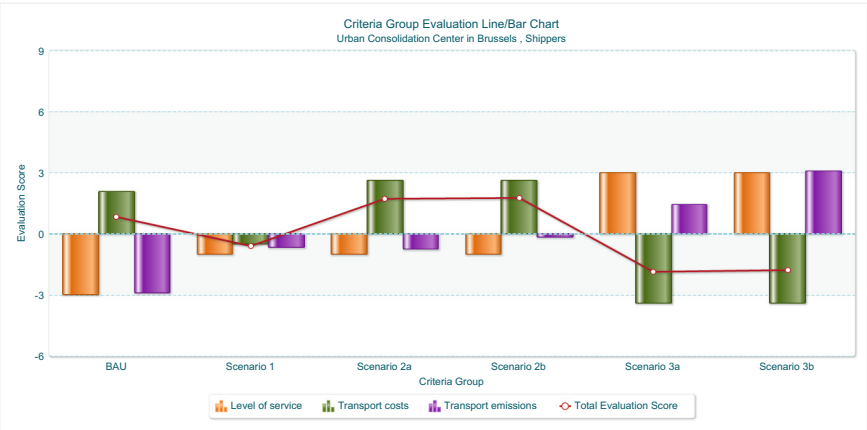


Fig. 10. MAMCA software: shipper's alternatives evaluation analysis

and the result of these changes are demonstrated in the multi-actor analysis chart. Hence, we can investigate how changing the weights of some of the criteria for one or more of the groups will affect the results in the multi-actor view. In the second case, evaluation analysis, we can adjust the "Alternatives Evaluation Analysis" chart of some of the stakeholder groups, and the result of this will be shown in the "Multi-Actor Analysis" chart. Changing the noise level of Scenario 2b for Citizens results in receiving slightly positive score for this scenario. Thus, we realize if the noise problem of the Scenario 2b will be solved, this scenario will have positive scores from all the stakeholders.

5 Discussion and Future Work

The use of the MAMCA software has made interactive workshops possible. In these interactive workshops, the focus is on the discussion, as well as on the weights of the criteria and the evaluation. Directly a common understanding of the problem and better insight is gained about the objectives and concerns of the other stakeholders. Finally, some template and guided projects, such as the city distribution game with 5 stakeholder groups, namely the shippers, the logistic service providers, the receivers, the citizens and the authorities can be used for training and exercise purposes.

A future work in the MAMCA development is to incorporate the "participation" concepts, by further enhancing the implemented "Participant" concept. The nature of the problem could be defined by stakeholders, others than the public officials and organized interests. Participation addresses the need for a dynamic approach facilitating mutual decision making accessible by a wider community. Herewith decision making goes beyond the existing political structures and dialogs.

Additional future work in the MAMCA software is integrating co-creation and co-design into the planning and decision-making process. Identification of alternatives with high levels of cooperation requires an interface facilitating identification and integration of the co-created alternatives into the evaluation process. Co-creation incorporates the formal structures of planning and evaluation, with the generally more informal process of identifying the alternatives.

6 Conclusion

In this paper, the MAMCA software was introduced, in order to aid groups in decision-making process and provide support to all the stakeholders. Software follows the steps of the MAMCA methodology, involving stakeholders explicitly in the appraisal process. Using the MAMCA software in interactive workshops, further enhances a mutual understanding between the stakeholders.

The softwares introduced in the related work section have either AHP or PROMETHEE for evaluation. The MAMCA software provides a choice of both of the algorithms. In the reviewed softwares, several stakeholder groups can be involved, but the criteria are the same for all the groups. In the MAMCA software, each stakeholder group has their own criteria. A MCA model is built for each stakeholder and aggregated in the final step. The "participant" concept to include large groups' opinions in the decision-making process, through a survey, has been introduced in this paper.

In this contribution, a four-step procedure has been proposed as an alternative to the original five-step PROMETHEE algorithm. The proposed four-step procedure provides the same net scores for the alternatives as the original PROMETHEE. The benefit of the proposed procedure is the performance indicators received for each of the alternatives on the criteria set. Several visualizations such as "Alternatives Evaluation Analysis" providing an overview on

performance of the alternative on the criteria set with positive and negative scores, "Multi-Actor Analysis" and "Multi-Actor Box Chart Analysis" are further contributions of this paper.

References

1. Behzadian, M., Kazemzadeh, R.B., Albadvi, A., Aghdasi, M.: PROMETHEE: a comprehensive literature review on methodologies and applications. Eur. J. Oper. Res. **200**(1), 198–215 (2010)
2. Bouwman, M.E., Moll, H.C.: Environmental analyses of land transportation systems in The Netherlands. Transp. Res. Part D: Transp. Environ. **7**(5), 331–345 (2002)
3. Brans, J.P.: L'ingenierie de la Decision, l'Laboration d'Instruments d'Aidea la Decision. Universite Laval, Colloque sur l'Aidea la Decision. Faculte des Sciences de l'Administration (1982)
4. Brans, J.P., Mareschal, B.: The PROMCALC & GAIA decision support system for multi-criteria decision aid. Decis. Support Syst. **12**(4–5), 297–310 (1994)
5. Brans, J.P., Vincke, P.: Note-A preference ranking organisation method: (The PROMETHEE method for multiple criteria decision-making). Manage. Sci. **31**(6), 647–656 (1985)
6. Brans, J.P., Vincke, P., Mareschal, B.: How to select and how to rank projects: the PROMETHEE method. Eur. J. Oper. Res. **24**(2), 228–238 (1986)
7. De Brucker, K.: Sociaal-Economische Evaluatie van Overheidsinvesteringen in Transportinfrastructuur: Kritische Analyse van het Bestaande Instrumentarium Ontwikkeling van een Eclectisch Evaluatie-Instrument. Garant, Leuven (1998)
8. Dooms, M., Verbeke, A., Macharis, C., SJegers, R.: De Zaak DHL. Het Regional Hub Project van DHL in Brussel-Nationaal: Een Socio-Economische en Maatschappelijke Evaluatie. Garant, Leuven (2006)
9. Geldermann, J., Zhang, K.: Software review: "Decision Lab 2000". J. Multi Criteria Decis. Anal. **10**(6), 317 (2001)
10. Hannan, M.T., Freeman, J.: Structural inertia and organizational change. Am. Sociol. Rev. **49**, 149–164 (1984)
11. Hayez, Q., De Smet, Y., Bonney, J.: D-Sight: a new decision making software to address multi-criteria problems. Int. J. Decis. Support Syst. Technol. (IJDSST) **4**(4), 1–23 (2012)
12. Ishizaka, A., Labib, A.: Analytic hierarchy process and expert choice: benefits and limitations. OR Insight **22**(4), 201–220 (2009)
13. Lebeau, P.: Towards the Electrification of City Logistics. Vrije Universiteit Brussel, Brussel (2016)
14. Leyva-Lopez, J.C., Fernandez-Gonzalez, E.: A new method for group decision support based on ELECTRE III methodology. Eur. J. Oper. Res. **148**(1), 14–27 (2003)
15. Macharis, C.: Strategic modeling for intermodal terminals: socio-economic evaluation of the location of barge/road terminals in flanders. Vrije Universiteit Brussel (2000)
16. Macharis, C.: The importance of stakeholder analysis in freight transport: the MAMCA methodology. Eur. Transp. (Transporti Europei) **25**(26), 114–126 (2004)
17. Macharis, C., Ampe, J.: The Use of Multi-Criteria Decision Analysis (MCDA) for the Evaluation of Transport Projects: A Review. Euro XXII Prague Book of Abstracts (2007)

18. Macharis, C., Brans, J.-P., Mareschal, B.: The GDSS promethee procedure. J. Decis. Syst. **7**(4), 283–307 (1998)
19. Macharis, C., Springael, J., De Brucker, K., Verbeke, A.: PROMETHEE and AHP: the design of operational synergies in multi-criteria analysis.: strengthening PROMETHEE with ideas of AHP. Eur. J. Oper. Res. **153**(2), 307–317 (2004)
20. Macharis, C., Turcksin, L., Lebeau, K.: Multi-Actor Multi-Criteria Analysis (MAMCA) as a tool to support sustainable decisions: state of use. Decis. Support Syst. **54**(1), 610–620 (2012)
21. Nijkamp, P., Van den Bergh, C.J.M., Soeteman, F.J.: Regional sustainable development and natural resource use. World Bank Econ. Rev. **4**(1), 153–188 (1990)
22. Pugnaire, P.V.: Multicriteria Decision-Aid (1992)
23. Saaty, T.L.: Group decision making and the AHP. In: Golden, B.L., Wasil, E.A., Harker, P.T. (eds.) The Analytic Hierarchy Process, pp. 59–67. Springer, Heidelberg (1989)
24. Sirikijpanichkul, A., Ferreira, L.: Multi-objective evaluation of intermodal freight terminal location decisions (2005)
25. Tzeng, G.H., Lin, C.W., Opricovic, S.: Multi-criteria analysis of alternative-fuel buses for public transportation. Energ. Policy **33**(11), 1373–1383 (2005)
26. Vreeker, R., Nijkamp, P., Ter Welle, C.: A multi-criteria decision support methodology for evaluating airport expansion plans. Transp. Res. Part D: Transp. Environ. **7**(1), 27–47 (2002)
27. Żak, J.: The MCDA methodology applied to solve complex transportation decision problems. In: 13th Mini-EURO Conference (2002)
28. Żak, J., Thiel, T.: Multiple criteria evaluation of the development scenarios of the mass transit system. In: Selected Proceedings of the 9th World Conference on Transport Research (2001)

Dealing with the Complexity of Stakeholder Interaction in Participatory Transport Planning

Michela Le Pira[1(✉)], Giuseppe Inturri[1], Matteo Ignaccolo[1], and Alessandro Pluchino[2]

[1] University of Catania, Catania, Italy
{mlepira,ginturri,matig}@dica.unict.it
[2] University of Catania and INFN Catania, Catania, Italy
alessandro.pluchino@ct.infn.it

Abstract. Including an active participation of citizens and stakeholders from the beginning of transport decision-making processes is widely recognized as a precondition to avoid the failure of projects/policies/plans as a consequence of a lack of consensus. Appropriate methods and tools are needed to support participation processes towards well-thought and shared solutions. In this paper quantitative methods, stakeholder interaction and simulation models are used to guide and reproduce a participatory experiment aimed at consensus building about mobility management strategies. Analytic Hierarchy Process (AHP) has been used to elicit stakeholder preferences, different voting methods have been used to aggregate the individual preferences, group interaction has been performed via a facilitated dialogue to reach a consensus among stakeholders and an agent-based model (ABM) has been used to simulate the same consensus building process. Besides the social network of stakeholders has been analyzed to gain insights on its influence on the consensus formation.

The results of this integrated procedure, applied in a pilot experiment with University students as stakeholders, provide useful suggestions on how to use different methods and guide effective and efficient participation processes aimed at consensus building.

Keywords: Stakeholder engagement · Group interaction · Analytic Hierarchy Process · Agent-based model · Social network analysis

1 Stakeholder Engagement in Transport Planning

Transport planning is a decision-making process aimed at defining and implementing transport systems operations [30]. It is the result of a technical and political process that inevitably favours some interests at the expense of others within a community. The consequence is that transport decisions have the potential to generate a big reactivity of people and groups of interests and stop their implementation, even when small decisions are made at a very fine scale, such as restricting the use of a street to pedestrians or changing the parking rules. This phenomenon is known as the so called "Not In My Back Yard" (NIMBY) syndrome [7].

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_3

The traditional pure rational approach where the decision is the optimal solution with respect to one objective (measured by some utility indicator), as a result of the analysis and choice made by transport experts through quantitative method, has shown many limitations. Even if the transport plan is meant to increase the net welfare, the benefits will never be equally distributed within the different groups of the community [18]. Including an active participation of citizens and stakeholders from the beginning of the transport decision-making process is widely recognized as a precondition to avoid the failure of a project as a consequence of a lack of consensus [2, 7, 25, 41]. Finding the best trade-off between the "most shared" solution (based on consensus building) and the "optimal" one (based on the results of technical evaluations) is now considered a key issue for the success of a transport plan or project.

In this respect, Le Pira et al. [23] propose a new participatory transport planning framework (Fig. 1) able to include (a) different levels of involvement (as presented in [16]), (b) different actors, i.e. experts, stakeholders and citizens that contribute with different degrees of competence and interest (as explained in [19]) to the "bounded-rationality" decision-making process, where quantitative evaluations of the planning alternatives are integrated within a participation process (as introduced in [7]).

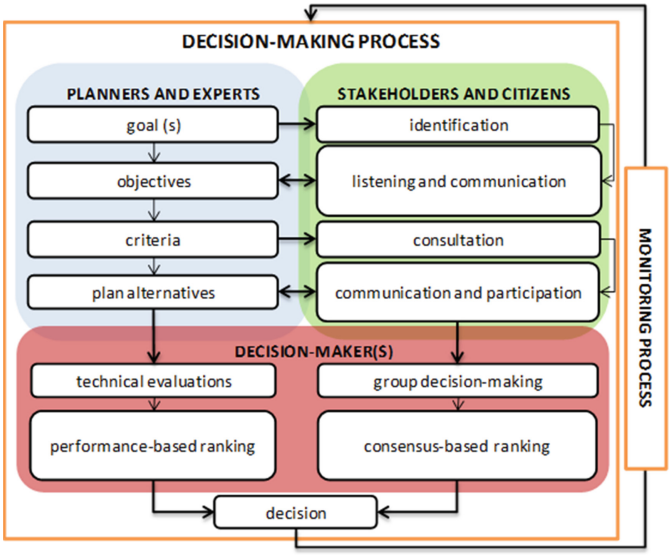


Fig. 1. Framework of the participatory decision-making process in transport planning [23].

Planners and experts are in charge of analysing and modelling the transport system from the overall goal of the plan (e.g. improving the transport system) to the objectives (e.g. social, environmental and economic sustainability), evaluation criteria (e.g. air pollution) and plan alternatives to be submitted for the final choice.

Stakeholders and citizens are engaged in all the planning steps to share objectives, evaluations criteria and alternatives. All relevant stakeholders should be identified at the

beginning, from institutions to transport users, citizen associations and transport operators and also media and financial institutions [7]. Listening and communication is the first step of the engagement process, where planners and experts understand interests and expectations of stakeholders and communicate to them the basic goals that originated the decision to make the plan: the aim is to identify a set of plan specific objectives. Based on objectives, planners and experts define the evaluation criteria and stakeholders and citizens are consulted to understand their preferences, in terms of weights of the chosen criteria and objectives. Finally, they take part in the definition and evaluation of the plan alternatives, in a communication and participation procedure. Decision-makers (i.e. public authorities) will make the choice supported by a performance-based ranking (from the planners and experts) and a consensus-based ranking of plan's alternatives (from stakeholders and citizens). The transport planning framework can be applied to rank different alternative plans, each made of a package of measures, or to rank priorities of different alternative measures of the same plan. This decision-making process is synthetized in the framework of Fig. 1.

The complexity of transport decisions, also raised by the multidimensional impacts of the measures, requires a Multi-Criteria Decision-Making (MCDM) approach to select the best solution or to rank the alternatives. Including stakeholders' opinions into the participation process calls for methods to group individual stakeholders' preferences into a collective decision.

In this respect, the Multi Actor Multi Criteria Analysis (MAMCA) has been proposed in [26] as a sequence of Multi Criteria Analyses (MCAs), each carried out with a different stakeholder group, and tested in different transport planning contexts (e.g. [27, 39]). The aim of MAMCA is to gain insights in what the stakeholders' objectives are and how an alternative contributes to these objectives. With this input, the alternatives could be altered or adapted to the different stakeholders' needs. In this respect, the MAMCA methodology does not intend to make a final decision based on a consensus building process, but aims to support stakeholder engagement and give the decision-maker insights in the stakeholders' needs.

A variety of qualitative and quantitative methods have demonstrated that interaction and deliberation can change stakeholders' mind about public policy problems [33] and favour the success of a group decision-making process. Moreover, the method to aggregate individual opinions into a collective one is important, because different methods can lead to different solutions.

Based on these premises, the chapter presents a procedure based on MCDM methods, stakeholder interaction and simulation models to guide and reproduce a participatory experiment aimed at consensus building about mobility management strategies. The aim is to tackle the complexity of stakeholder interaction in participatory processes aimed at finding a convergence of opinions towards a shared decision.

Analytic Hierarchy Process (AHP) has been used to elicit stakeholder preferences, different voting methods have been used to aggregate the individual preferences, group interaction has been performed via a facilitated dialogue to reach a consensus among stakeholders and an agent-based model (ABM) has been used to simulate the same consensus building process.

The results of this integrated procedure, applied in a pilot experiment with University students as stakeholders, provide useful suggestions on how to use different methods and guide effective and efficient participation process aimed at consensus building.

The chapter, built on the participation experiment presented in [23], adds more details about the setting up of the procedure, performs a social network analysis of the stakeholders involved and include the results of additional agent-based simulations useful to gain insights in the stakeholder network structure and its capability to support effective and efficient consensus building processes.

The remainder of the paper is organized as follows: Sect. 2 introduces the materials and methods used in the study; Sect. 3 illustrates the case study and its related results; in Sect. 4 the results are discussed and some general conclusions are provided.

2 Materials and Methods

The study takes its clue from the framework of Fig. 1 where stakeholders express their preferences and exchange their opinions interacting with each other and with the experts (e.g., transport planners). It is assumed that stakeholders express their preference on prefixed alternative transport measures and that these are turned in priority rankings via appropriate MCDM methods. In this task, stakeholders might be supported with technical evaluations on the impact each alternative has on the evaluation criteria, in such a way that their priority rankings result as the combination of their needs, previous knowledge and provided information. The order (or priority) assigned by stakeholders to alternatives can be obtained via MCDM methods from pairwise comparisons with the other alternatives or from numerical scores assigned to each of them. In both cases, the ordered list can be turned into a binary vector whose components assume the value of +1 if the generic alternative A precedes B in the list or –1 if the opposite occurs. Then, a preference aggregation procedure can be applied to derive a final collective ranking, which has to be: transitive, i.e. if alternative A is preferred to B and B to C, then A has to be preferred to C; consistent, i.e. it should derive from logical – non-random - judgments; shared, i.e. it should reflect the individual preferences at a reasonable level (or at a good degree of consensus).

The research reproduces a participation process through an experiment in which university students acted as “sophisticated stakeholders” [29] of the mobility services provided by the university. Collective preference rankings are calculated from the collected data by different aggregation procedures and used to evaluate to what extent interaction contributes to achieve a more shared decision.

Next subsections present the methods used in this study to set up and reproduce the participatory process based on MCDM methods, voting procedures, consensus measures, social network analysis and an agent-based simulation model.

2.1 Analytic Hierarchy Process

The Analytic Hierarchy Process (AHP) was developed by Thomas Saaty [35]. It is based on pairwise comparisons through the building of matrices to derive priority scales and weights at the different levels of a structured decision problem.

It can be synthesized in 4 steps:

1. Problem structuring.
2. Priority calculation through pairwise comparisons.
3. Consistency check.
4. Sensitivity analysis.

The use of AHP in transport planning is based on the decomposition of a decision-making problem into a tree structured decisions' hierarchy that contains: the general goal of the plan, a set of specific objectives represented by evaluation criteria (and possible sub-criteria) and finally the decision alternatives aimed at achieving the general goal (Fig. 2). In general, the problem structure derives from brainstorming sessions with experts or from the analysis or similar problems.

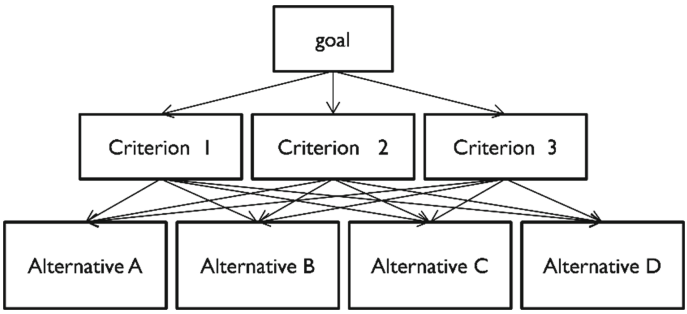


Fig. 2. Basic hierarchy of a decision problem in transport planning.

A set of pairwise comparison matrices is built by comparing couples of elements at the same level, with respect to the elements of the upper level. The pairwise comparison expresses a judgment on a qualitative scale that is turned into a quantitative one [35] (Table 1). The numerical scale goes from 1 to 9, widely used because it is more detailed than the smaller ones (e.g., 1–5) and more appropriate than the larger ones (e.g., 1–100), which are more difficult to be used by decision-makers. Being a ratio scale (and not an interval scale), it allows to aggregate more judgments with the same measurement unit.

At each level of the problem structure the pairwise matrices can be transformed into a set of local priority vectors with different methods [36], e.g. the approximate method, the eigenvalue method, the geometric mean method [15]. Finally, a ranking of alternatives is obtained by combining all the levels into a global priority vector.

It is known that pairwise comparisons can lead to some inconsistency, meaning that individual judgments can be affected by lack of rationality and violate the consistency condition of the matrix. For instance, given 3 alternatives A, B, C, all terms a_{ij} of the pairwise comparison matrix must satisfy the relation $a_{AC} = a_{AB} \cdot a_{BC}$. This means that

Table 1. Scale of pairwise comparisons (adapted from [35])

Intensity of importance[a]	Definition	Explanation
1	Equal importance	Two elements contribute equally to the objective
3	Moderate importance	Experience and judgment slightly favour one element over another
5	Strong Importance	Experience and judgment strongly favour one element over another
7	Very strong importance	One element is favoured very strongly over another, it dominance is demonstrated in practice
9	Extreme importance	The evidence favouring one element over another is of the highest possible order of affirmation

[a]the values 2,4,6,8 can be used to express intermediate values, 1.1, 1.2, etc. for elements that are very close in importance

judgments must be transitive and mutually correlated. The inconsistency can be measured (and, therefore, monitored) through the comparison between a Consistency Index derived by the matrix elements with the one obtained by purely random judgments [35]. In general, an inconsistency less than 10% is accepted.

A sensitivity analysis can be carried out to test the robustness of the results, in terms of elasticity of the final ranking to the criteria weights.

AHP is widely used in transport planning and management [11, 28, 31, 38]. When AHP is used to elicit single decision-maker opinions, the only condition to respect is judgments' consistency; when its use is extended to group decision-making an appropriate procedure to aggregate the individual judgments has to be defined.

There are four ways to combine multiple preferences into a consensus rating [15], according to the level of aggregation (from judgments or from priorities) and the type of aggregation (mathematical or based on consensus vote).

The consensus vote can be used with "synergistic" groups that agree on the values of the matrices or on the priority vectors. In general, if a consensus cannot be reached, a mathematical aggregation can be adopted.

An important issue is the level of the decision process to which the aggregation is made [15]:

- Aggregation of Individual Judgments (AIJ), i.e. the elements of each stakeholder matrix are aggregated into a group matrix,
- or Aggregation of Individual Priorities (AIP), i.e. a group priority vector is calculated from the individual vectors.

In any case the final result consists of an "averaged" ranking that could not be representative of the individual points of view. Therefore, the problem of preference aggregation remains still open. While mathematical aggregation implies transparency of the calculation and clarity of results even if it could not reflect the individual preferences, consensus vote is a more democratic and fair way to find a group ranking, though it

could be supported only by a relative majority and therefore having a low rate of acceptance.

According to the authors, the optimal solution should be based on a mixed procedure that combines mathematical aggregations with consensus building, through an interaction process among stakeholders that allows a convergence of opinions to increase the acceptability of the final results while at the same time guaranteeing transparency of the decision process, as described in Table 2.

Table 2. Group aggregation procedures [22]

	Transparency and reproducibility	Fairness of the process	Probability of acceptance
Mathematical aggregation	✓		
Consensus vote		✓	
Consensus building process with the help of mathematical aggregation	✓	✓	✓

2.2 Voting Methods

The problem of combining individual preferences into a collective decision has been widely studied by the social choice theory [1]. The origin of voting or preference aggregation methods is often attributed to Condorcet and Borda, having introduced the former the pairwise comparisons method or pairwise majority rule (PMR), the latter the scoring method or Borda rule [5].

With the Borda rule the ranking of alternatives is obtained on the basis of the sum of different scores assigned to reflect each individual preference order of alternatives (e.g. if, for 3 alternatives A, B and C the individual ranking is A>B>C, then A will score 3, B will score 2 and C will score 1).

With the PMR the ranking is obtained by computing how many times each alternative in a pair is preferred to the other one. In particular, the pairwise preferences of each individual list are coded as components of a binary vector assuming the values of +1 and −1. Finally, the collective preference list is derived by applying a majority rule to the binary vectors. Given n alternatives, the number of possible pairs is $n^*(n-1)/2$. For example, for the alternatives A, B, C and D ($n = 4$), there are six pairs (4*3/2 = 6): AB, AC, AD, BC, BD, CD.

Nevertheless, the aggregation of single preference lists by the PMR does not exclude the possibility of intransitive collective lists as result, falling in the so called "Condorcet paradox". It was studied for the first time in 1785 by the Marquis de Condorcet [10] who demonstrated that, for n>2, the collective social preference order can be intransitive even if the individual preference orders are transitive. The final consequence is the impossibility of taking a consistent decision.

In this respect, the impossibility theorem by Arrow [1] states that no method can simultaneously satisfy five axioms (universal domain; ordering; weak Pareto principle; independence of irrelevant alternatives (IIA); non-dictatorship).

The PMR can violate the ordering axiom, but it is considered one of the fairest social rules [9, 34] particularly with respect to the Borda rule, though this can violate the axiom of independence of irrelevant alternatives (IIA). In this case, the choice between the two most voted alternatives might be influenced by the preferences obtained by the less voted one, giving space to "the risk of tactical and opportunistic voting" [14]. More details on these methods and the pros and cons of using them can be found in Le Pira [18].

In this study, both the PMR and the Borda rule will be used to aggregate the individual rankings derived from AHP and the results, in particular with regard to the degree of consensus, will be compared with those derived from the traditional AHP techniques.

2.3 Consensus Measure via Overlap

In this study a simple indicator will be used to measure the degree of consensus towards the collective decision and, in particular, to see to what extent different aggregation methods of individual preferences or the level of stakeholder interaction may affect the degree of achieved consensus towards the final decision [20, 23, 24].

If n is the number of alternatives and $m = n(n-1)/2$ is the number of the possible couples of alternatives (i.e. the number of components of each binary vector), the overlap between two lists is defined as:

$$O_{ij} = \frac{1}{m} \sum_{k=1}^{m} V_i^k V_j^k \tag{1}$$

where V_{ik} and V_{jk} are the k-th components of the two binary vectors V_i and V_j representing the preference lists of stakeholders S_i and S_j. From this definition follows that $O_{ij} \varepsilon [-1;1]$; if S_i and S_j have the same opinion, then $V_i = V_j$ and $O_{ij} = 1$; if all the homologous components V_{ik} and V_{jk} have opposite signs, then $O_{ij} = -1$; if V_i and V_j are uncorrelated, then $O_{ij} = 0$. The overlap of two stakeholders' opinions about the preference order of n alternatives can be interpreted as the scalar product of the two binary vectors with m components, that is the degree of alignment of two opinions in a m-dimensional space. If N is the number of stakeholders involved in the decision, the concept of overlap can be extended to represent the average similarity between the collective list c and the N individual ones:

$$\bar{O}_{i,c} = \frac{1}{N} \frac{1}{m} \sum_{i=1}^{N} \sum_{k=1}^{m} V_i^k V_c^k \tag{2}$$

This average overlap can be considered as an indicator of the average degree of consensus towards the collective decision, since it accounts for the similarity between each stakeholder individual list and the collective one deriving from aggregation as described in the previous section.

2.4 Social Network Analysis

A social network of stakeholders is represented by a graph consisting of nodes (i.e. the social agents) and links (i.e. the relationships among them). Representing stakeholders

in social networks can be helpful to have a clear insight on the importance of actors involved in the decision-making process and the potential interactions among them.

Social networks fall within the category of complex networks, whose structure is irregular, complex and dynamically evolving in time [4] and adequate methods are needed to study their structure and dynamics. Social Network Analysis (SNA) allows to quantify the social importance of a given individual in a network via centrality indexes and understand the potential problems due to topology [37]. The use of SNA in the field of stakeholder engagement can simply consist of stakeholder mapping or extend to centrality measures [13, 32].

In this study, the in-degree centrality of a node, i.e. the number of ties directed to a node, will be used to have a measure of stakeholder centrality and to make inference about the outcome of their interaction.

2.5 Agent-Based Model of Stakeholder Interaction

Agent-based modelling (ABM) is a computer technique simulating "the observed world in terms of actors (agents) characterized by certain rules (behaviour) that depend on the state of the environment, the state of the agent and its spatial location. Each agent is represented as an independent computerised entity capable of acting locally in response to stimuli or to communicate with other agents" [40]. ABM has been widely used to reproduce transport problems and interaction among transport stakeholders. For more details, the reader can refer to the papers in the special issues of Transportation Research C on Agents in Traffic and Transportation edited by Bazzan, Klügl and Ossowski [3, 17]. ABMs are suitable to simulate complex systems and to reproduce the actions of single agents characterized by particular properties and internal complexity. In this context, social dynamics and, in particular, the emerging collective phenomena that derives from social interaction are widely investigated via opinion dynamics models [8], giving birth to a new branch of physics, i.e. the statistical physics of social dynamics.

ABM and opinion dynamics models have been widely used to investigate social interaction in stakeholder networks and reproduce typical participatory decision-making processes in transport planning [20, 21, 24]. In this study, the ABM described in Le Pira et al. [24] is used to reproduce the group interaction process in terms of opinion dynamics in networks of stakeholders.

The main aim of the model is to understand what role interaction plays in escaping from intransitive cycles (i.e. the "Condorcet paradox" described in Sect. 2.2) and if it favours the convergence of opinions towards a final decision, i.e. a collective list reflecting quite appreciably the individual preferences. The model consists of several routines, from the creation of the network of stakeholders to the simulation of their opinion exchanges until a transitive and shared collective decision is obtained. A list of ordered alternatives is initially randomly assigned to all stakeholders, to represent their individual preferences. Then, each stakeholder is allowed to know the lists of his neighbours (i.e. the directly connected nodes in the network) and, at each interaction step, he decides to change his preferences' list according to the overlap between his list and the one of the majority of his neighbours, weighted by their influence. Then, a collective

list is obtained aggregating all individual lists by PMR and a check of its transitivity is carried out. The algorithm goes on until a transitive list is found. Generally, the first transitive list found with this method does not have a high average overlap (i.e. degree of consensus), therefore the interaction is repeated in order to find new more shared transitive solutions. The average overlap in general presents a growing trend as far as the interaction goes on, until it reaches a stationary state, corresponding to its maximum value. The final list is assumed as the transitive "most shared" collective solution, appreciably reflecting the individual preferences.

The results from the real participation experiment will be compared with those from the simulations that consider different types of stakeholder networks, in order to see to what extent the topology can contribute in increasing the efficiency of the interaction process.

3 Case Study

A pilot participation experiment with University students as stakeholders was carried out with the aim to test the combined approach where: AHP was used to elicit stakeholder preferences and to aggregate them (Sect. 2.1), together with the PMR and the Borda rule (Sect. 2.2); the overlap was used to evaluate the degree of consensus of the collective ranking before and after interaction (Sect. 2.3); a social network analysis was performed to have insight in the network of stakeholders (Sect. 2.4); the agent-based model was used to simulate the same participation process with different network topologies (Sect. 2.5).

The participation experiment involved 17 University students enrolled in a master degree program in Transport Engineering at the University of Catania (Italy). The participated process consists of a two-round interactive AHP to elicit their preferences about mobility management strategies to be adopted in their University with the aim to investigate to what extent interaction among them could change and possibly align their opinions.

In particular, students answered a questionnaire with questions split in four parts:

1. critical issues related to the mobility system and accessibility to the University sites,
2. objectives that the University Mobility Manager (as decision-maker) should pursue to improve mobility and accessibility,
3. priorities about six mobility management alternatives,
4. pairwise comparisons of the six alternatives.

Besides, they were asked to indicate two of their peers (involved in the experiment) that they considered competent with respect to the specific subject. This was done to recreate a social network of the students.

They provided their opinions (in terms of ranking) about the main critical issues regarding mobility, i.e.:

A. low reliability, punctuality and speed of public transport systems
B. lack of coordination between urban and suburban public transport systems
C. high levels of road congestion to access the city

D. lack of infrastructures and facilities for non-motorized mobility

Then, they also expressed their opinions (in terms of ranking) about the primary objectives that the Mobility Manager should pursue, i.e.:

I. reduce the number of trips by private vehicle
II. incentivize car pooling initiatives
III. reduce total km-travelled

Six mobility management alternatives were considered as part of the University Travel Plan [6]:

(1) establishment of public transport (PT) lines dedicated to students and staff
(2) facilitation for using local public transport
(3) better management of the University parking spaces
(4) car pooling promotion
(5) bike-sharing service dedicated to University members
(6) rescheduling of working and studying hours and telecommuting

Part 1, 2 and 3 of the questionnaire were meant to know stakeholders' general opinions about the topic, while part 4 allowed to elicit their preferences and derive priority ranking of the alternative solutions via AHP. In this respect, the students were asked to:

- express pairwise comparison judgments about the six alternatives
- motivate the priority ranking (from AHP) and share their opinions through interaction with the others and
- reformulate their pairwise comparison judgments after interaction.

AHP was applied before and after interaction to derive a priority ranking of alternatives for each student using the geometric mean method [15], on the basis of their judgments of preference between all couples of alternatives in a "local context", i.e. the alternatives were compared upon a unique criterion [12]. Only in the first-round questionnaire they were asked to express their opinion in terms of priority about the 6 alternatives (Sect. 3 of the questionnaire) and this was done to test the consistency of their judgments via comparison with the AHP results.

Then the students were trained about how to use AHP to derive priority vectors to increase the awareness of the relevance of MCDM and AHP in transport decisions, and soon after they were asked to motivate their preferences. In one step of interaction all-to-all they supported their ideas in front of the others and finally repeated the pairwise comparisons of the alternatives (Fig. 3). In general, the process of interaction - together with the training session on AHP - led to an increase in the convergence of opinions (in terms of average overlap with the collective ranking) and a decrease in the inconsistency of pairwise judgments, as it will be shown in the following section.

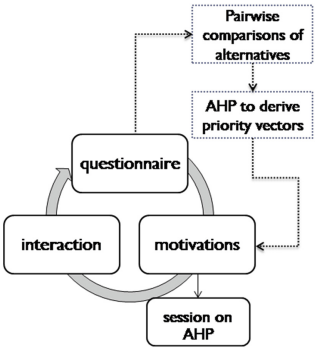


Fig. 3. Description of the participation experiment (two-round interactive AHP).

3.1 Results of the Experiment

The results of the part 1, 2 and 3 of the questionnaire regarding students' general opinions are reported in Table 3.

To apply the Borda rule a score was assigned to each element related to the position it occupies in the rankings (decreasing values from 4 to 1 from the 1st choice to the last one). Analyzing these results, it is clear that the first problem perceived by the students is related to public transport, both in terms of quality of service (low reliability, punctuality and speed) (A) and in terms of lack of coordination between urban and suburban systems (B). The other two issues are considered less critical, i.e. high levels of road congestion (C) and lack of infrastructures for non-motorized mobility (D). The first ranked objective regards the reduction of the number of trips by private vehicle (I), followed by incentivizing car pooling initiatives (II) and reducing total km-travelled (III). For what concerns the priority of the measures, alternative 2 and 3, i.e. facilitating the use of local public transport (2) and better managing of the University parking spaces (3) are considered with a high priority, followed by car pooling promotion (4) and bike sharing service (5), while the less important are the establishment of public transport (PT) lines dedicated to students and staff (1) and rescheduling of working and studying hours and telecommuting (6). These results confirm that, according to the limited number of students involved in this pilot, the priorities for the Mobility Manager should be promoting public transport and increase and efficient use of private cars.

Table 3. Results of students' general opinions

Elements	Ranking with Borda	Priority
Criticalities	A>C>B=D	Improve the quality of service of public transport systems and the coordination between urban and suburban systems
Objectives	I>II>III	Reduce the number of trips by private vehicle and incentivize car pooling initiatives
Measures	2>3>4=5>1=6	Facilitate the use of local public transport and better manage the University parking spaces

Table 4 summarizes the results of the part (3) of the questionnaire (i.e. pairwise comparisons), in terms of collective preference orders and related average overlap calculated before and after interaction with AHP aggregation methods, i.e. aggregation of Individual Judgments (AHP-AIJ) and Aggregation of Individual Priorities (AHP-AIP), with the PMR and with the Borda rule.

Table 4. Results of collective orders and average overlap before and after interaction using different aggregation methods

Aggregation method	Collective Preference Order		Average Overlap	
	Before	After	Before	After
PMR	2>3>4>5>6>1	2>3>4>5>**1>6**	0.43	0.59
Borda	2>3>4>5>6>1	2>3>4>5>**1>6**	0.43	0.59
AHP-AIJ	2>3>4>5>6>1	2>3>4>5>**1>6**	0.43	0.59
AHP-AIP	2>3>4>5>1>6	2>3>4>**1>5**>6	0.42	0.55

The rankings obtained with all the methods are quite the same, except from the one derived from AHP-AIP where the alternatives 1 and 5 are inverted. PMR, Borda and AHP-AIJ lead to the same results (before and after interaction) and the collective list resulting from them shows the maximum values of overlap. This can be considered a sound result, both because different aggregation procedures lead to the same collective list and for the good overlap related to it. Besides, the results are very similar to the ones obtained by assigning a priority to the alternatives in Table 3. In any case, with all the methods there is a general increase in the convergence of opinions and this confirms the efficacy of interaction in the group decision-making process. Figure 4a shows the overlap of each stakeholder list with the collective list before (1) and after interaction (2). The Consistency Ratio (CR) is used to evaluated judgments' consistency in terms of the deviation of the individual pairwise comparisons from random judgments (Fig. 4b). After the first questionnaire, the majority of stakeholders' judgments (15 out of 17) resulted inconsistent (CR>0.10, according to Saaty, 1980), but after the second questionnaire – and after the explanation of AHP – they were only 6 (out of 17), even if the average consistency didn't increase with the same proportion.

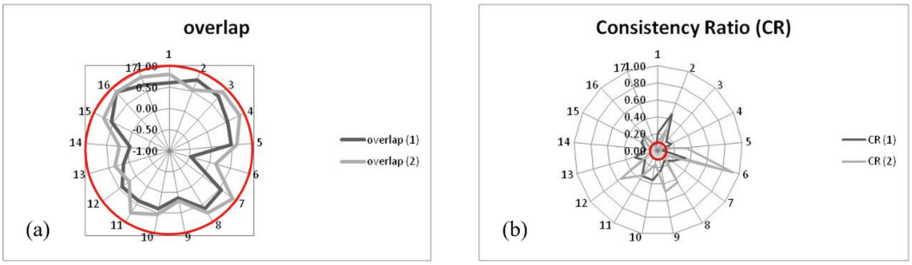


Fig. 4. Overlap (a) and consistency ratio (CR) (b) for the 17 students before (1) and after (2) interaction; the red circles represent the maximum reachable overlap (O(i, c) = 1) in (a) and the limit of consistency in (b) (CR ≤ 0.10) (Le Pira et al. 2015) [23].

An interesting analysis can be performed by considering the social network of stakeholders. In this respect, by asking to indicate two peers (involved in the experiment) that they considered competent it was possible to build their social network (Fig. 5a). As can be noticed, the in-degree centrality of the students (that is related to the number of times they were indicated by other students) does not reflect the values of overlap (Fig. 5b). The most central nodes (6, 9 and 10) that should be also the most influential in the network, show low values of overlap with respect to the others, meaning that the final decision derived from aggregation do not fully satisfy them. This result can be ascribed to the type of interaction that was performed during the experiment, i.e. a guided discussion all-with-all with fixed and limited time to express their opinion that maybe hindered the influence of the most important participants.

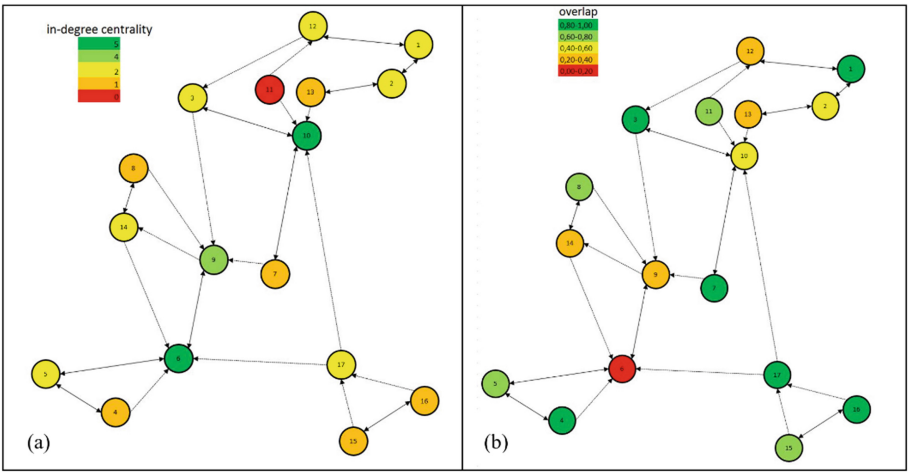


Fig. 5. Social network of the students: (a) in-degree centrality; (b) overlap after interaction.

3.2 Simulation Results

We reproduced the experiment with the ABM, considering a network with exactly the same number of stakeholders (17 nodes) (for more details see [23]). The network is "fully connected", i.e. each node is linked with all the others. This can represent quite well the interaction process among students, where each of them expressed his opinion in front of the others. Actually, the model simulates a repeated interaction, starting from an intransitive collective list, where at each step of interaction stakeholders can change their opinion and a new PMR is evaluated, until average overlap reaches a stationary state. The simulations are run 500 times starting from different initial conditions to have a statistics of the events (Fig. 6).

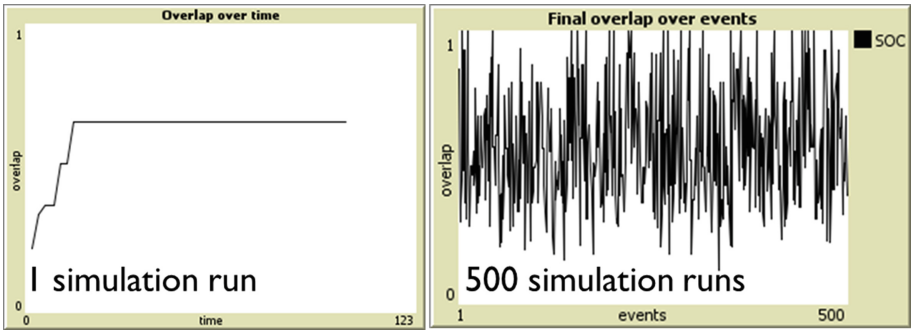
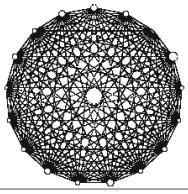
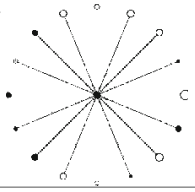
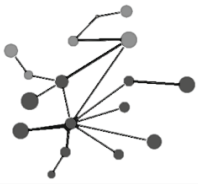


Fig. 6. Results of the simulation after 1 and 500 runs.

Results are presented in Table 5. The final overlap is very similar to the one obtained with PMR from the participation experiment (i.e. 0.59), but the ratio of overlap to number of links is very small. The latter can be considered as a measure of the interaction efficiency, which takes into account the "cost" of each interaction. This suggests that other topologies, e.g. a star network (where only one central node is linked with all the others), could be more efficient (see also [24]). The same simulations are therefore repeated for this kind of network and, as shown in Table 5, while the average overlap remains the same, the average overlap/links is sensitively higher than before. In this study, another network is considered, drawing inspiration from the results of the social network analysis performed in Sect. 3.1. This network is a scale free tree with fully

Table 5. Results of agent-based simulation using different networks

Network	Overlap	Overlap/links
Fully connected 	**0.57**	**0.004**
Star 	**0.57**	**0.04**
Scale free tree with fully connected hubs 	**0.51**	**0.03**

connected hubs described in Le Pira et al. [20]. It has a power law degree distribution, the high-connected nodes (i.e. the hubs) are all linked with each other and the communities of each hub form trees, representing typical hierarchical structures where the information from the nodes converges to the hubs. Although it is a prototypical network, it can be considered similar to the artificially-built students' network of Fig. 5, where three "hubs" were connected to the other nodes, even though there are triangles among them, so the communities are not trees (resulting in 34 links instead of 17). The results of the simulations performed with this network show lower values of overlap with respect to the previous ones, while the interaction efficiency is quite good. This confirms to some extent the considerations done in the previous section (Fig. 5): if interaction had taken place spontaneously (and maybe according to the identified social network) probably the final overlap would have been lower due to the influence of the "hubs" that showed different preferences with respect to the average.

It is also worthy of notice that different initial conditions lead to the same result: while the real experiment was carried out with a quite "homogeneous community" of stakeholders, the model starts from a random assignment of initial preferences among stakeholders, i.e. reproducing "heterogeneous communities". This implies that the same result is obtained after just one step of "real" interaction among students, and much more steps of repeated interaction in the model, until the final overlap becomes stable. In our case, the homogeneity of stakeholders simplified the carrying out of the experiment, but further experiments could consider different groups of stakeholders to reproduce more realistic situations and to see to which extent it is possible to reach a good convergence of opinions.

4 Discussion and Conclusions

In this chapter a procedure based on MCDM methods, stakeholder interaction and simulation models has been presented to guide and reproduce a participatory experiment aimed at consensus building about transport decisions.

The procedure has been tested in a pilot participation experiment with students as stakeholders discussing alternative mobility management strategies.

The rankings obtained with all the aggregation methods are quite the same, except from the one derived from AHP-AIP where the alternatives 1 and 5 are inverted. PMR, Borda and AHP-AIJ lead to the same results (before and after interaction) and the collective list resulting from them shows the maximum values of overlap. This suggests that one should use different aggregation procedures to test the stability of the results derived from individual judgments.

The judgment consistency condition is not fully respected, even if it improves after interaction. This can be ascribed to a mixed behavior of the student-stakeholders, whose preferences were conditioned both by their experience of students using the transport services to access the University sites and their expertise in the field of transport management. The improvement of the judgment consistency after interaction can also be ascribed to a learning effect about how to use AHP.

By performing a social network analysis of the students, it appears that the most central ones were not able to influence the final outcome, resulting in lower values of their overlap with the collective decision. This can be ascribed to the type of interaction process chosen for the experiment, where a guided discussion all-with-all with fixed and limited time for each student to express his opinion might have hindered the influence of the most important participants.

The nature of the pilot study allows to make limited generalization of the results for policy recommendations even if it is clear that students think that the priorities for the Mobility Manager should be promoting public transport and optimize the use of private cars.

Finally, the outcome of the real experiment has been compared with the analogous one obtained through agent-based simulations. The final results suggest that ABM can be useful, on one hand, to anticipate the behavior of a real network of stakeholders with a given topology in terms of degree of consensus; on the other hand, to select the more appropriate topology in terms of efficiency of the interaction process. Indeed, other topologies, e.g. a star network (where only one central node is linked with all the others), resulted to be more efficient than the real stakeholder network (i.e. a fully connected network).

In conclusion, in the framework of participatory decision-making processes in transport planning, the role of quantitative methods is important to elicit stakeholder preferences and to aggregate them. Interaction is fundamental for the success of the participation process because it allows to reach more shared decisions; agent-based simulation can be a very useful tool both to reproduce and manage a real process of stakeholder interaction. The procedure presented in this chapter is suitable to support participation process where a limited number of stakeholders with some competence are involved and can intensively and equally interact with each other. Further research will explore other complex transport decision-making contexts, characterized by multiple criteria and multiple stakeholders, adapting the procedure to the specific case and the number of stakeholders involved.

References

1. Arrow, K.J.: Social Choice and Individual Values. Yale University Press, New Haven (1951)
2. Banister, D.: The sustainable mobility paradigm. Transp. Policy **15**, 73–80 (2008)
3. Bazzan, A.L.C., Klügl, F., Ossowski, S.: Agents in traffic and transportation: exploring autonomy in logistics, management, simulation, and cooperative driving. Transp. Res. Part C **13**, 251–254 (2005)
4. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.-U.: Complex networks: structure and dynamics. Phys. Rep. **424**, 175–308 (2006)
5. Borda, J.C. de: Memoire sur les elections au scrutiny. Historoire de l'Academie Royals des Sciences, Paris (1781)
6. Caprì, S., Greco, A.: MOMACT (Mobility Management dell'Ateneo di CaTania): Piano Spostamenti Casa-Università: Documento Strategico (2009). http://www.momact.unict.it/sites/default/files/MOMACT_PSCU2009.pdf
7. Cascetta, E., Pagliara, F.: Public engagement for planning and designing transportation systems. Procedia Soc. Behav. Sci. **87**, 103–116 (2013). SIDT Scientific Seminar 2012

8. Castellano, C., Fortunato, S., Loreto, V.: Statistical physics of social dynamics. Rev. Mod. Phys. **81**, 591–646 (2009)
9. Dasgupta, P., Maskin, E.: The Fairest Vote of All. Sci. Am. **290**(3), 64–69 (2004)
10. de Condorcet, M.: Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix. Imprimerie Royale. France, Paris (1785)
11. De Luca, S.: Public engagement in strategic transportation planning: An analytic hierarchy process based approach. Transp. Policy **33**, 110–124 (2014)
12. Escobar, M.T., Moreno-Jiménez, J.M.: Aggregation of individual preference structures in AHP-group decision making. Group Decis. Negot. **16**, 287–301 (2007)
13. García Melón, M., Estruch Guitart, V., Aragonés-Beltrán, P., Monterde-Roca, B.: Social network analysis in participatory environmental decision making. The case of Spanish wetland La Albufera. In: 12th International Symposium on the Analytic Hierarchy Process). Multi-criteria Decision Making, Kuala Lumpur, Malasia (2013)
14. Giansanti, A.: Remarks on the Condorcet's paradox. In: AIP Conference Proceedings, vol. 965(1), pp. 308–314 (2007)
15. Ishizaka, A., Nemery, P.: Multi-Criteria Decision Analysis. Methods and Software. Wiley, Chichester (2013)
16. Kelly, J., Jones, P., Barta, F., Hossinger, R., Witte, A., Christian, A.: Successful Transport Decision-Making – A Project Management and Stakeholder Engagement Handbook. European Commission, Bruxelles (2004)
17. Klügl, F.: A validation methodology for agent-based simulations. In: ACM symposium on Applied Computing, pp. 39–43 (2008). ISBN:978-1-59593-753-7
18. Le Pira, M.: Towards participatory decision-making processes in transport planning: an agent-based approach. Ph.D. thesis. University of Catania, Italy (2015). https://www.researchgate.net/publication/292643361_Towards_participatory_decision-making_processes_in_transport_planning_an_agent-based_approach
19. Le Pira, M., Ignaccolo, M., Inturri, G., Garofalo, C., Pluchino A., Rapisarda, A.: Agent-based modelling of Stakeholder Interaction in Transport Decisions. In: 13th World Conference on Transport Research, Rio de Janeiro, Brasil (2013). ISBN:978-85-285-0232-9
20. Le Pira, M., Ignaccolo, M., Inturri, G., Pluchino, A., Rapisarda, A.: Finding shared decisions in stakeholder networks: an agent-based approach. Physica A Stat. Mech. Appl. **466**, 277–287 (2017)
21. Le Pira, M., Ignaccolo, M., Inturri, G., Pluchino, A., Rapisarda, A.: Modelling stakeholder participation in transport planning. Case Stud. Transp. Policy **4**(3), 230–238 (2016)
22. Le Pira, M., Inturri, G., Ignaccolo, M.: Combined expert, stakeholder and citizen involvement for priority setting of cycling mobility strategies using analytic hierarchy process. In: International Conference on Traffic and Transport Engineering (ICTTE) (2016). ISBN: 978-86-916153-3-8
23. Le Pira, M., Inturri, G., Ignaccolo, M., Pluchino, A.: Analysis of AHP methods and the Pairwise Majority Rule (PMR) for collective preference rankings of sustainable mobility solutions. Transp. Res. Procedia **10C**, 852–862 (2015)
24. Le Pira, M., Inturri, G., Ignaccolo, M., Pluchino, A., Rapisarda, A.: Simulating opinion dynamics on stakeholders' networks through agent-based modeling for collective transport decisions. Procedia Comput. Sci. **52**, 884–889 (2015)
25. Litman, T.: Transportation cost and benefit analysis: techniques, estimates and implications [Second Edition] (2009). http://www.vtpi.org/tca/
26. Macharis, C.: The importance of stakeholder analysis in freight transport. Europ. Transp. Trasporti Europein **25–26**, 114–126 (2005)

27. Macharis, C., De Witte, A., Turcksin, L.: The Multi-Actor Multi-Criteria Analysis (MAMCA) application in the Flemish long-term decision making process on mobility and logistics. Transp. Policy **17**, 303–311 (2010)
28. Mahmoud, M., Hine, J.: Using AHP to measure the perception gap between current and potential users of bus services. Transp. Plan. Technol. **36**(1), 4–23 (2013)
29. Mau-Crimmins, T., de Steiguer, J.E., Dennis, D.: AHP as a means for improving public participation: a pre-post experiment with university students. For. Policy Econ. **7**, 501–514 (2005)
30. Ortuzar, J., Willumsen, L.G.: Modelling Transport. Wiley, New York (2011)
31. Piantanakulchai, M., Saengkhao, N.: Evaluation of alternatives in transportation planning using multi-stakeholders multi-objectives AHP modelling. In: Proceedings of the Eastern Asia Society for Transportation Studies, vol. 4, October 2003 (2003)
32. Prell, C., Hubacek, K., Reed, M.: Stakeholder analysis and social network analysis in natural resource management. Soc. Nat. Resour. Int. J. **22**, 501–518 (2009)
33. Quick, K., Narvaez, G. E., Saunoi-Sandgren E.: Changing minds through deliberation: citizens' accounts of their changing policy problem definitions and preferences. In: International Research Society for Public Administration, Ottawa, Canada (2014)
34. Raffaelli, G., Marsili, M.: A Statistical Mechanics Model for the Emergence of Consensus (2005). arXiv:cond-mat/0403030 [cond-mat.stat-mech]
35. Saaty, T.L.: The Analytic Hierarchy Process. McGraw Hill, New York (1980)
36. Saaty, T.L., Hu, G.: Ranking by eigenvector versus other methods in the analytic hierarchy process. Appl. Math. Lett. **11**(4), 121–125 (1998)
37. Scott, J.: Social Network Analysis. SAGE Publications Ltd, London (2013)
38. Sivilevičius, H., Maskeliūnaite, L.: The criteria for identifying the quality of passengers' transportation by railway and their ranking using AHP method. Transport **25**(4), 368–381 (2010)
39. Vermote, L., Macharis, C., Hollevoet, J., Putman, K.: Participatory evaluation of regional light rail scenarios: a Flemish case on sustainable mobility and land-use. Environ. Sci. Policy **37**, 101–120 (2014)
40. Voinov, A., Bousquet, F.: Modelling with stakeholders. Environ. Model Softw. **25**, 1268–1281 (2010)
41. Wefering, F., Rupprecht, S., Buhrmann, S., Bohler-Baedeker, S.: Guidelines. Developing and Implementing a Sustainable Urban Mobility Plan (2014). http://www.rupprecht-con

A Methodology of Redesigning and Evaluating Medium-Sized Public Transportation Systems

Jacek Żak[1(✉)] and Maja Kiba-Janiak[2]

[1] Poznań University of Technology, Piotrowo 3, 60-965 Poznań, Poland
jacek.zak@put.poznan.pl
[2] Wroclaw University of Economics,
3 Nowowiejska Street, 58-500 Jelenia Gora, Poland

Abstract. The chapter presents a generic methodology of carrying out a comprehensive redesign of a public transportation system. The proposed paradigm can be applied by municipal authorities (local governments) and management of urban transportation systems to upgrade local transportation. The authors propose two major stages of this methodology: 1. A heuristic construction of the public transportation system redesign variants/scenarios; 2. Multiple criteria evaluation of these scenarios. In the first stage 6 redesign concepts of the public transportation system are developed. The variants are constructed heuristically, based on "common sense", best practices and authors' expert knowledge in the field. In the second stage, the evaluation of redesign scenarios is defined as a multiple criteria ranking problem. A consistent family of criteria is defined, preferences of the decision maker (DM) and stakeholders are identified and modeled, and finally computational experiments are performed. In the computational phase selected Multiple Criteria Decision Making/Aiding (MCDM/A) methods, such as: AHP and ELECTRE III/IV are applied. The analysis is carried out based on a real world case study in a medium sized city in Poland. The generated results are compared with the intuitive decisions made by the officers of the local Town Hall.

Keywords: Multiple criteria evaluation · Public transportation system · Redesign · AHP method · Electre III/IV method

1 Introduction

The growing wealth of societies and continuous increase of citizens' aspirations towards the higher comfort of life result in a growing overall number of individual cars used in different countries [17]. This phenomenon generates, in turn, a higher number of passengers' movements that translates into increased congestion and substantial travel delays [16]. The situation is negatively reinforced by natural tendencies, such as: searching for dwellings at the suburbs of the metropolitan areas, called urban sprawl [6], growing mobility strongly correlated with the prosperity of the society or intensive development of commercial malls in the cities. The increased number of cars used, in particular, within the metropolitan areas, generates also other negative side-effects such as: increased levels of air pollution, noise and vibration, shrinking of the green and

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_4

recreational areas in the cities, that are replaced by transportation infrastructure and discomfort of pedestrians walking in the dense flows of cars or struggling to squeeze between the "crowds" of vehicles parked virtually everywhere (on the streets and pedestrian paths, in the continuously spreading designated parking areas, and/or on the bike routes).

The solution to these problems might be an efficient, effective and integrated system of public transportation that satisfies the passengers' expectations to carry out a smooth and comfortable door-to-door travel between an origin and destination. Unfortunately, in many metropolitan areas public transportation systems do not fulfil these requirements. In many cases, subway systems are overcrowded, public trams and buses are featured by poor technical condition, there are no convenient connections between private, individual movements and public transportation solutions (poor Park and Ride systems), timetables of different transportation modes are not integrated, local railways/city trains are poorly developed and not reliable (delayed and/or operating out of schedule), the passengers' comfort of travelling in many transportation means is not satisfactory [51]. As a result, one observes a significant increase of the motorization index with a simultaneous fall in a public transportation usage [33]. Forecasts of the European Commission indicate that by 2050 passengers' individual movements will increase by roughly 50% while the utilization of public transportation will further deteriorate [17].

This observation indicates that concrete actions should be undertaken to change this trend and encourage people to use public transportation more frequently and with more trust and desire. The authors of this chapter claim that many urban, public transportation systems require comprehensive redesign resulting in the improvement and enhancement of the service delivered. Redesign of the public transportation system is a complex process that consists in introducing substantial changes in several of its critical components [11]. The redesign of the public transportation system may involve route changes (extensions, eliminations, and reconstructions), relocation of stops, construction of integrated multi-modal transfer terminals, fleet reassignment, better coordination of schedules, and many others. As a result, different variants – transportation solutions for the public transportation system can be constructed. These variants, corresponding to the public transportation system alternative development scenarios, are to enhance the standard of travel, improve overall efficiency of the public transportation system and ensure the improvement of its major performance characteristics.

The proposed variants should be evaluated and the best/most desired option should be selected. There are several approaches and methodologies that allow to carry out such an evaluation. These include: Cost-Benefit Analysis (CBA) [12] and its variations, such as Cost-Effectiveness Analysis (CEA), Cost Utility Analysis (CUA); also the Economic Impact Analysis (EIA) and Social Return on Investment Analysis (SRoIA) [32], Regional Economic Impact Study (REIS), Environmental Impact Assessment (EIA), and Multiple Criteria Analysis (MCA) [13], often called Multiple Criteria Decision Making/Aiding (MCDM/A) [18, 29, 34, 53, 55].

Based on the comprehensive literature review the authors of this chapter would risk a statement that Cost Benefit Analysis (CBA) and its variations as well as Multiple Criteria Decision Making/Aiding (MCDM/A) may be considered as the most frequently used methodologies in the analysis and assessment of urban public transportation solutions [9, 32, 35, 46].

CBA is a technique/methodology used by researchers, analysts and decision-makers (mostly governmental bodies) to appraise the efficiency of certain transportation policies, projects or activities based on the aggregation and comparison of their overall costs and benefits expressed in monetary terms/units [7]. In CBA the analysed costs and benefits are adjusted for the time value of money, thus all cash flows (positive and negative) are discounted. Finally, they are aggregated into one generic index, called Net Present Value (NPV) that allows for balancing all costs and benefits associated with a certain variant (solution). The higher the surplus of financially expressed benefits the better the overall evaluation of the considered project. Several authors claim that CBA is theoretically unambiguous and methodologically consistent [5, 7]. The opponents of this methodology criticize the fact that CBA expresses all, often quite diversified, aspects and components of the analysis in financial terms and uses aggregated financial indexes while comparing different transportation options [5, 32, 35].

MCDM/A is a methodology that involves a comprehensive, multiple–dimensional analysis and evaluation of various variants, actions and solutions, including transportation ones, with respect to a number of criteria. In the MCDM/A many subjective, frequently contradictory interests and expectations of various stakeholders may be taken into account. This feature of MCDM/A is particularly important during the redesign process of the public transportation system, which requires an assessment of designed variants from different perspectives. The MCDM/A–based redesign of the mass transit system allows to considering the interests of different stakeholders and finding solutions that balance these interests accordingly. Similarly to the analysis of other comprehensive, urban, public transportation projects the assessment of the public transportation system redesign scenarios requires that the interests of the following major groups of stakeholders should be taken into consideration [54, 56]: local authorities (including the Board of the Urban Transportation System), operators of the public transportation system, travellers/users of the urban transportation system (including passengers of the mass transit system); local business units and other institutions and local community (residents). In many cases the expectations of these groups are in contrast and the MCDM/A methodology helps to find a comprise solution that satisfies them to a certain extend.

As opposed to CBA in MCDM/A both financially-oriented measures, i.e. such characteristics that can be expressed in financial terms (including the aspect of the value of money in time), and those that require non-financial interpretation e.g. of social, environmental or technical character, can be taken into account. All considered criteria can be expressed in their natural units and do not need to be aggregated into one universal, financial index. These features constitute major asset of the MCDM/A methodology [2, 18, 50] and contribute to its popularity in many areas, including transportation [9, 13, 43].

Interesting comparisons of both approaches are presented in the works of Annema et al. [1], Beria [5] and Lee [35]. Annema et al. [1] discuss the usefulness of CBA and MCDM/A for the assessment of transportation solutions from a politicians' perspective. They find out that politicians use CBA for the assessment of transportation solutions and activities but in a non-decisive manner and they consider its aggregate outcome (the composite result of CBA) pretentious. At the same time politicians seem to be interested in such transportation appraisal tools that would clearly show the

important political trade-offs resulting from any transportation policy, project or action. Thus, the authors of the above mentioned chapter claim that politicians appreciate the principles of MCDM/A but they do not apply them in a proper manner. Lee [32] shows a comprehensive overview of different transportation projects appraisel techniques applied at the US market. He investigates the suitability of both approaches in concrete situations and presents their strengths and weaknesses. Beria et al. [5] claim that CBA is more rigorous, transparent and formal than MCDM/A – based evaluation of transportation projects. At the same time, the authors indicate certain weaknesses of CBAsuch as: high labor intensity and costs as well as controversial character of financially – oriented aggregation of certain intangible magnitudes.

As mentioned above MCA or MCDM/A methodology has gained more and more appreciation and popularity among transportation researchers, recently. In the last 20 years, several successful applications of this methodology in transportation planning, analysis and control were reported. These include the works of: Chang and Shyu [9], Czyżak and Żak [10], De Brucker et al. [13], Filcek and Żak [19], Lee [32], Novak et al. [37], Satty [43], Tabucanon and Lee [46], Tan et al. [47], Żak [55], Żak and Kruszyński [58], Żak and Thiel [60], Żak [54], Żak et al. [59].

Various MCDM/A tools have been applied to analyze and optimize transportation. Gercek et al. [21] have applied Analytic Hierarchy Process (AHP) [42] to evaluate different rail transit projects for the European side of Istanbul. Hsu [24] have used a fuzzy Delphi AHP to the evaluation of the mass transit system in Kaohsiung. Ergun et al. [15] have carried out a comprehensive multiple criteria assessment of transportation alternatives for Istanbul metropolitan area. Lee [32] have shown a comprehensive review of various methods and approaches, including the multiple criteria ones used in transportation. Tan et al. [47] have presented a multiple objective formulation of a vehicle routing problem with time windows and an efficient, evolutionary algorithm of solving this problem. Tuzkaya and Onut [48] have applied a fuzzy ANP (Analytic Network Process) method to examine different modes for transporting freight between Turkey and Germany. De Brucker et al. [13] have applied MAMCA method (the combination of AHP and Promethee methods) [35] to a joint multiple criteria and group-oriented evaluation of diversified, complex transportation projects.

Żak has presented in his individual and co-authored works [10, 52, 54, 55, 60] a variety of decision models, multiple criteria methods and applied approaches to solving complex, multiple criteria decision problems arising in transportation. In some of his publications [10, 52, 54, 55] multiple criteria programming techniques have been utilized to optimize the mass transit system. Both deterministic and non-deterministic formulations of the multiple criteria vehicle assignment and vehicle scheduling problems for the existing public transportation network based on buses and trams have been presented. In the non-deterministic (fuzzy) case FLIP – Fuzzy Linear Integer Programming method [10, 52] have been applied to solve the vehicle assignment and scheduling problems under uncertainty. In the deterministic case the LBS – Light Beam Search method [28, 54] have been used to solve multiple criteria network and process optimization problems. In other works [28, 52, 57] a multiple objective hybrid evolutionary metaheuristic procedure, called Pareto Memetic Algorithm (PMA) has been applied to solve complex multiple objective optimization problems in passengers' transportation systems, such as: crew scheduling and vehicle assignment. In many

cases Żak and his co-authors [54, 56, 60] have formulated complex, transportation decision problems as multiple criteria ranking problems and have solved them with the application of different multiple criteria ranking methods, including ELECTRE and AHP. Żak and Thiel [60] and Żak [54] have applied Electre III/IV method to evaluate and rank several development scenarios of a medium-sized mass transit system. Żak and Fierek (2007) have compared the computational results generated by Electre III/IV and AHP methods for the assessment of alternative solutions integrating the urban transportation system. Żak and Kruszyński [58] have investigated the suitability of Electre III/IV and AHP methods for multiple criteria and multiple level evaluation of various urban transportation projects and their impact on the satisfaction of strategic, tactical and operational goals (objectives) of the metropolitan area.

In this chapter the authors have proposed a generic methodology of an MCDM/A-based redesign of the public, urban transportation system. They have split the developed procedure into two major stages, including: 1. A heuristic construction of the public transportation system redesign variants/scenarios; 2. Multiple criteria evaluation of these scenarios. All components/steps performed in these stages have been described. They have focused their analysis on a comprehensive, multiple criteria evaluation of the generated variants and have shown how the evaluation of the urban public transportation system should be performed. They have defined a universal, consistent family of criteria that can be applied to evaluate different redesign variants of any public transportation system. They have also presented how to model the DM's and stakeholders' preferences and how to generate rankings of the designed variants. In the computational phase, they have proven the applicability of two MCDM/A methods: AHP and ELECTRE III/IV to the evaluation of the public, urban transportation systems. The presented chapter is the extension of their previous work, presented in 2014 at the EWGT meeting in Seville [27]. The original work was more condensed and referred to the analysis of the transformation of the tram system, only. The current research is broader. It presents different redesign variants and covers the transformation of the whole public transportation system.

The authors have made an attempt to answer the following research questions:

- What is the most suitable methodology of redesigning the urban, public, transportation system? Why should it have a heuristic and multiple – criteria character?
- Can one apply both AHP and Electre III/IV methods to the evaluation of the urban, public transportation redesign variants?

The chapter is composed of 4 sections. The first one provides general background for the performed analysis. The second section introduces the readers into the concepts and basic principles of MCDM/A methodology, with particular emphasis on description of AHP and ELECTRE III/IV methods. A comprehensive Sect. 3 presents the general concept and the proposed methodology of the redesign of the considered urban, public transportation system. This section includes: problem statement, description of the redesign variants, definition of the evaluation criteria and all phases of computational experiments. In the last section, the authors formulate conclusions and define further directions of their research. The chapter is supplemented by a comprehensive list of references.

2 Multiple Criteria Decision Making/Aiding (MCDM/A)

2.1 Basic Concepts and Terms of MCDM/A

Historically, the background of Multiple Criteria Decision Making/Aiding (MCDM/A) has been shaped by the theoretical works of a mathematician Cantor and economists Edgeworth and Pareto in the second half of the 19th century [30]. These initial works were followed by landmark research carried out at the beginning of the 20th century by: Ramsey, van Neumann and Morgenstern, and Nash [18, 30]. Also the research carried out in the first half of the 20th century by Samuelson and Simon [30] contributed substantially to the development of MCDM/A. In the 1950s, Koopmans introduced the concept of "an efficient vector" and Kuhn and Tucker [30] formulated the problem of vector maximization and developed formal conditions for the existence of an efficient solution, paving the way for multiple objective mathematical programming [30]. In the 1960s and 1970s the split of the MCDM/A into two major methodological streams was initiated. This division shaped two major schools of MCDM/A: 1. The American one, based on Multi-Attribute Utility Theory [30] and 2. The European (French) one focused on the Outranking Relation [38–41, 50].

Currently MCDM/A can be considered as a mature but still dynamically developing field which aims at giving the DM tools and methods that enable him/her to solve a complex decision problem, where several – often contradictory – points of view must be taken into account [50]. MCDM/A does not yield "objectively best" solutions, because the solutions which are the best from all points of view, simultaneously, do not exist. Instead of that MCDM/A focuses on Pareto optimal and compromise solutions [2, 14, 45, 50, 53] that allow taking into account the trade-offs between criteria and the DM's preferences.

MCDM/A methodology is concentrated on solving the so called multiple criteria decision problem, i.e. a situation in which, having defined a set of actions/variants/solutions $\boldsymbol{V}$ and a consistent family of criteria F the DM tends to [50]:

- determine the best subset of actions/variants/solutions in $\boldsymbol{V}$ according to $\boldsymbol{F}$ (choice problem),
- divide $\boldsymbol{V}$ into subsets representing specific, pre-defined classes of actions/variants/solutions, according to concrete classification rules (sorting problem),
- rank actions/variants/solutions in $\boldsymbol{V}$ from the best to the worst, according to $\boldsymbol{F}$ (ranking problem).

In this chapter a multiple criteria ranking problem is considered.

The MCDM/A methodology distinguishes major stakeholders of a decision process: a decision-maker (DM), an analyst and interveners (stakeholders). In the MCDM/A – based decision making process the interests of these stakeholders are considered and the interaction between them takes place. Due to the complex nature of a decision problem, the decision process is strongly supported by different computer – based, mathematical and computer science methods, called MCDM/A methods. In general, they can be divided into 3 major families [18, 38, 40, 50, 53]:

- The methods of American inspiration based on the utility function;
- The outranking methods of the European origin;
- Interactive methods.

The first group includes such methods as: Analitic Hierarchy Process – AHP [42, 44] and UTilités Additives – UTA method [26]. Among methods based on European school, one can distinguish: Electre I-IV [4, 40, 41] and Promethee I and II [8]. The last group contains methods focused on the dialogue and interaction with the DM, including: STEP/STEM [3], GDF [20], VIG [31] and LBS [28]. In this chapter the authors apply and compare two MCDM/A methods: AHP and Electre III/IV.

Short description of the AHP method

The AHP method, developed by Saaty [42–44], is a computational algorithm that belongs to a family of multiple criteria ranking methods based on a multi-attribute utility theory [29]. It allows ranking a finite set of variants $\boldsymbol{A}$ from the best to the worst based on their evaluation with respect to a family of criteria $\boldsymbol{F}$. The AHP procedure supports structuring and solving a complex, multiple criteria, and hierarchical decision problem. The hierarchy of the decision problem is constructed through the definition of the overall objective, evaluation criteria, sub-criteria and variants.

The critical component of AHP method is the definition of the decision maker (DM) preferences based on pairwise comparisons of analyzed elements of the hierarchy- criteria and variants (on particular criteria) with the application of a grading scale from 1 to 9 [42, 44]. Each grading number, called relative weight $\boldsymbol{w_r}$ represents relative strength of the compared element against another and has a compensatory character, i.e.: the value characterizing the less important element (1/2, 1/5, 1/9) is the inverse of the value characterizing the more important element in the compared pair (2, 5, 9).

The algorithm of the AHP method focuses on finding a solution for a, so called, eigenvalue problem [42] on each level of the hierarchy and it can be divided into the following stages [44]:

Stage 1. Identification of the decision problem. Development of its hierarchy
Stage 2. Definition of preferences of DM (relative weights $\boldsymbol{w_r}$) in the form of matrices of pairwise comparisons of all elements of the hierarchy
Stage 3. Investigation of the consistency level of the preferential information given by the DM at each level of hierarchy (matrices of relative weights $\boldsymbol{w_r}$). Calculation of the consistency indexes $\boldsymbol{CI}$ ($\boldsymbol{CI} < 0.1$)
Stage 4. Computation of a set of vectors containing normalized values of absolute weights wa for criteria, sub-criteria and variants, summing up to 1 (100%)
Stage 5. Aggregation of the absolute weights wa by an additive utility function. Computation of the utility of each variant $\boldsymbol{i}$ – $\boldsymbol{U_i}$ which defines its position in the final ranking.

The final ranking of variants is constructed based on the computed values of the variants' utilities. More detailed description of the AHP method can be found in several publications [42–44, 49].

Basic features of the ELECTRE III/IV method

Electre III/IV method belongs to a common family of Electre methods. It was developed by Roy [40] in mid-1970s as an extension and upgrade of the previously proposed Electre I and Electre II methods [4]. The Electre III/IV algorithm is a multiple criteria ranking procedure, based on the outranking relation, which enables ordering of a finite set of variants from the best to the worst. The ranking of variants is generated as a result of their evaluation on a family of criteria $\boldsymbol{F}$, and with the application of preferential information submitted by the DM. The computational algorithm of the Electre III/IV method consists of three stages [40, 50, 53]:

- **Stage 1.** Definition of a set of variants $\boldsymbol{A}$ and a consistent family of criteria $\boldsymbol{F}$ combined with construction of an evaluation matrix. This stage is also focused on the definition of the DM's preference model. The preferential information is defined in the form of criteria weights - $\boldsymbol{w}$ (importance of criteria) and the indifference - $\boldsymbol{q}$, preference - $\boldsymbol{p}$ and veto - $\boldsymbol{v}$ thresholds. The thresholds define the sensitivity of the DM through the distinction of the following intervals of preference between variants on each criterion: indifference (up to $\boldsymbol{q}$), weak preference (between $\boldsymbol{q}$ and $\boldsymbol{p}$), (strong) preference (between $\boldsymbol{p}$ and $\boldsymbol{v}$) and incomparability (beyond $\boldsymbol{v}$). The principle that $\boldsymbol{q < p < v}$ applies.
- **Stage 2**. Determination of the outranking relation $\boldsymbol{S(a, b)}$, which indicates the extent to which "$\boldsymbol{a}$ outranks $\boldsymbol{b}$" overall. This relation is expressed by the degree of credibility $\boldsymbol{d(a, b)}$, being an equivalent of the global concordance indicator $\boldsymbol{C(a, b)}$ weakened by the discordance indexes $\boldsymbol{D_j(a, b)}$. The global concordance indicator $\boldsymbol{C(a, b)}$ is constructed in the concordance test, while the discordance indexes $\boldsymbol{D_j(a, b)}$ for each criterion $\boldsymbol{j}$ are computed in the discordance test. The values of $\boldsymbol{d(a, b)}$ are from the interval [0,1].
- **Stage 3**. Exploitation of the outranking relation $\boldsymbol{S(a, b)}$. Based on the computation of $\boldsymbol{d(a, b)}$ the method establishes two preliminary rankings (complete preorders) using a classification algorithm (distillation procedure). During this procedure, one can obtain a descending and an ascending preorder. In the descending distillation, the ranking process starts from the selection of the best variant, which is placed at the top of the ranking. In the ascending distillation, the variants are ranked in the inverse order. The final ranking is generated as an intersection of the above mentioned complete preorders. It can be presented either in the form of the ranking matrix or in the form of the outranking graph. The following situations can be distinguished there: indifference $\boldsymbol{(I)}$, preference $\boldsymbol{(P)}$, lack of preference $\boldsymbol{(P^{-})}$ and incomparability $\boldsymbol{(R)}$.

More information about the Electre methods can be found in the works of Roy and others [18, 39, 40, 50].

3 The Redesign of the Mass Transit System

3.1 Problem Statement

The necessity of the redesign of the urban public transportation system in a medium-sized metropolitan area (Gorzow city in Poland) has been revealed through its comprehensive analysis. The system, based on buses and trams, has the following features [27]:

- It is substantially imbalanced in terms of the modal split; the dominating component is bus transportation sub-system (85% of the vkm covered).
- As opposed to the bus sub-system (36 lines), the tram sub-system is underdeveloped. It is composed of 3 lines, only.
- The bus sub-system covers the area of 80 km^2 and it connects major parts of the city, while the tram sub-system is very fragmented and concentrated in the central and eastern part of the city only. It covers the area of 25 km^2.
- Both bus and tram sub-systems are old-fashioned and highly depreciated, although trams represent much poorer technical condition than buses. The average age of the bus fleet is 14 years, while the average age of tram fleet equals 47 years. These characteristics result in the low reliability of the means of transport and poor level of the comfort of travel.
- Buses are better equipped and adapted for passengers with special needs (disabled, the elderly and parents with small children) while trams do not satisfy the adequate standard of travel.
- Bus sub-system is featured by higher accessibility, while the tram system is less user-friendly. In the major housing estate, inhabited by about 40,000 people, bus stops are located very close to the main passengers' flows, while tram stops are located at a substantial distance (more than 500 m).
- The urban public transportation system offers a limited number of direct connections between major parts of the city. Thus, the level of passengers' transfers is relatively high.
- The system is not properly integrated. The integration measures, such as: multimodal terminals and stops, Park & Ride system, common tariff and ticketing system, passengers' information system, and common, separate bus – tram tracks are not in place.
- The public transportation system generates substantial on-going expenditures for fleet and infrastructure repairs and maintenance. Many vehicles operate underutilized which translates into a high level of empty vehicle-kilometres, in particular in the tram sub-system.

The above mentioned weaknesses of the urban public transportation system create a sense of passengers' reluctance towards public transportation services. The passengers complain mainly about the fragmentation, low accessibility, low reliability and unsatisfactory standard of travel (long travel times, delays, crowdedness, dirty vehicles). The operator of the public transportation system is also concerned due to low reliability, high utilization costs and low level of popularity of the mass transit system among passengers. Local authorities share the view of both passengers and the operator

and realize that the current level of efficiency and standard of travel offered by the urban transportation system is not satisfactory.

Thus, the City Board (the decision maker – DM) made a decision of redesigning the public transportation system to improve its overall efficiency and substantially increase the quality standard of service. The idea of the redesign consisted in the elaboration/definition of alternative transportation solutions corresponding to specific variants/scenarios of the urban, public transportation system development.

The proposed solutions of redesigning the urban public transportation system have been developed heuristically by experts (transportation engineers, administrative officers and scientists). These solutions constitute 7 variants, from V-0 to V-6. Variant V-0 corresponds to an existing public transportation system (projected for 2025), while variants V1, V2 to V6 correspond to various restructuring strategies of this system, projected into the same time horizon [22]. They introduce less and more radical changes in the urban public transportation system depending on the extent of its transformation.

In order to develop the variants expectations of the DM and different stakeholders have been examined through the survey research and interviews. The interests of local authorities, operator, passengers and local residents (community) have been taken into account. Thus, the redesign has been based on satisfying the needs and requirements of the above mentioned groups of DM and stakeholders. The DM declared that the main objective of this quite comprehensive and radical transformation of the urban public transportation system is the improvement of the quality of life in the city and enhancement of efficiency and effectiveness of a public transportation system.

3.2 Description of Variants

Variant 0 (V0) represents the existing public transportation system. The system is based on trams and buses, and delivers service to approximately 125,000 inhabitants. It covers the area of 110 km2, including: 85% of the area covered by buses and 15% of the area covered by trams. The bus sub-system is composed of 36 routes of the total length of 432 km. It is spread within the whole city. At the same time, the tram sub-system includes only 3 routes with the total track length of 25 km. There tram system connect inhabitants from the West-South and North and East-North parts of the city with the city centre. There are 69 buses and 35 tramcars operating in the system. The bus fleet has an average age of 16 years, while the tramcars have been in use for more than 28 years. The means of transport are highly depreciated and their failure rate is very high. The overall number of vehicle-kilometres covered by the system in variant V0 is 5 981 and the modal split reflects the following proportion: 84% (5002 vkm) for buses and 16% (913 vkm) for trams. Some 66% of all tramcars and buses in the public transportation system are adapted for the disabled and the elderly. In the public transportation system, the overall utilization rate of the vehicles is four passengers per vehicle-kilometre. In this variant. investment costs are the lowest and includes only modernization of transportation means and roads. However, operating costs are quite high and are similar to variants: V3 and V4.

Variant 1 (V1) called the evolutionary option, introduces substantial but not drastic changes in the public transportation system. The newly designed solutions focus on reaching one of the largest housing estates in the city and the regional hospital by tram transportation system (e.g. Dekerta Street, Czartoryskiego Street, Górczyńska Street, Okulickiego Street,). On the one hand, this is a convenient solution for local residents. On the other hand, it may cause some difficulties related to the removal of the eastern belt of allotments. Moreover, the planned route extensions run in direct contact with residential buildings. The introduction of the new tram route, connecting the city center with the hospital, generates certain difficulties in the neighbourhood of the hospital where traction would be directly adjacent to the roadway. Variant V1 covers the same area as Variant V0 but the share of area covered by trams increases in this variant to 25%. The bus sub-system is composed of 33 routes with the total length of 393 km, while the tram sub-system includes 5 routes whose total track length equals 37 km. There are 63 buses and 50 tramcars operating in the system. The bus fleet has an average age of 15 years, while the tram fleet is more than 20 years old. The overall number of vehicle-kilometres covered by the system is 6,402 and the modal split is 70% (4,476 vkm) for buses and 30% (1926 vkm) for trams. The share of tramcars and buses adapted for the disabled and the elderly in the total number of vehicles of the public transportation system is at the level of 81%. The overall utilization rate of the public transportation means in this variant is three passengers per vehicle-kilometre. The investment costs are much higher than in Variant V0 and include mainly costs of tram routs extensions and transportation infrastructure modernization. The extension of the tram system causes in this variant an increase in operating costs, which are higher than in the variant V0.

Variant 2 (V2) is a tram-oriented solution representing more radical changes in the public transportation system than those introduced in variant V1. It is proposed in this variant to further expand the tram transportation system in the city. Variant V2 may be considered the more advanced transformation of variant V1. In addition to the changes proposed in variant V1, variant V2 is featured by two extra tram routes running to the city center from the north-eastern and the western parts of the largest housing estate (e.g. Dekerta Street, Czartoryskiego Street, Górczyńska Street, Okulickiego Street, Piłsudskiego Street, Roosvelta Street), respectively. It must be mentioned, that the planned revitalization of the city centre may exclude these additional routes from utilization for a period of time. Variant V2 covers the same area as variant V0 but the tram sub-system coverage increases to 35%. The bus sub-system is composed of 31 routes with the total length of 389 km, while the tram sub-system includes 7 routes whose total track length amounts to 39 km. There are 62 buses and 50 tramcars operating in the system. The average age of the bus and tram fleet is 14 and 20 years, respectively. The overall number of vehicle-kilometers covered by the system is 6447 and the modal split is 69% (4432 vkm) for buses and 31% (2015 vkm) for trams. The number of tramcars and buses adapted for the disabled and the elderly in the total number of vehicles of the public transportation system is the same as in variant V1. The overall utilization rate of public transportation system in variant V2 is the same as in variant V0. The investment costs and annual operating costs are insignificantly higher than in variant V1.

Variant 3 (V3) is a conservative redesign solution featured by a marginal transformation of the public transportation system. It is a limited edition of variant V1, which is simple to implement, but not necessarily beneficial for the passengers. This variant introduces a basic extension by one additional route of the existing tram transportation network to the largest housing estate (Piłsudskiego Street, Górczyńska Street). Variant V3 covers the same area as variant V0, including the tram coverage that remains in this variant at the level of 15%. The bus sub-system is composed of 34 routes with the total length of 415 km, while the tram sub-system includes 4 routes whose total track length is 31 km. 66 buses and 43 tramcars operate in the public transportation system envisaged by variant V3. The average age of the bus fleet is 15 year, while the age of the tramcars is 23 years. The overall number of vehicle-kilometers covered by the system is 6188 and the modal split is 76% (4691 vkm) for buses and 24% (1497 vkm) for trams. There are 81% of all tramcars and buses in the public transportation system adapted for the disabled and the elderly. The overall utilization rate of the public transportation system in variant V3 is the same as in variants V0 and V2 and amounts to four passengers per vehicle-kilometer. The investment costs are lower than in V1 and V2 and include the extension of the tram subsystem. The operating costs are 1,6 mln higher than in variant V0 but almost 2 mln lower than in variants V1 and V2.

Variant 4 (V4) represents a concept of a bus-intensive public transportation system in which the tram system is further reduced. It is assumed in this variant the tram routes running through the city center are replaced by an eco-friendly bus transportation system running nearby the previous tram line (8 electric buses). Therefore, the problem of movements within the city center during the revitalization period is avoided. Elimination of two tram lines brings both benefits and difficulties. Currently both tram lines go through the city centre which could be adapted for pedestrians. The disadvantage of this solution is greater dystance to bus stops (about 300 metres) to the existing tram stations. The buses in this variant run paralel street (Kosynierów Gdyńskich street) compared to the previous tram tracks (Chrobrego Street, Mieszka I Street). This may cause demotivate residents to use public transport. Variant V4 covers the same area as variant V0 but the tram coverage decreases in this variant to 5%. The bus sub-system is composed of 37 routes with the total length of 480 km, while the tram sub-system includes only 1 route whose total track length is 18 km. A total of 77 buses and 28 tramcars operate in the system. The bus fleet has an average age of 14 years. At the same time, the tramcars are more than 23 years old. The overall number of vehicle-km covered by the system is 5981 and the modal split is 91% (5451 vkm) for buses and 9% (530 vkm) for trams. In variant V4 there are 87% of all tramcars and buses adapted for the disabled and the elderly. The overall utilization rate of the public transportation system in this variant is higher than in variant V2 and V3 and equals 5 passengers/vehicle-km. The investment costs in this variant are insignificantly lower than in V0 however annual operating costs are higher than in V0 and lower than in V3.

Variant 5 (V5) can be called the environmentally friendly option. It is a combination of variant V2 and extended version of variant V4, thus it represents very radical changes in the public transportation system. It is proposed in this variant to expand the tram transportation system in the same way as in variant V2 (access by tram to the

streets such as: Dekerta, Czartoryskiego, Górczyńska, Okulickiego, Piłsudskiego, Roosvelta) and introduce a substantial number of eco-friendly electric buses (replacing all currently used, traditional city buses) in the whole public transportation system. As a result, it is anticipated that variant V5 will face the same difficulties as variant V2 but at the same time it may generate substantial environmental benefits. Variant V5 covers the same area as variant V2. The bus sub-system is composed of 31 routes with the total length of 389 km, while the tram sub-system includes 7 routes whose total track length amounts to 39 km. There are 62 electric buses and 50 tramcars operating in the system. The medium-sized electric buses have a capacity of 61 passengers each. Their riding/driving range distance-wise amounts to roughly 45 km which corresponds to approximately 5 h of driving time-wise. The average age of the tram fleet is 20 years, while all the electric buses are brand new. The overall number of vehicle-km covered by the system is 6447 and the modal split is 69% (4432 vkm) for buses and 31% (2015 vkm) for trams. The number of tramcars and buses adapted for the disabled and the elderly in the total number of vehicles of the public transportation system is the same as in variant V2. The overall utilization rate of public transportation system in V5 is 5 passengers per vehicle-km and is the same as in variant V4. The investment costs in this variant are very high and equal 506 mlnzl. This amount includes purchasing of new electric buses, modernization of existing tram subsystem, etc. At the same time, the projected operating costs are much lower than in the previous variants. The reduction of operating costs in variant V5 results from the lower energy consumption costs for electric buses [26].

Variant 6 (V6) called the balanced/sustainable solution, represents the holistic approach to the transformation of the urban public transportation system and features the most radical changes in its existing sub-systems. It is inspired by a bicycle – oriented connection of the suburbs and city center implemented in Shanghai, China. In addition to the changes proposed in variant V5, the public bicycle rental system is introduced in variant V6 for those citizens whose dwellings are located outside the administrative boundaries of the city (metropolitan suburbs). Residents who live in these areas are slightly handicapped as far as accessibility to the regular mass transit system is concerned. Therefore, many of them have to use private/individual cars to get to their destinations (work, school). Thus, the introduction of a public bicycle rental system operating primarily as a feeder sub-system for the public transportation system in the city could significantly improve the connection between the city centre and its remote areas.

It is assumed that rented bicycles are available at public bicycle rental stations (PBRS-s), located at the suburbs of the metropolitan area in three housing estates: Łupowo (PBRS 1), Kłodawa (PBRS 2), Wawrów (PBRS3). Each station is equipped in 15 public, locked bicycles and 5 spare locking stands. These suburban bicycle stations correspond to the centrally located bicycle stations: PBRS 4, PBRS 5 and PBRS 6 respectively to ensure good integration of the public transportation solutions and convenient door-to-door service for passengers. PBRS 4 is corresponding to the PBRS 2 and is located 3,3 km away from it at the major bus stop, closest to the local medical center (Dekert Hospital). PBRS 5, corresponding to the PBRS 1 (distance of 4,6 km) is placed at one of the tram loops (Wieprzyce). PBRS 6 located 1,4 km away from PBRS 3 is placed at the major bus loop (Ustronie), which provides access to 8 bus routes.

Variant V6 assures the integration of the bicycle-bus-tram transportation corridors, which substantially increases the mobility of those residents whose dwellings are located at the suburbs and far away from the public transportation system. As described in variant V0 those areas have irregular and rare (one vehicle per every two/three hours) connection with a city center. Thus, it is assumed that public bicycle rental system should facilitate residents their access to the regular public transportation system.

The bicycle rental system should also influence on the overall modal split in the city. In variant V6 the total number of vehicle-km covered by the public transportation system is 7005 vehicle-km, including 4553 vehicle-km for buses (65%), 2172 vehicle-km for trams (31%) and 280 vehicle-km for bicycle (4%). The utilization rate in this variant is the higest and amounts to 6 passengers per vehicle-km. In addition, variant V6 is featured by the highest investment costs that equal 507 mln PLN. This sum includes purchasing of new electric buses, bicycles, bicycle stations and modernization of existing tram subsystem. The operating costs in variant V6 are at the similar level as in variant V5.

3.3 Definition of the Family of Criteria

Based on a thorough literature review and taking into consideration various aspects and interests of different stakeholders (see Sect. 3.1) the authors have constructed a consistent family of 11 criteria (C1–C11) that allow the evaluation of the proposed redesign variants of the public transportation system [32, 36, 437, 56]. They are as follows:

- **C1 – Accessibility**, measuring the aggregated "closeness" and availability of the public transportation system (particular variants) for different categories of users, maximized criterion. It is composed of such sub-characteristics as: density of the transportation network, availabilty of vehicles (including those adopted for elderly and disabled), accessibility to/from the suburban areas; maximized criterion.
- **C2 – Reliability**, charactering the overall trust of passengers towards public transportation system. It covers such aspects as timetable realization and timeliness; maximized criterion.
- **C3 - Quality of the Means of Transport**, defined as an aggregated measure of the standard of vehicles used in the public transportation system (particular variants). It is composed of such sub-characteristics as share of brand new and modernized vehicles and age of trams and buses; maximized criterion.
- **C4** - **Waiting time**, defined as an average passenger waiting time in the entire public transportation system (particular variants) during the peak hours; minimized criterion.
- **C5** - **Directness of Connections**, defined as a share (percentage) of tram and bus lines/routes providing direct connection between the suburbs and the city centre in the total number of tram and bus lines operating in the public transportation system (particular variants); maximized criterion.

- **C6 -Investment Costs**, formulated as a total capital expenditure associated with the implementation of transportation solutions characterizing particular variants; minimized criterion.
- **C7 –Operating Costs**, defined as annual operating costs of the public transportation system (particular variants); minimized criterion.
- **C8 – Efficiency**, characterizing the overall productivity (output) of the public transportation system and level of its utilization. It concentrates on the analysis of the efficiency of vehicles, including: fleet pproductivity (number of passengers per vehicle-km) and fleet utilization; maximized criterion.
- **C9 – Social Sustainability**, measuring the oveall, aggregated social quality of the public transportation system (particular cariants), including such aspects as safety and security of inhabitants, the impact of the public transportation system on their health, comfort of life and mobility, maximized criterion.
- **C10 – Environmental Friendliness**, characterizing the overall environmental sustainability of the public transportation system (particular variants), including such aspects as level of hazardous emissions and share of "green vehicles" used.
- **C11 – Nuisance of variants' implementation** [points], characterizing overall level of diffulties associated with the implemantaion of particular variants, including such aspects as: spatial difficulties, residents' inconvenience and technical constraints, minimized criterion.

It is worth pointing out that the number of criteria is large and exceeds the manageable level of 7 ± 2 characteristics. The authors have made an effort to reduce it but such a reduction have resulted in constructing a "handicapped family of criteria" which has not satisfied the conditions of consistency [18, 50]. The elimination of certain characteristics did not allow for covering all the required aspects of evaluation and interests of all considered stakeholers. In the authors' opinion the proposed 11 criteria guarantee a complete evaluation of a public transportation system, including its redesign variants. They cover such aspects as: social (accessibility, social sustainability, nuisance of variants' implementation); safety and quality of travel (reliability, waiting time, directness of connections); technical (quality of the means of transport); economic (investment and operating costs, efficiency) and ecological - environmental friendliness (the impact of a public transportation system on environment, share of vehicle-km covered by tramcars and electric buses). In addition, they represent the interests of the following stakeholders: criteria C1, C2, C3, C4 and C5 express passengers' expectations, criteria C9, C10 and C11 correspond to the interests of the local community (residents), while criteria C6, C7 and C8 define points of view of local authorities and the operator. All the criteria guarantee a complete evaluation of a public transportation system, including its redesign variants and assure an appropriate standard/quality of generated decisions.

6 criteria: C4, C5, C6, C7, C9, C11 are formulated as single, aggregated measures, while the 5 remaining ones: C1, C2, C3, C8, C10 include 2-4 sub-criteria. Criterion C10 is a particular case, which consists of 2 sub-criteria and 4 sub-sub-criteria. The names of sub-criteria and values of all characteristics for 7 considered variants are presented in the evaluation matrix – Table 1. All data presented in Table 1 are computed based on the forecasts for 2025. In the majority of cases (criteria:

C1(except C1.3); C2, C3, C4, C5, C6, C7, C8, C10) they are defined based on raw data collected in the public transportation system and characterizing its operations. These characteristics are defined as concrete quantitavive ratios. In case of 3 characteristics: C1.3, C9 and C11 subjective impressions of the DM and stakeholders have been recognized through interviews and guided discussions. The DM and stakeholders have expressed their subjective evaluations in qualitative, linguistic terms, such as: "poor", "average", "good", "very good" and "excellent", looking at different aspects of each considered characteristic. These linguistic evaluations have been transferred to a quantitative scale of 1–6 or 1–10 points (depending on the considered characteristic). All quantified evaluations given by the representatives of the DM and stakeholders have been aggregated into a final integer score in a brainstorming session.

Table 1. The evaluation matrix for seven redesign variants (V0 – V6) of the public transportation system

Criteria		Variants						
		V0	V1	V2	V3	V4	V5	V6
C1. Accessibility	C1.1. The total length of the tram and bus transportation network to the area of the city [km/km2], maximized	5,31	5,09	4,98	5,19	5,79	4,98	4,98
	C1.2. The share of public means of transport adapted for the disabled and the elderly, [%], maximized	80%	81%	81%	81%	87%	81%	81%
	C1.3. Accessibility to the suburban areas with the use of public bicycle rental system, [points], maximized	3,00	3,00	3,00	3,00	3,00	3,00	6,00
	C1.4. Total average daily number of vehicles operating in the city per 1000 inhabitants, [-], maximized	0,60	1,16	1,16	0,88	0,37	1,16	1,16
C2. Reliability	C2.1. Share of rides carried out according to a timetable, [%], maximized	95%	96%	96%	95%	95%	96%	96%
	C2.2. Share of rides carried out in a timely manner, [%], maximized	82%	85%	85%	84%	81%	85%	85%
C3. Quality of the means of transport	C3.1. Share of brand new vehicles introduced to the system, [%], maximized	14%	26%	26%	21%	13%	26%	26%
	C3.2. Share of modernized vehicles as percentage of the total number of vehicles, [%], maximized	17%	16%	16%	17%	11%	16%	16%
	C3.3. Average age of tramcars, [years], minimized	28	20	20	23	23	20	20
	C3.4. Average age of buses [years], minimized	16	15	15	14	14	14	14
C4. Waiting time, [minutes], minimized		4	4	3,5	3,5	5	3,5	3,5
C5. Directness of connections, [-], maximized		8%	13%	17%	11%	3%	17%	17%
C6. Investment Costs, [mlnzl], minimized		129	395,1	406,8	246,2	185,3	506	507
C7. Operating Costs [mlnzl], minimized		41,0	44,2	44,5	42,6	40,6	20,9	20,9

(*continued*)

Table 1. (*continued*)

Criteria			Variants						
			V0	V1	V2	V3	V4	V5	V6
C8. Efficiency	C8.1. Number of passengers per vehicle-km in the public transportation system, [-], maximized		4	3	4	4	5	5	6
	C8.2. Fleet utilization rate, [%], maximized		19%	24%	27%	20%	20%	27%	29%
C9. Social sustainability [points], maximized			2	6	7	3	2	8	9
C10. Environmental friendliness	C10.1. The impact of the public transportation system on environment	C10.1.A. Average level of COemission, [g/km], minimized	1,76	1,41	1,38	1,57	1,86	0,00	0,00
		C10.1.B. Average level of HC emission, [g/km], minimized	0,43	0,33	0,33	0,38	0,43	0,00	0,00
		C10.1.C. Average level of NOx emission, [g/km], minimized	2,71	2,11	2,06	2,38	2,72	0,00	0,00
		C10.1.D. Average level of PM emission [units/kWh], minimized	92	65	63	74	92	0	0
	C.10.2. Share of vehicle-km covered by tramcars and electric buses in all vehicle-km, [%], maximized		15%	30%	31%	24%	18%	100%	100%
C11. Nuisance of variants' implementation [points], minimized			1	8	9	3	2	9	9

3.4 Computational Experiments

Computational experiments have been conducted with the application of Electre III/IV and AHP methods (described in Sect. 2), implemented in the "Electre III/IV Toolkit" and "Make It Rational" software packages, respectively. The experiments carried out with the application of the AHP method has been based on the source, raw data, presented in the evaluation matrix (Table 1). Thus, in this case all criteria and sub-criteria have been utilized in its original form. At the same time the application of the Electre III/IV method has required that all sub-criteria have been aggregated into their corresponding superordinate criteria. Thus, in this case only the "pure" criteria defined without any sub-criteria have been characterized by the original, raw data. For all the remaining criteria that have contained some sub-criteria the original data have been normalized [23] and transformed into a 0–1 interval. Different formulas of normalization have been applied for maximized and minimized sub-criteria. Afterwards, all normalized values of sub-criteria have been aggregated with the application of the

Table 2. The Evaluation Matrix based on the transformed data, adjusted to the requirements of ELECTRE III/IV method

Criteria	Variants						
	V0	V1	V2	V3	V4	V5	V6
C1. Accessibility [points: 0–1)	0,30	0,62	0,66	0,47	0,30	0,65	0,75
C2. Reliability [points: 0–1]	0,1	1	1	0,6	0,3	1	1
C3. Quality of the means of transport [points: 0–1]	0,30	0,72	0,74	0,55	0,13	0,74	0,74
C4. Waiting time [minutes]	4	4	3,5	3,5	5	3,5	3,5
C5. Directness of connections [%]	8	13	17	11	3	17	17
C6. Investment costs [mlnzl.]	129	395,1	406,8	246,2	185,3	506	507
C7. Annual operating costs [mlnzl.]	41,0	44,2	44,5	42,6	40,6	20,9	20,9
C8. Efficiency [points: 0–1]	0,17	0,48	0,59	0,36	0,46	0,5	1
C9. Social sustainability [Points: 1–9]	2	6	7	3	2	8	9
C10. Environmental friendliness [points: 0–1]	0,01	0,22	0,23	0,13	0,01	1,00	1,00
C11. Nuisance variants' implementation [Points: 1–9]	1	8	9	3	2	9	9

weighted averages to generate the final, normalized values of criteria. This procedure has been applied to criteria C1, C2, C3, C8 and C10. The transformed data, constituting the content of the Evaluation Matrix, used in the computational experiment with the application of ELECTRE III/IV method, are presented in Table 2.

In the next step the models of DM's preferences for the Electre III/IV and AHP methods have been developed. In both cases, they have included two major components of the preference identification, i.e.: definition of the importance of criteria and the definition of the DM's sensitivity towards the changes of the criteria values. The respective models of the DM's preferences are presented in Tables 3, 4, 5 and 6. In the Electre III/IV method the importance of the individual criteria have been determined by weights – w on the linear scale of 0–20 points (Table 3), while in the AHP method it has been defined with the application of relative weights w_r on the scale of 1 to 9 points (Table 4), through pairwise comparisons of criteria. As described in Sect. 2.2 the grading system in the AHP method has a compensatory character, i.e. the more important criteria in the compared pairs are characterized by integer numbers larger than 1 (2 to 9 points), while the less important characteristics are described by fractional values (1/9 to 1/2 points). Value 1 is reserved for the equally important elements.

In both preference models the most important criteria are: C2, C1 that refer to the quality of the public transportation service and C6, C7 that represent financial – economic parameters. The former define the interests of passengers while the latter the expectations of local authorities and operator. The least significant criteria are C11 and C5, concerning nuisance of the variants' implementation and directness of connections, respectively. Seven criteria (C1–C3; C5, C8–C10) are maximized (Preference direction

Table 3. The model of the DM's preferences characteristic for the Electre III/IV method

Criterion	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
Weight (w)	18	20	12	8	4	15	16	14	10	6	2
Preference direction	+1	+1	+1	−1	+1	−1	−1	+1	+1	+1	−1
Indifference (q)	0,12	0,1	0,1	0,5	2	15	0,5	0,1	1	0,1	1
Preference (p)	0,3	0,3	0,3	2	5	50	10	0,3	4	0,4	4
Veto (v)	0,53	0,7	0,65	5	10	300	30	0,7	8	0,9	9

Table 4. The model of the DM's preferences characteristic for the AHP method. Pairwise comparisons of criteria

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	W_r (%)
C1	1	0.50	4.00	5.00	7.00	2.00	2.00	3.00	5.00	6.00	8.00	0.188
C2	2.00	1	5.00	7.00	8.00	3.00	3.00	4.00	6.00	7.00	9.00	0.266
C3	0.25	0.20	1	3.00	5.00	0.33	0.33	0.50	2.00	3.00	5.00	0.064
C4	0.20	0.14	0.33	1	2.00	0.25	0.25	0.33	0.50	2.00	3.00	0.033
C5	0.14	0.12	0.20	0.50	1	0.17	0.17	0.20	0.33	0.50	2.00	0.02
C6	0.50	0.33	3.00	4.00	6.00	1	1.00	2.00	4.00	5.00	7.00	0.128
C7	0.50	0.33	3.00	4.00	6.00	1.00	1	2.00	4.00	5.00	7.00	0.128
C8	0.33	0.25	2.00	3.00	5.00	0.50	0.50	1	3.00	4.00	6.00	0.087
C9	0.20	0.17	0.50	2.00	3.00	0.25	0.25	0.33	1	2.00	4.00	0.043
C10	0.17	0.14	0.33	0.50	2.00	0.20	0.20	0.25	0.50	1	3.00	0.027
C11	0.12	0.11	0.20	0.33	0.50	0.14	0.14	0.17	0.25	0.33	1	0.015

Table 5. The model of the DM's preferences characteristic for the AHP method. Pairwise comparisons of variants for a selected criterion - C6 – Investment costs

	V0	V1	V2	V3	V4	V5	V6
V0	1	6.00	7.00	4.00	2.00	9.00	9.00
V1	0.17	1	2.00	0.25	0.20	7.00	7.00
V2	0.14	0.50	1	0.33	0.20	3.00	3.00
V3	0.25	4.00	3.00	1	0.50	3.00	3.00
V4	0.50	5.00	5.00	2.00	1	8.00	8.00
V5	0.11	0.14	0.33	0.33	0.12	1	1.00
V6	0.11	0.14	0.33	0.33	0.12	1.00	1

described by value +1 in Table 3) while the remaining 4 (C4, C6, C7 and C11) are minimized (Preference direction described by value −1 in Table 3).

The sensitivity of the DM in the Electre III/IV method has been described by the values of thresholds: q, p, v, while in the AHP method by pairwise comparisons of variants on particular criteria. In the Electre III/IV method the thresholds for criteria: C4, C5, C6, C7, C9, C11(those without sub-criteria) have been expressed in the

Table 6. The model of the DM's preferences characteristic for the AHP method. Pairwise comparisons of variants for a selected criterion – C4. Waiting time

	V0	V1	V2	V3	V4	V5	V6
V0	1	1.00	0.50	0.50	4.00	0.50	0.50
V1	1.00	1	1.00	0.50	3.00	0.50	0.50
V2	2.00	1.00	1	1.00	6.00	1.00	1.00
V3	2.00	2.00	1.00	1	6.00	1.00	1.00
V4	0.25	0.33	0.17	0.17	1	0.17	0.17
V5	2.00	2.00	1.00	1.00	6.00	1	1.00
V6	5.00	1.00	1.00	3.00	8.00	1.00	1

original units of these criteria, while the thresholds for the remaining criteria C1, C2, C3, C8 and C10 (those containing sub-criteria) have been defined in relative normalized values (see Table 3). In the AHP method, the DM's sensitivity is defined for all variants on all criteria and sub-criteria on the 1 to 9 scale. The results of these comparisons are presented in Tables 5 and 6 for criteria: C6 – Investment costs and C4 – Waiting time, respectively. Again, the preferred variants are characterized by integer numbers larger than 1 (2 to 9 points), while the non-preferred variants are described by fractional values (1/9 to 1/2 points).

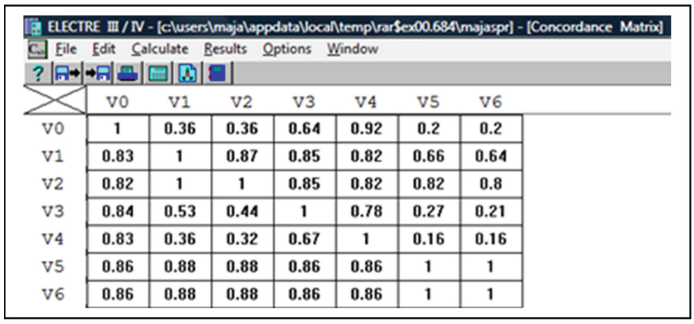
Based on the above described models of preferences all steps of the computational procedures of Electre III/IV and AHP methods (described in Sect. 2) have been performed. The application of ELECTRE III/IV method required the following calculations:

- Concordance indicators ***C(a, b)***, presented as a concordance matrix.
- Discordance indexes $\boldsymbol{D_j(a, b)}$.
- Outranking relation ***S*** expressed by the degree of credibility ***d(a, b)*** and presented in the form of Credibility matrix.

Figure 1 presents the first stage of these computations, i.e. the image of the Concordance matrix. As indicated in Fig. 1 the values of concordance indicators ***C(a, b)*** belong to the <0, 1> interval. They define the degree to which a outperforms b counting only the arguments "for", i.e. the values of those criteria at which a is better than b. Based on the generated results one can easily conclude that variants V5 and V6 are featured by the highest values of concordance indicators ***C(a, b),*** ranging between 0, 86 and 1. At the same time the variants characterized by the lowest values of concordance indicators ***C(a, b),*** ranging between 0,21 and 1 and 0,16 and 1 are V3 and V4, respectively.

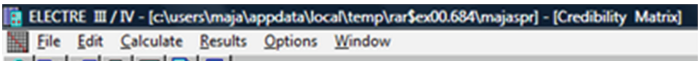
In the next step the discordance indexes $\boldsymbol{D_j(a, b)}$ have been computed and their values have been combined with the values of concordance indicators ***C(a, b).*** As a result the outranking relation ***S(a, b)*** expressed by the degree of credibility ***d(a, b)*** have been calculated. The values of ***d(a, b)*** presented in the form of the Credibility matrix indicate the extent to which "a outranks b" overall. Both arguments "for" and "against" are taken into account (Fig. 2).

Based on the generated results one can see, for instance, that variant V0 has no chances to outperform variants V1, V2, V5 and V6 [$\boldsymbol{d(a, b) = 0}$], while variants V5 and



	V0	V1	V2	V3	V4	V5	V6
V0	1	0.36	0.36	0.64	0.92	0.2	0.2
V1	0.83	1	0.87	0.85	0.82	0.66	0.64
V2	0.82	1	1	0.85	0.82	0.82	0.8
V3	0.84	0.53	0.44	1	0.78	0.27	0.21
V4	0.83	0.36	0.32	0.67	1	0.16	0.16
V5	0.86	0.88	0.88	0.86	0.86	1	1
V6	0.86	0.88	0.88	0.86	0.86	1	1

Fig. 1. Concordance Matrix generated by Electre III/IV method, containing the concordance indicators ***C(a, b)*** for all considered variants V0, V1, …, V6 of the redesign of the urban, public transportation system



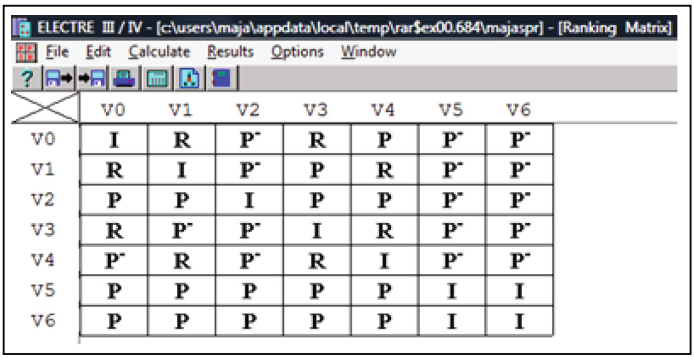
	V0	V1	V2	V3	V4	V5	V6
V0	1	0	0	0.64	0.92	0	0
V1	0.65	1	0.87	0.85	0.82	0.39	0.32
V2	0.41	1	1	0.85	0.82	0.82	0.79
V3	0.84	0.53	0.016	1	0.78	0	0
V4	0.83	0	0	0.67	1	0	0
V5	0	0.88	0.88	0.86	0	1	1
V6	0	0.88	0.88	0.86	0	1	1

Fig. 2. Credibility Matrix generated by Electre III/IV method, containing the degrees of credibility ***d(a, b)*** for all considered variants V0, V1, …, V6 of the redesign of the urban, public transportation system

V6 are featured by high values of ***d(a, b)***, ranging between 0, 86 and 1 for the majority of pairwise comparisons between variants. As opposed to variants V5 and V6, variant V4 is characterized by low values of ***d(a, b)*** [4-fold ***d(a, b) = 0***].

In the final stage of the computational procedure the outranking relation ***S(a, b)*** is exploited and two preliminary rankings (descending and ascending complete preorders) are generated. The final ranking is an intersection of the above mentioned complete preorders. It can be presented either in the form of the Ranking matrix or in the form of the outranking graph. In the analyzed case Fig. 3 presents the Ranking matrix while Fig. 4 demonstrates its corresponding outranking graph.

In the Ranking matrix the following situations can be distinguished: indifference (***I***), preference (***P***), lack of preference (***P-***) and incomparability (***R***). They are equivalents of the following situations in the outranking graph: the variants are placed in the same box - (***I***); one variant precedes another variant (***P***); one variant follows its counterpart (***P-***); variants are not interconnected by any arrow (***R***).



	V0	V1	V2	V3	V4	V5	V6
V0	I	R	P-	R	P	P-	P-
V1	R	I	P-	P	R	P-	P-
V2	P	P	I	P	P	P-	P-
V3	R	P-	P-	I	R	P-	P-
V4	P-	R	P-	R	I	P-	P-
V5	P	P	P	P	P	I	I
V6	P	P	P	P	P	I	I

Fig. 3. Ranking matrix generated by Electre III/IV method containing the final relationships between all considered variants V0, V1, …, V6 of the redesign of the urban, public transportation system

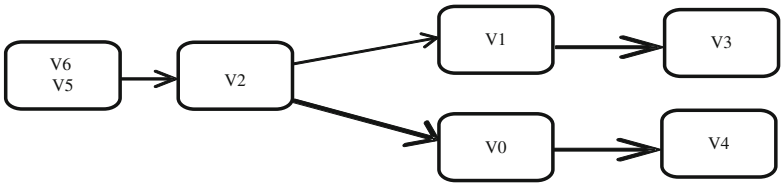


Fig. 4. Final ranking of the redesign variants of the public transportation system generated generated by Electre III/IV

As presented in the Ranking matrix two variants V5 and V6 substantially outperform the remaining variants. Their rows are featured by a large number of (***P***) relationships, which indicates that these two variants outperform all their competitors, except themselves (***I)*** relationship between V5 and V6. This status is confirmed by the top position of V5 and V6 in the outranking graph and their placement in the same box. At the same time the weakest variants V3 and V4 are characterized by a large number of (***P-***) and (***R***) relationships in the Ranking matrix. These two variants are incomparable one against another (relationship ***R***). In the graphical terms, this corresponds to the bottom position of both variants and their placement in separate boxes, not connected by any arrow.

The application of AHP method algorithm has corresponded to the generation of the following components:

- Consistency indexes ***CI*** for each matrix of relative weights w_r at each level of the hierarchy (criteria, sub-criteria and variants).
- A set of vectors containing normalized, absolute values of weights w_a for criteria, sub-criteria and variants.
- Utility of each variant i – U_i.

Initially, for each matrix of relative weights w_r (see Tables 4, 5 and 6) the consistency indexes ***CI*** have been computed. For all 25 matrices, the values of ***CI-s*** have fallen in the range between 0.0 and 0.12. For those situation where ***CI –s*** exceeded value 0.1 the modeling of preferences has been repeated. After two such iterations the satisfactory effect has been obtained and the values of ***CI-s*** for all considered matrices of relative weights w_r have reached the levels lower than 0.1 (between 0.0 and 0.06). These results have indicated that the appropriate levels of consistency of preferential information have been reached.

In the next step, the so-calle, eigenvalue problem has been solved and the absolute values w_a for criteria, sub-criteria and variants have been computed. The results of these computations for all variants V0, V1, …, V6 evaluated on all criteria C1, C2, …, C11 are presented in Table 7. For each criterion, the values of w_a sum up to 1. They indicate what is the position of each variant in the assessment by each criterion. It is visible that the strongest variants (e.g. V6 and V5) are characterized by high values of w_a while the weakest variants (e.g. V3 and V4) receive low scores of w_a on the majority of criteria. Based on the values of w_a the final utilities U_i of all variants have been calculated. They are presented in the last column of Table 7. The values of U_i define the position of each variant in the final ranking and indicate what is its relative strengths against other variants and share in reaching the final goal of the analysis. As presented in Table 7 the overall utilities U_i of variants V6 and V5 are the highest (overall scores of 0.21 and 0.20, respectively). At the same time, variants V3 and V4 are featured by the lowest values of utilities U_i (overall scores of 0.09 and 0.10, respectively). The utility based assessment characteristic for the AHP method results in the linear ranking of variants (see Fig. 5) and eliminates incomparability between them.

Table 7. Absolute weights w_a (against all considered criteria C1, C2,…, C11) and final utilities U_i of all considered variants V0, V1, …, V6 of the redesign of the urban, public transportation system, generated by the AHP method

Variants	Absolute weights w_a for criteria											Final utilities U_i
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	
V0	0,05	0,06	0,08	0,1	0,04	0,39	0,06	0,06	0,03	0,05	0,4	0,11
V1	0,14	0,2	0,18	0,11	0,12	0,1	0,03	0,16	0,11	0,07	0,05	0,14
V2	0,18	0,2	0,19	0,18	0,25	0,06	0,04	0,22	0,17	0,07	0,03	0,15
V3	0,08	0,08	0,12	0,19	0,07	0,14	0,05	0,05	0,05	0,05	0,2	0,09
V4	0,15	0,06	0,05	0,03	0,02	0,25	0,06	0,07	0,03	0,03	0,26	0,1
V5	0,18	0,2	0,19	0,19	0,25	0,03	0,38	0,22	0,27	0,36	0,03	0,2
V6	0,22	0,2	0,19	0,2	0,25	0,03	0,38	0,22	0,34	0,37	0,03	0,21
Σ	1	1	1	1	1	1	1	1	1	1	1	1

Looking at the results of computational experiments carried out with the application of both methods one may easily discover that they are not identical but similar. The differences between final rankings generated by both methods result from axiomatic differences between Electre III/IV and AHP methods. At the same time, thanks to careful modeling of the DM's preferences, both methods do not generate contradictory results.

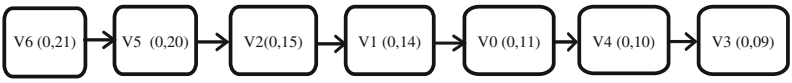


Fig. 5. Final ranking of the redesign variants of the public transportation system generated generated by AHP methods

Both rankings produce the same winning variants, placed at the top of the classifications, including: V5, V6 and V2. These are the strongest and highly recommended solutions. Variant V6 is the most desired option among them, based on both computational experiments. In the ranking generated by AHP method it holds the first position alone, while in the ranking generated by Electre III/IV method it shares it with variant V5. It is important to mention, that the utilities of those two variants (V5 and V6) computed by the AHP algorithm are very close to each other, which confirms the results obtained by Electre III/IV method, suggesting that V5 and V6 are indifferent.

The remaining variants are also placed in similar positions in both rankings. At the next level of classification one can place variants V1 and V0, which are incomparable in the ranking generated by Electre III/IV method, while placed sequentially (V1 as preferred against V0) in the ranking generated by AHP method. They can be considered as moderate, acceptable solutions, having both strong and weak points but not being able to compete with the winners of the rankings. The last group of solutions includes variants V3 and V4. They are classified as the weakest or poorest candidates for urban, public transportation system transformation and are not recommended for implementation. In the ranking generated by Electre III/IV method variants V3 and V4 are incomparable, while in the ranking produced by AHP method V4 is preferred against V3, although the distance between these two variants in terms of their utilities is very short.

Some of the above mentioned differences between the final rankings result from methodological assumptions of the two considered methods. The natural feature of Electre III/IV method is that it applies the concept of incomparability ***(R)*** and thus, classifies variants V1 and V2, V1 and V4, V0 and V3 as well as V3 and V4 as incomparable. AHP method that does not recognize the incomparability between variants and thus, for the same cases connects the variants by the preference ***(P)*** and non-preference ***(P^-)*** relationships.

Based on the generated results the authors of this chapter (acting as analysts) recommend variants V5 and V6 as the best forms of transformation of the urban, public transportation system. They leave the final decision concerning the implementation of a concrete variant to the DM. Both variants are characterized by the similar features and their scores on the majority of criteria are the same. Both are environmentally friendly and sustainable transportation solutions. Although they generate high investment expenditures (above 500 mlnzl each) they also reduce substantially operating costs of the public transportation system (by roughly 50%). V5 receives the highest scores on 12 out of 20 criteria and sub-criteria, while V6 is the leader on 16 out of the 20 considered characteristics. V5 is highly assessed on such criteria as: C2. Reliability, C5. Directness of connections, C7. Operating costs, C10. Environmental friendliness. Thus, it can be considered as an environmentally friendly and passenger – oriented

variant that guarantees high transportation standards. V6 is highly assessed on such measures as: C1. Accessibility, C2. Reliability, C5. Directness of connections, C7. Operating costs, C8. Efficiency, C10. Environmental friendliness. In these circumstances it can be also denominated an environmentally friendly and passenger – oriented variant, with a special emphasis on efficiency and resource utilization. Variant V6 is the most efficient solution.

Two variants at the bottom part of the rankings V3 and V4 are featured by many weak points. Variant V4 presents the worst results on those criteria that are passenger oriented, including: C1. Accessibility, C3. Quality of the means of transport, C4. Waiting time, C5. Directness of connections, C10. Environmental friendliness. It can be concluded that this variant is the least passenger friendly option. Variant V3 is the least effective and efficient transportation solution. It has obtained the lowest scores on the following criteria: C7. Annual operating costs, C8. Efficiency, C9. Social sustainability, C10. Environmental friendliness.

4 Conclusions

The chapter presents the generic methodology that has been developed to support DM-s in redesigning and evaluating a public transportation system. The proposed approach has a universal/generic character and can be denominated by the Procedure of MCDM/A-based redesign of the public, urban transportation system. It can be split into two general stages, including:

1. A heuristic construction of the public transportation system redesign variants/scenarios.
2. Multiple criteria evaluation of these variants.

As proven by this article the redesign of the urban public transportation system should have a heuristics and multiple criteria character. The authors support the heuristic design of the variants, as a method that utilizes human creativity in the design phase and does not impose any strict boundaries on the process of generating alternative options. The authors claim that in the phase of designing/constructing the variants the analysts/designers should use their intuition, common sense and multidisciplinary skills. They should take into consideration various aspects, including: technical, economic, social, safety-oriented and environmental as well as satisfy the interests of the DM and different stakeholders, including passengers, operator, local authorities and residents (local community).

The important component of the proposed approach is a comprehensive, multiple criteria evaluation of the generated variants (phase 2). In this phase, the multiple criteria assessment of variants permits to check how efficient and satisfactory the design process was. It allows verifying whether the variants satisfy critical aspects and interests of important stakeholders. The authors have shown that the evaluation of the redesign variants of the urban public transportation system can be formulated as a multiple criteria ranking problem. They have shown how the ranking process should be performed. A universal, consistent family of criteria, suitable for the evaluation of different redesign variants of any public transportation system, has been defined.

The process of modelling the DM's and stakeholders' preferences has been presented and computational experiments have been shown. As a result, the rankings of the designed variants have been generated. In the computational phase, in order to select the best/compromise solution the authors have applied two MCDM/A ranking methods: Electre III/IV and AHP. 7 considered variants (V0–V6) developed heuristically in the first stage of the proposed approach have been evaluated according to 11 criteria and ranked from the best to the worst.

In the authors' opinion this holistic approach based on two phases guarantees generation of a rational, balanced and sustainable transportation solution, called a compromise variant. The authors strongly recommend in the first phase an intuitive construction of the redesign variants, based on expert knowledge, brain storming and using unconstrained, imaginative thinking. At the same time, they claim that a ridged, systematic verification of the generated concepts is required in the second phase. Thus, they advise in that phase a multiple criteria evaluation of the considered variants.

The authors also indicate that in the evaluation phase different multiple criteria ranking methods can be applied. In their opinion, both AHP and Electre III/IV methods have generated satisfactory computational results. In their opinion both methods can be applied while evaluating the variants of the urban, public transportation transformation (redesign). However, they would like, to point out certain advantages and disadvantages of both methods. These were described in the authors' previous work [27].

Additional methodological advantages of the MCDM/A-based redesign of the urban, public transportation system, proposed in this chapter, are as follows:

- It helps both designers, analysts and DM-s to look at the problem at stake from different perspectives and handle it in a similar manner in both phases (redesign and evaluation).
- The proposed paradigm allows for solving a complex decision situation in which a wide group of stakeholders should be involved and a compromise solution must be found.
- In the assessment phase both quantitative and qualitatitve criteria can be applied; different subjective, verbally axpressed and umbiguously defined measures can be applied.
- In both phases of the proposed approach social, environmenttal, safety and sustainability aspects can be inserted into the the decision model, which is critical for local authorities and EU policies.

From a practical point of view, the results of this chapter can be concluded as follows:

- The authors (acting as analysts) support the radical and investment intensive transformation of the urban, public transportation system characteristic for variants V5 and V6. They claim that these variants, being the winners of the generated rankings, are both sustainable, environmentally friendly and efficient financial – wise.
- The MCDM/A-based redesign of the urban, public transportation system has generated the following benefits:

  - Implementation of variants V5 or V6 results in 50% reduction of annual operating costs; 20-30% increase of the fleet utilization rate; substantial improvement of the environmental standard (zero emission of harmful pollutants) and significant improvement of the quality of life for residents.
  - Variants V5 or V6 are based on high level of travel comfort for passengers. Thanks to their introduction the average waiting time in the public transportation system is reduced by 12,5%; the levels of accessibility and directness of connections are doubled and the reliability of the system is notably improved.
  - This public transportation system proposed in variants V5 and V6 is modern, functional and user friendly for elderly and disabled people.
  - Both winning variants V5 and V6 are consistent with EU objectives related to passenger transportation in a metropolitan area. They are sustainable, competitive and cost- efficient.
  - In addition, variant V6 substantialy facilitates the access to the city centre for residents living in suburbia. The proposed connection by public bicycle rental system to the public transportation network improves mobility of the people who do not have private cars and live outside the city.
- Although the winning variants are investment-intensive, they bring substantial costs reduction, including operating, environmental and social costs. This results in a realtively short period of the investment return (10 years).

Further research can be conducted in the following directions:

- Extension of the redesign concepts towards such solutions as: integration of private/individual transportation with public transportation (Park & Ride System); application of autonomous (driverless vehicles) and introduction of shared means of transport.
- Application of other MCDM/A methods in order to rank various redesign variants of an urban, public transportation system and compare the generated results.
- Application of the combined MCDM/A and Group Decision Making (GDM) methodologies for the selection of the most rational variant of the public transportation system redesign.

References

1. Annema, J., Mouter, N., Razaei, J.: Cost-benefit Analysis (CBA), or multi-criteria decision-making (MCDM) or both: politicians' perspective in transport policy appraisal. Transp. Res. Procedia **10**, 788–797 (2015). 18th Euro Working Group on Transportation, EWGT 2015, 14-16 July 2015, Delft, The Netherlands
2. Belton, V., Stewart, T.: Multiple Criteria Decision Analysis. An Integrated Approach. Kluwer Academic Publishers, Norwell (2003)
3. Benayoun, R., Montgolfier, J., Tergny, J., Laritchev, O.: Linear programming with multiple objective functions: STEP method (STEM). Math. Program. **1**, 366–375 (1971)

4. Benayoun, R., Roy, B., Sussman, B.: ELECTRE: Une méthode pour guider le choix en presence de points de vue multiples. Note de travail 49, SEMA-METRA International, Direction Scientifique (1966)
5. Beria, P., Maltese, I., Mariotti, I.: Multicriteria vs. Cost Benefit Analysis: a comparative perspective in the assessment of sustainable mobility. Eur. Transp. Res. **4**, 137–152 (2012)
6. Bhatta. B.: Analysis of Urban Growth and Sprawl from Remote Sensing Data. 17 Advances in Geographic Information Science. Springer, Berlin (2010). doi:10.1007/978-3-642-05299-6_2
7. Boardman, A., Greenberg, D., Vining, A., Weimer, D.: Cost-Benefit Analysis: Concepts and Practice. Pearson, New York (2011)
8. Brans, J., Vincke, P., Mareschal, B.: How to select and how to rank projects: the PROMETHEE Method. Eur. J. Oper. Res. **24**, 228–238 (1986)
9. Chang, Y.-H., Shyu, T.-H.: A fuzzy multicriteria model to evaluate the privatization of the public bus operations. J. Adv. Transp. **29**(1), 63–79 (1995)
10. Czyżak, P., Żak, J.: A model of an urban transportation system formulated as a multiobjective mathematical programming problem under uncertainty. J. Adv. Transp. **29** (1), 43–62 (1995)
11. Daganzo, C.: Fundamentals of Transportation and Traffic Operations. Pergamon Press, New York (1997)
12. Damart, S., Roy, B.: The uses of cost-benefit analysis in public transportation decision-making in France. Transp. Policy **16**(4), 200–212 (2009)
13. De Brucker, K., Macharis, C., Verbeke, A.: Multi-criteria analysis in transport project evaluation: an institutional approach. Eur. Transp./Trasporti Europei **47**, 3–24 (2011)
14. Ehrgott, M.: Multicriteria Optimization. Springer, Berlin (2005)
15. Ergun, M., Iyinam, S., Iyinam, A.: An assessment of transportation alternatives for Istanbul Metropolitan City for Year 2000. In: Proceedings of the 8th Meeting of the Euro Working Group Transportation, The Rome Jubilee 2000 Conference: Improving Knowledge and, Tools for Transportation and Logistics Development, Rome, September 11–14, 2000, pp. 183–190 (2000)
16. European Commission: A Sustainable Future for Transport: Towards an Integrated, Technology-led and User-Friendly System, Brussels (2009)
17. European Commission: Commission Staff Working Document Accompanying the White Paper - Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system, Brussels, March 28, 2011, SEC, 391 final, 12-14 (2012). http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=SEC:2011:0391:FIN:EN:PDF
18. Figueira, J., Greco, S., Ehrgott, M.: Multiple Criteria Decision Analysis: State Of The Art Surveys. Springer, Boston, Dordrecht, London (2005)
19. Filcek, G., Żak, J.: The multiple criteria optimization problem of joint matching carpoolers and common route planning. In: Świątek, J., Wilimowska, Z., Borzemski, L., Grzech, A. (eds.) Information Systems Architecture and Technology: Proceedings of 37th International Conference on Information Systems Architecture and Technology – ISAT 2016 – Part III. AISC, vol. 523, pp. 225–236. Springer, Cham (2017). doi:10.1007/978-3-319-46589-0_18
20. Geoffrion, A., Dyer, J., Feinberg, A.: An interactive approach for multi-criterion optimalization with an application to the operation of an academic department. Manage. Sci. **19**(4), 357–368 (1972)
21. Gercek, H., Karpak, B., Kilincaslan, T.: A multiple criteria approach for the evaluation of the rail transit networks in Istanbul. In: Presentation of the 8th World Conference on Transport Research, Antwerp, July 12–17, 1998, Working paper (1998)
22. Gromadzki, M.: The development concept of the tram sub-system in the Gorzow mass transit system, consulting report. Public Transport Consulting, Reda-Gorzow (2013)

23. Hillier, F., Lieberman, G.: Introduction to Operations Research. McGraw-Hill, New York (2005)
24. Hsu, T.H.: Public transportation system project evaluation: a fuzzy delphi AHP. In: Presentation of the 8th World Conference on Transport Research, Antwerp, July 12–17, 1998, Working paper (1998)
25. http://www.srodowisko.abc.com.pl/czytaj/-/artykul/elektryczny-autobus-drozsza-cena-tansza-eksploatacja (2016)
26. Jacquet-Lagrèze, E., Siskos, Y.: Assessing a set of additive utility functions for multicriteria decision making: the UTA method. Eur. J. Oper. Res. **10**(2), 151–164 (1982)
27. Janiak, M.K., Żak, J.: Multiple criteria evaluation of different redesign variants of the public tram system. Transp. Res. Procedia **3**, 690–699 (2014). 17th Meeting of the EURO Working Group on Transportation, EWGT 2014, 2-4 July 2014, Sevilla, Spain
28. Jaszkiewicz, A., Słowiński, R.: The 'Light Beam Search' approach - an overview of methodology and applications. Eur. J. Oper. Res. **113**(2), 300–314 (1999)
29. Keeney, R., Raiffa, H.: Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Cambridge University Press, Cambridge (1993)
30. Koksalan, M., Wallenius, J., Zionts, S.: Multiple Criteria Decision Making: From Early History to the 21st Century. World Scientific, Singapore, Hackensack, London (2011)
31. Korhonen, P., Laakso, J.: A visual interactive method for solving the multicriteria problem. Eur. J. Oper. Res. **24**(2), 277–287 (1986)
32. Lee, D.: Methods for evaluation of transportation projects in the USA. Transp. Policy **7**, 41–50 (2000)
33. Lee, R., Rivasplat, C.: Metropolitan transportation planning in the 1990 s: comparisons and contrasts in New Zealand, Chile and California. Transp. Policy **8**, 47–61 (2001)
34. Levine, J., Underwood, S.: A multiattribute analysis of goals for intelligent transportation system planning. Transp. Res. C **4**(2), 97–111 (1996). Elsevier
35. Macharis, C., De Witte, A., Ampe, J.: The multi-actor, multi-criteria analysis methodology (MAMCA) for the evaluation of transport projects: theory and practice. J. Adv. Transp. **43** (2), 183–202 (2009)
36. Masser, I., Sviden, O., Wegener, M.: Transport planning for equity and sustainability. Transp. Plann. Technol. **17**, 319–330 (1993)
37. Novak, D., Sullivan, J., Scott, D.: A network-based approach for evaluating and ranking transportation roadway projects. Appl. Geogr. **34**, 498–506 (2012)
38. Roy, B.: Main sources of inaccurate determination, uncertainty and imprecision in decision models. In: Munier, B., Shakun, M. (eds.) Compromise, Negotiation and Group Decision, pp. 43–62. Reidel Publishing Company, Dordrecht (1988)
39. Roy, B.: Decision-aid and decision making. Eur. J. Oper. Res. **45**, 324–331 (1990)
40. Roy, B.: Multiple Criteria Decision Aiding (in Polish). Wydawnictwo Naukowo-Techniczne, Warsaw (1990)
41. Roy, B.: The outranking approach and the foundations of ELECTRE methods. In: Bana e Costa, C.A. (ed.) Readings in Multiple Criteria Decision Aid, pp. 155–183. Springer, Berlin (1990)
42. Saaty, T.: The Analytic Hierarchy Process. McGraw- Hill, New York (1980)
43. Saaty, T.: Transport planning with multiple criteria: the analytic hierarchy process applications and progress review. J. Adv. Transp. **29**(1), 81–126 (1995)
44. Saaty, T.: Decision making with the analytic hierarchy process. Int. J. Serv. Sci. **1**(1), 83–98 (2008). doi:10.1504/IJSSci.2008.01759
45. Steuer, R.: Multiple Criteria Optimization: Theory, Computation and Application. Wiley, New York (1986)

46. Tabucanon, M., Lee, H.-M.: Multiple criteria evaluation of transportation system improvement projects: the case of Korea. J. Adv. Transp. **29**(1), 127–143 (1995)
47. Tan, K.C., Lee, T.H., Chew, Y.H., Lee, L.H.: A multiobjective evolutionary algorithm for solving vehicle routing problem with time windows. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernatics, Washington D.C., October 5–8, 2003, pp. 361–366 (2003)
48. Tuzkaya, U., Onut, S.: A fuzzy analytic network process based approach to transportation-mode selection between Turkey and Germany: a case study. Inform. Sci. **178**(15), 3133–3146 (2008)
49. Vaidya, O., Kumar, S.: Analytic hierarchy process: an overview of applications. Eur. J. Oper. Res. **169**, 1–29 (2006)
50. Vincke, P.: Multicriteria Decision – Aid. Wiley, New York (1992)
51. Vuchic, V.: Urban Transit Systems and Technology. Wiley, New York (2007)
52. Żak, J.: The multiobjective fuzzy linear fractional model of the mass transit system in poznan. Arch. Transp. **14**(4), 97–123 (2002)
53. Żak, J.: Multiple Criteria Decision Aiding in Road Transportation (in Polish), Monograph No 394. Poznan University of Technology Publisher, Poznan (2005)
54. Żak, J.: The methodology of multiple criteria decision making/aiding in public transportation. J. Adv. Transp. **45**(1), 1–20 (2011)
55. Żak, J.: The methodology of multiple criteria decision making/aiding as a system-oriented analysis for transportation and logistics. In: Świątek, J., Tomczak, J.M. (eds.) ICSS 2016. AISC, vol. 539, pp. 265–284. Springer, Cham (2017). doi:10.1007/978-3-319-48944-5_25
56. Żak, J., Fierek, S: Design and evaluation of alternative solutions for an integrated urban transportation system. In: Selected CD Proceedings of the World Conference on Transport Research, Berkeley, 24–28 June (2007)
57. Żak, J., Jaszkiewicz, A., Redmer, A.: Multiple criteria optimization method for the vehicle assignment problem in a bus transportation company. J. Adv. Transp. **43**(2), 203–243 (2009)
58. Żak, J., Kruszyński, M.: Application of AHP and ELECTRE III/IV methods to multiple level, multiple criteria evaluation of urban transportation projects. Transp. Res. Procedia **10**, 820–830 (2015)
59. Żak, J., Redmer, A., Sawicki, P.: Multiple objective optimization of the fleet sizing problem for road freight transportation. J. Adv. Transp. **45**(4), 321–347 (2011)
60. Żak, J., Thiel, T.: Multiple criteria evaluation of the development scenarios of the mass transit system. In: Park, C.-H., Cho, J., Oh, J., Hayashi, Y., Viegas, J. (eds.) CD-Selected Proceedings of the 9th World Conference on Transport Research, Seoul, July 22–27, 2001 (2001)

A Framework for Solving Real-Time Multi-objective VRP

Oren E. Nahum(✉) and Yuval Hadas

Department of Management, Bar-Ilan University, Max ve-Anna Webb Street, 5290002 Ramat Gan, Israel
{oren.nahum,yuval.hadas}@biu.ac.il

Abstract. One of the most important logistics problems in the field of transportation and distribution is the Vehicle Routing Problem (VRP). In general, VRP is concerned with the determination of a minimum-cost set of routes for distribution and pickup of goods for a fleet of vehicles, while satisfying given constraints. Today, most VRPs are set up with a single objective function, minimizing costs, ignoring the fact that most problems encountered in logistics are multi-objective in nature (maximizing customers' satisfaction and so on), and that for both deterministic and stochastic VRPs, the solution is based on a pre-determined set of routes. Technological advancements make it possible to operate vehicles using real-time information. Since VRP is a NP-Hard problem, it cannot be solved to optimality using conventional methods; therefore, the paper presents a heuristic framework for solving the problem. In real-time dynamic problems, a solution is given based on known data, as time progresses, new data are added to the problem, and the initial solution has to be re-evaluated in order to suit the new data. This is usually done at pre-defined time intervals. If the time intervals are small enough, thus, at each time interval the amount of information added is limited. Therefore, the new solution will be similar to the previous one. Due to the fact that the result is a solution set, not a single solution, and one solution is to be selected within a short time window, it is necessary to automatically select a single solution. For that, a framework, based on traditional and evolutionary multi-objective optimization algorithms, which incorporate multi-criteria decision making methods, for solving real-time multi-objective vehicle routing problems is presented.

Keywords: Vehicle routing problems · Multi-objective · Multi-criteria decision making · Real-time

1 Introduction

One of the most important logistics problems in the field of transportation and distribution is the Vehicle Routing Problem (VRP) [6, 61, 62]. In general, VRP is concerned with the determination of a minimum-cost set of routes, usually the shortest ones, for distribution and pickup of goods for a fleet of vehicles, while satisfying given constraints. Since the problem was first introduced by Dantzig and Ramser [7] several extensions to the problem, with different types of "cost" and constraints were developed.

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_5

As of today, most VRPs are set up with the single objective of minimizing the cost of the solution, despite the fact that the majority of the problems encountered in industry, particularly in logistics, are multi-objective in nature. In fact, numerous aspects, such as balancing workloads (time, distance, etc.), can be taken into account simply by adding new objectives [27].

Moreover, traditionally, vehicle routing plans are based on deterministic information about demands, vehicle locations and travel times on the roads. Advancement of the technology in communication systems, the geographic information system (GIS) and the intelligent transportation system (ITS), make it possible to operate vehicles using the real-time information about travel times and the vehicles' locations [17]. What is likely to distinguish most VRPs today from equivalent problems in the past, is that information needed to come up with a set of good vehicle routes and schedules is dynamically revealed to the decision maker [48].

While traditional VRPs have been thoroughly studied, limited research has to date been devoted to multi-objective, real-time management of vehicles during the actual execution of the distribution schedule, in order to respond to unforeseen events that often occur and may deteriorate the effectiveness of the predefined and static routing decisions. Furthermore, in cases when traveling time is a crucial factor, ignoring travel time fluctuations (due to various factors, such as peak hour traveling time, accidents, weather conditions, etc.) can result in route plans that can direct the vehicles into congested urban traffic conditions. Considering time-dependent travel times as well as information regarding demands that arise in real time in solving VRPs can reduce the costs of ignoring the changing environment [21].

One point that was neglected, which its importance intensifies in multi-objective, real-time VRP, is the need for a quick and automated selection of a single solution from the non-dominated solution's set. For that, a framework that combines multi-objective VRP together with multi-criteria decision making (MCDM), is presented and assessed.

The rest of the paper is as follows. Section 2 provides a review on both multi-objective and dynamic VRPs as well as multi-criteria decision making methods. Section 3 describes a framework for solving real-time multi-objective VRPs. Section 4 describes results obtained from case study, and, finally, Sect. 5 concludes the paper.

2 Literature Review

The framework presented in this paper is for solving real-time (or dynamic), multi-objective VRPs. As shell be seen later in the paper, the framework incorporates multi-criteria decision methods while solving the problem. Therefore, the three topics, multi-objective VRP, dynamic VRP and multi- criteria decision methods are reviewed next in this chapter.

2.1 Multi-objective VRP

Most VRPs, frequently used to model real cases, are set up with a single objective (minimizing the cost of the solution), although the majority of the problems encountered in industry, particularly in logistics, are multi-objective in nature. According to Jozefowiez et al. [27], multi-objective VRPs are used mainly in three ways:

1. **Extending classic academic problems** – In this case, the problem definition remains unchanged, and new objectives are added. As an example of such an objective, we can consider the following: (1) *Driver workload* – an extension to VRP in which the balance of tour lengths is considered (to increase the fairness of the solution) [43, 60]. (2) *Customer Satisfaction* – an objective added to VRP with time windows [10] in order to improve customer satisfaction with regard to delivery dates [1, 16, 68].
2. **Generalizing Classic Problems** – Another way to use multi-objective optimization is to generalize a problem by adding objectives instead of one or several constraints and/or parameters [58].
3. **Studying real-life cases** - Multi-objective routing problems are also studied for a specific real-life situation, in which decision makers define several clear objectives that they would like to see optimized [2, 11, 19, 31, 66].

The different objectives studied in the literature can be presented and classified according to the component of the problem with which they are associated [27]. The following is a summary of the most common objectives.

1. **Objectives related to the tour**: (a) **Cost**: Minimizing the cost of the solutions generated is the most common objective, usually for economic reasons; however, other motivations are possible. For instance, in [45, 46], it is done to avoid damaging the product being transported. (b) **Makespan**: Minimizing the makespan ensures some fairness in solutions [30]. (c) **Balance**: Some objectives are designed to even out disparities between the tours [37].
2. **Objectives related to node/arc activity**: Most of the studies dealing with objectives related to node/arc activity involve time windows. Time windows are usually replaced by an objective that minimizes the number of violated constraints [4], the total customer and/or driver's wait time due to earliness or lateness [3, 9, 22], or both [13, 14].
3. **Objectives related to resources**: A common objective is the minimization of the number of vehicles, as in VRP with time windows (usually treated lexicographically) [41]. Goods-related objectives are used to take the nature of the goods into account (merchandise is perishable and we want to avoid its deterioration [45, 46]).

Over the last several years, many techniques have been proposed for solving multi-objective problems. These strategies can be divided into three general categories:

1. **Scalar methods** - The most popular is weighted linear aggregation. For multi-objective VRPs, weighted linear aggregation has been used with specific heuristics [18], local search algorithms [64], and genetic algorithms [65].
2. **Pareto methods -** Pareto methods use the notion of Pareto dominance directly. Pareto methods are used with evolutionary algorithms, local searches, heuristics, and/or exact methods [64, 65].
3. **Methods that belong to neither the first nor the second category** - These non-scalar and non-Pareto methods are based on genetic algorithms, lexicographic strategies, ant colony mechanisms, or specific heuristics [29].

2.2 Dynamic VRP

In many real-life applications relevant data changes during the execution of transportation processes and schedules have to be updated dynamically. Thanks to recent advances in information and communication technologies, vehicle fleets can now be managed in real-time. In this context, Dynamic or real-time VRPs (DVRPs), are becoming increasingly important [49].

The most common source of dynamism in VRP is the online arrival of customer requests during the operation [49].

Travel time variations have been studied by Haghani and Jung [21]; Potvin et al. [47]; Fleischmann et al. [12]; Hu [23]; Hu et al. [24]; Ichoua et al. [26]. Malandraki and Daskin [35], who used a step function for that purpose, while Gendreau et al. [15]; Liao [34] proposed tabu search algorithms for solving the problem.

Finally, some more recent work considers dynamically revealed demands for a set of known customers [44, 55, 56] and vehicle availability [32, 33, 39], in which case the source of dynamism is the possible breakdown of vehicles.

2.3 Multi-criteria Decision Making

In most cases, when solving a multi-objective optimization problem, the result is a set of non-dominated solution (a set in which there is no solution that is better in all objectives from another solution in the set), from which the decision maker (DM) has to choose his preferred alternative.

Multi-criteria decision-making (MCDM) methods, some of which are listed below, are automated methods for selecting a preferred solution given a set of feasible solutions, while having conflicting criteria [8, 67]. MCDM methods also allow assigning the various solutions to different pre-defined classes and ordering them from best to worst [63].

The Max-Min method, for example, can be used when the DM wants to maximize the achievement in the weakest criterion. On the other hand, the Min-Max method can be used when the DM wants to minimize the maximum opportunity loss. Compromise Programming identifies the solution whose distance from the ideal solution (an artificial

solution consists of the upper bound, for maximization, of the criteria set) is minimum. The ELECTRE Method [50] compares two alternatives at a time and attempts to eliminate alternatives that are dominated using the outranking relationship. In the first version of this method, the result is a set of alternatives (called the kernel) that can be presented to the DM for selecting the preferred solution. The second version of this method provides a complete rank ordering of the original set of alternatives. The TOPSIS method [25] assumes that the preferred solution should simultaneously be closest to the ideal solution and farthest from the negative-ideal solution (an artificial solution consists of the lower bound, for maximization, of the criteria set). For every solution, TOPSIS calculates an index that combines both its closeness to the positive-ideal solution and its remoteness from the negative-ideal solution. The alternative that maximizes this index value is the preferred alternative. Multi-attribute utility theory (MAUT) [28] is based upon the assumption that every DM tries to optimize a utility function (not necessarily known at the beginning of the decision process). The global utility of an alternative is assessed using a utility function, composed of various criteria. Each criterion is assigned with a marginal utility score by the DM, which in a second phase, is aggregated to the global utility score. Each alternative is evaluated and ranked using the utility function.

Many MCDM methods require the use of relative importance weights of criteria, which are usually proportional to the relative value of unit changes in criteria value functions. A simple and common method for ranking criteria is the "weights from ranks" method. In this method, the DM ranks each criterion, r_i, in order of increasing relative importance (highest ranked criterion is rank as 1.) Next each the weight of criteria is defined as $\lambda_i = (k + r_i + 1) \Big/ \left(\sum_{j=i}^{k} k + r_j + 1 \right)$, when k is the number of criteria. While this method produces an ordinal scale, it not guarantee the correct type of criterion importance because ranking does not capture the strength of preference information [36].

When a large number of criteria are considered, it may be easier for the DM to provide pairwise ranking instead of complete ranking. As an example of such a method, consider the analytic hierarchy process (AHP) proposed by Saaty [51, 53]. With AHP, the decision problem is first structured as hierarchal levels. At the top level is the goal of the problem while subsequent levels represent criteria, sub-criteria, and so on with the last level representing the decision alternatives. Next, value judgments concerning the alternatives with respect to the next higher level sub-criteria are calculated based on available measurements. If measurements are not available, the calculation is made from pairwise comparison. After the value judgments of alternatives have been computed, composite values are determined by finding the weighted average values across all levels of the hierarchy. The analytic network process (ANP), a generalization of the AHP method that deals with dependencies, is another example of MCDM methodology [52]. ANP allows for more complex interrelationships among the decision levels and attributes than AHP. Two-way arrows represent interdependencies among attributes and attribute levels. The directions of the arrows signify dependence.

Arrows emanate from an attribute to other attributes that may influence it. The relative importance or strength of the impacts on a given element is measured on a ratio scale similar to AHP (using pairwise comparisons and judgment). A priority vector may be determined by directly asking the DM for a numerical weight but there may be less consistency, since part of the process of decomposing the hierarchy is to provide better definitions of higher level attributes. The ANP approach is capable of handling inter-dependence among elements by obtaining the composite weights through the development of a "supermatrix".

3 Posteriori Decision Making Framework for Solving Real-Time Multi-objective Vehicle Routing Problems

In a posteriori framework, a multi-objective algorithm is executed, followed by a decision making algorithm that automatically selects a preferred solution from the solution set.

For a given multi-objective VRP, let's assume that there exists an algorithm for solving the problem. It is then possible to solve the real-time version of this VRP, simple by re-solving the problem as soon as new information is available, using the existing algorithm. However, working in such a way is time consuming, and cannot guarantee that a proper solution will exist when needed. Since in most cases, new information may causes relatively small changes (if it is processed soon enough), it will be ideal if we can update the current solution, so it will reflect the new information, and at the same time, provide an optimal or near optimal solution. Population based algorithms and evolutionary algorithms are well suitable for this task, as the new information can be inserted into the current population, which will be considered as the initial conditions for the new updated problem.

Evolutionary Algorithms belong to the evolutionary computation field of study concerned with computational methods inspired by the process and mechanisms of biological evolution. Evolutionary Algorithms are concerned with investigating computational systems that resemble simplified versions of the processes and mechanisms of evolution, toward achieving the effects of these processes and mechanisms, namely the development of adaptive systems.

This section describes a simple framework, based on evolutionary or population based algorithms, for solving real-time multi-objective VRP. The framework is described using genetic algorithms [38, 57], that are a family of computational models inspired by evolution. These algorithms encode a potential solution to a specific problem on a simple chromosome-like data structure and apply recombination operators to these structures in order to preserve critical information.

An implementation of a genetic algorithm begins with a population of (typically random) chromosomes (solutions). One then evaluates these structures and allocated reproductive opportunities in such a way that these chromosomes which represent a better solution to the target problem are given more chances to 'reproduce' than those chromosomes which are poorer solutions. The 'goodness' of a solution is typically defined with respect to the current population.

Genetic Algorithm processes a number of solutions simultaneously. Hence, in the first step a population having P individuals is generated. Next, individual members of the population are evaluated to find the objective function value, which is mapped into a fitness function that computes a fitness value for each member of the population. Three main operators, reproduction, crossover and mutation, are used to create a new population. The purpose of these operators is to create new solutions by selection, combination or alteration of the current solutions that have shown to be good temporary solutions. The new population is further evaluated and tested until termination. Since the problem for which the framework is designed is a multi-objective problem, it is necessary to implement a multi-objective algorithm, such as VEGA [54] and SPEA2 [69], as the base of the framework.

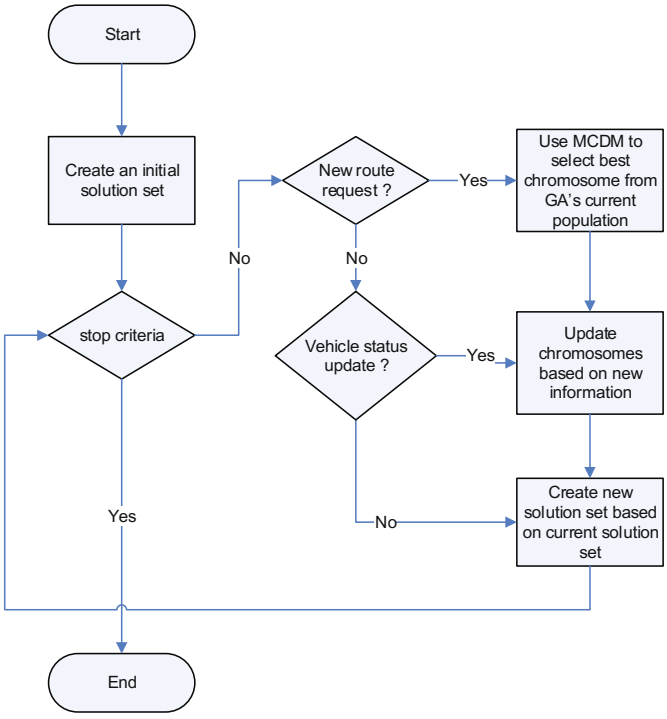


Fig. 1. A'posteriori decision making framework

Figure 1 illustrates the a'posteriori DM framework for solving a real-time multi-objective VRP.

As with any genetic algorithms, the first step of the framework is to generate an initial population (a set of chromosomes or solutions). This population can be created randomly, or using some king of a heuristic. Next, each of the population's

chromosomes is evaluated and given a fitness value. The fitness value calculation, is based on the multi-objective genetic algorithm implementation. Next an iterative section is executed until a stopping condition is met. In this framework, the stopping condition is that there are no customers that have to be served and all vehicles are at the depot. The first step of the iterative section is the request route operation. The request route operation selects a single chromosome from the current population, and assigns it routes to the vehicles. The request route operation is executed on two conditions: (1) Current time is defined as departure time. It is then possible that new vehicles have to leave the depot. The request routes operation is used to determine whether new vehicles have to leave the depot and their destinations. (2) A vehicle is at a customer, and the customer has been fully served. The vehicle needs to start its way towards the next customer. A request route operation is performed in order to determine the vehicle's destination, as it can be changed based on current information. Next, all chromosomes are updated based on new information, if any (new routes, customers' demands, travel times, etc.). If a chromosome has been changed, then its fitness must be re-calculated. Next, a new population is generated based on the current population, using crossover and mutation operations. The fitness of each new chromosome in the new population is also calculated.

The request route operation, as mentioned before, selects a single chromosome (or solution) from the current population (set of solution). In most cases, when solving a multi-objective optimization problem, the result is a set of non-dominated solution (a set in which there is no solution that is better in all objectives from another solution in the set), from which the decision maker (DM) has to choose his preferred alternative. This set of non-dominated solution can be obtained using various multi-objective optimization algorithms.

In an automated environment, however, a mechanism for choosing a preferred solution from a set of non-dominated solutions needs to be implemented. In this case, the request route operation uses a MCDM method, such as the ones described in the literature review, in order to rank the solution. Then, based on this ranking, a preferred solution is selected.

4 A Case Study Example

To demonstrate the usage of the posteriori decision making framework, a real-time multi-objective vehicle routing problem has been designed. The full details of the problem, including the description of the various objectives and constraints, the mathematical model and the various multi-objective evolutionary algorithms used for solving it, are outside the scope of this paper, and can be found in [40, 41]. Generally, the problem is defined as a vehicle fleet that has to serve customers of fixed demands from a central depot. Customers must be assigned to vehicles, and the vehicles routed so that a number of objectives are minimized/maximized.

Based on a vast literature review, five objectives were selected: (1) Minimizing the total traveling time [35]; (2) Minimizing the number of vehicles [5]; (3) Maximizing

customers' satisfaction [1, 16, 68]; (4) Maximizing drivers' satisfaction [43, 60] and (5) Minimizing the arrival time of the last vehicle.

As a real-time problem, vehicles' routes are adjusted at certain times in a planning period. This adjustment considers new information about the travel times (which is stochastics and depends on the distance between two points and the time of day), current location of vehicles, new demand requests (that can be deleted after being served, or added since they arise after the initial service began) and more. This result in a dynamic change in the demand and traveling time information as time changes, which has to be taken into consideration in order to provide optimized real-time operation of vehicles.

Having several assumptions and limitations, such as a system with dynamic conditions (real-time variation in travel times and real-time service requests); all demands have specified service times and service time intervals; soft time windows for service around the desired service times are considered, and more, Nahum et al. [41] formulated the problem as a mixed integer linear programming problem on a network. However, since the problem is a NP-Hard problem, it cannot be solved to optimality using conventional methods, and therefore, the posteriori decision making framework was used.

The simplicity and generality of the framework makes it possible to use any population based algorithm (as long as it can solve the problem as a static problem) for solving the problem as a real-time problem. For that reason, three evolution algorithms have been developed for solving the problem. The first algorithm is an improved version of the vector evaluated genetic algorithm (VEGA) (which incorporates elitism). The second algorithm is an implementation of the SPEA2 algorithm. And, the third evolutionary algorithm is a combination of the vector evaluated technique and artificial bee colony algorithm. The algorithms were incorporated into the posteriori decision making framework, while the multi-criteria decision making method used by the framework is the TOPSIS method.

The results of the three algorithms were compared using a case study. The case study is based on two transportations networks, each based on real-file information, each with different characteristics. The first network is based on metropolitan Tel-Aviv's urban road network. In this network, there are 45 customers (their locations are based on the stores' locations of "Mega Ba'ir" – a large super-market chain store in Israel) (not including the depot). The second network is based on Israel's interurban road system. In this network, there are 34 customers (not including the depot). The 34 customers are located in major Israeli cities. For both networks, "Google Maps", was used (1) to determine the shortest distance (based on actual network) between every two customers, and (2) to collect traveling time (at different times of the day) for each edge in the network. The traveling time information was later used in order to calculate a log-normal travel-time distribution function for each path [20]. Each customer was also associated with a time window. The time windows were randomly generated according to the following assumptions: (1) The minimum possible time window start time, PSTW, is equal to 8:00 am plus the time it takes to get from the depot to the customer (when leaving the depot at 8:00 am). It is assumed that the distance from the depot to the customer is known and the travel speed is 15 km per hour for the first network and 70 km per hour for the second. (2) The time window start time, STW, is

based on possible time window start time and is a random value within the range of PSTW to PSTW + 1.5 (plus one and a half hour). (3) The time window end time, ETW, is based on the time window start time and is a random value within the range of STW + 0.5 to STW + 3. Each customer is also associated with a randomly generated demand, in the range of 10 to 50, similar to the demands used in Solomon's instances [59].

In each test problem, half the customers are considered as customers with unknown demands. These are the customers with the latest time window start time. Each unknown demand is revealed to the simulation at least two hours prior to the beginning of the time window.

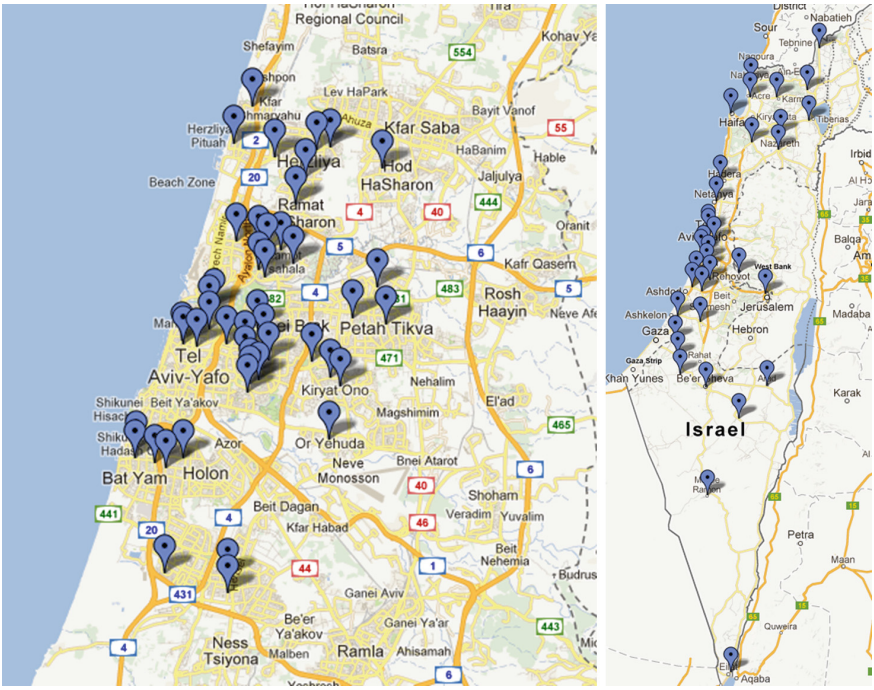


Fig. 2. Urban network (left), having 45 customers in the greater Tel-Aviv metropolitan area, and interurban network (right), having 39 customers in major cities in Israel.

In order to perform the case study, simulation was used. The simulation is based on two processes running in parallel, the algorithm process and the simulation process, which exchange information between each other.

In simulating a full-day of operation, several assumptions are made: (1) The planning period (the time that the algorithm runs before the first vehicle has to leave the depot) starts at 7:00 am and ends at 8:00 am. (2) Service starts at 8:00 am, when the first vehicle leaves the depot, and ends when the last vehicle returns to the depot. (3) During the planning period, new information about customer demands is not

acceptable. (4) The workday is divided into 24 time intervals, each one hour long, starting at 0:00 am. (5) For each edge in the transportation network, the travel time is given using log-normal distribution functions for each time interval. (6) Information about real travel times is known two hours in advance (i.e., for the next two time intervals), and is updated 15 min before the beginning of the hour. (7) Every half an hour on the hour, new vehicles that have to leave the depot, leave the depot on their way to their first customers (this can happen due to new customer demands or due to route splitting). (8) New customer demands are acceptable only if there is at least one vehicle who has not completed its route. (9) If all vehicles are either at the depot or driving to the depot, the algorithm stops working (end of the case study). (10) The capacity of a single vehicle is equal to 200 units, as in Solomon's instances.

The simulation process simulates an entire work today. It does so by handling each of the vehicles, collecting data about travel times and new customers' demands.

The three algorithms were compared using a number of case studies, based on the above mentioned real-world transportation networks (urban and interurban), with two different approaches for prioritizing customers' requests (equal or demand size bases priority) and two different customer satisfaction functions.

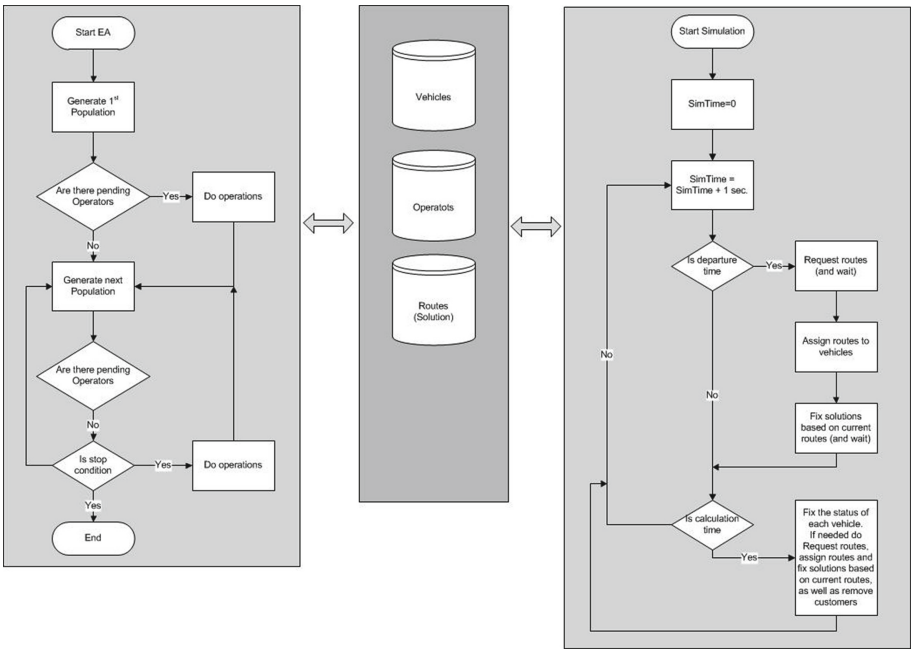


Fig. 3. The relationship between the algorithm process and the simulation process

As an example, during the execution of the SPEA2 algorithm, for the urban network when both travel times and customers' demands are unknown (desired real-world situation) (see [40]), there has been 56 times that a set of routes had to be selected from

the current set of solutions provided by the evolutionary algorithm (see Table 1). That means that in average, every four minutes and fifty seconds, the decision maker has to choose from about thirty-four non-dominated solutions. In such a case, the decision maker spends most of his time choosing preferred routes. By incorporating a multi-criteria decision making model into the multi-objective algorithm, the decision maker does not have to carry out the task of selecting preferred routes, which is now performed automatically by the algorithm.

Table 1. Results

Time	Size of pareto front	Routes				
		Total	Unchanged	Changed	New	Removed
07:45:00	44	7	–	–	7	–
08:00:00	54	7	3	4	0	0
08:30:00	84	7	7	0	0	0
08:55:00	89	8	5	2	1	0
08:56:19	92	9	5	3	1	0
09:00:00	42	10	5	4	1	0
09:20:05	87	12	5	5	2	0
09:26:44	10	13	7	5	1	0
09:28:15	35	14	12	1	1	0
09:30:00	19	14	14	0	0	0
09:34:43	72	16	12	2	2	0
09:37:45	79	17	11	5	1	0
09:51:52	55	18	7	10	1	0
09:52:19	37	19	16	2	1	0
09:59:55	57	21	14	5	2	0
10:00:00	11	21	21	0	0	0
10:01:58	60	21	21	0	0	0
10:03:39	87	22	20	1	1	0
10:03:51	16	22	20	1	1	1
10:04:06	17	22	22	0	0	0
10:05:49	15	22	22	0	0	0
10:07:47	20	22	21	1	0	0
10:20:41	64	21	18	3	0	1
10:28:35	72	21	18	3	0	0
10:30:00	43	20	17	3	0	1
10:32:13	74	20	17	3	0	0
10:34:44	48	19	19	0	0	1
10:35:14	78	19	19	0	0	0
10:35:39	45	18	18	0	0	1
10:37:35	52	18	18	0	0	0
10:39:48	48	17	17	0	0	1

(*continued*)

Table 1. (*continued*)

Time	Size of pareto front	Routes				
		Total	Unchanged	Changed	New	Removed
10:40:02	6	18	16	1	1	0
10:47:15	10	17	17	0	0	1
10:48:43	7	16	16	0	0	1
10:48:58	6	15	15	0	0	1
10:49:45	8	15	15	0	0	0
10:54:18	7	14	14	0	0	1
10:54:34	4	13	13	0	0	1
10:55:37	6	12	12	0	0	1
11:00:00	5	11	11	0	0	1
11:01:13	5	11	11	0	0	0
11:03:00	7	10	10	0	0	1
11:05:30	8	10	10	0	0	0
11:07:15	6	9	9	0	0	1
11:13:05	10	8	8	0	0	1
11:16:10	7	7	7	0	0	1
11:22:40	5	7	7	0	0	0
11:24:15	7	6	6	0	0	1
11:30:00	9	5	5	0	0	1
11:42:26	120	5	5	0	0	0
11:45:37	24	4	4	0	0	1
11:57:37	6	3	3	0	0	1
11:57:47	2	2	2	0	0	1
12:00:00	1	1	1	0	0	1
12:06:05	1	1	0	0	0	0

Table 2 is an example of some of the results obtained using the framework, and it shows the results for the fifth strategy (a situation in which both travel times and customers' demands are unknown, a desired real-world situation) for the urban network, using the three evolutionary algorithms with two different customer satisfaction functions (in the first, all customers are equally important, and in the second, the importance of a customer is relative to its demand). For each objective function, the best value is colored in red. From the results, it is clear the best values are either obtained using the Improved VEGA algorithm or the SPEA2 algorithm.

As mentioned, three algorithms were compared using a number of case studies. An analysis of the results obtained using the various algorithms shows that in 70% of the cases, best solutions were obtained using the improved VEGA algorithm. In 75% of the cases, best solutions were obtained using the SPEA2 algorithm. As for the VE-ABC algorithm, it provided the best solutions in 70% of the cases[1]. Furthermore, when

[1] A (best) solution obtained using more than one algorithm, is counted separately for each algorithm.

Table 2. Comparison of the 5th strategy used in all three algorithms

Customer's priority	Objective function	Algorithm		
		Imp. VEGA	SPEA2	VE-ABC
Equal	1	92.834	**82.654**	96.584
	2	**17.021**	18.293	19.226
	3	0.929	**0.454**	2.907
	4	**0.95**	35.596	46.852
	5	21.461	**19.375**	20.881
Demand based	1	**93.123**	98.894	100.742
	2	**18.567**	19.367	18.718
	3	5.356	**0.659**	0.785
	4	23.922	**0.152**	0.463
	5	**20.654**	20.656	20.831

comparing the results of two strategies: non-prioritized customers versus prioritized customers (based on the demand), interesting results are obtained. All the algorithms provide better results for the prioritized strategy on 60% of the cases, with the remaining cases equal for both strategies.

The results show that all three algorithms provide better solutions when each customer is assigned a different priority. The results also show that when the VEGA algorithm is used, it can provide solutions equal in quality to those obtained from more sophisticated and more recent algorithms. This is important, since the VEGA algorithm offers several advantages: its simplicity of implementation; its running speed compared with other algorithms (and as a result, the number of iterations in a given time period); and its capacity for modifications.

5 Summary

This paper presents a framework for solving real-time multi-objective VRP. The framework is based on evolutionary algorithms, which are well suitable for solving this kind of problems, since the previous solution can be considered as an initial solution for the updated problem, while there is no need to start the calculation of the new routes from the beginning. When a driver has to drive to a new customer, the current solution of the algorithm is used in order to define the driver's new destination. Since the result of the algorithm is a set of non-dominated solutions, as in the case of multi-objective, a multi-decision method is used for automatically choose the preferred alternative.

The advantage of the framework is illustrated using a case study. An example based on the case study, shows how frequent a route has to be chosen, and that the number of non-dominated solution from which the route has to be chosen is relatively high. In such a case, the decision maker spends most of his time choosing preferred routes. Incorporating a multi-criteria decision making model into a multi-objective algorithm, automates the process of selecting a preferred route, such that the decision maker handle other tasks.

Using a case study, which was solved using various evolutionary algorithm incorporated into the framework it was shown, based on real information, that all algorithms provide better solutions when each customer is assigned a different priority. Moreover, for the case study it was found that all algorithms provide relatively the same results. This means, that for real world conditions, we can use relatively simple algorithms and still get results similar to state of the art algorithms.

References

1. Afshar-Bakeshloo, M., Mehrabi, A., Safari, H., Maleki, M., Jolai, F.: A green vehicle routing problem with customer satisfaction criteria. J. Ind. Eng. Int. **12**(4), 529–544 (2016)
2. Anbuudayasankar, S.P., Ganesh, K., Lenny Koh, S.C., Ducq, Y.: Modified savings heuristics and genetic algorithm for bi-objective vehicle routing problem with forced backhauls. Expert Syst. Appl. **39**(3), 2296–2305 (2012)
3. Baran, B., Schaerer, M.: A multiobjective ant colony system for vehicle routing problem with time windows. In: Paper presented at the 21st IASTED International Conference on Applied Informatics, Innsbruck, Austria, 10–13 February 2003
4. Barkaoui, M., Berger, J., Boukhtouta, A.: Customer satisfaction in dynamic vehicle routing problem with time windows. Appl. Soft Comput. **35**, 423–432 (2015)
5. Corberan, A., Fernandez, E., Laguna, M., Marti, R.: Heuristic solutions to the problem of routing school buses with multiple objectives. J. Oper. Res. Soc. **53**(4), 427–435 (2002)
6. Cordeau, J.-F., Laporte, G., Savelsbergh, M.W., Vigo, D.: Vehicle routing. In: Handbooks in Operations Research and Management Science, vol. 14, pp. 367–428 (2007)
7. Dantzig, G.B., Ramser, J.H.: The truck dispatching problem. Manage. Sci. **6**(1), 80–91 (1959)
8. Ehrgott, M.: Multicriteria Optimization. Springer, Heidelberg (2005)
9. El-Sherbeny, N.: Resolution of a vehicle routing problem with multi-objective simulated annealing method. Ph.D. Dissertation. Faculte Polytechnique de Mons (2001)
10. El-Sherbeny, N.A.: Vehicle routing with time windows: An overview of exact, heuristic and metaheuristic methods. J. King Saud Univ. Sci. **22**(3), 123–131 (2010)
11. Faccio, M., Persona, A., Zanin, G.: Waste collection multi objective model with real time traceability data. Waste Manag. **31**(12), 2391–2405 (2011)
12. Fleischmann, B., Gnutzmann, S., Sandvoß, E.: Dynamic vehicle routing based on online traffic information. Transp. Sci. **38**(4), 420–433 (2004)
13. Geiger, M.J.: Genetic algorithms for multiple objective vehicle routing. In: Paper Presented at the MIC 2001-4th Metaheuristics International Conference, Porto, Portugal (2001)
14. Geiger, M.J.: Genetic algorithms for multiple objective vehicle routing (2008). Arxiv preprint arXiv:0809.0416
15. Gendreau, M., Guertin, F., Potvin, J.-Y., Taillard, É.: Parallel tabu search for real-time vehicle routing and dispatching. Transp. Sci. **33**(4), 381–390 (1999)
16. Ghannadpour, S.F., Noori, S., Tavakkoli-Moghaddam, R., Ghoseiri, K.: A multi-objective dynamic vehicle routing problem with fuzzy time windows: Model, solution and application. Appl. Soft Comput. **14**, 504–527 (2014)
17. Ghiani, G., Guerriero, F., Laporte, G., Musmanno, R.: Real-time vehicle routing: Solution concepts, algorithms and parallel computing strategies. Eur. J. Oper. Res. **151**(1), 1–11 (2003)

18. Gong, Y.-J., Zhang, J., Liu, O., Huang, R.-Z., Chung, H.S.-H., Shi, Y.-H.: Optimizing the vehicle routing problem with time windows: A discrete particle swarm optimization approach. IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.) **42**(2), 254–267 (2012)
19. Gupta, R., Singh, B., Pandey, D.: Multi-objective fuzzy vehicle routing problem: A case study. Int. J. Contemp. Math. Sciences **5**(29), 1439–1454 (2010)
20. Hadas, Y., Ceder, A.: Improving bus passenger transfers on road segments through online operational tactics. Transp. Res. Rec. **2072**(2072), 101–109 (2008)
21. Haghani, A., Jung, S.: A dynamic vehicle routing problem with time-dependent travel times. Comput. Oper. Res. **32**(11), 2959–2986 (2005)
22. Hong, S.-C., Park, Y.-B.: A heuristic for bi-objective vehicle routing with time window constraints. Int. J. Prod. Econ. **62**(3), 249–258 (1999)
23. Hu, T.Y.: Evaluation framework for dynamic vehicle routing strategies under real-time information. Artif. Intell. Intell. Transp. Syst. **1774**(1774), 115–122 (2001)
24. Hu, T.Y., Liao, T.Y., Lu, Y.C.: Study of solution approach for dynamic vehicle routing problems with real-time information. Transp. Netw. Model. **1857**, 102–108 (2003)
25. Hwang, C.L., Yoon, K.: Multiple Attribute Decision Making: Methods and Applications: A State-of-the-Art Survey, vol. 13. Springer, New York (1981)
26. Ichoua, S., Gendreau, M., Potvin, J.Y.: Vehicle routing with time-dependent travel times. Eur. J. Oper. Res. **144**, 379–396 (2003)
27. Jozefowiez, N., Semet, F., Talbi, E.-G.: Multi-objective vehicle routing problems. Eur. J. Oper. Res. **189**(2), 293–309 (2008)
28. Keeney, R., Raiffa, H.: Decisions with multiple objectives: preferences and value tradeoffs. Interfaces **7**(4), 115–117 (1977)
29. Kovacs, A.A., Parragh, S.N., Hartl, R.F.: The multi-objective generalized consistent vehicle routing problem. Eur. J. Oper. Res. **247**(2), 441–458 (2015)
30. Lacomme, P., Prins, C., Prodhon, C., Ren, L.: A multi-start split based path relinking (MSSPR) approach for the vehicle routing problem with route balancing. Eng. Appl. Artif. Intell. **38**, 237–251 (2015)
31. Lacomme, P., Prins, C., Sevaux, M.: A genetic algorithm for a bi-objective capacitated arc routing problem. Comput. Oper. Res. **33**(12), 3473–3493 (2006)
32. Li, J.Q., Mirchandani, P.B., Borenstein, D.: A Lagrangian heuristic for the real-time vehicle rescheduling problem. Transp. Res. Part E: Logistics Transportation Rev. **45**(3), 419–433 (2009)
33. Li, J.Q., Mirchandani, P.B., Borenstein, D.: Real-time vehicle rerouting problems with time windows. Eur. J. Oper. Res. **194**(3), 711–727 (2009)
34. Liao, T.Y.: A tabu search algorithm for dynamic vehicle routing problems under real-time information. J. Transp. Res. Board **1882**, 140–149 (2004)
35. Malandraki, C., Daskin, M.S.: Time dependent vehicle routing problems: Formulations, properties and heuristic algorithms. Transp. Sci. **26**(3), 185–200 (1992)
36. Masud, A.S.M., Ravindran, A.R.: Multiple criteria decision making. In: Ravindran, A.R. (ed.) Operations Research and Management Science. Handbook: Taylor & Francis Group, LLC (2008)
37. Matl, P., Hartl, R.F., Vidal, T.: Equity Objectives in Vehicle Routing: A Survey and Analysis (2016). arXiv preprint arXiv:1605.08565
38. Mitchell, M.: An Introduction to Genetic Algorithms. MIT Press, Cambridge (1996)
39. Mu, Q., Fu, Z., Lysgaard, J., Eglese, R.W.: Disruption management of the vehicle routing problem with vehicle breakdown. J. Oper. Res. Soc. **62**(4), 742–749 (2010)
40. Nahum, O.E.: The Real-Time Multi-Objective Vehicle Routing Problem. (Ph.D. Thesis), Bar-Ilan University, Ramat-Gan, Israel (2013)

41. Nahum, O.E., Hadas, Y., Spiegel, U.: Multi-objective vehicle routing problems with time windows: A vector evaluated artificial bee colony approach. Int. J. Comput. Inf. Technol. **3** (1), 41–47 (2014)
42. Nahum, O.E., Hadas, Y., Spiegel, U., Cohen, R.: The real-time multi-objective vehicle routing problem - case study: Information availability and the quality of the result. In: Paper presented at the Transportation Research Board (TRB) 93rd Annual Meeting, Washington DC, USA (2014)
43. Norouzi, N., Tavakkoli-Moghaddam, R., Ghazanfari, M., Alinaghian, M., Salamatbakhsh, A.: A new multi-objective competitive open vehicle routing problem solved by particle swarm optimization. Netw. Spat. Econ. **12**(4), 609–633 (2012)
44. Novoa, C.M., Storer, R.: An approximate dynamic programming approach for the vehicle routing problem with stochastic demands. Eur. J. Oper. Res. **196**(2), 509–515 (2009)
45. Park, Y.B., Koelling, C.P.: A solution of vehicle routing problems in a multiple objective environment. Eng. Costs Prod. Econ. **10**(2), 121–132 (1986)
46. Park, Y.B., Koelling, C.P.: An interactive computerized algorithm for multicriteria vehicle routing problems. Comput. Ind. Eng. **16**(4), 477–490 (1989)
47. Potvin, J.Y., Xu, Y., Benyahia, I.: Vehicle routing and scheduling with dynamic travel times. Comput. Oper. Res. **33**(4), 1129–1137 (2006)
48. Psaraftis, H.N.: Dynamic vehicle routing: Status and prospects. Ann. Oper. Res. **61**(1), 143–164 (1995)
49. Ritzinger, U., Puchinger, J., Hartl, R.F.: A survey on dynamic and stochastic vehicle routing problems. Int. J. Prod. Res. **54**(1), 215–231 (2016)
50. Roy, B.: The outranking approach and the foundations of ELECTRE methods. Theor. Decis. **31**(1), 49–73 (1991)
51. Saaty, T.L.: A scaling method for priorities in hierarchical structures. J. Math. Psychol. **15** (3), 234–281 (1977)
52. Saaty, T.L.: Analytic network process. Encyclopedia of Operations Research and Management Science, pp. 28–35. Springer, Heidelberg (2001)
53. Saaty, T.L.: Decision making with the analytic hierarchy process. Int. J. Serv. Sci. **1**(1), 83–98 (2008)
54. Schaffer, J.D.: Multi-objective optimization with vector evaluated genetic algorithms. In: Paper presented at the 1st International Conference on Genetic Algorithms, Carnegie-Mellon University, Pittsburgh, USA (1985)
55. Secomandi, N.: Comparing neuro-dynamic programming algorithms for the vehicle routing problem with stochastic demands. Comput. Oper. Res. **27**(11), 1201–1225 (2000)
56. Secomandi, N., Margot, F.: Reoptimization approaches for the vehicle-routing problem with stochastic demands. Oper. Res. **57**(1), 214–230 (2009)
57. Sivanandam, S.N., Deepa, S.N.: Introduction to Genetic Algorithms. Springer, Heidelberg (2007)
58. Sivaramkumar, V., Thansekhar, M., Saravanan, R., Amali, S.M.J.: Multi-objective vehicle routing problem with time windows: improving customer satisfaction by considering gap time. In: Proc. Inst. Mech. Eng. Part B: J. Eng. Manuf. 1–6 (2015)
59. Solomon, M.M.: Algorithms for the vehicle routing and scheduling problems with time window constraints. Oper. Res. **35**(2), 254–265 (1987)
60. Tavakkoli-Moghaddam, R., Alinaghian, M., Salamat-Bakhsh, A., Norouzi, N.: A hybrid meta-heuristic algorithm for the vehicle routing problem with stochastic travel times considering the driver's satisfaction. J. Ind. Eng. Int. **8**(1), 1–6 (2012)
61. Toth, P., Vigo, D.: The Vehicle Routing Problem. SIAM, Philadelphia (2001)
62. Toth, P., Vigo, D.: Vehicle Routing: Problems, Methods and Applications. SIAM, Philadelphia (2014)

63. Vincke, P.: Multicriteria Decision-Aid. Wiley, New York (1992)
64. Wang, C., Zhao, F., Mu, D., Sutherland, J.W.: Simulated annealing for a vehicle routing problem with simultaneous pickup-delivery and time windows. In: Prabhu, V., Taisch, M., Kiritsis, D. (eds.) APMS 2013. IAICT, vol. 415, pp. 170–177. Springer, Heidelberg (2013). doi:10.1007/978-3-642-41263-9_21
65. Wang, H.-F., Chen, Y.-Y.: A genetic algorithm for the simultaneous delivery and pickup problems with time window. Comput. Ind. Eng. **62**(1), 84–95 (2012)
66. Wen, M., Cordeau, J.-F., Laporte, G., Larsen, J.: The dynamic multi-period vehicle routing problem. Comput. Oper. Res. **37**(9), 1615–1623 (2010)
67. Żak, J., Kruszyński, M.: Application of AHP and ELECTRE III/IV methods to multiple level, multiple criteria evaluation of urban transportation projects. Transp. Res. Procedia **10**, 820–830 (2015)
68. Zhang, J.L., Wang, W.L., Zhao, Y.W., Cattani, C.: Multiobjective quantum evolutionary algorithm for the vehicle routing problem with customer satisfaction. Math. Probl. Eng. **2012**, 1–19 (2012)
69. Zitzler, E., Laumanns, M., Thiele, L.: SPEA2: Improving the strength pareto evolutionary algorithm. In: Paper Presented at the Eurogen (2001)

Multiple Criteria Evaluation of Suppliers in Different Industries - Comparative Analysis of Three Case Studies

Jacek Żak[1(✉)] and Barbara Galińska[2]

[1] Poznań University of Technology, 3 Piotrowo street, 60-965 Poznań, Poland
jacek.zak@put.poznan.pl
[2] Łódź University of Technology, 266 Piotrkowska street, 90-924 Łódź, Poland
barbara.galinska@p.lodz.pl

Abstract. The chapter presents the comparative analysis of the solution procedures of the suppliers' selection problem in different industries. The authors carry out a multiple criteria evaluation and ranking of three distinctive categories of suppliers, i.e.: meat suppliers in a food industry, logistics service providers (LSPs) in a household chemistry industry and suppliers of packaging and supplementary materials in a printing industry. The definition of variants (suppliers), family of evaluation criteria and modeling of the decision maker's (DM's) preferences are presented. The results of computational experiments performed with the application of Electre III/IV and AHP methods are demonstrated. The processes of selecting the most desired suppliers in the above mentioned cases are compared and different aspects of the proposed multiple criteria - based procedures are thoroughly discussed.

Keywords: Suppliers' selection problem · Ranking of suppliers · Multiple Criteria Decision Making/Aiding (MCDM/A) · Electre III/IV method · AHP method

1 Introduction

The process of procurement belongs to the major categories of logistics activities and has a critical impact on the operations of many manufacturing and service companies all over the world. Different entities, business units and firms order and purchase various categories of raw materials, components, semi-finished goods, final products, by-products, utilities and services. They carry out the sourcing of the required goods and services either on the local markets or internationally, which results in a different scale and organizational effort of the purchasing process.

In each procurement process suppliers play a critical role and are responsible for satisfying the universal logistics principles, widely known as the "Seven Rights Rule" [35]. They are obliged to deliver: right product, in a right condition and right quantity, at a right time, to right customer, in a right place and at right costs. Based on the modern concepts of Supply Chain Management [4] suppliers should be also featured by: agility, leanness, flexibility, cooperative spirit, compromise orientation, ability to solve problems, customers' service excellence. The suppliers should guarantee a short

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_6

and long – term deliveries of certain materials and products. Based on the modern logistics principles they should be involved in the manufacturing and distribution processes of their customers to generate the highest possible added value to the final consumer.

The efficiency and reliability of procurement and supply processes have a strong influence on the overall profitability and competitive position of individual companies and whole supply chains. They have also a strong impact on satisfaction and trust of their customers. Since in each supply chain the natural connection between its links is the supplier – customer relationship the customer companies search for the most efficient, reliable, safe, flexible, trustworthy and cost-effective arrangement of their deliveries. They penetrate different markets to find the suppliers that are featured by the above mentioned characteristics. The current tendency in the supply chain is that customers tend to find the suppliers that match their profile, represent the similar organizational culture and aspire to common overall goals and values [14, 34].

For the above mentioned reasons the selection of suppliers belongs to an important category of decision problems in modern logistics. It is a widely discussed topic in research reports and various categories of publications [1, 13, 14, 47]. It is also a topic of thorough practical analysis by business organizations. As presented by Żak [47] the selection of suppliers needs a comprehensive analysis and requires a lot of organizational effort. Different authors [10, 17, 41, 42] develop rules, methodologies and paradigms to assist the decision makers (DMs) in a proper selection of suppliers. Some of those rules have universal, global character others are branch specific and local – oriented. Different methods, procedures and algorithms have been developed to recognize strengths and weaknesses of suppliers, evaluate their experience and market position, assess their organizational capabilities and compatibility with the customer [7, 31]. Based on the opinions of different authors one may conclude that the suppliers' evaluation process should include [4, 14, 43] different functional segments of the supplying company and various aspects of their activities. Some researchers [14, 34] claim that, based on the modern rules of the supply chain management, the evaluation of suppliers should include the analysis of their: potential to build long-term, stable cooperation, cultural and organizational integrity with the customer, capabilities for innovation and development, reliability and trustworthiness in other partnerships, willingness to share risk and profit with the cooperating institution.

In this chapter the authors carry out a comparative analysis of multiple criteria evaluations of suppliers in three different industries. They evaluate meat suppliers in a food industry, logistics service providers (LSPs) in a household chemistry industry and suppliers of packaging and supplementary materials in a printing industry. In all three cases they formulate the suppliers' evaluation problems as a multiple criteria ranking problem and apply the universal procedure for solving the multiple criteria decision problems, based on the principles of Multiple Criteria Decision Making/Aiding (MCDM/A). They define the suppliers as the considered variants and construct the same evaluation criteria for all three cases. They present the process of modeling the DM's and stakeholders' preferences and carry out a series of computational experiments with the application of selected multiple criteria ranking methods (Electre III/IV and AHP). Based on the generated rankings they recommend the selection of the most desired supplier in each of three cases. Finally, they compare the generated results and

present their thorough analysis and discussion. In contrast to their previous works [1, 14, 47] they extend the analysis to three cases, featured by the analysis of different categories of suppliers in distinctive industries. As opposed to the approach presented by Żak [47] they do not present alternative formulations of the suppliers' evaluation problem in each of the cases but put their effort to define them in a very similar way. In all cases they define the suppliers as variants and apply the same evaluation criteria. They distinguish the same DMs and stakeholders in all three cases and model their preferences in a similar manner. In the phase of computational experiments they apply the same decision making/aiding methods.

The overall research goal of this chapter is to develop a universal, generic methodology of evaluating suppliers and selecting the most desirable ones regardless the considered business environments, supply conditions and external circumstances. The authors of this chapter claim that the suppliers' selection problem has a multiple criteria character, and thus develop the proposed approach based on the principles of Multiple Criteria Decision Making/Aiding [11, 38, 43, 46]. The challenge and the novelty of this work is to present the comparison of the multiple criteria evaluations of suppliers across different industries, where the nature of the selection process and the profile of suppliers are entirely different. The originality of this work consists also in the description and confrontation of all components of multiple criteria analysis of suppliers in different industries. To the best of the author's knowledge such a contribution has not been reported in the literature, so far.

The chapter is composed of 5 sections. In the first one the background of the analysis is presented and the principles of the suppliers' selection problem are discussed. In addition, objectives of the research are defined. The second section includes the description of the methodological background of research. The principles of Multiple Criteria Decision Making/Aiding, the applied MCDM/A methods and the analysis of the suppliers' selection problem are presented in this section. Section 3 is focused on the presentation of the new, original approach to solving the suppliers' selection problem. This section includes a proposed procedure/paradigm of evaluating and selecting suppliers in various industries. The developed approach is practically verified in Sect. 4, where the analysis of three case studies is carried out. Each of the case studies refers to the selection of suppliers of a specific profile and character in a different industry, including: meat suppliers in a food industry, logistics service providers in a household chemistry industry and packaging suppliers in a printing industry. This section includes the results of computational experiments generated with the application of Electre III/IV and AHP methods followed by their analysis. It contains a comprehensive comparison of the considered case studies and the corresponding decision processes focused on the selection of various suppliers in different industries. It is followed by final conclusions presented in Sect. 5, called Summary. The chapter is supplemented by a list of references.

2 Methodological Background of Research

2.1 General Characteristics and Basic Notions of the MCDM/A Methodology. Major Features of the Applied MCDM/A Methods

Multiple Criteria Decision Making/Aiding (MCDM/A) is a field of study that develops rules, tools and methods supporting the decision maker (DM) in solving complex decision problems, in which several – often contradictory – points of view must be taken into account [11, 38, 43, 46]. The methodology of MCDM/A has a universal character and can be applied in various cases when the DM solves a so called multiple criteria decision problem (MCDP).

MCDP is a situation in which, having defined a set of actions/variants/solutions A and a consistent family of criteria F the DM tends to [11, 38, 43, 46]:

1. determine the best subset of actions/variants/solutions in A according to F (choice problem),
2. divide A into subsets representing specific classes of actions/variants/solutions, according to concrete classification rules (sorting problem),
3. rank actions/variants/solutions in A from the best to the worst, according to F (ranking problem).

There are two major components of each MCDP, i.e. a set of actions/variants/solutions A and a consistent family of criteria F. The set of A can be defined directly in the form of a complete list or indirectly in the form of certain rules and formulas that determine feasible actions/variants/solutions, e.g. in the form of constraints [43, 46]. The consistent family of criteria F should be characterized by the following features [29, 30, 38]: it should provide a comprehensive and complete evaluation of A, each criterion in A should have a specific direction of preferences (minimized – min or maximized – max) and should not be related with other criteria in F. The domain of each criterion in F should be disjoint with the domains of other criteria.

In this chapter the multiple criteria evaluation of suppliers is defined as a multiple criteria ranking problem. The suppliers of services (transportation and logistics services) and goods/raw materials (meat, packages and supplementary materials) in three different industries (household chemistry industry, food industry and printing industry) constitute the considered variants A. In each case study suppliers are evaluated by a standardized, consistent family of criteria F and finally ranked from the best to the worst. The criteria evaluate different aspects of the considered variants, i.e.: organizational, market-oriented, technical, economic, social and safety-oriented, environmental. They represent the interests of different stakeholders, including: the customer (manager/owner), the supplier, the final consumer.

The above defined MCDPs are solved with the application of specific procedures, methods and decision making/aiding tools, which can be classified according to several criteria, including the manner of the preference aggregation [43–45]. Using this classification criterion one can distinguish:

1. The methods of American inspiration [20] based on the utility function; e.g. AHP [32, 33] or UTA [11].
2. The methods of the European/French origin, based on the outranking relation (e.g. Electre methods [29, 30], Promethee I and II [11].

In this chapter the Electre III/IV method and AHP method are applied to rank the suppliers in the three above mentioned industries. The evaluation of suppliers is carried out based on a universal 5-step procedure of solving the multiple criteria decision problem, described in the works of Roy [29, 30] and Figueira et al. [11].

When the MCDPs are solved it is practically impossible to select options or variants which are the best from all points of view simultaneously. Thus, multiple criteria methods do not yield "objectively best" solutions. Instead of that they focus on producing compromise solutions [11], i.e. such variants that take into account the trade-offs between criteria, balance the expectations of different stakeholders and satisfy the preferences of the DM.

Major features of Electre III/IV method

The Electre III/IV method belongs to a family of multiple criteria ranking procedures based on the outranking relation [29, 30]. It generates final rankings of a finite set of variants and orders them from the best to the worst, taking into account the following relationships between variants: indifference (I), preference (P) and incomparability (R).

To start the computational procedure of the Electre method the following input data is required: matrix of performances, which comprises the evaluation of each variant by each criterion and the DM's preference model constructed with the application of indifference q_j, preference p_j and veto v_j thresholds as well as the weight w_j, defined for each criterion. The thresholds define the sensitivity of the DM to the changes of the criteria values and the weight w_j, expresses the importance of each criterion.

The computational procedure starts with the calculation of concordance indicators ***C(a, b)*** for each pair of variants a and b and presented as a concordance matrix. ***C(a, b)*** expresses the extent to which the scores of a and b on all criteria are in concordance with the proposition a outranks b. Next, the discordance index $\boldsymbol{D_j(a, b)}$ is calculated for each criterion j, taking into account arguments represented by the values of criteria for which one cannot accept the statement that a outranks b. Finally the outranking relation S is constructed. S indicates the extent to which "a outranks b" overall. It is expressed by the degree of credibility ***d(a, b)***.

Based on ***d(a, b)*** the method establishes two preliminary rankings (complete preorders) using a classification algorithm (distillation procedure). During this procedure one can obtain a descending and an ascending preorder. In the descending distillation the variants are ranked from the best to the worst, while in the ascending distillation they are ranked in the inverse order. The intersection of two preorders gives the final ranking, which is usually presented in a graphical form. It corresponds to a relation matrix that includes final, overall, mutual relations between variants, expressed in the following form: indifference (I); preference (greater than –">"); non-preference or inverse of preference (less than –"<") and incomparability (R). Further details of the computational algorithm applied in Electre III/IV method can be found in the works of Roy [29, 30], Vincke [38], Figueira et al. [11].

Major features of Analytic Hierarchy Process (AHP) method

Similarly to Electre III/IV method the AHP is also a multiple criteria ranking procedure that ranks a finite set of variants from the best to the worst. It is focused on the hierarchical analysis of the decision problem. Through the definition of the overall objective, evaluation criteria, sub-criteria and variants the method constructs the hierarchy of the decision problem. As opposed to Electre III/IV method the AHP is based on a multi-attribute utility theory [20] and generates final rankings of variants based on their aggregated evaluations represented by a utility function. In the final rankings generated by the AHP method the possible relationships between variants are limited to: indifference (I) and preference (P). All variants evaluated by the AHP method are comparable.

The computational procedure of the AHP method starts with the definition of the DM's preferential information on each level of the hierarchy, based on the pair-wise comparisons of criteria, sub-criteria and variants. This model of preferences is expressed in the form of relative weights w_r [32], which represent relative strength of the compared element against another, expressed on the 1 to 9 point scale. All weights have a compensatory character, i.e.: the value that characterizes the less important element (1/2, 1/5, 1/9) is the inverse of the value assigned to the more important element in the compared pair (2, 5, 9).

In the next step of the AHP computational procedure the consistency of the above mentioned preferential information given by the DM on each level of the hierarchy, is investigated. For each matrix of relative weights w_r the consistency index CI is computed. If the value of CI is close to 0 the preferential information given by the DM is considered to be almost perfect. The acceptable level of CI is below 0.1. If this condition is not satisfied the redefinition of the preferential information is required.

The algorithm of the AHP method solves a, so-called, eigenvalue problem [32] on each level of the hierarchy. As a result a set of vectors containing normalized, absolute values of weights w_a for criteria, sub-criteria and variants is generated. The sum of the elements of the vector is 1 (100%). The absolute weights w_a are aggregated by an additive utility function. The utility of each variant i – U_i is calculated as a sum of products of absolute weights w_a on the path in the hierarchy tree (from the overall goal, through criteria and sub-criteria) the variant is associated with. The utility U_i represents the contribution of variant i in reaching an overall goal and constitutes its aggregated evaluation that defines its position in the final ranking.

The comparison of Electre III/IV and AHP methods

Both Electre III/IV and AHP are MCDM/A methods classified as multiple criteria ranking methods, thus they are suitable for solving the decision problems formulated as multiple criteria ranking problems, including the suppliers' selection problem. From a methodological point of view they represent different schools of MCDM/A. AHP method belongs to the American school of MCDM/A and it is based on the Multiattribute Utility Theory [20]. Thus, it aggregates the evaluations of variants according to different criteria into one Utility function whose values constitute the variants' final, overall scores. As opposed to that, Electre III/IV, being the MCDM/A method of the European/French origin is based on the binary outranking relation S. As a result, while

ranking the variants (a and b) it investigates the arguments "for" and "against" the statement that "a outranks b" (aSb).

The final rankings generated by both methods are different. The AHP method produces the quantitatively oriented, cardinal ranking, based on the values of Utilities of particular variants. Its form is linear (top – down) and based on two alternative, possible relationships between variants, i.e. indifference (I), when U(a) = U(b) and preference (P), when U(a) > U(b). The AHP does not accept the situations when variants are incomparable (R). The ranking generated by the AHP gives the DM the ability to recognize the distance between the variants. The Electre III/IV method generates the ranking in both graphical and tabular form. The final graph has a form of a hierarchical tree representing the following mutual relationships between variants: Indifference (I), Preference (P) and incomparability (R). In this case the DM cannot see the distance between variants.

Both methods are suitable for ranking the relatively small sets of variants, characterized by deterministic input information. They are both user-friendly and relatively easy to understand. Both methods help the DM to carry out the analysis of a complex decision problem in a systematic, rigid way and structure it in an appropriate manner. AHP method enables the DM to shape it in the form of a simple hierarchy, while Electre III/IV method allows for designing the tabular evaluations of all variants according to all criteria. Both methods are flexible enough to accept both quantitative and qualitative characteristics/criteria.

The critical advantage of AHP method is its ability to accept a comprehensive structure of criteria, including sub-criteria and sub-sub-criteria. Electre III/IV method does not have these capabilities. It allows for considering a simple structure of criteria, only. In return Electre III/IV method is featured by low labor-intensity, while AHP is very time consuming.

The authors have deliberately chosen these two methods, as suitable for solving the multiple criteria ranking problem but at the same time handling it in a different manner. They wanted to demonstrate how the representatives of two alternative schools of MCDM/A can deal with the suppliers' evaluation and selection problem. The choice of Electre III/IV and AHP methods has allowed the authors to compare the computational results generated by two methods based on different methodological (axiomatic) principles, alternative ways of defining and structuring criteria and different techniques of modeling the DM's preferences.

2.2 The Suppliers' Selection Problem

In general, a supplier, often called a vendor is a party/organization that supplies goods or services [40]. It is one of the five major forces of the Porter's [27] famous model of strategic analysis of any company/organizational unit and defining the supplier's competitive position and bargaining power. Based on this model the supplier interacts with the customer (manufacturing and/or service organization), an entity that competes at the target market. As described in this model the role of supplier is critical for the prosperity and competitive position of the target organization. Different categories of suppliers exist at each link of any supply chain. The supplier plays an important role in

it and its cooperation with customer, operational efficiency and quality standard of delivery contribute to the overall value added of the supply chain and satisfaction of the final consumer [4, 7].

For this reason the supplier selection problem is a widely discussed topic that has attracted the attention of many researchers [4, 13, 14, 18, 21, 22, 26, 43]. Different procedures have been developed to recognize strengths and weaknesses of suppliers, evaluate their experience and market position, assess their organizational capabilities and compatibility with the customer [7, 31]. There are many approaches to the suppliers' selection problem, reported in the literature. In general, they can be classified into the following groups [36]:

1. Suppliers' Elimination Methods based on critical criteria or constraints that must be satisfied [13, 48];
2. Multiple Criteria Ranking Methods (e.g.: AHP, Electre, ANP, Promethee, TOPSIS), both deterministic and non-deterministic (e.g. fuzzy methods) (e.g.: [10, 41]);
3. Other methods, including analytical methods (e.g. Total Cost of Ownership Method); optimization methods, probabilistic classification methods, hybrid approaches (e.g.: [10, 13]).

The typical suppliers' selection process is composed of four phases: the definition of the decision problem (suppliers' selection problem), the formulation of criteria, the qualification of potential suppliers and the final choice/selection of the best supplier [8]. The identification of suppliers/vendors and the definition of criteria are the most important phases of this process. The criteria, being the critical characteristics/variables in the suppliers' analysis must be objectively and carefully selected. Then the suppliers have to be identified, examined and assessed according to all considered criteria. Certain aggregation of the overall performance of vendors is required to select the most suitable party in the final stage of the procedure. As proved by many experimental studies the vendors are not expected to be the best on all criteria. A certain balance of their performance may be a key factor of their success [10, 15]. The supplier that delivers products of a very good quality might be too demanding cost-wise and thus, neglected by the customer. At the same time the vendor that can offer the products at low prices may be inefficient and unreliable delivery-wise and rejected due to its delivery performance. The trade-offs definitely exist.

In general, researchers agree that the evaluation and selection of suppliers should be supported by different quantitative techniques. They propose to use in the suppliers' selection process different computational procedures, from simple weighted techniques (empirical perspective) [39] to advanced mathematical programming methods - mathematical perspective [3, 28, 37, 41]. For the time being the above mentioned authors have not agreed on a single, universal path/process of selecting the suppliers. At the same time some of them have pointed out that the suppliers selection problem has a multiple criteria character and involves the trade-offs between conflicting qualitative and quantitative characteristics [3, 28, 41]. The authors of this chapter support this statement and extend it, claiming that the suppliers' selection problem should involve the analysis of various aspects of technical, economic, social, organizational, market-oriented and environmental character and the interests of different stakeholders (interveners).

The application of MCDM/A methods for the evaluation and selection of suppliers is becoming more and more popular. Guneri and Kuzu [17] have developed and applied a multiple – criteria oriented, fuzzy logic method to solve the suppliers' selection problem. Their method is based on generating a set of Pareto optimal variants (suppliers) and then calculating fuzzy suitability indices for them measuring their compatibility with the pre-defined expectations and aspirations of the decision maker (DM). The fuzzy indices are used to rank the variants from the best to the worst, which finally results in the selection of the best vendor. Yu and Wong [41] have used the fuzzy TOPSIS (fuzzy Technique for Order Preference by Similarity to Ideal Solution) method to select the most suitable supplier based on its reference to the ideal solution. The authors have applied in their approach a pre-selection algorithm, which has allowed them to filter some suppliers based on the definition of critical constraints and aspiration levels. Afterwards they have developed a mathematical model that has generated a fuzzy ranking of filtered suppliers based on their fuzzy, weighted distance to the ideal solution. Ertay et al. [10] have proposed a hybrid methodology for the suppliers' selection problem based on the application of both AHP and Electre methods. In the proposed selection process the fuzzy AHP (Analytic Hierarchy Process) method has been used to define the hierarchy of goals (criteria) and weight their importance and performance of variants (suppliers) on all criteria using the standardized AHP scale. Based on the generated evaluations of suppliers the Electre III method has been used to classify and rank the suppliers based on their performance on all considered criteria. Yücenur et al. [42] have proposed another hybrid decision model for the suppliers' selection problem in a real world environment. In their procedure they have applied a combination of the AHP and fuzzy ANP (Analytic Network Process) methods. The AHP- based methodology has allowed them to generate the importance of different decision criteria and performance of vendors on these criteria. The authors have used various universal characteristics applied in the selection of suppliers. The ANP method, being an extension and modification of the AHP procedure, has been used to generate the fuzzy overall score – utility of all variants (suppliers) and finally rank them from the best to the worst. In their analysis the authors have used the important characteristics of the ANP method, i.e. its ability to envisage interactions and dependencies between the elements of the proposed hierarchy. As a result the network structure of paths of influence (importance) with a source and a link has been applied. It has helped the DM-s (managers) to solve the real world case study focused on the selection of the most desired supplier.

The publications on multiple criteria evaluation of suppliers show that the authors apply in their analysis different families of criteria. Since criteria are the most important characteristics/variables of the suppliers' selection problem (as mentioned earlier) this inconsistency is surprising because it leads to different solutions of the considered decision problem. Based on the review of different articles it has become transparent that various authors present diversified views on the criteria formulation for the supplier's selection problem. Dickson [9], who was the first author to discuss the problem of criteria selection for the suppliers' evaluation problem, identified 23 different characteristics. He indicated that the cost and quality of the product delivered as well as delivery performance are the most important attributes. Burton [5] selected 10 criteria for the suppliers' evaluation. His perspective was very much a manufacturing oriented

one. The author considered such characteristics as: quality of the product delivered, standard of delivery, production facilities and capacity at the supplier's site, net price of the product delivered, technical capability of the supplier, its packaging ability, geographic location of the supplier, training assistance, supplier's management and organization and operational control. Monczka et al. [24] used another, however somehow similar set of 10 criteria, including: product quality, standard of delivery, performance history, supplier's production facilities and capacity, net price of the product delivered, financial condition of the supplier, its reputation and competitive position in the industry, its management and organization, labor relations record and geographical location. Bernard [2] applied 5 attributes, including: product quality, standard of delivery, product net price, supplier's management and organization and service. Chapman [6] has reduced the number of considered characteristics to 3 criteria only, i.e.: product quality, standard of delivery and supplier's production facilities and capacity. Yücenur et al. [42] proposed in the described above suppliers' evaluation model 4 major criteria, such as: service quality, cost, risk factors and supplier's characteristics. They defined them based on the analysis of 28 different characteristics, factors and attributes. Ordoobadi [25] and Labib [23] applied in the suppliers' selection processes, described in separate works, 4 criteria, including: product quality, standard of delivery, service quality and costs of delivery. These criteria were sub-divided into 12 sub-criteria. The approach of Labib [23] was a certain modification and extension of the model proposed by Ordoobadi [25]. Ertay et al. [10] proposed a suppliers' assessment model based on 6 main criteria, such as: reliability, responsiveness, flexibility, cost and financial conditions, assets and infrastructure, environment which are constructed with the application of 20 sub-criteria. The approach of these authors is a certain deviation of the comprehensive analysis of suppliers carried out by Huang and Keskar [19]. The latter researchers used in their suppliers' evaluation 7 main criteria, the same one as in the work of Ertay et al. [10] and, in addition, safety. They defined them based on the development of 101 sub-criteria.

Other authors also support the idea that the suppliers' selection problem has a multiple criteria character. They claim that the suppliers' evaluation process should include [4, 14, 43] the analysis of different functional segments of the supplying companies and various aspects of their activities. In their opinion the considered criteria should reflect the supplier's performance, organizational efficiency and market position. Some researchers [14, 34] add that the evaluation of suppliers should include the analysis of their potential to build long-term, stable cooperation with a customer as well as their cultural and organizational integrity with its client. They also indicate that such features as: capabilities for innovation and development, reliability and trustworthiness in other partnerships as well as willingness to share risk and profit with the cooperating institution seem to be equally important. As mentioned in the introduction, there are also universal logistics principles, well known as "Seven Rights" [35] that each supplier should satisfy. According to this popular paradigm the supplier should deliver: right product in a right condition and quantity at right time and cost to the right customer and place. These simple rules create concrete guidelines for the definition of criteria that are useful in the suppliers' evaluation process.

All in all, the above presented literature survey shows that there are many approaches to the evaluation of suppliers and their selection. Multiple criteria analysis

is becoming more and more popular in this field. At the same time the proposed approaches and methods are quite different. There is no universal and coherent family of criteria that could be recommended as a generic set of measures used for the evaluation of suppliers. Some of the presented approaches do not satisfy the principle rules of the MCDM/A methodology (e.g. consistency of the family of criteria – see Sect. 2.1) and the psychological concept of 7 ± 2 measures/characteristic/aspects that an average human being (DM) is able to handle [43]. In some cases the authors evaluating the suppliers have had difficulties with the proper definition of the participants of the decision making process, i.e.: DM and stakeholders and their interests. Based on the above mentioned weaknesses of the existing procedures of the suppliers' evaluation the authors of this chapter undertake the challenge to develop a universal/generic and consistent methodology for carrying out a comprehensive and objective evaluation of suppliers in different industries.

3 The Proposed Methodology of Selecting/Evaluating Suppliers in Different Industries

Based on the literature review presented in Sect. 2 the authors of this chapter propose the following universal procedure for evaluation the suppliers and selection of the most desirable entity (supplying) partner. The proposed approach is based on the universal procedure of solving the multiple criteria decision problems [11, 29, 30] and customized to specific features of the suppliers' selection problem. It includes the following stages:

1. Investigation of the decision situation and its verbal description. Definition of the overall objectives of analysis (selection of a certain supplier) and participants of the decision process (DM and stakeholders). Description of major interests, trade-offs and preferences.
2. The suppliers' selection problem structuring and mathematical modeling. Analysis of the parameters characterizing the decision problem, collecting data definition of the set of variants/feasible solutions A and construction of the consistent family of criteria F.

    - It is assumed that the suppliers' selection problem is defined as a multiple criteria ranking problem.
    - The variants A are defined as a set of the considered suppliers. They are denominated by certain symbols, names and/or abbreviations.
    - The family of criteria F must satisfy the conditions of consistency. The authors propose the following universal family of criteria for the supplier's evaluation:
        - K1 – Product Price/Cost and Payment Conditions;
        - K2 – Timeliness of Delivery/Supplier;
        - K3 – Reliability of Delivery/Supplier;
        - K4 – Cost of Delivery;
        - K5 – Accessibility of Supplier;
        - K6 – Customer Service Quality (during the supply process);
        - K7 – Market Position of the Supplier;

- K8 – Performance of the Supplier;
- K9 – Modernity of the Supplier.

3. Analyzing, modeling and aggregating the DM's and stakeholders' preferences (synthesis of the global model of preferences). Defining preferences, aspirations and expectations of the interveners through interviews, group discussions, survey forms.
4. Review, evaluation and selection of the appropriate multiple criteria methods (computational procedures, algorithms) to rank the suppliers from the best to the worst. Matching the specific characteristics of the suppliers' selection problem with appropriate multiple criteria methods. Testing the applied methods. It is recommended to utilize several MCDM/A methods to ensure the reliability and stability of the generated results (for the time being Electre III/IV and AHP methods).
5. Carrying out a series of computational experiments. Solving the suppliers' selection problem with an application of the global model of preferences and selected MCDM/A methods (Electre III/IV and AHP methods). Generating the rankings of suppliers. Review and analysis of the generated results.
6. Selection of the most desired variant – supplier, the winner in the generated rankings and the compromise solution that best matches the trade-offs and expectations of the DM and stakeholders. Recommendation of the selected supplier as the most desired/satisfactory partner. Implementation of the selected solution.

4 Practical Application of the Proposed Methodology

4.1 Phase 1 - Investigation of the Decision Situation and Its Verbal Description. Phase 2 - the Suppliers' Selection Problem Structuring and Mathematical Modeling

Case study I - Analysis of suppliers in the food industry

The first case study refers to the selection of meat suppliers for a small producer GARM of ready-made food products, such as: dumplings stuffed with meat, rump stakes and meat croquettes. Due to the fact that current level of GARM's sales has been relatively small (0.20 million PLN per year; 1 PLN = 0.23 EUR) the company has been focused on purchasing the raw material (pork meat) from a local retailer. In the recent months the demand for GARM's products has increased substantially, which has resulted in the need for ordering the meat in much larger quantities. The demand for pork meat has doubled from 300 kg per month to 600 kg per month. It has been forecast that the future annual demand for meat may reach the level of 10 000 kg (10 tones). This level of demand qualifies GARM for wholesale purchases and requires much larger and more frequent meat deliveries with higher regularity.

In such circumstances the manager/owner of GARM, acting as a decision maker (DM), has decided to terminate the contract with the existing meat supplier and select a new one. Due to the increased demand for meat and the resulting wholesale nature of meat purchases he has decided to investigate a group of slaughterhouses and/or meat processing plants as prospect suppliers of pork meat for GARM. In his opinion, the effect of scale and elimination of one link of the supply chain should result in certain

cost savings (cheaper meat) and higher quality of the delivered material (pork meat). The DM wants to carry out a comprehensive, multiple aspect, objective evaluation of all considered suppliers, including the existing one.

In the analysis 7 meat suppliers - variants, operating at the local market have been considered. The suppliers are denominated by symbolic names: D1, D2, D3, D4, D5, D6 and D7. They are featured by different size, experience and market reputation. In this set of suppliers 6 companies (D1, D2, … D6) represent the meat processing plants. Four of them have their own slaughterhouses and purchase the livestock from Polish farmers, while the two remaining ones cooperate with external slaughterhouses and purchase the preprocessed meat in the form of carcass (of slaughtered animals). The current supplier – D7, being the meat retail store, purchases the meat from a local slaughterhouse.

It is assumed that the variants should be evaluated by a set of criteria, representing different aspects of the suppliers' evaluation, including: economic, market-oriented, technical, social, organizational and environmental. The criteria may have both quantitative and qualitative character and should express the interests of various stakeholders, such as: the manager and owner of GARM, the supplier(s) and final consumer of GARM products.

Case study II – Evaluation of logistics services providers (LSPs) in the household chemistry industry

The suppliers' selection problem considered in the second case study consists in evaluating and ranking a set of logistics service providers (LSPs) that compete for transportation services to be delivered to a customer - manufacturer and distributor of cosmetics, washing detergents and laundry products. The customer - an international company CUS located in Warsaw, Poland delivers its products to 400 customers distributed all over Poland and generates an annual turnover of 400 million PLN (roughly - 92 million euro).

The management team of CUS, acting as a decision maker (DM), is not satisfied with the existing level of transportation services offered by the current carrier. Thus, they search for a new transportation company, that would overtake all activity of the current transport operator. The considered transportation operations are of a large scale. The annual distance to be covered is close to 5 million km. Average monthly shipments carried out on a regular basis correspond to 20 000–25 000 Euro pallets.

In these circumstances the management team of CUS wants to carry out a multi-dimensional analysis of the considered carriers - LSPs. They want to take into account several aspects of different nature and investigate the interests of different stakeholders. Thus, the decision problem has a multiple criteria character and it is formulated as a multiple criteria ranking problem.

In the analysis 8 LSPs - carriers, operating on both Polish and international market constitute the variants. They have different market experience and reputation. Their annual sales range between 7.5 million PLN and 182 million PLN. The smallest company employs 53 employees, while the largest almost 1000 people. The fleet of each carrier is composed of tractors and traitors, trucks and vans. The overall fleet size being in possession of each carrier is diversified and ranges between 0 and 200 vehicles. Some of the LSPs do not maintain their own fleet and subcontract transport

operations. The variants are denominated by their abbreviated names: D1, D2, D3, D4, D5, D6, D7 and D8. The analysis allows to compare the existing carrier – D7 with 7 prospect service providers.

Based on the expressed requirements and aspirations of the DM the variants should be evaluated by a family of criteria that represents different aspects of the LSPs' evaluation, including: organizational, technical, economic, market-oriented, social and environmental. The proposed criteria may have both quantitative and qualitative nature and they should express the interests and expectations of three major stakeholders: the management team and the owners of CUS, the LSP and the customers being serviced by the LSP.

Case study III - Selection of suppliers of packages and supplementary materials for a printing shop

In case study III the analysis is focused on the evaluation and selection of suppliers of supplementary materials (transportation boxes, pallets, wrappings) for a printing shop PROTOB specializing in manufacturing different categories of paper packaging, including various kinds of packages for a tobacco industry, such as: consumer packets of cigarettes, multipack wrappings of cigarettes, tobacco packaging (paper bags and boxes) and customized packages for cigars. The printing shop is located in Lodz, Poland and manufactures and distributes different categories of packages for 150 customers located in different regions of Central, Eastern and Western Europe. PROTOB generates an annual turnover of 130 million PLN (roughly - 30 million euro).

For shipment of its products PROTOB needs a variety of packaging units that are delivered by different suppliers. The critical packaging materials are transportation pallets. The suppliers of these materials have been selected by the PROTOB's Logistics Manager cooperating with a Competition Committee (CC), so far. The CC was responsible for inviting suppliers to submit their proposals for cooperation. The selection of suppliers was based either on a public/open tender or a closed market – oriented competition between invited supplying partners.

Due to the fact that the members of top management of PROTOB, including the Executive Director have not been satisfied with the current supplier of the packaging materials, they have decided to organize a new selection process, which terminate the contract with the current supplier of packaging materials and initiate cooperation with a new one. The process will be based on a comprehensive, multiple aspect and objective evaluation (multidimensional analysis). In the opinion of Top Management more extensive and reliable analysis is required. Thus, the top managers (including the Executive Director), acting as the decision maker – DM, have decided to carry out a thorough, broad and objective analysis of different suppliers of packaging materials to select the most reliable supplying partner that would deliver high quality materials at an appropriate price.

Based on the initial recognition of the market the PROTOB's top management decided to consider a group of 5 suppliers - variants, located in a proximity of the manufacturing site (not further than 125 km). The considered suppliers have different market experience and reputation. Their annual sales range between 0.35 million PLN and 2.5 million PLN. The smallest company employs 14 employees, while the largest

almost 70 people. The variants are denominated by the following symbols: D1, D2, D3, D4 and D5.

Based on the expressed requirements and aspirations of the DM the variants should be evaluated by a family of criteria that represents different aspects of the suppliers' evaluation, including: economic, market-oriented, technical, social, environmental and organizational. The proposed criteria may have both quantitative and qualitative nature and they should express the interests and expectations of major stakeholders: the management team of PROTOB, the supplier(s) and final consumer of products.

Definition of a consistent family of criteria for all three cases. Construction of the evaluation matrices

As presented in Sect. 3 in all three cases the same and universal families of criteria have been applied to guarantee the complete and coherent evaluation of suppliers, regardless the considered industry and characteristics of the supplying process. The proposed family of criteria, defined in Sect. 3, satisfies the expectations of the above mentioned groups of various stakeholders. It covers all aspects required by the DM-s and described in the particular case studies. Specifically, as presented below, technical aspects are represented by criteria K2, K8 and K9, economic aspects are associated with criteria K1 and K4, social issues are described by criterion K6, market-oriented dimensions correspond to criteria K3, K5 and K7 and environmental aspect is included in criterion K9.

Some of the criteria are individual, separate measures, others are composed of certain sub-criteria. They are either maximized (Max) or minimized (Min) depending on their character. The hierarchy/structure of the family of criteria is as follows:

- K1 – Product Price/Cost and Payment Conditions. This criterion characterizes the financial aspects of cooperation with the supplier, including:
  - K1.1 – Unit cost of the product delivered expressed in monetary units [PLN] per a unit of a product (pallet, kg, tone). It is minimized (Min).
  - K1.2 – Payment conditions defined as a length of the payment period for a service (invoice due date), expressed in the number of [Days]. Maximized sub-criterion (Max).
- K2 – Timeliness of Delivery/Supplier is a maximized (Max) criterion defined as a share (percentage) [%] of the deliveries carried out on time (within a predefined time span).
- K3 – Reliability of Delivery/Supplier is a criterion that characterizes aggregated accuracy and promptness of the supplier and supply process. It is composed of the following sub-criteria:
  - K3.1 – Share (percentage) of deliveries of products in appropriate quantity and condition (undamaged) – maximized (Max).
  - K3.2 - Share (percentage) of deliveries carried out as agreed (according to a predefined contract/schedule) – maximized (Max).
  - K3.3 – Quality of the product delivered expressed as a share (percentage) of deliveries in which the delivered product has had a satisfactory quality (according to the predefined standards) – maximized (Max).
- K4 – Cost of Delivery – Overall cost of the delivery process carried out according to the elaborated scheme (frequency, vehicle routes, type of vehicles used) in a

considered period (month, year), expressed in monetary units – [PLN]. This criterion is minimized (Min).

- K5 – Accessibility of Supplier is a magnitude/quantity that measures overall availability of the supplier's service and ease of reaching and contacting the supplier. It is composed of two sub-criteria:
  - K5.1 - Time-oriented accessibility expressed in [Days], measures the aggregated availability of the product and/or service in a certain moment of time- one week. It includes such components as: frequency of deliveries, operating hours of the supplier, 24/7 service – maximized (Max).
  - K5.2 - Geographical accessibility measures the physical availability of the product/service and the closeness between the supplier and customers, expressed as a distance in [Km] between the closest warehouses/depots of the supplier and the customer. Minimized sub-criterion (Min).
- K6 – Customer Service Quality (during the supply process). This criterion measures the supplier's attitude to its customers, defined based on their opinions and experience (expressed in questioner surveys). It is composed of two sub-criteria:
  - K6.1 – Level of customer support expressed in [Points] measures the level of assistance the suppliers provides to its customers. This sub-criterion includes such components as: provision of information, monitoring the supply process, problem solving attitude. It is maximized (Max).
  - K6.2 - Flexibility of the supplier measures its ability to react to urgent customer's requests, needs and expectations concerning order changes and adjustments. It is expressed in the number of [Days] representing the supplier's reaction time, and thus it is minimized (Min).
- K7 – Market Position of the Supplier. This criterion defines the overall reputation and market image of the supplier corresponding to customers' trust and supplier's dependability. It is composed of two sub-criteria:
  - K7.1 – Market experience of the supplier, which measures the lengths of the supplier's business operations and successful service to customers in the considered industry in [Years]. It is maximized (Max).
  - K7.2 – Market share, expressed in [%] in a considered industry as a generic measure of customer's satisfaction and loyalty. It is maximized (Max).
- K8 – Performance of the Supplier. This criterion measures the overall efficiency of the supplier and its business operations. It is composed of two major components, including the Human Resource productivity and Fixed Assets utilization.
  - K8.1 – Productivity/Efficiency of Human Resources is defined as a level of annual sales per employee generated by a supplier. It is maximized (Max) and expressed in monetary units [PLN].
  - K8.2 – Assets Turnover is constructed as a ratio/quotient of the supplier's annual sales and the average value of fixed assets in the considered year. This measure is a maximized (Max), dimensionless quantity [-].
- K9 – Modernity of the Supplier. This criterion measures the overall level of technological and organizational advancement of the supplier according to the up-to-date standards. It includes such components as: standard of technological know-how, quality and modernity of the machine stock and fleet, application of the advanced management methods and techniques (e.g. Lean Management;

Just-in-Time) and applied methods and technologies of environmental protection (e.g. methods or systems supporting sustainable development, water and gas purification technologies, reverse logistics techniques). The criterion is expressed in [Points] and maximized (Max).

It is worth mentioning that in case study II – evaluation of LSP-s - the nature of the selection process is slightly different than in case studies I and III. It results from the fact that in case study II, as opposed to case studies I and III, the "supplied good" is a service and not a physical product. Thus, it is characterized by such features as: intangibility, impermanence (inability to collect and store), irregularity (dispersion) of quality (difficulty with standardization) and simultaneity of production and consumption. Due to this features the interpretation of evaluation criteria applied in the processes of selecting service suppliers is different than in those cases where the selection refers to physical product suppliers. This is particularly transparent for the following criteria: K1, K3 and K4. In case study II one can apply either criterion K1 or criterion K4. It is worth emphasizing that sub-criterion K1.1 is a unit cost (e.g. vehicle-km cost) of the transportation service while criterion K4 is its global equivalent (total transportation/delivery costs). Thus, they are directly correlated and using both of them is useless. For the same reasons (service intangibility and simultaneity of production and consumption) one of the components of criterion K3, i.e. sub-criterion K3.3 does not have any sense for the selection of the LSP-s. It is recommended to eliminate this sub-criterion for any case when the analysis is focused on the selection of service providers.

In all three case studies the evaluation matrices based on the original raw data have been constructed. They have included all the evaluations of the suppliers, in respective industries, according to all the above described, nine criteria. In case studies I, II and III the evaluation matrices have been featured by the following sizes at the level of criteria: 9 by 7; 9 by 8 and 9 by 5, respectively. However, these sizes change if one takes into account all the considered sub-criteria and the above mentioned specific character of the selection of service suppliers. This results in the following dimensions of the evaluation matrices in the respective case studies: 16 by 7, 14 by 8 and 16 by 5. The example of the raw data evaluation matrix for case study I (selection of the meat suppliers) is presented in Table 1.

In all three case studies the raw data has been properly processed. Due to the fact that the selection of suppliers in different industries has been based on the application of two alternative multiple criteria ranking methods (Electre III/IV and AHP) the presented raw data has been handled and adjusted to the requirements of these methods. In all the computational experiments based on the application of AHP method the raw data, including the evaluations of variants/suppliers on all criteria and sub-criteria has remained unchanged. At the same time the raw data for the computational experiments with the application of Electre III/IV method required certain adjustments. For all instances in which single, separate criteria have been applied to evaluate suppliers (criteria K2, K4 and K9) all the evaluations remained unchanged. For all the remaining criteria (K1, K3, K5, K6, K7, K8), structured as quantities composed of sub-criteria, the sub-criterion evaluations have been normalized, i.e. transformed into 0–1 intervals and then aggregated (arithmetically or weighted averaged) within each criterion. As a

Table 1. The Evaluation Matrix based on raw data for case study I (selection of meat suppliers)

Criteria		Variants/Suppliers						
		D1	D2	D3	D4	D5	D6	D7
K1	K1.1. [PLN]	13.07	13.14	12.26	11.97	11.74	11.14	12.98
	K1.2. [Days]	1	1	7	1	1	1	1
K2 [%]		80	95	97	89	93	91	99
K3	K3.1. [%]	80	85	95	90	90	90	95
	K3.2. [%]	80	90	95	95	98	95	99
	K3.3. [%]	85	95	90	100	95	95	100
K4 [PLN]		500	350	150	250	300	115	25
K5	K5.1. [Days]	1	5	5	2	5	3	6
	K5.2. [KM]	100	70	30	50	60	23	5
K6	K6.1. [Points]	3	3	5	3	2	4	1
	K6.2. [Days]	2	1	1	1	1	1	1
K7	K7.1. [Years]	15	24	25	20	20	22	23
	K7.2. [%]	1.6	2.0	3.9	1.8	0.9	3.5	0.2
K8	K8.1. [PLN]	25 000	20 800	25 200	22 500	20 900	22 600	12 500
	K8.2. [-]	1.1	2.5	5.9	1.6	0.8	3.8	0.2
K9 [Points]		3	4	4	2	3	4	1

result for all criteria composed of sub-criteria standardized and normalized evaluations have been computed. Thus, the size of all evaluation matrices in the transformed form are equal 9 by 7, 9 by 8 and 9 by 5 for the considered case studies, respectively. Table 2 presents the transformed data for case study II – evaluation of LSPs used in the computational experiments with the application of Electre III/IV method.

4.2 Phase 3 - Analyzing, Modeling and Aggregating the DM's and Stakeholders' Preferences

This phase is focused on the definition of the model of preferences of major actors participating in the decision process. In all three considered cases the analyzed groups of interests have involved: final customer(s)/consumer(s), suppliers and managers/owners of the company evaluating and selecting the suppliers. The latter acting as decision makers (DMs).

All these groups of stakeholders have been interviewed and surveyed. They have been composed of several individuals each. For each person an individual preference profile has been defined. It has included two elements: his/her perception on the importance of each criterion and his/her sensitivity on the changes on the criteria values. The individual preferences have been aggregated into a common preference model for customers/consumers, suppliers and managers/owners (acting as DMs). Afterwards all three group models have been averaged and integrated into one, final model of preferences.

Table 2. The Evaluation Matrix based on transformed data, used in computational experiments with the application of Electre III/IV method for case study II (evaluation of LSPs)

Criteria	Variants/Suppliers							
	D1	D2	D3	D4	D5	D6	D7	D8
K1 [0–1]	0,62	0,35	0,44	0,55	0,84	0,85	0	0,65
K2 [%]	90	78	86	92	75	98	92	96
K3 [0–1]	0,37	0,12	0,54	0,77	0	1	0,31	0,88
K4 [PLN]	2,36	4,43	3,78	2,92	3,35	3,3	5,97	2,14
K5 [0–1]	0,63	0,11	0,45	0,88	0	0,96	0,23	0,9
K6 [0–1]	0,68	0,07	0,5	0,82	0	1	0,22	0,94
K7 [0–1]	0,19	0,74	0,15	0,09	0,13	0,58	0,04	0,08
K8 [0–1]	0,21	0,26	0,18	0,04	0,58	0,23	0,82	0,12
K9 [Points]	4	4,5	2,5	5	4	5	4,5	5

Different modeling techniques have been applied to construct the preference models characteristics for the Electre III/IV and AHP methods. As presented in Sect. 2 the Electre III/IV method preference model has been constructed with the application of indifference q_j, preference p_j and veto v_j thresholds as well as the weight w_j, defined for each criterion. The thresholds define the sensitivity of the DM to the changes of the criteria values and the weight w_j, expresses the importance of each criterion. The pair-wise comparisons of criteria, sub-criteria and variants/suppliers have been applied to generate the preference model characteristic for the AHP method. This model of preferences is expressed in the form of relative weights $\boldsymbol{w_r}$ on the 1 to 9 point scale. Each evaluation represents relative strength of the compared element against another. All weights have a compensatory character, i.e.: the value that characterizes the less important element (1/2, 1/5, 1/9) is the inverse of the value assigned to the more important element in the compared pair (2, 5, 9).

The proposed, final models of preferences characteristic for both Electre III/IV and AHP methods in case study III - selection of suppliers of packages and supplementary materials for a printing shop are presented in Tables 3 and 4.

Table 3. The final model of preferences characteristic for the Electre III/IV method applied in case study III - selection of suppliers of packages and supplementary materials for a printing shop

Preference information					
Criterion	Preference direction	Weight	Indifference threshold	Preference threshold	Veto threshold
K1	Increasing (Gain)	9,000	0,05	0,18	0,5
K2	Increasing (Gain)	5,000	0,03	0,1	0,3
K3	Increasing (Gain)	6,000	0,1	0,3	0,6
K4	Decreasing (Cost)	8,000	100	200	400
K5	Increasing (Gain)	4,000	0,15	0,3	0,6
K6	Increasing (Gain)	7,000	0,1	0,3	0,5
K7	Increasing (Gain)	3,000	0,1	0,3	0,5
K8	Increasing (Gain)	8,000	0,1	0,3	0,5
K9	Increasing (Gain)	2,000	0	1	3

Table 4. The final model of preferences characteristic for the AHP method applied in case study III - selection of suppliers of packages and supplementary materials for a printing shop

Criteria weights w_r									
	K1	K2	K3	K4	K5	K6	K7	K8	K9
K1	1	3	2	1	3	2	4	1	5
K2	1/3	1	1	1/2	1	1/2	2	1/2	2
K3	1/2	1	1	1/2	2	1	2	1/2	3
K4	1	2	2	1	3	1	3	1	4
K5	1/3	1	1/2	1/3	1	1/2	1	1/3	2
K6	1/2	2	1	1	2	1	3	1	3
K7	1/4	1/2	1/2	1/3	1	1/3	1	1/3	1
K8	1	2	2	1	3	1	3	1	4
K9	1/5	1/2	1/3	1/4	1/2	1/3	1	1/4	1

It is worth noticing that both models of preferences (characteristic for Electre III/IV and AHP methods) define financial and efficiency-oriented criteria: K1 – Product Price/Cost and Payment Conditions, K4 - Cost of Delivery and K8 – Performance of the Supplier, as the most important measures. At the same time the criteria that are the least important for the DM are as follows: K9 – Modernity of the Supplier and K7 - Market Position of the Supplier. The importance of criteria is expressed by their weights w in the model characteristic for Electre III/IV method (see Table 3) and by relative weights $\boldsymbol{w_r}$ presented in Table 4, which are characteristic for the AHP method. As presented in Tables 3 and 4 criterion K1 is featured by the weight w equal to 9, and by integer numbers >= 1, characterizing its relative weights $\boldsymbol{w_r}$. The weight w = 9 and weights w = 8 for criteria K1, K4 and K8, respectively (Table 3) correspond to their equal relative importance $\boldsymbol{w_r}$ = 1 (Table 4). A similar relationship is visible for other criteria. This model of preferences corresponds to the profile of the DM, who intends to reduce the delivery costs and select the cheapest and most productive supplier.

4.3 Phase 4 - Review, Evaluation and Selection of the Appropriate Multiple Criteria Methods

Based on different publications [12, 16] the authors of this chapter claim, that while selecting the most appropriate MCDM/A method for solving a specific multiple criteria decision problem one should respond to the following questions:

1. How the features of the method correspond to the type, scope and the specific character of the considered DP? This aspect requires the consideration of the following components:
   - The category of the DP (ranking, choice, classification).
   - The size of the set of the variants.
   - The character of the input and output information (deterministic, non-deterministic).
   - The way of expressing the mutual relationships (positions) between criteria and variants (ordinal scale, cardinal scale).

2. How easy and accurate can the method model the verbally expressed preferences of the DM? This results in the analysis of three sub-questions:
   - Does the preference modeling procedure applied in a certain method properly expresses the intended preferences of the DM?
   - Is preference modeling characterized by high labor intensity?
   - Is the preference modeling process clear and understandable to the DM?
3. What is the form and reliability of generated results (final rankings)? This aspect can be further investigated in two directions:
   - Is the form of final ranking generated by a certain method consistent with the DM's expectations and requirements?
   - Is the generated result (final ranking) reliable and compatible with the DMs overall, global preferences?

Based on this considerations the authors performed a thorough review and analysis of different MCDM/A methods, originating from both American and European schools of MCDM/A. Finally, two alternative multiple criteria ranking methods: Electre III/IV and AHP have been selected. In the authors' opinion they both well match the characteristics of the suppliers' selection problem, allow for flexible adjustments and modeling of the DM's and stakeholders preferences and generate reliable and complementary final rankings. In addition, as very representative and popular methods of two alternative schools of MCDM/A their axiomatic background is different which can lead to generation of reliable and objective final results.

4.4 Phase 5 - Computational Experiments

As described in previous sections all computational experiments in three considered case studies have been carried out with the application of two alternative multiple criteria ranking methods, i.e.: Electre III/IV and AHP. As a result, for the analyzed three case studies, six computational procedures have been performed and six final rankings have been generated, two in each case study.

In all three case studies all stages of the computational algorithms of Electre III/IV and AHP methods have been performed.

The application of Electre III/IV method required the following calculations:

1. Concordance indicators $C(a, b)$, presented as a concordance matrix.
2. Discordance indexes $D_j(a, b)$.
3. Outranking relation S expressed by the degree of credibility $d(a, b)$ and presented in the form of Credibility matrix.

The application of AHP method algorithm corresponded in all three case studies to the following computations:

1. Consistency indexes CI for each matrix of relative weights w_r at each level of the hierarchy (criteria, sub-criteria and variants).

2. A set of vectors containing normalized, absolute values of weights w_a for criteria, sub-criteria and variants.
3. Utility of each variant i – U_i.

For sake of clarity these stages are clearly demonstrated in case study I. In the remaining two case studies, only final outcomes of the suppliers' selection process are presented.

Case study I - Analysis of suppliers in the food industry

Table 5 presents the Credibility matrix for case study I – analysis of suppliers in the food industry. As one can see the degree of credibility ***d(a, b)*** that variant D1 outranks other variants (D2, D3, D4, D5, D6 and D7) is extremely low. At the same time variant D3 is featured by very high values of the degree of credibility ***d(a, b)***. This means that variant D3 substantially outranks other variants (D1, D2, D4, D5, D6 and D7).

Table 5. Credibility matrix generated in the computational procedure based on the application of Electre III/IV method in case study I - analysis of suppliers in the food industry

Credibility matrix							
Alternative	D1	D2	D3	D4	D5	D6	D7
D1	1,000	0,005	0,000	0,000	0,000	0,000	0,000
D2	0,000	1,000	0,000	0,335	0,278	0,000	0,047
D3	0,899	1,000	1,000	0,994	1,000	1,000	0,772
D4	0,000	0,431	0,002	1,000	0,867	0,166	0,000
D5	0,000	0,716	0,000	0,950	1,000	0,133	0,283
D6	0,563	0,986	0,420	0,994	1,000	1,000	0,493
D7	0,000	0,000	0,000	0,078	0,102	0,000	1,000

In the next step, based on Credibility matrix and the values of ***d(a, b)*** two preliminary rankings (complete preorders) have been generated by two classification algorithms, called descending and ascending distillations. The median and the final rankings have been produced as an intersection of two above mentioned preorders. Figure 1 shows the output of these computations, including the following results: Descending distillation, Ascending distillation, Median ranking and Final ranking.

As one can see the leader of the ranking is variant D3 – a large size meat processing plant with its own slaughter house, located relatively close to the customer (30 km), having substantial market experience (25 years) and large market share (3.9%). The current supplier – variant D7 – a small meat retailer receives very good overall scores. It is placed on the second/third place in the final and median rankings. Variant D7, similarly to variant D3 is featured by long market experience (23 years). As opposed to D3, as a small retailer, it has a marginal market share (0.2%). Variant's D7 major benefit is its close location to the customer (5 km). The variants placed at the bottom of the ranking are suppliers D2 and D4. Both variants are medium- sized meat processing plants with a substantial market experience of 24 and 20 years and a moderate market share (2.0% and 1.8%), respectively. The distance of both suppliers from the customer is relatively large (70 km and 50 km).

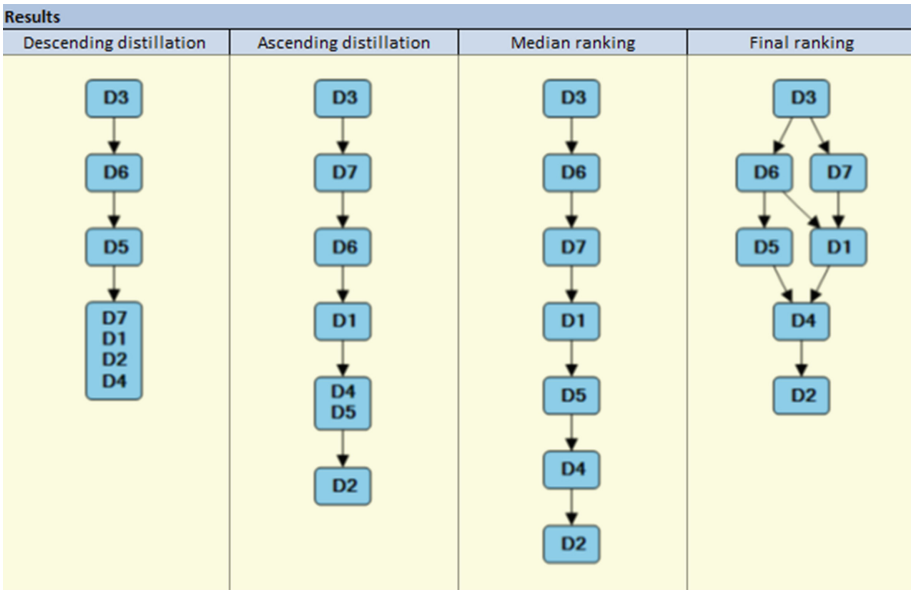


Fig. 1. The results of case study I - analysis of suppliers in the food industry, generated by a computational procedure based on the application of Electre III/IV method

It is worth mentioning that the final ranking presented in Fig. 1 in a graphical form corresponds to a relation matrix shown in Table 6. The relation matrix includes final, overall, mutual relations between variants (D1, D2,…, D7), expressed in the following form: indifference (I); preference (>); non-preference or inverse of preference (<); incomparability ("R"). As one can see from the relation matrix variant D3 is preferred (>) against all the remaining variants. This status corresponds to its top position in the final ranking (Fig. 1). One can also see in the relation matrix that two different variants are not indifferent (I), which is reflected in the graphical form by the fact that none of them are placed in the same rectangular box. In several cases the variants are incomparable between each other, e.g. variant D7 is incomparable (R) against variants

Table 6. Relation matrix generated in the computational procedure based on the application of Electre III/IV method in case study I - analysis of suppliers in the food industry

Relation matrix							
Alternative	D1	D2	D3	D4	D5	D6	D7
D1	I	>	<	>	R	<	<
D2	<	I	<	<	<	<	<
D3	>	>	I	>	>	>	>
D4	<	>	<	I	<	<	<
D5	R	>	<	>	I	<	R
D6	>	>	<	>	>	I	R
D7	>	>	<	>	R	R	I

D5 and D6. The incomparability between variants is expressed by lack of connection between them in the graphical form (see no connecting arrows between neither D7 and D5 nor D7 and D6 in Fig. 1).

Based on the AHP method algorithm the consistency indexes CI for each matrix of relative weights w_r at each level of the hierarchy (criteria, sub-criteria and variants) have been generated. In the analyzed case study 23 CI-s have been computed, including 1 for criteria level, 6 for the sub-criteria levels and 16 for variants compared against each criterion.

In the next step of the AHP method computational algorithm 23 normalized, absolute values of weights w_a for criteria, sub-criteria and variants have been produced. Due to space limitation the results of these calculations are not presented in this text.

Finally, utility of each variant/supplier i – U_i have been computed. The final, normalized score corresponds to the overall position of the variant in the ranking of suppliers. Table 7 presents the computed utilities of each variant – supplier with its absolute values. Figure 2 shows the classification of variants based on their generated utilities in the graphical form. Each variant – supplier, presented in the graph, is featured by the level of computed utility (from 0.240 – D1 to 0.718 – D3 in the absolute values). This corresponds to the range between 0.070 and 0.208 in the normalized values. The winner of the ranking generated with the application of AHP method is variant - supplier D3, followed by variants D6 and D7 (current supplier). The weakest supplier that occupies the bottom position of the ranking is variant D1. Based on the ranking generated by the AHP method algorithm one can also conclude about the distance between the variants – suppliers. It is transparent that utility U_i of variant D3 is substantially larger than utilities of the remaining variants. The difference between D3 and D6 is 0.056 and between D3 and D1 is 0.478 (in the absolute values). The important feature of the ranking is its ability to demonstrate the contribution of each criterion to the final score and position of each variant – supplier. Each distinctive colour on the graph represents the share of each criterion contribution in the utility U_i of each supplier.

Table 7. The values (absolute) of utility of each variant - supplier generated in the computational procedure based on the application of AHP method in case study I - analysis of suppliers in the food industry

	Alternative	Score
☑	D3	0.718
☑	D6	0.662
☑	D7	0.572
☑	D4	0.431
☑	D5	0.422
☑	D2	0.401
☑	D1	0.24

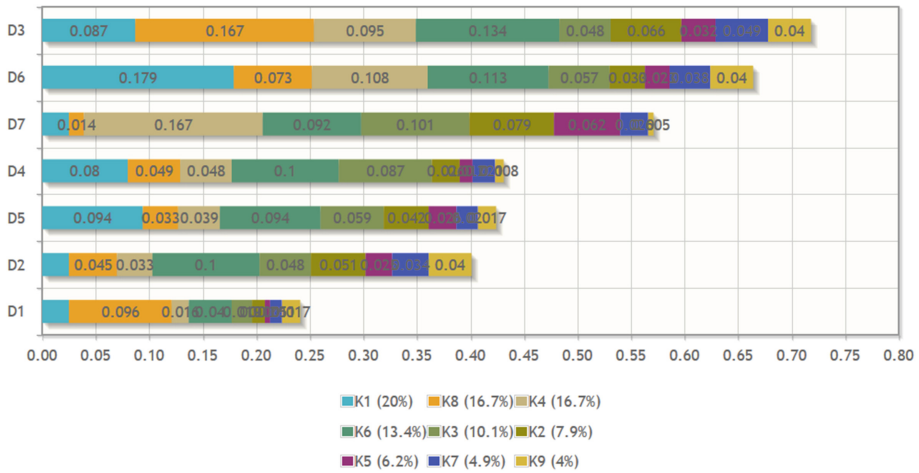


Fig. 2. Graphical representation of the final ranking generated with the application of AHP method in case study I - analysis of suppliers in the food industry

Case study II – Evaluation of logistics services providers (LSPs) in the household chemistry industry

As described above also in case study II all stages of the computational procedures based on the application of Electre III/IV and AHP methods have been performed. The final results of the analysis are presented in Figs. 3 and 4.

In case study II the leader of the ranking, generated with the application of Electre III/IV method, is variant D6 – a large size LSP, employing almost 1000 employees and generating an annual sales of more than 180 million PLN. It is featured by excellent accessibility (both time-oriented and geographical one), substantial market experience (23 years) and large market share (5.5%). The current supplier of transportation and logistics services – variant D7 – a medium sized LSP falls in the middle of the ranking. Variant D7 is featured by a reasonable but shorter than variant D6 market experience (17 years) and a marginal market share (0.7%). The major strength of variant D7 is its excellent performance (HR productivity and fixed assets turnover). The ranking is characterized by many instances of incomparability (R) between variants (e.g.: D1 and D2; D4, D5 and D7; D2 and D3). The variants placed at the bottom of the ranking are LSPs: D1, D2 and D3. Two of them: D1 and D3 are medium sized, Polish carriers with a substantial market experience of 24 years, each and a moderate market share (1.21% and 0.85%), respectively. They employ close to 200 employees each and generate annual sales of 40 and 28 million PLN, respectively. Both are featured by moderate accessibility. They are not characterized by critical disadvantages. Variant D2 is a large-sized, international LSP with a long market experience and tradition. It has been operating on the Polish market for 56 years. Unfortunately it is featured by many weaknesses, such as: low accessibility of supplier (criterion K5), low customer service (criteria K2, K3 and K6) and high delivery costs (criterion K1/K4).

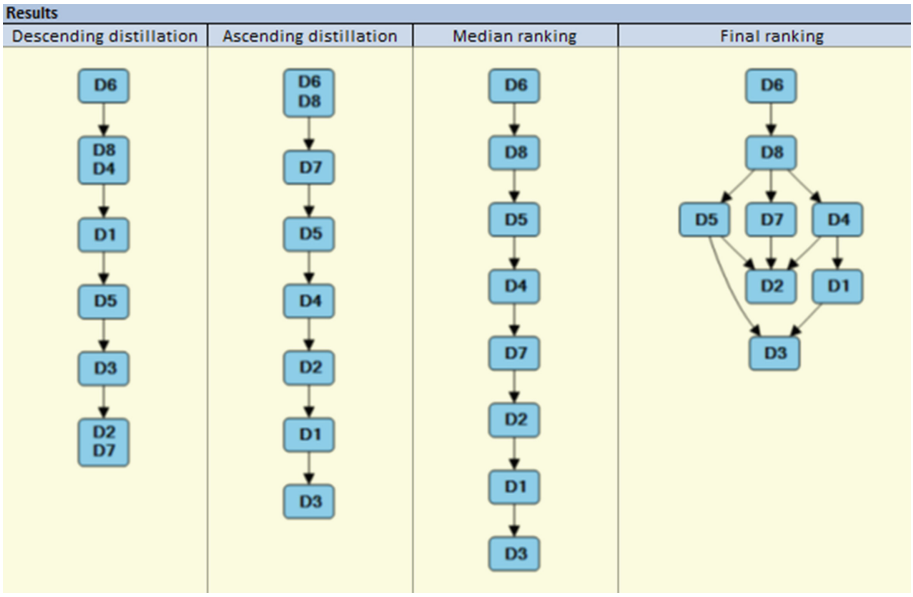


Fig. 3. The results of case study II - evaluation of logistics services providers (LSPs) in the household chemistry industry, generated by a computational procedure based on the application of Electre III/IV method

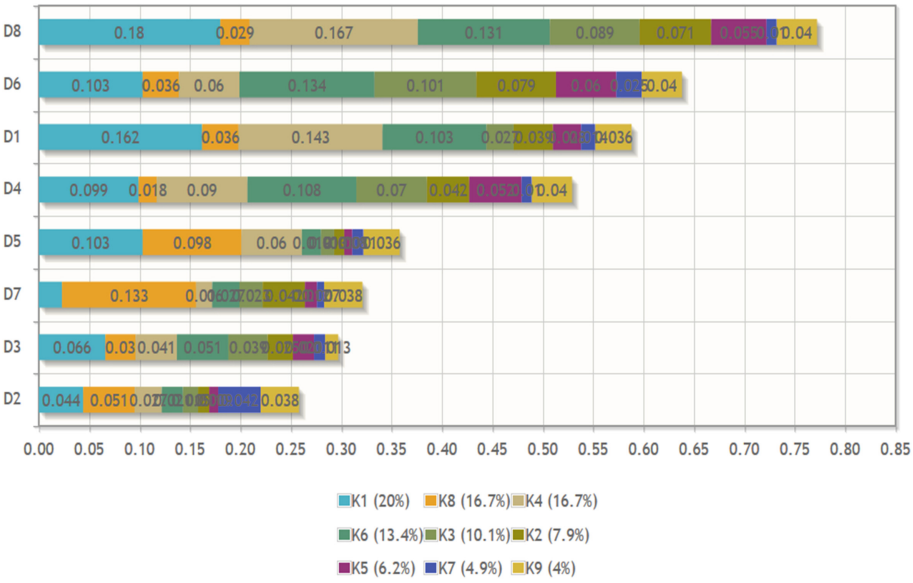


Fig. 4. Graphical representation of the final ranking generated with the application of AHP method in case study II - evaluation of logistics services providers (LSPs) in the household chemistry industry

The output of computational experiments based on the application of AHP method is similar to the results generated by the application of Electre III/IV method, with one exception the position of variant D1. The leader of the ranking, featured by the highest utility - $\boldsymbol{U_i}$ = 0.77 (in the absolute terms) and $\boldsymbol{U_i}$ = 0.204 (in the normalized terms) is variant D8, followed by LSPs D6 and D1. Variant D8 is a medium sized international carrier that employs 130 employees and generates an annual sales of 25 million PLN. It has a tiny market share (0.76%) and reasonable but relatively short market experience among its competitors (19 years). Variant D8 is very efficient cost-wise (criterion K1/K4) and it offers excellent customer service (criteria K2, K3 and K6). The surprising position of variant D1 can be clarified by the lack of its critical disadvantages (as described above) and several good and moderate characteristics. Variants D4, D5 and D7 are placed in the middle of the ranking (similar positions in both experiments) while the bottom positions are occupied by variants D3 and D2. Variant D2 honestly deserves its last position of the ranking due to many weaknesses described above.

As in case study I the ranking generated by the AHP method algorithm allows to draw some conclusions concerning the distance between variants. It is clear that variant D8 outperforms other variants substantially. Its utility is 20% larger than the utility of its closest competitor (variant D6) and 3 times larger than the utility of the loser of the ranking (variant D2). Based on the analysis of the contribution of particular criteria to the overall score of variants one may conclude that the winner of the ranking (variant D8) performs very well on criteria K1, K4 and K6, which belong to the most important measures for the DM and stakeholders.

Case study III - Selection of suppliers of packages and supplementary materials for a printing shop

Similarly to case studies I and II also in case study III all stages of the computational procedures based on the application of Electre III/IV and AHP methods have been performed. The final results of the analysis are presented in Figs. 5 and 6.

It is interesting that in case study III the final ranking of suppliers generated with the application of Electre III/IV method has a linear character, with a sequential order of variants – suppliers from the best to the worst. In the authors' opinion this exceptional situation (total elimination of indifference and incomparability relationships between variants) results from the fact that in case study III the differences between variants on particular criteria are transparent and relatively large. Thus, there is no preferential "ambiguity" while making pair-wise comparisons between variants.

In this case study the indifference and incomparability between variants does not exist. The leader of the ranking is variant D5 – a large sized manufacturing company, operating in Poland and being a daughter-company of a Czech corporation. D5 employs in Poland 70 employees and generates an annual sales of approximately 2.45 million PLN. It is featured by a relatively good accessibility (small distance to the customer – 67 km) and a substantial market experience (24 years). Its market share is large (9.2%). In the middle of the ranking is variant D2, which is featured by a reasonable market experience (25 years) and a moderate market share (3.3%). The major strength of variant D2 is its good customer service (criteria K2 and K3).

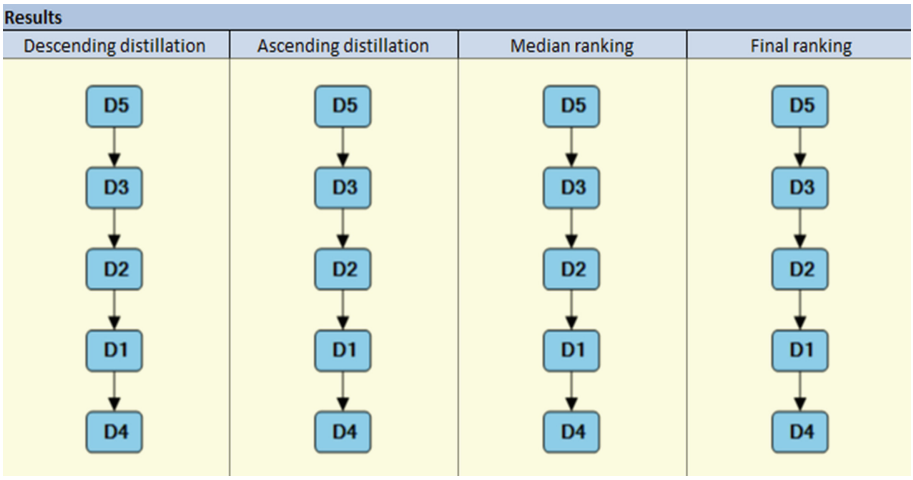


Fig. 5. The results of case study III - selection of suppliers of packages and supplementary materials for a printing shop, generated by a computational procedure based on the application of Electre III/IV method

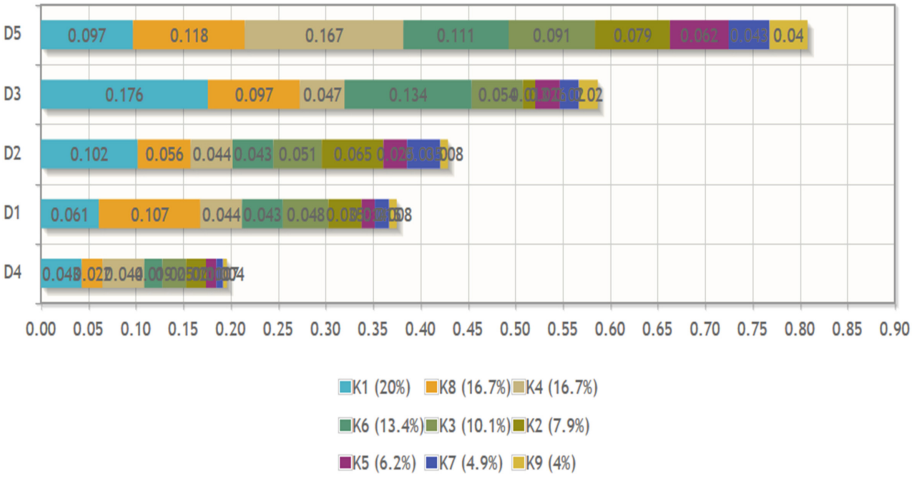


Fig. 6. Graphical representation of the final ranking generated with the application of AHP method in case study III - selection of suppliers of packages and supplementary materials for a printing shop

The variant placed at the bottom of the ranking – D4 is a small manufacturing company with a limited market experience of 16 years and a moderate market share (1.3%). It employs 14 employees and generates annual sales of 0.35 million PLN. It is featured by many weaknesses, such as: high product price (criterion K1), low accessibility of supplier (criterion K5) and low customer service (criteria K2, K3 and K6).

The results generated with the application of AHP method are identical to those produced by the application of Electre III/IV method. Again, the sequence of suppliers from the best to the worst is D5, D3, D2, D1 and D4. Variant D5, the leader of the ranking, substantially outperforms other variants. Its utility $\boldsymbol{U_i}$ = 0.81 (in the absolute terms) and $\boldsymbol{U_i}$ = 0.34 (in the normalized terms) is roughly 30% larger than the utility $\boldsymbol{U_i}$ of variant D3, placed second in the ranking and 4 times larger than the utility of supplier D4, placed the last in the ranking. As presented above variant D5 has many advantages that contribute to its final success. As presented in Fig. 6 the following criteria have the major impact on its leading position in the ranking: criterion K4 – Cost of Delivery (16.7%), criterion K8 – Performance of the Supplier (11.8%) and criterion K6 – Customer Service Quality (11.1%). This means that variant D5 offers well balanced supply service based on efficiency and high quality standards. The AHP-based analysis clearly shows what are the weakest points of variant D4. In comparison with other suppliers it is high product price and cost of delivery (criterion K1 and K4), low accessibility of supplier (criterion K5), low customer service (criteria K2, K3 and K6) and very low modernity of the supplier (criterion K9).

4.5 Phase 6 - Selection of the Most Desired Variant – Supplier

As conclusion of the suppliers' evaluation process in each case study the most satisfactory and desired variant is recommended. This step is carried out based on detailed analysis of the final rankings, generated by different MCDM/A methods in combination with the review of the Evaluation matrix and assessment of each variant – supplier on various criteria. As presented in Table 8 the winners and the losers of the rankings generated by various multiple criteria ranking methods may be similar but not identical.

Table 8. The comparison of final results generated in all three case studies

Features of the solutions	Case study I (Food industry)		Case study II (LSPs)		Case study III (Printing shop)	
	Electre method	AHP method	Electre method	AHP method	Electre method	AHP method
Leading variants/suppliers	D3; D6; D7	D3; D6; D7	D6; D8	D8; D6	D5; D3	D5; D3
Bottom variants/suppliers	D2; D4	D1; D2	D1; D2; D3	D2; D3	D4	D4
Recommended supplier	D3		D6; D8		D5	

In case study I the decision concerning the selection of variant D3 as a recommended supplier is obvious. Variant D3, as a leader of both rankings, substantially outperforms all the remaining suppliers. Variant D3 is a large size meat processing plant with its own slaughter house. It is a reasonable solution cost-wise (moderate

values of criteria K1 and K4) and excellent supplier in terms of market position (criterion K7), customer service quality (criterion K6), performance (criterion K8) and modernity (criterion K9). Supplier D3 has no transparent weaknesses.

In case study II the final selection of the best LSP is more complicated and less transparent. Since there are two different leaders of the rankings: variants D6 and D8, generated by the application of Electre III/IV and AHP methods, they may be considered as equally good for the DM. In both rankings they are placed at the top positions, in the inverse order. Thus, it ambiguous indeed to make the final recommendation and support one of these variants. Variant D6 is a large size LSP. It is featured by excellent accessibility (both time-oriented and geographical one), substantial market experience (23 years) and large market share (5.5%). In contrast to variant D6, its competitor - variant D8 is a medium sized international carrier. It has a tiny market share (0.76%) and reasonable but relatively short market experience among its competitors (19 years). Variant D8 is very efficient cost-wise (criterion K1/K4) and it offers excellent customer service (criteria K2, K3 and K6). Both suppliers are very good and match well the expectations of the DM. The authors of the chapter support both of them as excellent LSPs. Possibly, variant D8 should be recommended for the more cost/price – sensitive DMs.

In case study III both rankings are identical. The linear structure of the ranking generated by Electre III/IV method clearly corresponds to the graphical representation of the utility – based ranking, produced by the AHP method. Without any doubts the recommended solution is supplier D5, which is a large sized manufacturing company. It is featured by a relatively good accessibility (small distance to the customer – 67 km) and a substantial market experience (24 years). Its market share is high (9.2%). The major strength of variant D5 is its very good timeliness of Delivery (criterion K2), reliability of Delivery (criterion K3) and excellent modernity (criterion K9). What is interesting, the product price is here not as low as in the case of other suppliers.

5 Summary

The presented chapter is a comprehensive study concerning the evaluation and selection of suppliers in different industries. Based on the analysis of three case studies that refer to evaluation of: meat suppliers in a food industry, logistics service providers (LSPs) in a household chemistry industry and suppliers of packaging and supplementary materials in a printing industry the authors propose a universal and generic procedure of assessing and selecting suppliers. The proposed paradigm is based on the application of the principles of Multiple Criteria Decision Making/Aiding (MCDM/A). This research is an extension of the authors' previous works [1, 13, 14, 47].

The proposed approach is composed of the following stages:

1. Investigation of the decision situation and its verbal description.
2. The suppliers' selection problem structuring and mathematical modeling. Formulating the suppliers' selection problem as a multiple criteria ranking problem. Definition of variants A and a consistent family of criteria F.

3. Analyzing, modeling and aggregating the DM's and stakeholders' preferences.
4. Review, evaluation and selection of the appropriate multiple criteria methods (computational procedures, algorithms) to rank the suppliers from the best to the worst.
5. Carrying out a series of computational experiments. Solving the suppliers' selection problem with an application of the global model of preferences and selected MCDM/A methods.
6. Selection of the most desired variant – supplier, the winner in the generated rankings and the compromise solution that best matches the trade-offs and expectations of the DM and stakeholders.

These stages have a universal/generic character and can be applied to solve any category of the multiple criteria decision problems that can be formulated as ranking problems. At the same time the authors demonstrate specific features of the suppliers' selection problem within each stage of the proposed algorithm. This refers, in particular, to the definition of criteria (stage 2), original expression of preferences (stage 3) and final selection of the most desired supplier (stage 6).

In each of the above mentioned stages the authors generate an original output, including:

1. Ad 1 and 2. The authors formulate the suppliers' selection problem as a multiple criteria ranking problem and propose the following universal family of criteria for the supplier's evaluation:
   - K1 – Product Price/Cost and Payment Conditions;
   - K2 – Timeliness of Delivery/Supplier;
   - K3 – Reliability of Delivery/Supplier;
   - K4 – Cost of Delivery;
   - K5 – Accessibility of Supplier;
   - K6 – Customer Service Quality (during the supply process);
   - K7 – Market Position of the Supplier;
   - K8 – Performance of the Supplier;
   - K9 – Modernity of the Supplier.
2. Ad 3. They propose a model of preferences based on importance of criteria and sensitivity of the DM. They present its definition in the form characteristic for the Electre III/IV and AHP methods.
3. Ad 4 and 5. They recommend Electre III/IV and AHP methods as reliable, universal MCDM/A tools and prove their applicability for the suppliers' selection problem.
4. Ad 6. They show different forms of the final rankings and characterize the generated output (sometimes ambiguous). They provide guidelines for selecting the most satisfactory supplier.

In the authors' previous works [1, 13, 14, 47] they compared different concepts giving background for the definition of a consistent family of criteria. Their experience with defining different sets of evaluation criteria resulted in the definition of a universal, generic set of characteristics. The comparison of the previously used approaches and the newly proposed family of criteria is presented in Table 9.

Table 9. The comparison of evaluation criteria used in the authors' previous works and in this chapter (universal family of criteria) and compared with the Logistics Standard of "7 Right Concept"

Category/Profile of criteria	Major aspects of suppliers' evaluation and selection	Logistics standard of "7 Rights"	Authors' previous works		Universal family of criteria – case studies I, II and III
			Case study I – evaluation of meat suppliers	Case study II – evaluation of LSPs	
Financial	Cost of the product delivered	Right cost	Unit price of the product (Meat)	(-)	Product price/cost and payment conditions (K1)
	Cost of delivery	Right cost	Transportation cost	Cost of delivery. Financial conditions	Cost of delivery (K4)
Quality	Timeliness of deliveries	Right time	Timeliness of deliveries	Delivery time	Timeliness of delivery/supplier (K2)
	Reliability of delivery	Right product Right quantity Right place	Accuracy of order fulfilment (No. of claims/month) Risk of delivery/exposure to danger.	(-)	Reliability of delivery/supplier (K3)
	Availability of supplier. Accessibility to the delivery system	(-)	Frequency of delivery	Service complexity and availability	Accessibility of supplier (K5)
	Quality of the delivery service and the product delivered	Right condition Right customer	Flexibility of supplier Quality of the raw material	Service complexity and flexibility Quality of human resources	Customer Service quality (K6)
Customer service	Market position, image, experience of the supplier	(-)	(-)	Market experience; market share	Market position of the supplier (K7)
	Economic efficiency of the supplier	(-)	(-)	Fixed Assets Turnover Sales/employee	Performance of the supplier (K8)
	Quality and suitability of the delivery fleet	Right condition	Quality and modernity of the fleet	Fleet quality and suitability	Modernity of the supplier (K9)

As one can see in the authors' previous works several elements of "7 Rights Concept" have not been included in the analysis. At the same time in certain cases of suppliers' evaluation the "7 Right Concept" does not exhaust all the required aspects of the analysis. That is why the introduction of a generic family of criteria for the suppliers' selection problem constitutes an important contribution for transportation and logistics research. Further comments on this topic can be found in the article of Żak [47].

In conclusion, it is important to emphasize that the evaluation of suppliers has a multiple criteria character and can be formulated as a multiple criteria ranking problem. As proved in this chapter different multiple criteria ranking methods, including AHP and Electre III/IV can be applied to rank the suppliers. Due to axiomatic differences of various MCDM/A methods the generated rankings of suppliers may differ. However, the differences between rankings generated by the AHP and Electre methods are not substantial. That is why, the author suggests that in the evaluation of suppliers different MCDM/A methods should be used simultaneously.

It is also important to add that in various processes of suppliers' evaluation and selection similar bodies are involved. Usually, the stakeholders of the suppliers' analysis are: managers and owners of the customer (company that selects suppliers), second echelon customers (final consumers) and competing suppliers. The interest of these groups (often contradictory) should be taken into consideration while selecting suppliers. The decision makers (DMs) making final choice concerning the supplying partner are representatives of the customer (managers, owners). They take major responsibility for selecting the appropriate supplier cooperating with their company. Thus, the analysis presented in this chapter is very important for them and for all analytics supporting managers in selecting the suppliers.

In the authors' opinion further research should be carried out in two directions:

1. Application of alternative MCDM/A methods (Promethee, Mappac, Pragma, ANP, UTA) to the evaluation of different categories of suppliers and in-depth analysis of their suitability, strengths and weaknesses.
2. Further analysis of suppliers' selection processes in different industries. Comparison of evaluation criteria, aspects considered and interests of various stakeholders. This research should finally confirm the consistency and universality of the proposed family of criteria.

References

1. Baran, J., Żak, J.: Multiple criteria evaluation of transportation performance for selected agribusiness companies. Procedia Soc. Behav. Sci. J. **111**, 320–329 (2014)
2. Bernard, P.: Managing vendor performance. Prod. Inventory Manage. J. **30**(1), 1–7 (1989)
3. Boran, F.E., Genc, S., Kurt, M., Akay, D.: A multi-criteria intuitionistic fuzzy group decision making for supplier selection with TOPSIS method. Expert Syst. Appl. **36**(8), 11363–11368 (2009)
4. Brewer, A., Button, K., Hensher, D.: Handbook of Logistics and Supply-Chain Management. Pergamon, Amsterdam (2006)
5. Burton, T.T.: JiT/repetitive sourcing strategies: tying the knot with your suppliers. Prod. Inventory Manage. J. **29**(4), 38–41 (1988)
6. Chapman, S.N.: Just-in-time supplier inventory: an empirical implementation model. Int. J. Prod. Res. **27**(12), 1993–2007 (1989)
7. Coyle, J., Bardi, E., Langley, C.: Logistics Management. Polskie Wydawnictwo Ekonomiczne, Warszawa (2010)
8. De Boer, L., Labro, E., Morlacchi, P.: A review of methods supporting supplier selection. Eur. J. Purchasing Supply Manage. **7**(2), 75–89 (2001)

9. Dickson, G.W.: An analysis of vendor selection systems and decisions. J. Purchasing **2**, 15–17 (1966)
10. Ertay, T., Kahveci, A., Tabanlı, R.M.: An integrated multi-criteria group decision-making approach to efficient supplier selection and clustering using fuzzy preference relations. Int. J. Comput. Integr. Manuf. **24**(12), 1152–1167 (2011)
11. Figueira, J., Greco, S., Ehrgott, M.: Multiple Criteria Decision Analysis. State of the Art Surveys. Springer, New York (2005)
12. Frederic, A., Patrick, H.: Encyclopedia of Decision Making and Decision Support Technologies. Hershey (2008)
13. Galińska, B.: Koncepcja Global Sourcing. Teoria i Praktyka. Difin S.A (2015)
14. Galińska, B., Rybińska, K., Żak, J.: Multiple Criteria Evaluation of Suppliers in Food Industry. Logistyka, 2 (2015)
15. Gencer, C., Guerpinar, D.: Analytic network process in supplier selection: a case study in an electronic firm. Appl. Math. Model. **31**, 2475–2486 (2007)
16. Guitouni, A., Martel, J.: Tentative guidelines to help choosing an appropriate MCDA method. EJOR **109**(2), 501–521 (1998)
17. Guneri, A.F., Kuzu, A.: Supplier selection by using a fuzzy approach in just-in-time: a case study. Int. J. Comput. Integr. Manuf. **22**(8), 774–783 (2008)
18. Ho, W., Xu, X., Dey, P.K.: Multi-criteria decision making approaches for supplier evaluation and selection: a literature review. Eur. J. Oper. Res. **202**(1), 16–24 (2010)
19. Huang, S.H., Keskar, H.: Comprehensive and configurable metrics for supplier selection. Int. J. Prod. Econ. **105**, 510–523 (2007)
20. Keeney, R., Raiffa, H.: Decisions with Multiple Objectives. Preferences and Value Tradeoffs. Cambridge University Press, Cambridge (1993)
21. Kerkhoff, G.: Zukunftschance Global Sourcing. Wiley-Vch, Weinheim (2005)
22. Koppelmann, U.: Beschaffungsmarketing. Springer, Heidelberg (2004)
23. Labib, A.W.: A supplier selection model: a comparison of fuzzy logic and the analytic hierarchy process. Int. J. Prod. Res. **49**(21), 6287–6299 (2010)
24. Monczka, R.M., Giunipero, L.C., Reck, R.F.: Perceived importance of supplier information. J. Purchasing Mater. Manage. **17**, 21–29 (1981)
25. Ordoobadi, S.M.: Development of a supplier selection model using fuzzy logic. Supply Chain Manage. Int. J. **14**(4), 314–327 (2009)
26. Piontek, J.: Internationales Beschaffungsmarketing. Schäffer-Poeschel, Stuttgart (1993)
27. Porter, M.E.: Competitive Strategy. The Free Press, New York (1980)
28. Ravindran, A.R., Bilsel Ufuk, R., Wadhwa, V., Yang, T.: Risk adjusted multicriteria supplier selection models with applications. Int. J. Prod. Res. **48**(2), 405–424 (2010)
29. Roy, B.: Decision-aid and decision making. Eur. J. Oper. Res. **45**, 324–331 (1990)
30. Roy, B.: The outranking approach and the foundations of ELECTRE methods. In: Bana e Costa, C. (ed.) Readings in Multiple Criteria Decision Aid. Springer, Heidelberg (1990)
31. Rushton, A., Croucher, P., Baker, P.: The Handbook of Logistics and Distribution Management. Kogan Page, London (2006)
32. Saaty, T.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)
33. Saaty, T.: Transport planning with multiple criteria: the analytic hierarchy process applications and progress review. J. Adv. Transp. **29**(1), 81–126 (1995)
34. Sawicka, H., Żak, J.: Ranking of distribution system's redesign scenarios using stochastic MCDM/A procedure. Procedia Soc. Behav. Sci. **111**, 186–196 (2014)
35. Shapiro, R., Heskett, J.: Logistics Strategy. West Publishing, Minnesota (1985)
36. Shil, N.C.: Customized supplier selection methodology: an application of multiple regression analysis. Int. J. **11**(2), 58–70 (2010)

37. Talluri, S., Narasimhan, R.: Vendor evaluation with performance variability: a max-min approach. Eur. J. Oper. Res. **146**(3), 543–552 (2003)
38. Vincke, P.: Multicriteria Decision-Aid. Wiley, New York (1992)
39. Weber, C.A., Current, R.J., Benton, W.C.: Vendor selection criteria and methods. Eur. J. Oper. Res. **50**(2), 2–18 (1991)
40. Webster's Encyclopedic Dictionary of the English Language. Lexicon Publications (2005)
41. Yu, C., Wong, T.N.: A supplier pre-selection model for multiple products with synergy effect. Int. J. Prod. Res. **52**(17), 5206–5222 (2014)
42. Yücenur, G.N., Vayvay, Ö., Demirel, N.Ç.: Supplier selection problem in global supply chains by AHP and ANP approaches under fuzzy environment. Int. J. Adv. Manuf. Technol. **56**, 823–833 (2011)
43. Żak, J.: Multiple Criteria Decision Aiding in Road Transportation. Poznan University of Technology Publishers, Poznan (2005)
44. Żak, J.: Application of operations research techniques to the redesign of the distribution systems. In: Dangelmaier, W., Blecken, A., Delius, R., Klöpfer, S. (eds.) IHNS 2010. LNBIP, vol. 46, pp. 57–72. Springer, Heidelberg (2010). doi:10.1007/978-3-642-12494-5_6
45. Żak, J.: Decision support systems in transportation. In: Jain, L., Lim, C. (eds.) Handbook on Decision Making – Vol. 1: Techniques and Applications. Intelligent Systems Reference Library. Springer, Heidelberg (2010)
46. Żak, J.: The methodology of multiple criteria decision making/aiding in public transportation. J. Adv. Transp. **45**(1), 1–20 (2011)
47. Żak, J.: Comparative analysis of multiple criteria evaluations of suppliers in different industries. Transp. Res. Procedia **10**, 809–819 (2015)
48. Zusammenarbeit an der Fachhochschule Kiel: Leitfaden zur Internationalisierung der Beschaffung, Seminar Einkauf und Logistik (2006)

Aircraft Type Selection Problem: Application of Different MCDM Methods

Slavica Dožić(✉) and Milica Kalić

Faculty of Transport and Traffic Engineering,
University of Belgrade, 305 Vojvode Stepe, 11000 Belgrade, Serbia
{s.dozic,m.kalic}@sf.bg.ac.rs

Abstract. In order to make a proper choice when selecting aircraft type, planners can apply some of multiple criteria decision-making (MCDM) methods as an aid to decision making. In this paper, three MCDM methods, Analytic Hierarchy Process (AHP) and Fuzzy Analytic Hierarchy Process (FAHP) as widely used methods based on hierarchy and Even Swaps Method (ESM) that belongs to group of so-called dominance methods, are applied onto the same problem, under the same conditions and illustrated with a case study of a hypothetical airline. These methods are compared, as well as solutions they arrived at. Considering the difference among the methods, a sensitivity analysis is carried out in different ways. In the AHP and FAHP, the sensitivity of alternative ratings in respect to different pairwise comparisons of the alternatives is analyzed, showing that the methods are sensitive to this kind of changing. In the even swaps method, the objective ranking across alternatives is varied, showing that the ESM is not sensitive at all.

Keywords: Aircraft type selection · AHP · FAHP · Even Swaps Method

1 Introduction

One of the main factors that affects airline success is bringing supply and demand in observed market conditions and economic environment as closely together as possible. Moreover, airlines should both make profit and keep their customers (passengers) satisfied, while costs should be as low as possible. In order to be able to accomplish their mission in the market most suitably, airlines need appropriate methodological approach for the fleet planning process, corresponding fleet selection and permanent fleet management.

When it comes to airline fleet, it should be noted that the two main fleet features are fleet structure and fleet size. Fleet structure represents the number of different aircraft types in the fleet; thus, it could be single fleet (only one type of aircraft) and multi fleet (more than one aircraft type in the fleet). In case of single fleet, maintenance costs and cost of flight crew are lower in comparison to multi fleet. On the other hand, different aircraft types in the fleet enable airlines to match demand and supply more closely, make high passengers' load factors and increase their income. Therefore, it is very important to plan the fleet structure according to airline's needs. Fleet size represents the number of aircraft by different types and the total number of aircraft in the airline's fleet, and it also

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_7

needs to be determined according to the airline's requirements. Number of aircraft larger than required means lower fleet utilization, which further induces increase of costs. Number of aircraft lower than required means spill of the demand, i.e. reallocation of passengers to competitor's flights and missing out on the opportunity to earn money. Both fleet structure and fleet size must be determined properly in order to enable the airline to realize the planned schedule and generate a profit.

In order to retain its market position, an airline need to manage its fleet, which means that it should permanently monitor the fleet and decide how many aircraft and when would be acquired (bought or leased) and retired. The existence of a large number of different aircraft types, which would have certain mission and purpose and which could operate markets that have different air travel demand and characteristics, emphasizes the complexity of aircraft type selection problem. Aircraft should be chosen to be used in the future and to meet air travel demand in given market conditions (price and competition). In order to select appropriate aircraft type, planners very often have to balance multiple, usually conflicting criteria. Interests of both the airline and passengers must be considered, as well as operational requirements. In order to provide a satisfactory choice while dealing with multiple criteria, planners can apply some of the multiple criteria decision-making (MCDM) methods as an aid to decision making. Since the last decade has been very turbulent for the global airline industry, airlines - as one of the players - are faced with the inability to respond to market and demand changes adequately. Different infectious diseases, volcanic eruptions and terrorist attacks had direct impact on air travel demand. Since the airline industry is one of the main pillars of global economy, it was directly affected by the world financial crisis which has further resulted in job reduction, losses and bankruptcies. Therefore, certain airlines could still go bankrupt and could be replaced by new ones, while others could have an opportunity to confirm their market position. For all airlines, either new ones or airlines that are well positioned in the market, fleet planning and aircraft type selection are very important and always actual problems, which have motivated the authors to research them.

This paper proposes three different MCDM methods for the aircraft type selection: Analytic Hierarchy Process (AHP), Fuzzy Analytic Hierarchy Process (FAHP) and Even swaps method (ESM). According to [13, 14], AHP and FAHP are applied to various problems, and could be considered as widely used methods (over 150 papers are cited in each paper). Considering their application in different areas and their frequencies of appearance in the relevant literature we have been found reasonable to choose them for comparison. Furthermore, the AHP and FAHP are MCDM methods based on pairwise comparisons, while the ESM is a method that belongs to a group of so-called dominance methods, which is not widely used. It was interesting to compare these three methods due to their diversities. Sometimes it is possible to use each of the techniques, but sometimes the technique is determined on the basis of the data availability. The main contribution of the paper is that it indicates methods suitability depending on the data availability. Our goal is to show that different MCDM methods can be used as an effective solution for the same problem.

The introduction and review of the relevant literature are followed by the main issues related to aircraft type selection, pointing out the criteria that should be considered. After appropriate criteria selection, three MCDM methods - Analytic Hierarchy Process (AHP), Fuzzy Analytic Hierarchy Process (FAHP) and Even swaps method (ESM) - are applied to the same problem of regional aircraft type selection and their applicability is illustrated through a hypothetical airline case study.

2 Literature Review

Many researchers consider aircraft type selection problem in different ways. Papers related to aircraft selection/evaluation problem are given in chronological order in Table 1. Multi-attribute decision making based on hypothetical equivalents and inequivalents is used to establish a fleet with one type of aircraft to serve the routes on major cities among Asia Pacific countries and the United States [17]. Three criteria are used in order to choose an aircraft. The authors have neglected to consider the most important criteria related to economic and financial issues. In the study that presents the model for aircraft selection in the case of a Saudi Arabian airline with the base in Jeddah and Madniah [9, 10], specific aircraft types are chosen for consideration based on air travel demand and aircraft performance parameters for given route network. Cost efficiency is calculated using Excel application, and results are low seat mile cost and low trip cost per sector, which could help airline planner to choose the right aircraft. A systematic evaluation model is proposed for selection of an optimal training aircraft for Air Force Academy, mainly from the perspective of pilot drillmasters and trainees [19]. This is the single paper which threats military aviation. New fuzzy group MCDM approach is proposed for aircraft selection problem faced by Taiwan's domestic airline for its major routes [23]. The Analytic Network Process (ANP) is suggested to help choosing a middle range aircraft for Turkish Airlines [15]. Costs, time and physical attributes and others are considered as the main criteria. The three groups of criteria (financial, logistics and quality) in the multi-criteria decision aiding method named NAIADE (Novel Approach to Imprecise Assessment and Decision Environments) are proposed for the aircraft selection problem in regional charter flights in Brazil [7]. Four criteria and eight sub-criteria are selected to support aircraft evaluation using hybrid approach based on AHP and fuzzy set theory in final ranking [3]. As aircraft type selection is recognized as multi-criteria decision making, the AHP and ESM are applied [4, 6], and further compared [5] to research their solution sensitivity to different changes. The robust three-stage model developed in [6] involves approximate fleet composition (number of aircraft types), fleet sizing (number of aircraft per each type) and aircraft type selection (specific types of aircraft) based on fuzzy logic, heuristic and analytic approaches, and multi-criteria decision making, respectively.

Considering cited literature and Table 1 it can be seen that most of researchers employ different MCDM approaches. Ranking of alternatives is usually offered as result [3, 4, 5, 7, 15, 19, 23].

Table 1. The problem of aircraft type selection in the literature

Authors (year)	Approach	Parameters	Technique	Output
See et al. (2004) [17]	Multi-attribute methodology	Speed, range, number of passengers (pax)	Method of the hypothetical equivalents and inequivalents	Single robust optimal alternative (aircraft – a/c)
Harasani, W.I. (2006) [9], Harasani, W.I. (2008) [10]	Five step approach based on number data analysis	Number of pax per route, a/c performance parameters, cost efficiency (DOC) including fuel cost, maintenance cost, annual insurance rate, annual salaries paid, traffic allocation and scheduling	No specific technique	Opt. efficiency a/c - low seat mile cost and low trip cost per sector
Wang, T-C, Chang, T-H. (2007) [19]	MCDM approach/group decision	A/c performance parameters (power plant, stalling speed when flameout, max operating speed, max G limits and fuel capacity)	TOPSIS with a fuzzy environment	Ranking of the seven military aircraft
Yeh, C-H, Chang, Y-H. (2009) [23]	Fuzzy group MCDM approach	Technological advance (maintenance, pilot adaptability, a/c reliability, max range), social responsibility (passenger preference, noise level), economical efficiency (operational productivity, airline fleet economy of scale, purchasing price, DOC, consistency with corporate strategy)	New fuzzy MCDM algorithm	Ranking of the five aircraft
Ozdemir et al. (2011) [15]	MCDM approach	Cost (purchasing, operation and spare, maintenance, salvage), time (delivery time, useful life), physical attributes and others (dimensions, security, reliability, suitability for service quality)	Analytic Network Process (ANP)	Ranking of the three aircraft

(*continued*)

Table 1. (*continued*)

Authors (year)	Approach	Parameters	Technique	Output
Gomes et al. (2012) [7]	Fuzzy stochastic approach	Financial (acquisition cost, liquidity, operating cost), logistics (range, flexibility, cruising speed, replacement parts availability, landing and take-off distance), quality (comfort, avionics availability, safety)	NAIADE method	Ranking of the eight aircraft
Dožić, S, Kalić, M. (2014) [4]	MCDM approach	Price of aircraft, payment conditions, total cost per available seat miles (CASM), seat capacity, total baggage, MTOM	AHP	Ranking of the seven aircraft
Dožić, S, Kalić, M. (2015) [6]	Three stage approach with MCDM	Air travel demand, distance; airline schedule; aircraft seat capacity, price of aircraft, luggage per passenger, MTOM, unit trip costs	Fuzzy logic, heuristics and Even swaps method	Fleet structure, fleet size, the most appropriate aircraft
Dožić, S, Kalić, M. (2015) [5]	MCDM approach	Price of aircraft, payment conditions, total cost per available seat miles (CASM), seat capacity, total baggage, MTOM	AHP and Even swaps method	Ranking of a/c and single, the most appropriate a/c
Bruno et al. (2015) [3]	Hybrid approach with MCDM	Economic performance (operative costs/ (range*seats), aircraft price), technical performance (speed, autonomy), aircraft interior quality (seat comfort, cabin luggage compartment size), environmental impact (environmental pollution, noise)	AHP and Fuzzy Set Theory	Ranking of the three aircraft

It is interesting to note that fuzzy set theory is used in three papers [3, 6, 19], but in different ways and in different stages of modelling. In the given papers a variety of criteria/sub-criteria are used, as well. These criteria generally could be recognized as airline, passengers or environmental oriented. In this research, we keep the criteria from previous work [4, 5] that are the most relevant from the airline's perspective. Considering the literature mentioned above, we find out that none of the paper compares three MCDM methods, therefore, we decide to investigate this topic, and contribute by giving comparison based on the problem of aircraft type selection. Hence, AHP and FAHP which have found significant and successful applications in different fields [13, 14] are employed to help airline planners when choosing appropriate aircraft type. In order to learn the advantages and disadvantages of different MCDM methods and possibilities of their application depending on the data that are available, we compare AHP, FAHP and ESM. Considering the difference among the MCDM methods, the sensitivity analysis is carried out in different ways. In the AHP and FAHP, the sensitivity of alternative ratings with respect to different pairwise/fuzzy pairwise comparisons of the alternatives is analyzed, while in the ESM the objective ranking across alternatives is varied in order to learn solution sensitivity.

3 Main Issues in Aircraft Evaluation Process

Whereas profitability of an airline depends on the aircraft type selection, it is necessary to make right decision in the right time following a defined procedure that involves all relevant aspects and factors. Hence, aircraft's purpose and mission in the market (cargo, training or commercial passenger aircraft, for charter, low cost, regional or full service airline), should be examined and the aircraft should be evaluated accordingly. Thus, different issues that may affect aircraft type selection problem can be identified. According to [22], aircraft evaluation process includes five areas: consideration of design characteristics, physical performance, maintenance needs, acquisition costs, and operating economics.

Consideration of design characteristics includes aircraft's dimensions, weights, fuel capacity, type of power plant, systems, seating configuration, containers and pallets, bulk volume, and total volume. These characteristics may predetermine the set of airports that could be operated by a specific aircraft type due to the physical characteristics that an airport has (runway length, runway capacity, apron capacity, etc.) and characteristics that the aircraft needs to perform operations safely. Maximum Take-Off Mass (MTOM) of an aircraft is the maximum mass at which the take-off is allowed, due to structural or other limits. At the same time, it is the main unit for calculation of airport and navigation fees. Airport charges dependent on MTOM at Belgrade Airport Nikola Tesla are calculated based on price list in 2016, for different regional aircraft (Fig. 1). It is very important to emphasize that different fees based on MTOM are calculated so that each started ton is calculated as a whole ton.

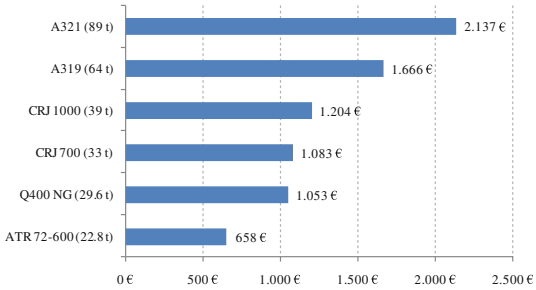


Fig. 1. Airport charges dependant of MTOM for the summer of 2016, Belgrade Airport

Total baggage (including overhead bins) and baggage per passenger are indicators that show how much space is available for cargo, which is not a negligible source of revenue for an airline. Depending on route types and categories of passengers that prevail on the route (business or leisure passengers) different volume of space is required and evaluated differently.

Physical performance includes payload-range diagrams, take-off and landing data, different speeds, runway requirements and noise performance, which could be limiting factors for operations at a particular airport as well. Aircraft range defines the maximum distance between take-off and landing which could be flown without refueling. Thus, depending on the aircraft range, the set of corresponding airports that could be served is defined (Fig. 2).

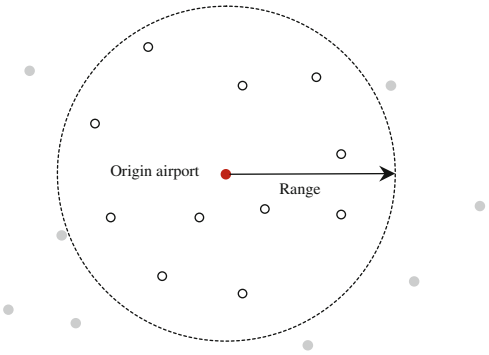


Fig. 2. Aircraft range and airport servicing area

Maintenance needs include spare parts availability, fleet commonality, product support, technical record keeping, and training support. According to [11], fleet commonality represents a concept introduced in order to enable airlines to minimize the number of aircraft types in the fleet and to adjust the fleet to route network at the same time. This concept brings advantages with regard to fleet flexibility in training and

rostering of aircrew and maintenance personnel (reduction of training costs and number of pilots), and also flexibility in spares and other equipment (single pool of spares and equipment for the aircraft family). Fleet commonality influences flight operations (lower number of reserve crew and flexible swaps), maintenance (less spare parts, lower costs, lower labour costs), aircraft servicing (standardized ground handling equipment, more cost-efficient), and aircraft capital (lower price for ordering several same type aircraft) [2].

Acquisition costs include the cost of aircraft with spare parts, ground equipment needed, maintenance and flight training and costs of money itself. Since airline costs can be divided into different cost categories, the costs relevant for fleet planning and aircraft selection could be categorized as acquisition costs (capital costs) and direct operating costs [11]. Capital costs, as shown in Fig. 3, include total investment needed to put an aircraft into operation. Therefore, these costs include aircraft price as well as any other costs and conditions related to the first appearance of the aircraft in the service. Bearing in mind that acquisition of a new aircraft requires a huge investment (aircraft produced by Airbus[1] 74–428, aircraft produced by Boeing[2] 80–400, and regional aircraft cost USD 20–50 millions), it is evident that a small savings of a few percent is not negligible for an airline. Although costs of acquiring an aircraft represent large capital expenditures for an airline, these costs are very often lowered by the appearance of used and attractive aircraft leasing options. Direct operational costs are dependent on aircraft type, distance and block time, therefore they affect aircraft type selection.

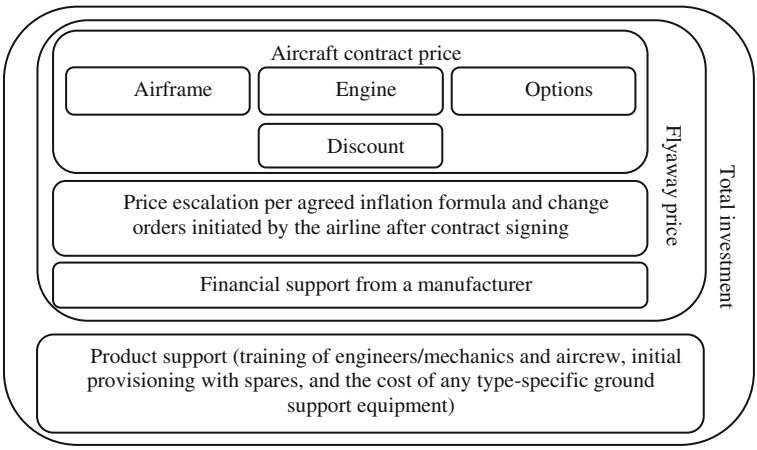


Fig. 3. Total investment for specific aircraft configuration

[1] http://www.airbus.com/presscentre/pressreleases/press-release-detail/detail/new-airbus-aircraft-list-prices-for-2015/, May 2016.

[2] http://www.boeing.com/company/about-bca/#/prices, May 2016.

Maintenance costs are an important factor in the cost structure. Aircraft should be permanently maintained in order to keep it safe for operations. Aircraft maintenance and monitoring of maintenance costs could be drivers for decisions related to aircraft replacement or retirement.

In order to compare operating costs from one airline to another, unit costs per available seat miles (CASM) are used. CASM are the cost of flying one aircraft seat for one mile [18] and may be calculated for a variety of costs, such as operating costs, total operating costs, or any other cost combination. Moreover, CASM may be calculated for an airline, or for an aircraft type operating on a specific distance. For example, total trip costs per available seat miles for regional aircraft vary from 13.5 c/ASM to 17 c/ASM [1].

Operating economics includes potential aircraft's contribution to the company's profitability (related to network structure, traffic flow and composition, existing traffic volumes, potential future growth, seating density, load factors and utilization). When it comes to operating economics, specific relations should be mentioned. Aircraft seat capacity (the number of seats offered by an airline to passengers on a certain flight represents a measure of potential income) is closely connected to costs. When average unit costs decrease with an increase in the quantity being produced, it can be said that economy of scale is achieved. Airlines offer capacity in line with the expected demand. The smaller the gap between capacity and expected demand is, the greater the load factor is. Greater load factor combined with appropriate revenue management and pricing policy offers airlines an opportunity to have a successful business. Therefore, it could be said that greater load factor is closely connected to greater profitability, as well. Finally, large capacity can be offered only if it is accompanied by high load factor. Airlines aim to reach as high load factor as possible, therefore it is very important to match offered capacity and rising demand.

Considering the above-mentioned issues that affect aircraft evaluation, one should choose criteria for aircraft type selection to reflect airlines' as well as some of passengers' perspectives. The most commonly considered criteria are related to technical and operational characteristics of the aircraft type, which could limit aircraft operations; hence it is reasonable to consider them as mandatory.

In the next section, we will describe different MCDM methods that will be applied for the same problem, in the case of a hypothetical airline.

4 Methods and Data

4.1 Analytic Hierarchy Process

Analytic Hierarchy Process (AHP) is an MCDM approach which divides the problem into a hierarchy of issues which should be considered [16]. It uses both quantitative and qualitative data translated into numbers and presents a theory of measurement through pairwise comparisons made using a scale of absolute judgments that represents the domination measure of one element over another with respect to a given attribute. In order to compare alternatives and criteria, a fundamental scale which indicates the intensity of importance on an absolute scale is introduced [16]. The scale consists of verbal judgments of preference ranging from equal to extreme (equal, moderate, strong,

very strong, extreme importance) with corresponding numerical judgments (1, 3, 5, 7, 9), and intermediate values between the two judgments, as well. Numerical judgments in the pairwise comparison matrix satisfy the reciprocal property: if an activity i has one of the above nonzero numbers assigned to it when compared with an activity j, then j has the reciprocal value when compared with i ($a_{ji} = 1/a_{ij}$). Pairwise comparison matrices for criteria and alternatives enable computing of local and global priorities as well as ranking of alternatives. Priorities from pairwise comparisons are calculated using eigenvector method.

4.2 Fuzzy Analytic Hierarchy Process

Criteria importance over each other cannot always be precisely express, and decision makers usually have to deal with uncertainty. Therefore we found a possibility to use fuzzy numbers in the AHP method, i.e. to apply Fuzzy Analytic Hierarchy Process (FAHP). Although there are different ways to derive priorities from fuzzy pairwise comparison matrix, one of the newer methodologies, the logarithmic fuzzy preference programming (LFPP) method [21], is applied in this paper. Table 2 shows the conversion of linguistic scale into triangular fuzzy scale.

Table 2. Triangular fuzzy conversion scale for pair of elements i and j

Linguistic scale	Triangular fuzzy scale	Triangular fuzzy reciprocal scale
Just equal	(1, 1, 1)	(1, 1, 1)
Equally important	(1/2, 1, 3/2)	(2/3, 1, 2)
Weakly important	(1, 3/2, 2)	(1/2, 2/3, 1)
Strongly important	(3/2, 2, 5/2)	(2/5, 1/2, 2/3)
Very strongly important	(2, 5/2, 3)	(1/3, 2/5, 1/2)
Extremely preferred	(5/2, 3, 7/2)	(2/7, 1/3, 2/5)

It should be mentioned that there are different triangular fuzzy scales in the relevant literature [12], and we decide to use scale according to [20]. The pairwise comparison matrix is filled out with fuzzy judgments instead of precise judgments. Fuzzy judgments reflect the vagueness and imprecision of human thought related to the problem considered. When comparing two criteria, i and j, the exact numerical ratio a_{ij} can be approximated with a fuzzy ratio "about a_{ij}", which is represented by the fuzzy number $\tilde{a}_{ij}$. It means that criterion i is between l_{ij} and u_{ij} times as important as criterion j with m_{ij} being the most likely times. Therefore, the fuzzy pairwise comparison matrix is given by (1):

$$\tilde{A} = (\tilde{a}_{ij})_{n\times n} = \begin{bmatrix} (1,1,1) & (l_{12},m_{12},u_{12}) & \cdots & (l_{1n},m_{1n},u_{1n}) \\ (l_{21},m_{21},u_{21}) & (1,1,1) & \cdots & (l_{2n},m_{2n},u_{2n}) \\ \vdots & \vdots & \vdots & \vdots \\ (l_{n1},m_{n1},u_{n1}) & (l_{n2},m_{n2},u_{n2}) & \cdots & (1,1,1) \end{bmatrix} \quad (1)$$

The elements of the matrix $\tilde{A}$ satisfy the reciprocal property, which means that $l_{ij} = 1/u_{ji}$, $m_{ij} = 1/m_{ji}$, $u_{ij} = 1/l_{ji}$ and $0 < l_{ij} \leq m_{ij} \leq u_{ij}$ for all $i, j = 1,\ldots, n$; $j \neq i$. Once the matrix is built, crisp priority vector using the LFPP method proposed by [21] can be computed.

4.3 Even Swap Method

Even swaps method (ESM) is an MCDM method which provides a reliable mechanism for making trade-offs [8]. It is based on the fundamental principle of decision making: if all alternatives are rated equally for a given objective, then that objective can be eliminated when choosing among the alternatives considered. The first step in this method is to create a consequences table in such a way that all defined criteria are listed down the left side, and possible alternatives along the top. The ESM enables description of criteria both in quantitative (by numbers) and qualitative terms (by words) in consequences tables. The next step is to create a ranking table where each criterion is ranked across the alternatives. The ranking table enables one to compare an alternative by all criteria and to find the alternative which is worse on some objectives and not better on all other objectives in comparison with the other alternatives - the dominated alternative. It can be eliminated from further consideration, which reduces the number of alternatives in the set of alternatives. Practical dominance (an alternative is worse or equal on some criteria and better in only one criterion) could also results in decrease of the number of alternatives, again, if it is possible. The alternative practically dominated can be eliminated as well, if the worse criterion is not of crucial importance, in the decision makers' opinion. When there are no more dominated or practically dominated alternatives, swaps can be made. ESM provides an opportunity to decrease the value of one of the criteria, while another one must increase by the equivalent value. In this way, different criteria could be adjusted in order to make its value equal across all alternatives, and finally, to eliminate it from further consideration. Progressive simplification of the problem is made using the dominance or practical dominance to eliminate alternatives and using trade-offs to equalize performances on a selected criterion allowing the elimination of that criterion. The objectives and alternatives are eliminated until one alternative dominates all others, or only one criterion for comparison of alternative remains.

4.4 Data

AHP, FAHP and ESM are used to choose the most suitable aircraft type from the set of aircraft considering selected criteria. These methodologies are applied to the case study of a hypothetical airline presented in previous researches [4–6]. In this paper, the focus is on the set of routes covered by small aircraft with capacity up to 100 seats [4, 6]. The determined set of aircraft consists of regional jets ERJ190, CRJ700, CRJ900 and CRJ1000, as well as turboprops ATR72-500, ATR72-600 and Q400NG [4–6].

Considering the main issues in the aircraft evaluation process, mentioned in the previous section, and bearing in mind that aircraft will operate short, regional routes, it

has been found reasonable to use the same criteria as in earlier authors' researches. Therefore, we considered aircraft seat capacity, aircraft price, total baggage, MTOM, payment conditions (advantages offered by manufacturers or leasing companies) and total cost per available seat miles – CASM [4–6]. All aircraft from the selected set are ranked with regard to the chosen criteria. The most suitable aircraft is the one the capacity of which best meets the estimated number of passengers per flight. Whereas airlines endeavor to lower their costs, it is expected that the lowest price of an aircraft and acceptable payment conditions are the most acceptable, while the highest price is not desirable. The more baggage per passenger is available, the more suitable the aircraft type is. In terms of MTOM, it is evident that the airline prefers lighter aircraft to heavier ones. Lower unit costs expressed by cents per available seat mile (ASM) make aircraft type more preferable for an airline.

Table 3. Pairwise comparison matrix for the first level and domination measures

	Seat capacity	Price	Total baggage	MTOW	Payment conditions	CASM
Seat capacity	1	0.25	3	0.5	0.25	0.25
Price	4	1	5	5	1	1
Total baggage	0.333	0.2	1	0.5	0.2	0.2
MTOW	2	0.2	2	1	0.2	0.2
Payment conditions	4	1	5	5	1	1
CASM	4	1	5	5	1	1

Domination measure of one aircraft over another with respect to seat capacity

	ATR72-500	ATR72-600	ERJ 190	Q400 NG	CRJ 700	CRJ 900	CRJ 1000
ATR72-500	1	1	4	2	1	3	4
ATR72-600	1	1	4	2	1	3	4
ERJ190	0.25	0.25	1	0.333	0.25	0.5	1
Q400NG	0.5	0.5	3	1	0.5	2	3
CRJ700	1	1	4	2	1	3	4
CRJ900	0.333	0.333	2	0.5	0.333	1	2
CRJ1000	0.25	0.25	1	0.333	0.25	0.5	1

Domination measure of one aircraft over another with respect to payment conditions

	ATR72-500	ATR72-600	ERJ 190	Q400 NG	CRJ 700	CRJ 900	CRJ 1000
ATR72-500	1	1	4	2	3	3	3
ATR72-600	1	1	4	2	3	3	3
ERJ190	0.25	0.25	1	0.333	0.5	0.5	0.5
Q400NG	0.5	0.5	3	1	0.5	0.5	0.5
CRJ700	0.333	0.333	2	2	1	1	1
CRJ900	0.333	0.333	2	2	1	1	1
CRJ1000	0.333	0.333	2	2	1	1	1

Tables 3 and 4 present the input data for the first hierarchy level and importance of alternatives with respect to seat capacity and payment conditions, for AHP and FAHP, respectively. Importance of alternatives with respect to other criteria is not presented due to the space limitation. Pairwise and fuzzy pairwise comparison matrices are created according to the authors' expert knowledge and experience as well as to the verbal (informal) conversation with the people from the airline industry. The fuzzy judgments (Table 2) are used in accordance with numerical judgments used in the AHP, as much as possible. The consistency ratio of matrices is checked, therefore all matrices are consistent.

Table 5 refers to data needed for ESM. In the first part of the Table 5 one can see real numerical data (taken from the manufacturers' official web sites). On the other hand, linguistic data related to payment conditions are assumed, due to unavailability of this kind of data.

Table 4. Fuzzy pairwise comparison matrix for the first level and domination measures

	Seat capacity	Price	Total baggage	MTOW	Payment conditions	CASM
Seat capacity	(1,1,1)	(1/3,2/5,1/2)	(1,3/2,2)	(1/2,2/3,1)	(1/3,2/5,1/2)	(1/3,2/5,1/2)
Price	(2,5/2,3)	(1,1,1)	(2,5/2,3)	(2,5/2,3)	(1/2,1,3/2)	(1/2,1,3/2)
Total baggage	(1/2,2/3,1)	(1/3,2/5,1/2)	(1,1,1)	(2/3,1,2)	(1/3,2/5,1/2)	(1/3,2/5,1/2)
MTOW	(1,3/2,2)	(1/3,2/5,1/2)	(1/2,1,3/2)	(1,1,1)	(1/3,2/5,1/2)	(1/3,2/5,1/2)
Payment conditions	(2,5/2,3)	(2/3,1,2)	(2,5/2,3)	(2,5/2,3)	(1,1,1)	(1/2,1,3/2)
CASM	(2,5/2,3)	(2/3,1,2)	(2,5/2,3)	(2,5/2,3)	(2/3,1,2)	(1,1,1)

Domination measure of one aircraft over another with respect to seat capacity

Seat capacity	ATR 72-500	ATR 72-600	ERJ190	Q400NG	CRJ700	CRJ900	CRJ1000
ATR 72-500	(1,1,1)	(1/2,1,3/2)	(2,5/2,3)	(1,3/2,2)	(1/2,1,3/2)	(3/2,2,5/2)	(2,5/2,3)
ATR 72-600	(2/3,1,2)	(1,1,1)	(2,5/2,3)	(1,3/2,2)	(1/2,1,3/2)	(3/2,2,5/2)	(2,5/2,3)
ERJ190	(1/3,2/5,1/2)	(1/3,2/5,1/2)	(1,1,1)	(2/5,1/2,2/3)	(1/3,2/5,1/2)	(1/2,2/3,1)	(1/2,1,3/2)
Q400NG	(1/2,2/3,1)	(1/2,2/3,1)	(3/2,2,5/2)	(1,1,1)	(1/2,2/3,1)	(1,3/2,2)	(3/2,2,5/2)
CRJ700	(2/3,1,2)	(2/3,1,2)	(2,5/2,3)	(1,3/2,2)	(1,1,1)	(3/2,2,5/2)	(2,5/2,3)
CRJ900	(2/5,1/2,2/3)	(2/5,1/2,2/3)	(1,3/2,2)	(1/2,2/3,1)	(2/5,1/2,2/3)	(1,1,1)	(1,3/2,2)
CRJ1000	(1/3,2/5,1/2)	(1/3,2/5,1/2)	(2/3,1,2)	(2/5,1/2,2/3)	(1/3,2/5,1/2)	(1/2,2/3,1)	(1,1,1)

Domination measure of one aircraft over another with respect to payment conditions

Payment conditions	ATR 72-500	ATR 72-600	ERJ190	Q400NG	CRJ700	CRJ900	CRJ1000
ATR 72-500	(1,1,1)	(1/2,1,3/2)	(2,5/2,3)	(1,3/2,2)	(3/2,2,5/2)	(3/2,2,5/2)	(3/2,2,5/2)
ATR 72-600	(2/3,1,2)	(1,1,1)	(2,5/2,3)	(1,3/2,2)	(3/2,2,5/2)	(3/2,2,5/2)	(3/2,2,5/2)
ERJ190	(1/3,2/5,1/2)	(1/3,2/5,1/2)	(1,1,1)	(2/5,1/2,2/3)	(1/2,2/3,1)	(1/2,2/3,1)	(1/2,2/3,1)
Q400NG	(1/2,2/3,1)	(1/2,2/3,1)	(3/2,2,5/2)	(1,1,1)	(1/2,2/3,1)	(1/2,2/3,1)	(1/2,2/3,1)
CRJ700	(2/5,1/2,2/3)	(2/5,1/2,2/3)	(1,3/2,2)	(1,3/2,2)	(1,1,1)	(1/2,1,3/2)	(1/2,1,3/2)
CRJ900	(2/5,1/2,2/3)	(2/5,1/2,2/3)	(1,3/2,2)	(1,3/2,2)	(2/3,1,2)	(1,1,1)	(1/2,1,3/2)
CRJ1000	(2/5,1/2,2/3)	(2/5,1/2,2/3)	(1,3/2,2)	(1,3/2,2)	(2/3,1,2)	(2/3,1,2)	(1,1,1)

Table 5. Consequences and ranking table for aircraft type choice

Consequences	ATR72-500	ATR72-600	E190	Q400NG	CRJ700	CRJ900	CRJ1000
Seat capacity	68	70	98	74	70	88	100
Price (mil. USD)	21.9	22.7	43	31.67	37	44.5	49.5
Total baggage (m^3)	13.75	15.13	32	17.9	18.3	20.32	23.6
MTOW (t)	22.5	22.8	47.8	29.6	33	36.5	39
Payment conditions	Excellent	Excellent	Poor	Very Good	Excellent	Good	Poor
CASM	16.4	16.4	15.4	15.6	>16.4	14.5	14.5
Ranking	ATR72-500	ATR72-600	E190	Q400NG	CRJ700	CRJ900	CRJ1000
Seat capacity	1	1	4	2	1	3	4
Price (mil. USD)	1	1	3	2	2	3	4
Total baggage (m^3)	6	5	1	4	4	3	2
MTOW (t)	1	1	6	2	3	4	5
Payment conditions	1	1	4	2	3	3	3
CASM	4	4	2	3	5	1	1

Payment conditions are subject to negotiation and can be real and available only when negotiation for ordering of aircraft is in progress. In the second part of Table 5, all alternatives are ranked throughout the six criteria. Ranking table enables decision maker to identify dominated/practically dominated alternatives, make trade-offs and progressively reduce the size of problem in order to reach the final solution.

The next section presents our results obtained by applying EMS, AHP and FAHP to the aircraft type selection problem.

5 Results and Discussion

5.1 Analytic Hierarchy Process

AHP is used [4, 5] to support selection of appropriate aircraft type from the set of seven regional aircraft, considering six criteria stated before. In this subsection we will summarize results from [4, 5]. Thus, the criteria are described by numerical, quantitative values, with the exception of payment conditions, which are defined qualitatively. ATR72-600 is chosen as the most appropriate aircraft in [4, 5]. AHP also gives as a result the final ranking of aircraft types (ATR72-600, ATR72-500, CRJ900, CRJ1000, Q400NG, CRJ700 and ERJ190, Table 6), which is used as the initial solution for sensitivity analysis.

Sensitivity of the solution (rank of alternatives) and consistency ratio in respect to different judgments in comparison matrices for the second level are analyzed in [5]. The sensitivity analysis with respect to the following four criteria was not conducted because CASM varies with the change of sector length (average sector length for the hypothetical airline is 200 miles, and CASM data are related to this sector length), MTOM and total baggage are characteristics of the aircraft which cannot be changed, and finally, lower price is always more desired than the higher one. Therefore, the numerical judgment of payment conditions and aircraft capacity is varied throughout

Table 6. Local and global priority weights

	Seat capacity (0.071)	Price (0.271)	Total baggage (0.043)	MTOW (0.075)	Payment conditions (0.271)	CASM (0.271)	Final priority vector
ATR72-500	0.227	0.250	0.037	0.278	0.263	0.065	0.1947
ATR72-600	0.227	0.250	0.056	0.278	0.263	0.065	0.1954
ERJ190	0.051	0.082	0.347	0.033	0.052	0.172	0.1037
Q400NG	0.134	0.144	0.090	0.176	0.093	0.107	0.1197
CRJ700	0.227	0.144	0.090	0.114	0.110	0.042	0.1082
CRJ900	0.083	0.082	0.148	0.073	0.110	0.274	0.1437
CRJ1000	0.051	0.050	0.232	0.048	0.110	0.274	0.1346

all alternatives because they are subject to changes. For example, for an airline it is possible to negotiate payment conditions with the manufacturer or a leasing company, while the value of aircraft capacity criteria could be more or less close to the demand.

In the initial experiment, excellent payment conditions are offered for the ATRs, very good for Q400NG, good for CRJs and poor payment conditions are offered for ERJ190 [5]. Considering four different categories of payment conditions (excellent, very good, good and poor) as 4 permutations of 4 elements, 24 experiments in total (4!) were carried out for this criterion (the initial one and additional 23 experiments). The 24 experiments carried out show that priority of an alternative in the final priority vector decreases with the decrease of the domination measure of one aircraft over another with respect to payment conditions, for all aircraft types.

Four different aircraft capacities were assumed in the initial experiment [5]. The capacity of ATRs and CRJ700 matches the estimated passenger number best. These aircraft capacities are followed by Q400NG, then CRJ900, and finally CRJ1000 and ERJ900 which do not satisfy the airline needs appropriately. Regarding the aircraft capacity criterion, it is possible to carry out (theoretically) 24 experiments in total, but only the initial experiment and additional 8 experiments are reasonable [5]. Therefore, 9 selected experiments were considered in sensitivity analysis for the aircraft capacity criterion, by changing the judgments for aircraft capacity, while other numerical judgments were kept constant. It is shown in [5] that the increase in demand that requires larger aircraft capacity caused the changes in the final ranking of aircraft. The initial ranking of alternatives is changed, thus the last two aircraft exchange their positions. The most inappropriate aircraft is CRJ700 instead of ERJ190, while the first five aircraft preserved their positions. When the increase in demand requires aircraft of the largest capacity, final ranking of aircraft is as follows: ATR72-500, ATR72-600, CRJ1000, CRJ900, ERJ190, Q400NG and CRJ700.

5.2 Fuzzy Analytic Hierarchy Process

The criteria importance over each other cannot always be precisely expressed, and airline planners in charge of fleet planning usually have to deal with uncertainty. Whereas the FAHP method has found significant and successful applications in

different fields [13], the authors decided to employ it for appropriate aircraft type selection. To obtain the initial solution, fuzzy pairwise comparison matrices are created. LFPP method is employed to derive priorities from fuzzy pairwise comparison matrix. We used the Microsoft Excel Solver in order to obtain optimal solution, and measure inconsistency. Whereas the consistency condition is satisfied, the ranking in the initial solution is as follows: ATR72-600, ATR72-500, Q400NG, CRJ900, CRJ1000, CRJ700 and ERJ190 (Table 7).

Table 7. Local and global LFPP priorities

	Seat capacity (0.0960)	Price (0.2368)	Total baggage (0.0839)	MTOM (0.1099)	Payment conditions (0.2368)	CASM (0.2368)	Global LFPP priorities
ATR72-500	0.1921	0.2045	0.0830	0.2129	0.2126	0.1035	0.1720
ATR72-600	0.1921	0.2045	0.0830	0.2471	0.2126	0.1035	0.1758
ERJ190	0.0740	0.0947	0.2602	0.0750	0.0819	0.1482	0.1140
Q400NG	0.1394	0.1567	0.1109	0.1663	0.1542	0.1482	0.1497
CRJ700	0.2219	0.1567	0.1109	0.1138	0.1129	0.0722	0.1240
CRJ900	0.1014	0.0947	0.1990	0.0924	0.1129	0.2122	0.1360
CRJ1000	0.0791	0.0883	0.1530	0.0924	0.1129	0.2122	0.1285

The sensitivity analysis is performed in the same way as in the case of the AHP, for the two criteria. The results are again obtained by using Microsoft Excel Solver, and they are as follows. If aircraft ATR72-500 and ATR72-600 are the aircraft with the most acceptable payment conditions in comparison to other aircraft types, then the ATR72-600 is the best ranked aircraft irrespective of payment conditions for other aircraft (Fig. 4, experiments 0–5). If the most acceptable payment conditions are offered for Q400NG, it will be the best ranked aircraft no matter what payment conditions are offered for the other aircraft (Fig. 4, experiments 6–11).

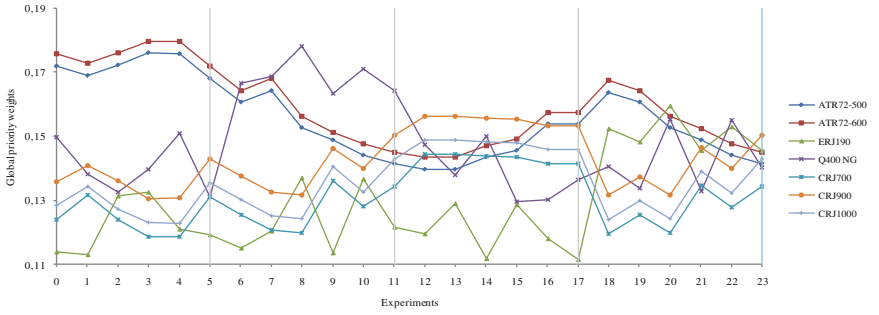


Fig. 4. Aircraft rankings with respect to payment conditions changes

If Bombaridier's CRJs (CRJ700, CRJ900 and CRJ1000) have the highest value of fuzzy pairwise judgment for payment conditions, in the final aircraft ranking the aircraft CRJ900 will be the best ranked aircraft in four experiments (12, 13, 14, and 15).

The most suitable aircraft in the experiments 16 and 17 will be ATR72-600 (Fig. 4, experiments 12–17), in the case when ATRs are in the second place with respect to payment conditions. If aircraft ERJ190 is the aircraft with the best payment conditions, in the final ranking this aircraft will be top-ranking only in the experiment 20, in which ERJ 190 is followed by Q400NG, ATRs and CRJs, respectively, according to fuzzy pairwise judgment with respect to payment conditions (Fig. 4, experiments 18–23). Aircraft ATR72-600 has the best ranking in the experiments 18, 19 and 21, while the top-ranked aircraft in the experiments 22 and 23 are Q400NG and CRJ900 respectively. The three aircraft ATR72-500, CRJ700 and CRJ1000, are not sensitive to changes of fuzzy judgment for single property payment conditions, thus these aircraft are never the final choice. Fuzzy judgment for payment conditions, will affect the final ranking of aircraft, which is expected because payment conditions are one of the most influential criteria (they have the highest priority in the priority vector for the first level, [4, 5]). It can be observed that the aircraft ATR72-600 is the most acceptable aircraft in most experiments (in 11 of 24 experiments).

Regarding the changes in the aircraft capacity, while other fuzzy judgments are kept constant, the initial experiment (denoted by 0) and 8 additional, possible experiments were carried out (Fig. 5). It can be observed that the increase in demand that requires larger aircraft capacity caused the changes in the final ranking of aircraft. The initial ranking of alternatives is changed, thus the first two aircraft are always the first ones, while the other aircraft exchange their positions. The most inappropriate aircraft are ERJ190 and CRJ700 (Fig. 5) which are always the last two.

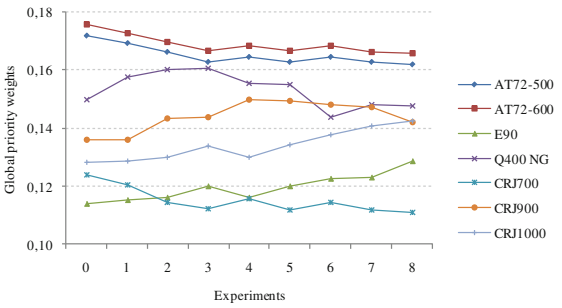


Fig. 5. Aircraft rankings with respect to aircraft capacity changes

5.3 Even Swap Method

Unlike the two previously applied methods, ESM does not rank alternatives, but rather only offers the most appropriate solution. In the case with our example, as it is shown in [5], the aircraft type ATR72-600 appears to be the most appropriate one. The sensitivity analysis performed corresponds to the sensitivity analysis for the second hierarchy level in the case of AHP and FAHP methods. As in AHP, initial experiment was carried out as well as 23 additional experiments, showing that the final solution is not sensitive to changes of payment conditions [5]. The most suitable aircraft type in all experiments

was always the same – ATR72-600. Speaking of solution sensitivity when the demand is changing, it is concluded that the ESM is not sensitive to this kind of changes either, and the choice is always ATR72-600 [5].

5.4 Discussion

Considering the chosen methods we show that all three methods can be successfully applied to the problem of aircraft type selection suggesting the same aircraft, ATR72-600, as the most suitable solution in the initial experiments. Considering the three presented MCDM methodologies and sensitivity analysis presented, it can be concluded that they differ in the data they require. Moreover ESM is suitable for cases when decision makers have complete data and can rank all alternatives across selected criteria. When decision makers are not able to make rankings and when only the pairwise comparison is available, AHP befits better. Finally, when decision makers do not have precise pairwise comparisons, FAHP is the appropriate method. Thus, ESM requires the most specific data, AHP needs pairwise comparison of the data, while FAHP can use imprecise pairwise comparison.

With regards to sensitivity analysis presented in this section, we demonstrate that solutions offered by the AHP and FAHP are sensitive to the changes of payment conditions. The solution's sensitivity to changes in demand can be observed, but only in aircraft ranking, not for the top ranked aircraft. The ESM is not sensitive at all to any kind of changes, and always gives the same solution. However, this conclusion that regards sensitivity of the solution refers to presented example and maybe would be changed in different one. It could depend on the airline policy (goals and priorities that could influence alternatives and criteria selection) and people involved in the process of decision making.

It should be mentioned that MCDM includes a certain level of subjectivity and decisions could depend on decision maker experience. In that sense, trade-offs which are made in the ESM, as well as pairwise comparison matrices in AHP and FAHP could be influenced by experience of decision maker.

6 Conclusion

This paper applies the three MCDM methodologies - AHP, FAHP, and ESM, to the same problem under the same conditions with the identical sets of alternatives and criteria, offering the same solutions (chosen aircraft type). The paper also presents a sensitivity analysis, underlining the rank of alternatives' sensitivity to the changing importance of aircraft types with respect to different criteria. The initial experiment and additional 23 experiments were carried out applying AHP, FAHP and ESM, by changing the criterion of payment conditions. It is shown that the AHP and FAHP are sensitive to this kind of change, while the ESM is not sensitive at all. The solutions obtained by AHP, show that the final priority weight for specific aircraft type decreases with the decrease of the domination measure of one aircraft over another with respect to payment conditions, for all aircraft types.

Influence of changes in travel demand, which further affect the required aircraft capacity, is also presented in this paper through the experiments conducted by applying the AHP, FAHP and ESM. The initial experiment and 8 additional experiments are analyzed and it is concluded that the changes in the required aircraft capacity influence the final ranking of aircraft types derived by AHP and FAHP. On the contrary, the most suitable aircraft according to the ESM is always ATR72-600.

The three methods, the AHP, FAHP and ESM, can be successfully used for aircraft type selection problems. The great advantage of these methods is that each of them can use both quantitative and qualitative data. With regards to data required, it can be seen that ESM requires the most specific data, AHP needs pairwise comparison of the data, while FAHP can use imprecise pairwise comparison. Having in mind that selection of final, the most appropriate aircraft is not dependent on the chosen method, decision makers have an opportunity to choose the method according to his own preferences or with regards to practical issues.

Decision making related to aircraft selection is process which involves people from different department in airlines that have different points of view. In order to encompass this diversity in thinking influenced by positions in airline as well as by educational background, group decision making could be employed in the future. It would include different opinion concerning fuzzy scale, as well. The interviews could be conducted with people involved in aircraft selection problem to get pairwise/fuzzy pairwise comparison matrices based on airline representatives' expert opinions. Different fuzzy numbers could be used as well, in order to learn sensitivity of solutions with regards to this kind of changing.

Acknowledgements. This research has been supported by the Ministry of Education, Science and Technological Development, Republic of Serbia, as part of the project TR36033 (2011–2017).

References

1. Aircraft Commerce, Aircraft Analysis & Fleet Planning - Large turboprops & RJs take on jetliners, vol. 64, pp. 30-38 (2009)
2. Brüggen, A., Klose, L.: How fleet commonality influences low-cost airline operating performance: Empirical evidence. J. Air Transp. Manag. **16**, 299–303 (2010)
3. Bruno, G., Esposito, E., Genovese, A.: A model for aircraft evaluation to support strategic decisions. Expert Syst. Appl. **42**, 5580–5590 (2015)
4. Dožić, S., Kalić, M.: An AHP approach to aircraft selection process. Transport. Res. Procedia **3**, 165–174 (2014)
5. Dožić, S., Kalić, M.: Comparison of two MCDM methodologies in aircraft type selection problem. Transport. Res. Procedia **10**, 910–919 (2015)
6. Dožić, S., Kalić, M.: Three stage airline fleet planning model. J. Air Transp. Manag. **46**, 30–39 (2015)
7. Gomes, L.F.A.M., Fernandes, J.E.M., Soares de Mello, J.C.C.B.: A fuzzy stochastic approach to the multicriteria selection of an aircraft for regional chartering. J. Adv. Transpor. **48**, 223–237 (2014)

8. Hammond, J.S., Keeney, R.L.: Smart Choices: A Practical Guide to Making Better Decisions. Harvard Business School Press, Boston (1999)
9. Harasani, W.I.: Evaluation and selection of a fleet of aircraft for a local airline. JKAU. Eng. Sci. **17**, 3–16 (2006)
10. Harasani, W.I.: Evaluation and selection of a fleet of aircraft located in Madniah, Saudi Arabia. In: 12th ATRS Conference, Athens, Greece (2008)
11. Holloway, S.: Straight and Level: Practical Airline Economics. Ashgate Publishing Limited, Ashgate Publishing Company, Farnham, Brookfield (2008)
12. Ishizaka, A., Nguyen, N.H.: Calibrated fuzzy AHP for current bank account selection. Expert Syst. Appl. **40**, 3775–3783 (2013)
13. Kubler, S., Roberta, J., Derigentb, W., Voisinb, A., Traona, Y.L.: A state-of the-art survey & testbed of fuzzy AHP (FAHP) applications. Expert Syst. Appl. **65**, 398–422 (2016)
14. Omkarprasad, S., Vaidya, S.K.: Analytic hierarchy process: An overview of applications. Eur. J. Oper. Res. **169**, 1–29 (2006)
15. Ozdemir, Y., Basligil, H., Karaca, M.: Aircraft selection using analytic network process: A case for Turkish airlines. In: Proceedings of the World Congress on Engineering, London, pp. 1155–1159 (2011)
16. Saaty, T.L.: The Analytic Hierarchy Process, Planning, Priority Setting, Resource Allocation. McGraw-Hill, New York (1980)
17. See, T.-K., Gurnani, A., Lewis, K.: Multi-attribute decision making using hypothetical equivalents and inequivalents. J. Mech. Design. **126**, 950–958 (2004)
18. Vasigh, B., Taleghani, R., Jenkins, D.: Aircraft Finance: Strategies for Managing Capital Costs in Turbulent Industry. J. Ross Publishing, New Castle (2012)
19. Wang, T.C., Chang, T.H.: Application of TOPSIS in evaluating initial training aircraft under a fuzzy environment. Expert Syst. Appl. **33**, 870–880 (2007)
20. Wang, Y., Jung, K.A., Yeo, G.T., Chou, C.C.: Selecting a cruise port of call location using the fuzzy-AHP method: A case study in East Asia. Tourism Manage. **42**, 262–270 (2014)
21. Wang, Y.M., Chin, K.S.: Fuzzy analytic hierarchy process: A logarithmic fuzzy preference programming methodology. Int. J. Approximate Reasoning **52**, 541–553 (2011)
22. Wensveen, J.G.: Air Transportation: A Management Perspective. Ashgate Publishing Limited, Ashgate Publishing Company, Farnham, Brookfield (2007)
23. Yeh, C.H., Chang, Y.H.: Modeling subjective evaluation for fuzzy group multicriteria decision making. Eur. J. Oper. Res. **194**, 464–473 (2009)

Urban Transportation and City Logistics

Urban Transportation and City Logistics

Yuval Hadas

Department of Management, Bar-Ilan University, Max ve-Anna Webb Street,
5290002 Ramat Gan, Israel
yuval.hadas@biu.ac.il

Urban population is constantly increasing, from 33% of the total population in 1960 to 53% in 2015 [10], meaning that more challenges are present to society, in and around cities. This is also true for logistics, with the notion of the global city [7], in which the global economy can be represented in flows, which are concentrated in and around some cities. Transportation and logistics are highly affected by those trends, as passengers and freight flow are increasing, resulting with congestion, pollution, accidents, etc. Thus, travelers in urban areas are being encouraged to create a mode shift from an excessive use of their cars and to increase the use of public transport and other active modes [9]. In order to increase the shift towards public transport, it is essential to provide better accessibility and connectivity [2, 6], as well as continuous priority to efficient urban transportation modes [4, 5]. Such priorities can be set not only for passengers, but for sustainable vehicle routes [8] as well. Smooth transfers are another direction to increase transportation's level of service [1, 3], which can also be used for freight cross-docking, and city distribution centers.

The provision of efficient transportation modes (public transport), sustainable modes (cycling), and the introduction of models aiming at supporting the design of innovative transportation systems are required in order to address those challenges. This section is devoted to research addressing the above-mentioned issues.

Seven high-quality papers were accepted for this section, based on a rigorous review process, that includes independent reviews by selected researchers, as well as the evaluation by the guest editors. These papers cover a wide range of urban transportation and city logistics: (1) transit network design, (2) railway network design, (3) tourists' transportation, (4) cycling network design, (5) the accessibility to service centers, (6) transit assignment, and (7) sustainable transportation.

The chapter by Seyedabrishami, Nazemi, and Zarrinmehr, entitled "**Transit Network Design Problem: An Expansion of the Route Generation Algorithm**", introduce an expansion algorithm for the transit network design problem, based on route generation. This algorithm contributes to a better design of transit network, which help increasing transit share of urban trips and reducing traffic congestion in urban networks. Previous research proposed the Route Generation Algorithm (RGA) in which shortest paths with the highest demands are selected and then expanded by insertion of new nodes in order to increase transit demand coverage. This paper expands RGA by introducing a new algorithm namely Extended Route Generation Algorithm (ERGA) with further details in node insertion scheme. In contrast to the conventional RGA in which all nodes are examined to be inserted between nodes of origin-destination pairs, the algorithm inserts only adjacent nodes to the shortest paths between selected origin-destinations.

The chapter by Canca, De-Los-Santos, Laporte, and Mesa, entitled "**The Railway network design, line planning and capacity problem: An adaptive large neighborhood search metaheuristic**" propose a model for the Railway Network Design and Line Planning (RNDLP) problem, integrating the first stages in the Railway Planning Process. The network design problem incorporates costs relative to the network construction, proposing a set of candidate lines. The line planning problem determines optimal frequencies and consequently train operations, taking into account rolling stock, personnel and fleet acquisition costs. The model consider the existence of an alternative transportation mode for each origin-destination pair in the network, and passengers choose their transportation mode according to their own utility. A computationally tractable algorithm, an Adaptive Large Neighborhood Search (ALNS) algorithm, was developed, which can handle the RNDLP problem.

Ruiz, and Seguí-Pons in the chapter entitled "**Diagnostic of the balance and equity of public transport for tourists and inhabitants**", propose a different aspect to public transportation, namely, the balance and equity of services for tourists and commuters. The significant development of tourism as an economic activity in some cities has led to the need to integrate potential demand for public transport generated by tourists into planning decisions. Public transportation is considered a critical factor of tourism competitiveness and plays a fundamental role in the promotion and maintenance of a sustainable tourist destination. Ensuring equity of public transport services requires specific methodological tools for diagnosis and optimization. A method for evaluating potential public bus transportation demand including both residents and tourists is presented. The method uses geographic information systems and statistics, namely, the Gini co-efficient, to diagnose imbalances between supply and demand and the degree of equity in transport services.

The Design of Leisure Devoted Cycling Networks was investigated by Giovannini, Malucelli, and Nonato in the chapter entitled "**On the Design of Leisure Devoted Cycling Networks**". This work also address tourists mobility in the form of cycling. The authors extend the design of cycling networks in order to capture the demand generated by tourism. Leisure intended cycling networks are the topic of this work. The authors revise the critical issues in cycling network design, present a design methodology for cycle tourist networks that integrates the specific features of the problem into a combinatorial optimization framework, and test the approach on realistic data comparing with previous contributions on this topic.

Shnaiderman, in the chapter "**The Influence of Transportation Service Level on a Municipal Service Center's Costs: A Numerical Study based on Supply Chain Management Models**" implemented supply chain management models, that consider surplus and shortage, in order to formulate the relation between urban centers' costs and transportation level of service. This paper deals with scheduling of appointments between providers and customers (with reservations or walk-in ones) in municipal service centers. In order to improve the service level and reduce the uncertainty of the number of customers' demand, a free transportation service from the customers' locations to the service center and back is operated. An optimal transportation service level (TSL) is set in order to minimize the provider's total idle time and overtime on the one hand, and the transportation service's operation cost on the other hand.

A stochastic transit assignment model was developed by Codina and Rosell. The chapter entitled "**A stochastic version of the strategy-based congested transit assignment model and a technique by smoothing approximations**" develops a stochastic version for the strategy based congested transit assignment problem. This stochastic version takes into account stochastic mean waiting times of passengers at stops and in-vehicle travel times. The model is formulated as a stochastic variational inequality derived from the formulation of the deterministic version of the problem. Furthermore, a consistent smoothing approximation to the deterministic model is developed and it is shown that this approximation provides an alternative way of solving the deterministic model. It is also shown that both, the stochastic model and the smoothed approximation, can be solved by means of an adaptation of a path based method for the asymmetric traffic assignment problem.

The final chapter, entitled "**Evaluation of CO_2 Emission Reduction from Vehicles by Information Provision Using Driving Simulator**" by Matsumoto and Ishiguro, evaluates CO_2 emission reduction from vehicles at intersections when information is provided to the drivers. Reduction of CO_2 emissions has become an important social issue all over the world. Therefore, drivers information provision is proposed in this paper to decrease unnecessary vehicle movements and to reduce the amount of CO_2 emissions while approaching signalized intersections. The system provides recommended speed information and accelerator-off indication, in which the vehicle could pass through the intersection or shorten the idling time if it would follow the provided information. Driving experiments were conducted with a driving simulator where the proposed information system was introduced. Multiple regression analysis showed that providing the recommended speed information and the accelerator-off indication reduced the amount of CO_2 emissions significantly.

References

1. Ceder, A., Hadas, Y., McIvor, M., Ang, A.: Transfer synchronization of public transport networks. Transp. Res. Record: J. Transp. Res. Board **2350**, 9–16 (2013)
2. Hadas, Y.: Assessing public transport systems connectivity based on Google Transit data. J. Transp. Geogr. **33**, 105–116 (2013)
3. Hadas, Y., Ceder, A.: Optimal coordination of public-transit vehicles using operational tactics examined by simulation. Transp. Res. Part C **18**, 879–895 (2010)
4. Hadas, Y., Ceder, A.A.: Optimal connected urban bus network of priority lanes. Transp. Res. Record: J. Transp. Res. Board **2598**, 49–57 (2014)
5. Hadas, Y., Nahum, O.E.: Urban bus network of priority lanes: a combined multi-objective, multi-criteria and group decision-making approach. Transp. Policy **52**, 186–196 (2016)
6. Hadas, Y., Ranjitkar, P.: Modeling public-transit connectivity with spatial quality-of-transfer measurements. J. Transp. Geogr. **22**, 137–147 (2012)
7. O'Connor, K.: Global city regions and the location of logistics activity. J. Transp. Geogr. **18**, 354–362 (2010)

8. Pamučar, D., Gigović, L., Ćirović, G., Regodić, M.: Transport spatial model for the definition of green routes for city logistics centers. Environ. Impact Assess. Rev. **56**, 72–87 (2016)
9. Rabl, A., de Nazelle, A.: Benefits of shift from car to active transport. Transp. Policy **19**, 121–131 (2012)
10. United Nations Population Division: Urban population. The World Bank (2016)

Transit Network Design Problem: An Expansion of the Route Generation Algorithm

Iran Khanzad, Seyedehsan Seyedabrishami(✉), Mohsen Nazemi, and Amirali Zarrinmehr

Civil and Environmental Engineering Department, Tarbiat Modares University, Tehran, Iran
seyedabrishami@modares.ac.ir

Abstract. The problem of public transportation routes network design deals with establishing the configuration of transit routes in a transportation network. This problem is recognized as one of the most complicated problems in transportation planning. An appropriate design of transit network will help increasing transit share of urban trips and reducing traffic congestion in urban networks. Previous research proposed the Route Generation Algorithm (RGA) in which shortest paths with the highest demands are selected and then expanded by insertion of new nodes in order to increase transit demand coverage. This paper extends RGA by introducing a new algorithm namely Extended Route Generation Algorithm (ERGA) with further details in node insertion scheme. A heuristic algorithm is proposed, tested in a medium-size network, and applied on a real-size network. In contrast to the conventional RGA in which all nodes are examined to be inserted between nodes of Origin-Destination (O-D) pairs, the algorithm inserts only adjacent nodes to the shortest paths between selected (O-D) pairs. Moreover, the proposed algorithm restricts the distance between each pair of nodes, not to be greater than two times of the shortest path length between the two nodes and the travel time of each generated route in the problem. Also, the number of common links between proposed routes of ERGA has been restricted due to the specifications of case study network.

Keywords: Transit routes network design · Route generation · Node insertion

1 Introduction

Many urban traffic problems such as traffic congestion, car accidents, and air pollution are due to the high share of private vehicles travelling in urban trips. Private vehicle demand management policies such as encouraging the use of the public transportation offer the most effective solutions to reduce the aforementioned problems. Enhancing accessibility and quality of the public transportation service will increase its utility and consequently a modal shift from private mode to public mode [15]. Public transportation routes network design is the first step towards enhancing the quality of public transport service. Network design and network accessibility has an important role in public transport demand management.

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_8

Formulating the problem is the primary step of Transit Routes Network Design Problem (TRNDP) and is followed by adopting an appropriate solution method. The problem is non-convex and non-linear with both discrete and continuous variables and constraints [1]. Due to the nature of such a problem, it is complicated to formulate the TRNDP with a mathematical programming approach and there are a limited number of studies in this regard [5, 12]. Therefore, non-exact optimization techniques such as heuristic and meta-heuristic approaches are utilized to solve the problem [10, 11]. For example, Poorzahedy and Safari [22] proposed a meta-heuristic solution aiming to minimize the total travel time of the users of the network. Mandle [17] presented a heuristic algorithm in order to minimize the travel time and maximize the route directness.

The main challenge of transit network planning is to find an efficient system that provides benefits for the users and the operator at the same time. From the users' viewpoint, efficient system is met by providing direct trips (i.e. trips with no transfer between routes) with less travel time and from the operator's viewpoint, efficiency is obtained by gaining as much profit as possible which will be provided by attracting more ridership [14]. Nevertheless, in recent years most researchers minimized the cost of the users and the operator simultaneously as the objective of the problem. Fan and Machemehl [8, 9] and Zhao and Zeng [29] used the total welfare as the objective of the problem. Ceder and Israeli [6], Nikolic and Teodorovic [20], and Tom and Mohan [25] considered the number of transfers, load profile, fleet size, and total cost (summation of users' and operator's costs) as the objective function. Maximizing the transit ridership was also explicitly taken into account as an objective function in the recent literature [27, 28]. Among the wide range of efforts in TRNDP, constraints are usually put on in vehicle travel time, waiting time, fleet size, and bus load factor, and network length [22].

Some existing studies redefined the TRNDP as the two following sub-problems: (1) Candidate route generation (2) route/network improvement. Heuristic algorithm (typically shortest-path-based algorithm) is used to generate candidate routes [15]. Bagloee and Ceder [3] and Lee and Vuchic [8, 9, 16] used the shortest path as the heuristic candidate route. These shortest paths will be improved by adding new nodes to the path attempting to satisfy certain criteria defined in the problem. Several studies used heuristic and meta-heuristic algorithms to improve generated candidate routes in first step. Pattnaik et al. [21], Chakroborty and Wivedi [7], Ngamchai and Lovell [19], and Tom and Mohan [25] used genetic algorithm and Yang et al. [26], Hu et al. [13] used ant colony algorithm to find the most appropriate routes among the primary candidate routes.

Baaj and Mahmassani [2] proposed a heuristic algorithm namely route generation design algorithm (RGA) which considers both interests of the operator and the users [18]. RGA starts with generating the candidate route between nodes by finding the shortest paths and improves them by inserting the pre-selected nodes. Following the study of Baaj and Mahmassani [21], Mauttone and Urquhart [18] presented a constructive algorithm based on RGA named pair insertion algorithm (PIA). The objective function of the problem was defined in two separate parts for the interests of operator and users. PIA inserts nodes to the shortest paths in pairs in the way that the best value for sum of user and operator costs is achieved.

In this paper, we focus on users' point of view in terms of route directness and a threshold for the ratio of transit route travel times to the shortest path travel times. Furthermore, operator's point of view in expanding transit routes is considered as maximizing demand covered by a public transportation route and minimizing the overlap between network routes. The following aspects of the problem are addressed in this research: (1) The solution algorithm has been applied to a real network to design public transportation, (2) Some practical constraints such as (a) Maximum and minimum route length based on the average length of the network links (b) Directness of proposed routes (between each separated node pairs of routes) (c) Maximum overlapping of routes, have been taken into account in the proposed algorithm.

2 Definitions, Model Formulation, and Solution Algorithm

In the following, the notations used in the formulation are defined.

$G = (N, E)$	graph "G" with "N" and "E" sets of nodes and edges, respectively
r	any feasible transit route between each origin and destination
R	the set of all "*proposed transit routes*" by the proposed algorithm
feasible routes set (*prs*)	The set of all feasible transit routes between each O-D pair
Adjacency matrix	a (zero/one) matrix that displays the adjacency between network nodes
$Z(r)$	the objective function defined as the ratio of route "r" transit travel time to its direct transit demand (a trip with no transfer) covered by route "r"
$t(r)$	the transit travel time of route "r"
$d(r)$	the direct transit demand covered by route "r"
t_{min}	the minimum allowable transit travel time for the "*subroute*"
t_{max}	the maximum allowable transit travel time for the "*subroute*"
$t_{(subroute,i,j)}$	the transit travel time from node "i" to node "j" taking "*subroute*"
$t_{(shortestpath,i,j)}$	the shortest path transit travel time between node "i" and node "j"
m	the maximum allowable transit travel time ratio (defined by the modeler)
l	the number of links in "*candidate route*" which overlap with links of other "*proposed transit routes*" in "R"
n	the maximum number of allowable overlaps (defined by the modeler)
D_{min}	the minimum transit demand coverage of all "*proposed transit routes*" in "R" (defined by the modeler)
candidate route	a generated transit route which meets criteria 2 (allowable subroute travel time) and 3 (with the best objective function after inserting all "*candidate nodes*" sequentially
candidate node	a common node present in both "*shortest path's adjacent node*" and "*candidate route's adjacent node*" sets
shortest path's adjacent node	a network node (except "*candidate route*'s" nodes) with one edge distance from shortest path's nodes

(*continued*)

(*continued*)

candidate route's adjacent node	a network node with one edge distance from "*candidate route*'s" nodes
subroute	a newly generated transit route by inserting a new "*candidate node*" to a "*candidate route*"
proposed transit route (*pr*)	a "*candidate route*" which meets criteria 4 and 5
adjacent node	a "*candidate route*'s" node with one edge distance from "*candidate node*"
overlapping links	links that are common between the resulted routes of problem
Note:	Since the focus of this study is on transit routes network design, the use of "travel time" and "travel demand" refers to "transit travel time" and "transit travel demand", respectively.

2.1 Model Formulation and Assumptions

The proposed model is designed to be implemented on bi-directional graphs shown by $G = (N, E)$ where "N" is the set of nodes and "E" is the set of edges in the network. Routes (r) are consisted of bi-directional links and the output of the proposed model and the relative algorithm is a set of proposed transit routes collected in set "R". The assumption of demand coverage in this research is to cover direct demand defined as passenger trips with no transfer. Furthermore, transit dwell times at stops is not considered in travel times.

The objective function in Eq. (1), in a general scheme, is defined to optimize the efficiency of the transit network in terms of minimizing the ratio of travel time to demand coverage for each transit route. Constraints (2) and (3) are defined for "*subroute*" acceptance with respect to its travel time. Constraints (4) and (5) are defined to check the functionality of the "*candidate route*" by limiting the number of joint links or overlaps and checking travel time compared to the minimum travel time, respectively. Constraint (6) is defined to terminate the process when the proposed transit network, consisted of transit network routes in "R", satisfies a predefined demand. Problem input data are zero/one matrix representation of graph "G" of the network, link travel times, and Origin-Destination (O-D) demands and the output is set R consisted of the proposed routes with less Z-values.

$$\min Z(r) = \frac{t(r)}{d(r)} \qquad \forall r \in prs \tag{1}$$

subject to:

$$t_{min} < t_{(subroute)} < t_{max} \qquad \forall (subroute) \in prs \tag{2}$$

$$\frac{t_{(subroute,i,j)}}{t_{(shortestpath,i,j)}} \leq m \qquad \forall (i,j) \in subroute,\ \forall (subroute) \in prs \tag{3}$$

$$l < n \tag{4}$$

$$t_{(candidate\ route)} \geq t_{min} \qquad \forall (candidate\ route) \in prs \tag{5}$$

$$\sum_{(pr) \in R} d(pr) \geq D_{min} \tag{6}$$

2.2 Expanded Route Generation Algorithm

TRNDP is a complicated problem and exact solution methods can rarely be applied to solve it [1]. Heuristic methods, as "non-exact" methods, are used to solve hard problems such as TRNDP. Similar to constructive algorithms of [3, 15], the route generation algorithm in this study, namely Expanded Route Generation Algorithm (ERGA) consists of two basic parts: the first part starts with the shortest path for the selected O-D pair and the second part expands this route investigating for better routes. Such expansion is based on the proposed heuristic algorithms for route expansion as explained in Figs. 1 and 2.

ERGA starts with selecting a pair of O-D with the maximum demand and finding the shortest path between these two nodes as the initial route called "*candidate route*". Then the value of $Z(r)$ as in Eq. (1) for this shortest path is calculated and the "*candidate nodes*" set is determined when "*shortest path's adjacent node*" and "*candidate route's adjacent node*" sets are identified. The *candidate node* having the greatest demand with the "*candidate route*'s" nodes is selected to be inserted in the route first. The newly generated routes by inserting this "*candidate node*", as explained in node insertion algorithms, are called "*subroutes*". "*Subroutes*" should first meet criteria 2 and 3. If these criteria were met, the $Z(r)$ is calculated for these *subroutes* and the "*subroute*" with the minimum $Z(r)$ is selected and is compared with the "*candidate route*". If the selected "*subroute*" had a lower $Z(r)$, this "*subroute*" replaces the former "*candidate route*". Then the new "*candidate nodes*" are identified for the updated *candidate route* and the process continues until there exists no "*candidate node*" to be inserted in the "*candidate route*". If this "*candidate route*" meets criteria 4 and 5, it becomes the "*proposed transit route*" for the pertinent (o, d) and will be added to set R. Otherwise, the demand of the selected O-D will be neglected to find the next O-D pair. Since it is assumed that the *proposed transit route* covers all the nodes' demands on its way, the demand matrix is updated by setting the demands of the "*proposed transit*

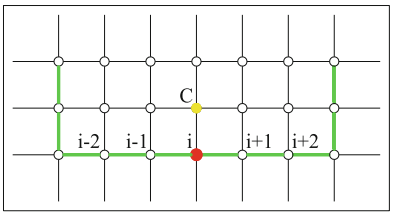
route's" nodes to zero. Finally, ERGA stops selecting next O-D pair, only if sum of the demands covered by "*proposed transit routes*" in R is greater than D_{min}.

2.3 Node Insertion Algorithm

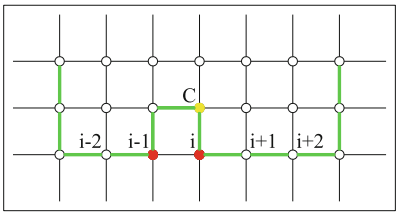
ERGA adopts a heuristic approach for node insertion in route expansion step. Compared to previous approaches suggested for transit route design algorithm, route expansion is enhanced in terms of considering the concept of adjacent nodes exclusively by defining two different procedures for node insertion algorithm. The initial "*candidate route*" is expanded by inserting *candidate nodes*. The insertion of "*candidate nodes*" is performed based on procedures I and II as follows:

1 - Procedure I: In this procedure the selected "*candidate node*" C is inserted between the nodes of the "*candidate route*" that are in one-edge or two-edges distance from the "*adjacent node*". Figure 1a shows the position of "*candidate node*" C the corresponding "*adjacent node*" i. Insertion algorithm of "*candidate node*" C in the "*candidate route*" for possible eight cases is illustrated through Fig. 1b to i. In the first case, i.e. *a*, no node insertion is carried out. The other seven cases can be described as follows:
   b. Inserting "*candidate node*" C between "*adjacent node*" and the node with one-link distance before that on the "*candidate node*" (i and $i - 1$)
   c. Inserting "*candidate node*" C between "*adjacent node*" and the node with one-link distance after that on the "*candidate route*" (i and $i + 1$)
   d. Inserting "*candidate node*" C between "*adjacent node*" and the node with two-links distance before that on the "*candidate route*" (i and $i - 2$)
   e. Inserting "*candidate node*" j between "*adjacent node*" and the node with two-links distance after that on the *candidate route* (i and $i + 2$)
   f. Inserting "*candidate node*" C between the nodes with one-link distance before and after "*adjacent node*" on the "*candidate route*" ($i - 1$ and $i + 1$)
   g. Inserting "*candidate node*" C between the nodes with two-links distance before and after "*adjacent node*" on the "*candidate route*" ($i - 2$ and $i + 2$)
   h. Inserting "*candidate node*" C between the nodes with one-link distance before and two-links distance after "*adjacent node*" on the "*candidate route*" ($i - 1$ and $i + 2$)
   i. Inserting *candidate node*. C between the nodes with two-links distance before and one-link distance after "*adjacent node*" on the "*candidate route*" ($i - 2$ and $i + 1$)

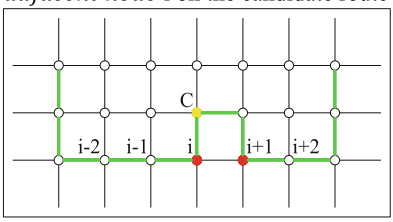
2 - Procedure II: In this procedure the selected "*candidate node*" C with one-edge distance from more than one nodes of the "*candidate route*" is inserted between each pair of the nodes of the *candidate route* which are adjacent to the "*candidate*



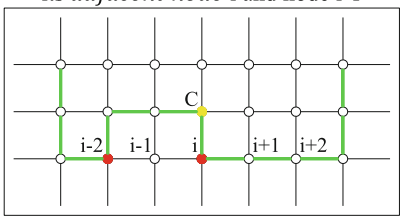
a. Vicinity of *candidate node* C to its *adjacent node* i on the candidate route



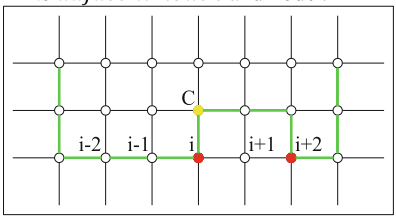
b. Inserting *candidate node* C between its *adjacent node* i and node i-1



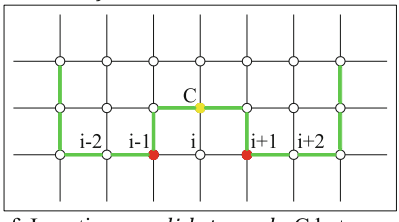
c. Inserting *candidate node C* between its *adjacent node i* and node *i*+1



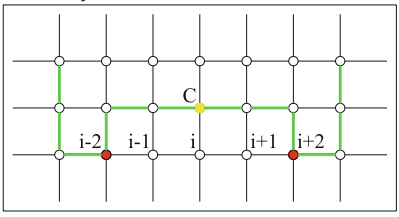
d. Inserting *candidate node* C between its *adjacent node* i and node i-2



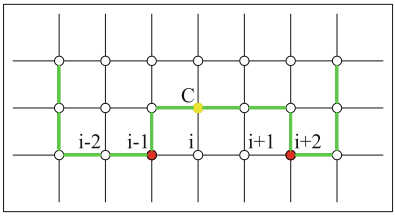
e. Inserting *candidate node* C between its *adjacent node* i and node i+2



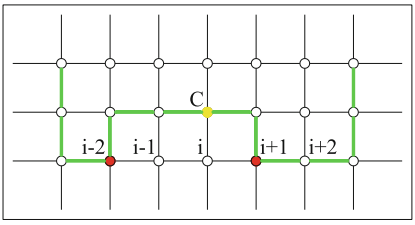
f. Inserting *candidate node* C between node *i*-1 and node *i*+1



g. Inserting *candidate node C* between node *i*-2 and node *i*+2



h. Inserting *candidate node C* between node *i*-1 and node *i*+2

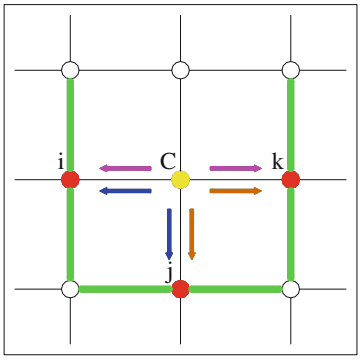


i. Inserting *candidate node C* between node *i*-2 and node *i*+1

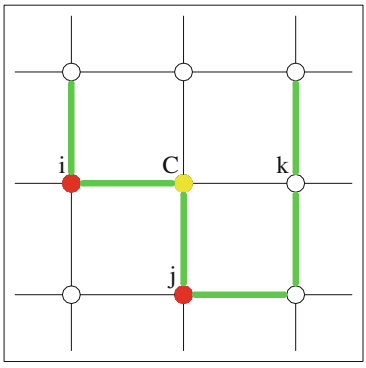
Fig. 1. Insertion cases of *candidate node C* in *candidate node* in procedure I

node". Figure 2 illustrates the steps of this procedure. *i-j-k* is a part of the "*candidate route*" and "*candidate node*" *C* is eligible to be inserted in the route. Figure 2b shows that node C is inserted between "*adjacent nodes*" *i* and *j* and expands a part of the "*candidate route*" to *i-C-j-k*. In Fig. 2c "*candidate node*" *C* is inserted between "*adjacent nodes*" *j* and *k* and this part of the "*candidate route*" is expanded to *i-j-C-k*, when in Fig. 2d "*candidate route*" *C* is inserted between "*adjacent nodes*" *i* and *k* in such a manner that "*adjacent node*". *j* is eliminated and the "*candidate route*" has become *i-C-k*.

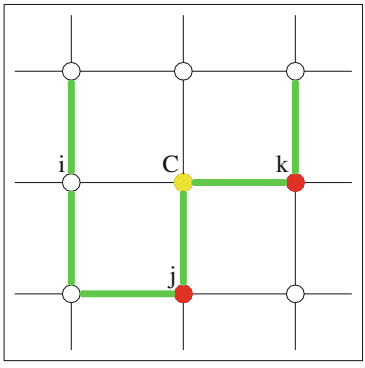
Figure 3 shows the ERGA algorithm in more details.



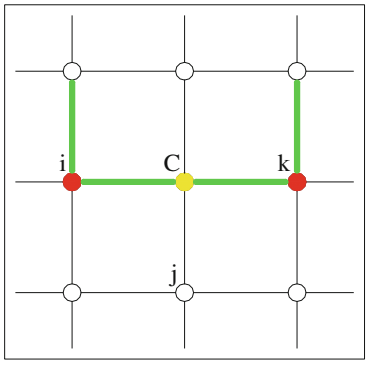
a. *Candidate node* C and *adjacent nodes* i, j, and k on the *candidate route*



b. Inserting *candidate node* C between *adjacent nodes* i and j



c. Inserting *candidate node C* between *adjacent nodes j* and *k*



d. Inserting *candidate node C* between *adjacent nodes i* and *k*

Fig. 2. Insertion algorithm of "*candidate node*" C with one-edge distance from more than one node of the "*candidate route*"

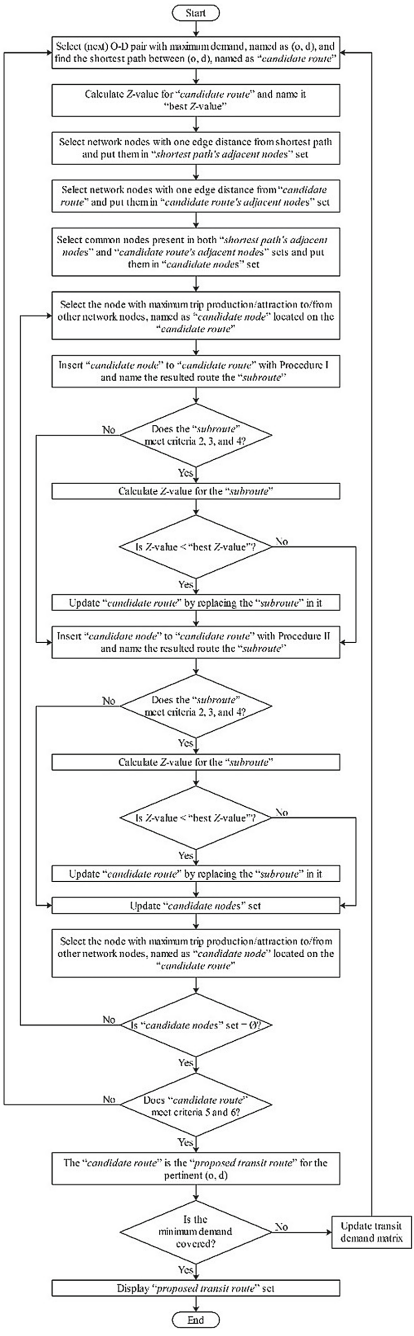


Fig. 3. ERGA algorithm

3 Application of the Proposed Algorithm and Discussion of Results

To illustrate the application of the ERGA algorithm, a medium-size test network as well as a network of a real-size are used in this section to report the results. The algorithm is coded in MATLAB 11 and consists of two main modules of route generation and node insertion. MATLAB is linked to a MS Excel file to read input data and to store generated routes as output. The models were run on a 64-bit machine using 2 GB memory.

Two types of input data are required; the first type is the basic supply-side data used for network coding, while the second type is related to the demand-side revealing the trip production/attraction of each zone. Specifically, the required input data format is as follows:

- Graph representation of network (nodes and directed links) in form of a zero-one Matrix and their associated free flow travel times (T_0)
- Zones, centroids, and connectors and associated O-D demand matrix
- A matrix representing the vicinity of zones (with zero/one values)
- Transit O-D demand matrix

The required input data for Sioux-Falls test network are referred to Bar-Gera [4]. Also, data for Ardebil network were obtained from comprehensive transportation studies of Ardebil city.

3.1 Sioux-Falls Test Network

Sioux-Falls test network was initially used to verify the performance of the proposed algorithm. This simple test network has 24 nodes (as zones) and 76 bi-directional links with symmetric travel time and demand matrix. Two different scenarios were defined to verify the performance of the whole process. In the first scenario after selecting O-D with maximum demand and generating shortest path between them, candidate nodes were inserted between all nodes of the shortest path and the subsequently created subroute, without any constraint. In the second scenario, candidate nodes were inserted taking into consideration the proposed constraints (2) to (5), as discussed in Sect. 2 of the paper. Table 1 shows the considered amounts for proposed constraints in this problem. As we can see in Table 2, Figs. 4 and 5, scenario 2 results in more practical routes with almost the same demand coverage compared to scenario 1.

Table 1. Considered amounts for proposed constraints for Siouxfalls network

Proposed parameters	t_{min} (min)	t_{max} (min)	m	n	D_{min} (%)
Considered amount	20	30	1.5	4	30

Table 2. Defined scenarios for Sioux-Falls test network

Scenario number	Considered constraints	Number of routes	Demand coverage (%)	*overlapping ratio*[a]
1	Inserting candidate nodes to the subroute without any constraint	2	33.7	0.043
2	Inserting candidate nodes to the subroute considering constraints (2) to (5).	4	33.0	0.110

[a] $overlapping\ ratio = \frac{\text{travel time of } overlapping\ links}{\text{all resulted routes travel time}}$

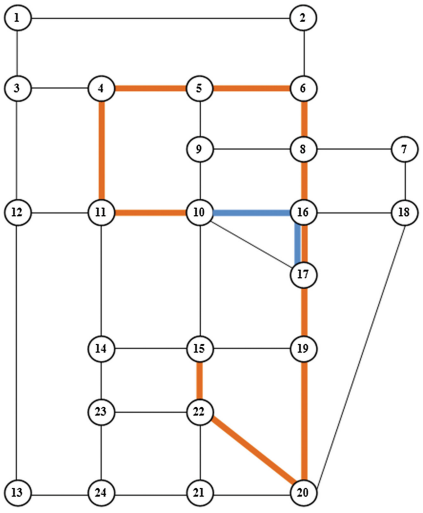


Fig. 4. Resulted routes for Sioux-Falls network without any constraint

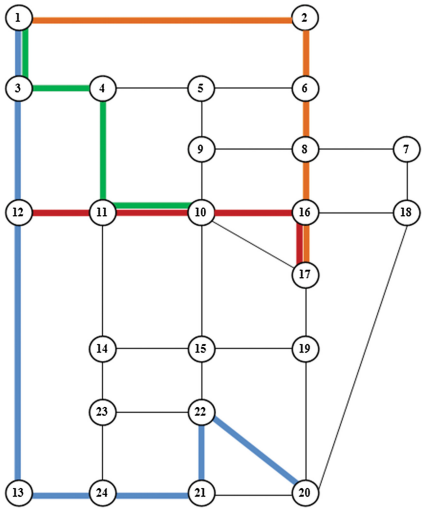


Fig. 5. Resulted routes for Sioux-Falls network considering constraints (2) to (5)

3.2 Ardebil City Case Study

In this case study, we used the urban network of Ardebil city, the capital and the largest city of Ardebil province located in Northwestern Iran. It has a dense population of over 480000 people with a considerable daily trip rate, customarily with work and educational purposes. This network was formerly coded in VISUM software for the comprehensive transportation planning studies pertaining to year 2013, and was also used as the study network in other researches [23, 24].

Ardebil city network consists of: 881 nodes, 2196 links, and 100 traffic area zones, as depicted in Fig. 6, with 110000 total transit demand trips. We used the transportation analysis zones, defined in the regional comprehensive transportation planning studies, for our model. The transit O-D matrix was an existing daily O-D matrix extracted from the classic four-step model taken from the aforementioned studies, based on regional

socio-economic data. The directed graph of Ardebil network was regenerated in ArcMap 10.3 associated with links' free flow travel times. We built zone adjacent matrix utilizing Proximity toolset (Polygon Neighbors) which creates a table with statistics based on polygon contiguity and coincident edges. To have a network representation of the routes, we used Network Analyst extension of ArcMap 10.3; given the zone numbers for each route, the shortest path on the network's links is generated by finding the projection of each zone's centroid on the closest link on the network and connecting them together considering directional links.

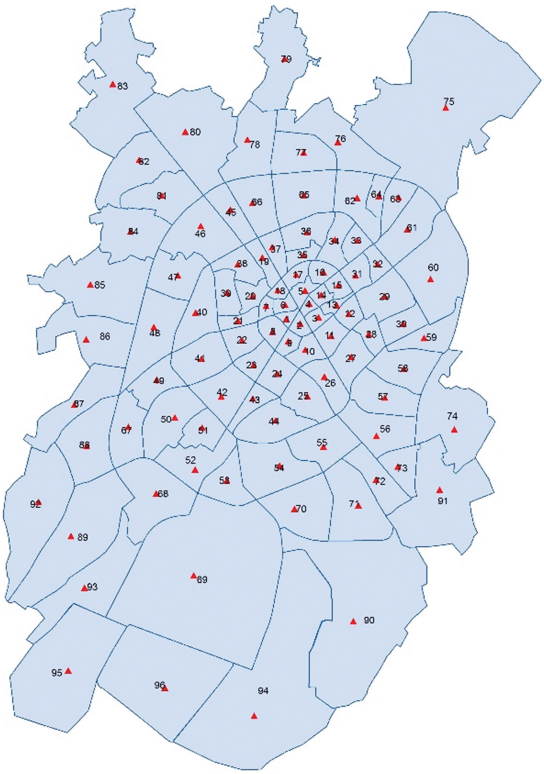


Fig. 6. Ardebil network layout

3.3 Results over Ardebil Network

Table 3 shows the values adopted for proposed constraints in this problem. The output of ERGA algorithm implemented on Ardebil network can be seen in Table 4. There are 4 routes generated and their travel time and demand coverage are presented. To have a

Table 3. Considered amounts for proposed constraints for Ardebil network

Proposed parameters	t_{min} (s)	t_{max} (s)	m	n	D_{min} (%)
Considered amount	1500	2000	2	8	10

Table 4. Optimal routes between each O-D pair for daily transit demand

Origin to destination	The feasible route	Value of t travel time (seconds)	Value of d demand coverage (trips)
92 to 69	92-89-68-53-54-69	1860	1467.8
2 to 83	2-6-17-37-66-80-81-82-83	1984	2901.2
2 to 85	2-1-8-21-39-40-41-48-47-85	1893	3219.4
63 to 74	63-61-29-30-59-58-57-74	1645	2108.8

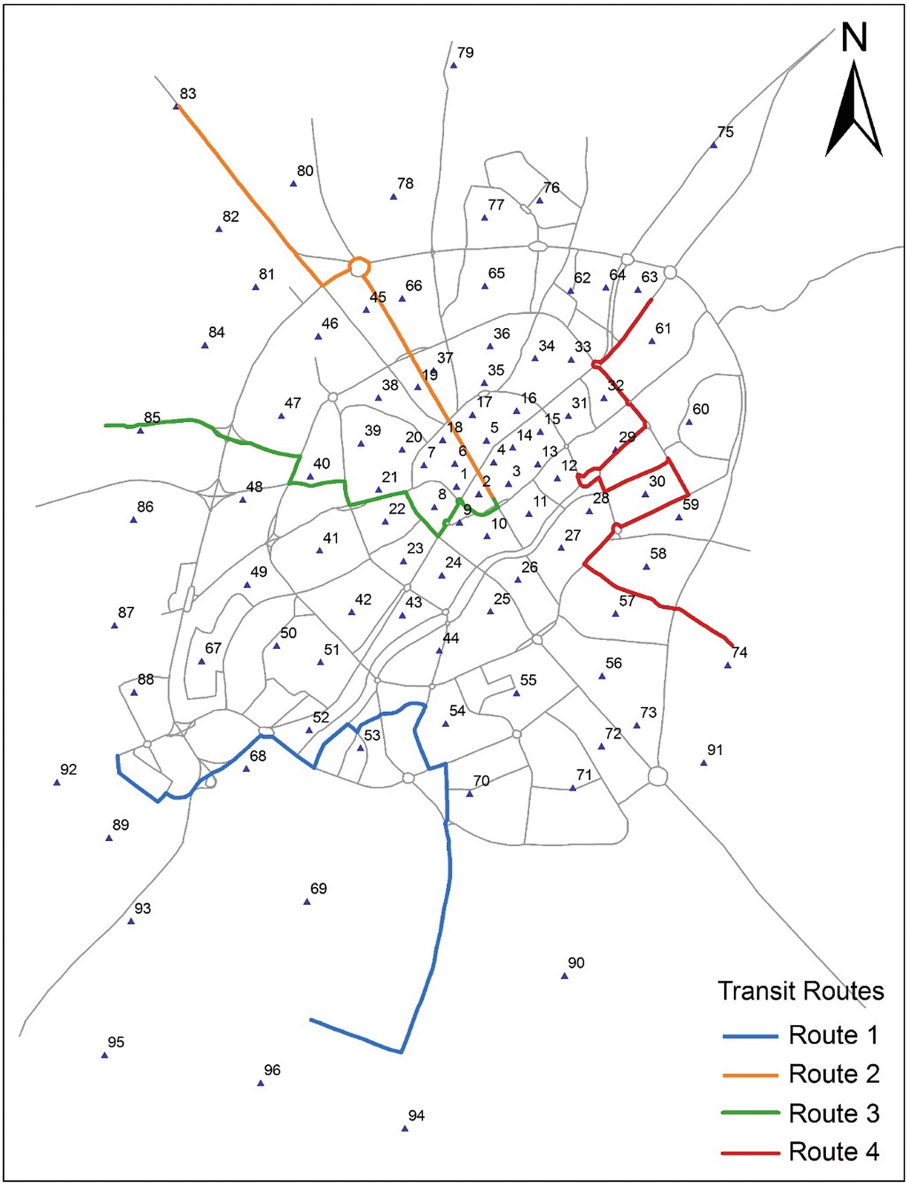


Fig. 7. Generated routes for daily transit demand

network representation of these routes, we used Network Analyst toolbox of ArcMap 10.3. Given the zone numbers for each route, the shortest path on the network's links is appeared. Figure 7 illustrates the graphical representation of the output.

4 Conclusion

In this research, a new method for transit routes network design problem has been studied and applied on a real network. The algorithm was intended to be easy to implement and less demanding in terms of both data requirements and analytical sophistication as compared with previous methods. The proposed algorithm, expansion of route generation algorithm (ERGA), is an extension to the formerly approach presented in [15]. ERGA produces transit routes in two phases. In the first phase, initial routes cover the maximum demand in the network and in the second phase, routes are expanded by inserting some candidate nodes into the shortest path. This algorithm considers new criteria for route acceptance after inserting candidate nodes and some criteria for adding expanded routes to network.

The results, reported over medium-to-large networks of Sioux-Falls and Ardebil, showed that the algorithm can be easily applied to real-size networks and generate feasible routes.

Acknowledgement. The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The authors gratefully acknowledge the contributions of Aavand Consulting Engineers Company for generously providing data and reports of the comprehensive transportation studies of Ardebil city.

References

1. Baaj, M.H., Mahmassani, H.S.: An AI-based approach for transit route system planning and design. J. Adv. Transp. **25**(2), 187–209 (1991)
2. Baaj, M.H., Mahmassani, H.S.: Hybrid route generation heuristic algorithm for the design of transit networks. Transp. Res. Part C Emerg. Technol. **3**(1), 31–50 (1995)
3. Bagloee, S.A., Ceder, A.A.: Transit-network design methodology for actual-size road networks. Transp. Res. Part B Methodol. **45**(10), 1787–1804 (2011)
4. Bar-Gera, H.: Transportation network test problems. http://www.bgu.ac.il/~bargera/tntp. Accessed 20 Sep 2011
5. Barra, A., et al.: Solving the transit network design problem with constraint programming. In: 11th World Conference in Transport Research-WCTR 2007 (2007)
6. Ceder, A., Israeli, Y.: Creation of objective functions for transit network design (1997)
7. Chakroborty, P., Wivedi, T.: Optimal route network design for transit systems using genetic algorithms. Eng. Optim. **34**(1), 83–100 (2002)
8. Fan, W., Machemehl, R.B.: Optimal transit route network design problem with variable transit demand: genetic algorithm approach. J. Transp. Eng. **132**(1), 40–51 (2006)
9. Fan, W., Machemehl, R.B.: Using a simulated annealing algorithm to solve the transit route network design problem. J. Transp. Eng. **132**(2), 122–132 (2006)

10. Fan, L., Mumford, C.L.: A metaheuristic approach to the urban transit routing problem. J. Heuristics **16**(3), 353–372 (2010)
11. Farahani, R.Z., et al.: A review of urban transportation network design problems. Eur. J. Oper. Res. **229**(2), 281–302 (2013)
12. Guan, J., Yang, H., Wirasinghe, S.: Simultaneous optimization of transit line configuration and passenger line assignment. Transp. Res. Part B Methodol. **40**(10), 885–902 (2006)
13. Hu, J., Shi, X., Song, J., Xu, Y.: Optimal design for urban mass transit network based on evolutionary algorithms. In: Wang, L., Chen, K., Ong, Y.S. (eds.) ICNC 2005. LNCS, vol. 3611, pp. 1089–1100. Springer, Heidelberg (2005). doi:10.1007/11539117_148
14. Johar, A., Jain, S., Garg, P.: Transit network design and scheduling using genetic algorithm–a review. Int. J. Optim. Control Theor. Appl. (IJOCTA) **6**(1), 9–22 (2016)
15. Kepaptsoglou, K., Karlaftis, M.: Transit route network design problem: review. J. Transp. Eng. **135**(8), 491–505 (2009)
16. Lee, Y.-J., Vuchic, V.R.: Transit network design with variable demand. J. Transp. Eng. **131** (1), 1–10 (2005)
17. Mandl, C.E.: Evaluation and optimization of urban public transportation networks. Eur. J. Oper. Res. **5**(6), 396–404 (1980)
18. Mauttone, A., Urquhart, M.E.: A route set construction algorithm for the transit network design problem. Comput. Oper. Res. **36**(8), 2440–2449 (2009)
19. Ngamchai, S., Lovell, D.J.: Optimal time transfer in bus transit route network design using a genetic algorithm. J. Transp. Eng. **129**(5), 510–521 (2003)
20. Nikolić, M., Teodorović, D.: A simultaneous transit network design and frequency setting: computing with bees. Expert Syst. Appl. **41**(16), 7200–7209 (2014)
21. Pattnaik, S., Mohan, S., Tom, V.: Urban bus transit route network design using genetic algorithm. J. Transp. Eng. **124**(4), 368–375 (1998)
22. Poorzahedy, H., Safari, F.: An ant system application to the bus network design problem: an algorithm and a case study. Pub. Transp. **3**(2), 165–187 (2011)
23. Seyedabrishami, S., Nazemi, M., Shafiei, M.: Off-line calibration of a macroscopic dynamic traffic assignment model: iterative demand-supply parameters estimation. In: 17th International IEEE Conference on Intelligent Transportation Systems (ITSC). IEEE (2014)
24. Shafiei, M., Nazemi, M., Seyedabrishami, S.: Estimating time-dependent origin–destination demand from traffic counts: extended gradient method. Transp. Lett. **7**(4), 210–218 (2015)
25. Tom, V., Mohan, S.: Transit route network design using frequency coded genetic algorithm. J. Transp. Eng. **129**(2), 186–195 (2003)
26. Yang, Z., Yu, B., Cheng, C.: A parallel ant colony algorithm for bus network optimization. Comput. Aided Civ. Infrastruct. Eng. **22**(1), 44–55 (2007)
27. Zarrinmehr, A., Saffarzadeh, M., Seyedabrishami, S.: A local search algorithm for finding optimal transit routes configuration with elastic demand. Int. T. Oper. Res. (2016). doi:10.1111/itor.12359
28. Zarrinmehr, A., Saffarzadeh, M., Seyedabrishami, S., Nie, Y.M.: A path-based greedy algorithm for multi-objective transit routes design with elastic demand. Public Transp. **8**(2), 261–293 (2016)
29. Zhao, F., Zeng, X.: Optimization of user and operator cost for large-scale transit network. J. Transp. Eng. **133**(4), 240–251 (2007)

The Railway Network Design, Line Planning and Capacity Problem: An Adaptive Large Neighborhood Search Metaheuristic

David Canca[1(✉)], Alicia De-Los-Santos[2], Gilbert Laporte[3], and Juan A. Mesa[2]

[1] Department of Industrial Engineering and Management Science, Universidad de Sevilla, Seville, Spain
dco@us.es
[2] Departament of Applied Mathematics II, Universidad de Sevilla, Seville, Spain
{aliciasantos,jmesa}@us.es
[3] Canada Research Chair in Distribution Management, HEC Montréal, Montreal, Canada
gilbert.laporte@cirrelt.ca

Abstract. In this chapter, we propose a model for the Railway Network Design and Line Planning (RNDLP) problem, integrating the two classical first stages in the Railway Planning Process. The network design problem incorporates costs relative to the network construction, proposing a set of candidate lines. The line planning problem is in charge of determining optimal frequencies and consequently train operations, taking into account rolling stock, personnel and fleet acquisition costs. Both problems are intertwined because the line design influences the selection of frequencies and the corresponding fleet size. We consider the existence of an alternative transportation mode for each origin-destination pair in the network. In this way, the rapid railway mode competes against the alternative mode for a given certain demand, represented by a global origin-destination matrix. Passengers choose their transportation mode according to their own utility. Since the problem is computationally intractable for realistic size scenarios, we develop an Adaptive Large Neighborhood Search (ALNS) algorithm, which can handle the RNDLP problem. As illustration, the ALNS performance is demonstrated in an artificial instance using estimated data from literature.

Keywords: Railway Rapid Transit · Network design · Line planning · Adaptive Large Neighborhood Search

1 Introduction

Railway Rapid Transit (RRT) is a high-capacity public transport that usually operates on an exclusive right-of-way in urban areas. The general aim of a RRT

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_9

system is to improve the mobility of the population in big cities and metropolitan areas, but other purposes like decreasing private traffic congestion and pollution has become relevant nowadays. RRT systems are the most effective transportation mode since in a very short time they can carry considerable more people at a higher speed than other public transportation modes.

The RRT planning process is a very complex task involving strategic, tactical, operational and real-time decisions. Among these decisions are the selection of the location of stations and the connections between them, the itinerary and the frequency of the lines, the capacity of the trains, the timetable, the scheduling of the trains, the crew and other staff planning, and the management of delays and disruptions. Several agents are implicated in this process. They can be grouped into local authorities and transportation agencies, potential travellers and construction and operating companies. These problems are complex due to several factors: large-scale, uncertainty in data, different criteria to be taken into consideration because of the different viewpoints of the involved agents, competition with other transportation modes and computational complexity of the optimization counterpart problems. For these reasons a sequential approach was traditionally proposed for the whole RRT planning process.

Knowing the current mobility patterns and the predictions over a period of time, the first phase consists in choosing from an underlying network the location of the access and the egress points to the system, and the links between pairs of them. The traditional non-optimization methodology is based on the selection of a set of corridors, combining them, and choosing the best combination according to several criteria. However, this approach could eliminate in a very early phase good alignments that do not are considered any more. The second step consists in selecting the itinerary of the lines and the frequency of them in each period of time of the day, day of the week and season of the year. An issue closely related with the frequency is the determination of trains' capacity. The first idea is that the higher the capacity is, the lower the required frequency. However, in the presence of competing modes the relationship between capacity and frequency is non-linear and, even more, it becomes non-continuous. Based on the knowledge of the hourly demand the timetable is designed. Then, the last step of the sequential process consists in the scheduling of rolling stock and personnel.

In this work we integrate the two first steps adding the determination of the capacity of the trains. Thus, we consider the problem of simultaneously determining the line network design, the frequency of the lines and the capacity of the trains, considering also a competing transportation mode.

RRT network design [15,26], can be classified depending on whether a single or multiple alignments are to be planned, and whether they are completely new or extensions of already existing ones. The main criteria used to design rapid transit alignments are described in [24]. For the problem of locating one single alignment, a tabu search was proposed in [19] in order to maximize the population covered, [7] proposed a bicriterion model for the location of a rapid transit line minimizing construction cost and passenger travel time, [6] developed a two-phase heuristic for the problem of designing an alignment in a urban

context maximizing the population coverage, [30] presented a heuristic for the construction of a rapid transit alignment maximizing trip coverage, and [29] addressed the problem of locating a metro line in a historical city maintaining a minimum distance between the alignment to be designed and protected buildings. In [32] the problem of locating the stations, determining the headway and the fare of a transit line in a linear corridor when maximizing the profit is addressed using a heuristic algorithm.

Regarding the multiple alignment problem, [28] solved the rapid transit network design problem of maximizing the trip coverage by using simulated annealing (SA), in [23] the infrastructure railway network design problem as well as its robust version are solved by using a Greedy Randomized Adaptive Search Procedure (GRASP), [31,35] used the demand coverage as objective function. Based on an a priori geometric configuration, in [31] a metro network design is proposed under the criteria of maximizing population coverage and minimizing construction cost. With a similar methodology, but considering traffic capture instead of population coverage, [27] proposes a mixed integer tractable model formulation. Regarding papers dealing with the extension of existing networks we highlight [3], that proposed a model and a heuristic for the problem of expanding the infrastructure of a railway network.

The second phase of the railway planning process is line planning, in which a set of itineraries or lines is selected from the resulting network after the first phase or from a line pool. Moreover, the frequency of each line at each period of time is determined. Usually in this phase the capacity of the trains is supposed to be known. The line planning problem has been tackled in several papers, among them [8,13] propose branch-and-cut algorithms to select lines from a previously generated set of candidate lines (line pool). In [9] linear and non-linear integer programming models are proposed for the line planning problem with minimum cost. In [25] lines can have different halting patterns. [21] considers the problem of designing the frequencies of a regional metro with elastic demand by minimizing the total cost. For more information on line planning we refer to the recent reviews by [4,33,38].

As mentioned, we integrate the first two phases of the planning process taking also into account aspects of capacity and personnel cost. We assume there is a competing mode of transportation. The number of trips captured by the railway rapid transit system, which is determined by using a logit modal split function, depends on the difference between the utilities of both transportation modes. The utility of each pair of origin-destination demand for the rapid transit system depends on the fare, waiting time-cost at stations, in-vehicle time-cost and transfer time-cost between lines. The objective function is the profit, that is the difference between the revenue and the total cost. The revenue depends on the number of passengers captured by the railway network (RN). The total cost includes construction cost, fixed and variable operating costs, crew cost and the fleet acquisition cost as a function of the rolling-stock required to cover the demand captured by the RN system.

A related model was proposed in [11]. However, in this paper the variables defining the flow of passengers differ from those of [11]. Here, for each origin-destination pair, we use binary variables for the use of links instead of variables representing fractions of the demand, as in the former model. As consequence, the solutions proposed by this model are less favourable to the service provider but more favourable to the travellers than in the previous one, because in this situation, passengers follow the best path to reach their destination give rise to more loaded trains. In [12] we analyse the RNDLP problem in terms of complexity by comparing the exact solution of small instances using commercial solvers against an ALNS. In that work train capacity is determined by means of integer variables defined in order to obtain the number of carriages needed in each line. The modal split considers only the travel time as a key factor to select the railway or an alternative mode instead of passengers utility. Others aspects like railway fare do not influence the mode choice. As illustration, the ALNS is applied to designing a RRT network for the city of Seville.

Since the RNDLP problem is obviously NP-hard, a metaheuristic is needed to solve medium and large instances. In this paper an adaptive large neighborhood search (ALNS) metaheuristic is applied.

The remainder of the chapter is structured as follows. The next section introduces a non-linear mixed integer model for the RNDLP problem that simultaneously determines the most convenient network topology and the most appropriate set of lines, determining line frequencies and selecting a specific train model for each one in presence of an alternative transportation mode. Section 3 presents an ALNS algorithm designed to manage real-size instances of the RNDLP problem. Section 4 illustrates the computational performance of the ALNS considering different experiments in a medium-size artificially generated instance of the RNDLP using estimated values of time data from literature. The last section provides some conclusions and point out some still open questions.

2 Description of the RNDLP Problem

Consider a set $N = \{1, \ldots, n\}$ of potential nodes for locating stations and a set of arcs $A \subseteq N \times N$ representing potential connections between nodes. Both sets define a potential graph used as a basis for the building of the railway rapid transit network. We define the edge set $E = \{\{i, j\} : i, j \in N, i < j, (i, j) \text{ or } (j, i) \in A\}$. Thus, the underlying network is topologically described as a graph $G_E = G(N, E)$. The alternative transportation mode network (private car), competing with the railway rapid transit system, is represented by an undirected graph $G_{E'} = G(N, E')$. As is usual in the network design, there exists an upper bound C_{max} on the total construction cost of the railway network RN.

Let $W = \{w_1, \ldots, w_{|W|}\} \subseteq N \times N$ be the set of ordered origin-destination OD pairs $w = (o^w, d^w)$, where o^w and d^w represents the origin and destination of pair w, respectively. Without loss of generality, we assume that all trips occur between nodes belonging to the underlying network, that is, potential stations acts as demand origins and destinations. The expected number of passengers

g_w associated with each OD pair $w \in W$, as well as the corresponding utility U_w^{ALT} of pair w using the alternative mode are known. Let d_{ij} be the length of edge $\{i, j\}$, and λ, the average speed of trains measured in km/h. In order to obtain applicable results, we work with a discrete set $\mathcal{H}$ of headways which are measured in minutes. Note that if $h \in \mathcal{H}$, then the line *frequency* is equal to $60/h$, measured in number of trains per hour. We consider a parameter γ representing the maximum number of lines that can circulate on any edge of the network. This is a topological constraint frequently used in order to not over saturate some open tracks, which would result in excessively long headways (low frequencies). We assume known train capacities, according to different models $m \in M$ available in the market with capacities K_{train}^m. We consider all trains of a line operate at the same capacity, i.e., each line is operated by an specific model. The transfer time is considered as the sum of two terms: the time spent between platforms uc_i, which is supposed to be known, and the average waiting time for taking the next train of the line to transfer. The last term can be approximated as the average headway of the line to transfer. For each line, the main variable to be determined is the headway. As previously mentioned, no a priori line pool is defined, and since a constructive approach is followed, a lower and an upper bound, N_{min} and N_{max}, on the number of stations of each line are considered [10]. In order to compute the expected utility of pair w, parameters β_{tt}, β_{tr}, β_{wt}, are used to denote respectively the monetary cost of travel time, transfer time and waiting time at the initial station (see [21]). The index $\ell \in \mathcal{L}$ denotes each line, being $\mathcal{L}$ a set used to describe the possible lines and $|\mathcal{L}| = \mathcal{L}^{\max}$.

The considered objective function obj_{NET} maximizes the net profit, expressed as the difference between the revenue obj_{REV} and the total system cost obj_{SC}. In order to calculate the revenue, we consider two parameters: the first one, ξ, denoting the passenger fare and the second one, η, a public subsidy per trip [5,17]. The fare is also considered in the modal split model as part of the passengers' utility. We also consider a parameter t_{final} representing the time horizon employed to finance the construction of the network and to amortize the rolling stock investment. The number of years spent to build the network is denoted by $t_{initial}$. Obviously, $t_{initial} \leq t_{final}$. Furthermore, in order to obtain a realistic model, we have incorporated a discount rate r. Finally, we denote by h_{year} the number of hours during which a train operates in a year.
Therefore, the revenue can be expressed as:

$$obj_{REV} = \sum_{k=t_{initial}}^{t_{final}-1} \frac{1}{e^{rk}} \left[(\xi + \eta) \sum_{w \in W} g_w \cdot \text{ pp of pair } w \text{ using the RN} \right],$$

where *pp* represents the proportion of passengers.

The system cost obj_{SC} is composed of three main terms as follows (for more details see [11]).

$$obj_{SC} = obj_{BC} + obj_{OC} + obj_{FAC}.$$

The first term obj_{BC} corresponds to the cost for building stations and edges. This term is obtained considering two parameters: c_{ij} and c_i, corresponding to the cost of the built stretch on edge $\{i, j\}$ or the constructed station i, respectively. For the sake of simplicity, we assume that both, edge and station construction costs, are independent of the number of lines traversing edges or reaching stations. Concretely, this cost can be computed as:

$$obj_{BC} = \frac{1}{t_{final}} \sum_{k=0}^{t_{final}-1} \frac{1}{e^{rk}} \left[\sum_{\{i,j\} \in Eb} c_{ij} + \sum_{i \in Nb} c_i \right],$$

where Eb and Nb are the constructed-edge set and the constructed-station set.

The second term obj_{OC} is the operating cost, which includes fixed obj_{FOC} and variable costs obj_{VOC}:

$$obj_{OC} = obj_{FOC} + obj_{VOC}.$$

The term obj_{FOC} is related to maintenance and overheads of rails $ORlc_{ij}$ and stations $OStc_i$, measured in monetary units per year

$$obj_{FOC} = \sum_{k=t_{initial}}^{t_{final}-1} \frac{1}{e^{rk}} \left[\sum_{\{i,j\} \in Eb} ORlc_{ij} + \sum_{i \in Sb} OStc_i \right].$$

The variable cost takes into account the cost of operating trains as well as the crew cost $cost_{crew}$ per train and year, which is closely related with line headways. The operating cost $cost^m_{train}$ of a train of model m per unit of length, is given. Therefore, the variable cost is defined as:

$$obj_{VOC} = \sum_{k=t_{initial}}^{t_{final}-1} \frac{1}{e^{rk}} \left[(h_{year} \cdot \lambda) \sum_{\ell \in \mathcal{L}} FS_\ell (\sum_{m \in M} cost^m_{train}) - cost_{crew} \sum_{\ell \in \mathcal{L}} FS_\ell \right],$$

in which FS_ℓ is the required fleet of line ℓ, measured as the number of trains required per hour. This number can be expressed as a function of the headway (h_ℓ) and the length of the line:

$$FS_\ell = \lceil 120/(h_\ell \lambda) \sum_{i,j \in E_\ell} d_{ij} \rceil, \ \ell \in \mathcal{L}.$$

The last term in the system cost expression is the fleet acquisition cost obj_{FAC} which is determined by the investment in trains Inv^m_{train}. Note that the size of the rolling stock and the choice of specific train models will be a consequence of the passenger demand and the frequency of the lines. A parameter χ is used in order to include a set of reserve trains that will be needed as consequence of train maintenance operations. So, this cost can be represented as

$$obj_{FAC} = \sum_{k=t_{initial}}^{t_{final}-1} \frac{1}{e^{rk}} \left[\frac{\chi}{(t_{final} - t_{initial})} \sum_{\ell \in \mathcal{L}} FS_\ell (\sum_{m \in M} Inv^m_{train}) \right].$$

The line capacity imposes an upper bound on the maximum number of passengers that each line can carry per hour on each edge:

$$h_\ell \sum_{w \in W} g_w \cdot (pp\ of\ pair\ w\ traversing\ \{i,j\}\ in\ line\ \ell) \leq 60 \cdot K^m_{train}.$$

The modal split is described by means of a logit model in order to determine the volume of passengers captured by the RN, similarly to [34]. The logit function compares the expected passengers' utility U^{RN}_w with the corresponding utility U^{ALT}_w in the competing mode.

$$f^{RN}_w = \frac{1}{1 + e^{(\alpha - \beta(U^{ALT}_w - U^{RN}_w))}}, w \in W.$$

The expected passengers' utility has three terms that correspond to the waiting time at stations ($u^{RN,wt}_w$), in-vehicle time ($u^{RN,tt}_w$), and transfer times ($u^{RN,tr}_w$).

$$u^{RN,tt}_w = \frac{60}{\lambda} \sum_{\ell \in \mathcal{L}} \Big(\sum_{\{i,j\} \in E} (pp\ of\ w\ traversing\ \{i,j\} in\ line\ \ell) \cdot d_{ij} \Big), w \in W,$$

$$u^{RN,tr}_w = \sum_{\ell \in \mathcal{L}} \sum_{\ell' : \ell' \neq \ell} \sum_{i \in N} (pp\ of\ w\ transferring\ from\ \ell\ to\ \ell'\ in\ i) \cdot \Big(\frac{h_{\ell'}}{2} + uc_i \Big),\ w \in W,$$

$$u^{RN,wt}_w = \frac{1}{2} \sum_{\ell \in \mathcal{L}} \sum_{j : \{o^w, j\} \in E} h_\ell \cdot (pp\ of\ w\ traversing\ \{o^w, j\}\ \ in\ line\ \ell), w \in W.$$

Parameters β_{tt}, β_{tr} and β_{wt} represent the value of ridding, transferring and waiting time respectively (see [21]).

$$U^{RN}_w = \xi + \beta_{tt} \cdot u^{RN,tt}_w + \beta_{tr} \cdot u^{RN,tr}_w + \beta_{wt} \cdot u^{RN,tw}_w, w \in W.$$

Finally, the Railway Rapid Transit Network Design and Line Planning (RNDLP) consists of choosing the line to be constructed ($\mathcal{L}_c$, the stations of each line (N_ℓ), the edges of each line (E_ℓ), the headway of each line to be constructed (h_ℓ), and the kind of train to be assigned to each line (M_ℓ), such that the net profit is maximized:

$$max\ obj_{NET} := (obj_{REV} - obj_{SC}).$$

3 An Adaptive Large Neighborhood Search Metaheuristic for the RNDLP Problem

The RNDLP is an NP-hard non-linear problem. In [12] we formulated and solved a quite similar Network Design and Line Planning problem. The main differences with respect to the problem here treated are on the modal split sub-model and the line capacity constraints. In that paper we demonstrated the inability of the state-of-the-art commercial solvers to solve real size instances. Moreover,

we compared an ALNS (very similar to the one considered in this chapter) with fully linearised versions of the RNDLP for small instances (linear MIP solvers cannot manage large instances), reporting the superiority of the ALNS technique. As consequence, in order to tackle real size instances efficiently, for the current problem, we develop an ALNS metaheuristic which provides a powerful algorithmic framework capable of simultaneously handling the network design and line planning problems. The ALNS metaheuristic was introduced by [37] to solve variants of the vehicle routing problem. Basically, this algorithm tries to improve iteratively an initial solution using destroy and repair operators. The ALNS belongs to the category of large scale neighborhood search techniques defined in [1] but only examines a relatively low number of solutions. The main difference between the original work of [37] and the proposed by [39] concerns the probability of selecting an operator. Concretely, in our ALNS implementation, we consider several destroy and repair operators which are independently applied as in [14].

3.1 The ALNS Metaheuristic

As mentioned, the ALNS starts with an initial solution which is modified at each iteration by means of operators. This initial solution is formed by one line or a set of connected lines randomly defined but holding the problem constraints. We remark that a line is characterized by two different terminal stations, the intermediate stops or itinerary, the headway and the capacity of each train. In this situation, when the initial solution is defined, we can compute the amount of people travelling on the current network, the corresponding construction and operation costs and, consequently, the associated profit, which represents the quality of the solution. This calculation is done using a heuristic local search algorithm (for more details see [18]). Given a network configuration, the local search solves the problem of maximizing the net profit of a line plan by selecting the headway and the train model of each line, assuming that all passengers interested to travel in the RN can be transported.

At each iteration, one or two lines are randomly modified by means of an operator. An operator is a heuristic method which modifies a solution, keeping feasibility conditions. We consider six operators: two destroy operators, two repair operators and two operators combining destroy and repair operations. The repair operators insert new lines or extend existing ones. The destroy operators remove part of a line or a full line. Concerning the combined operators, the first one eliminates an existing line and then, inserts a new one. The second one removes part of a line and extends a randomly selected line.

In order to apply these operators, the lines are randomly selected and the operators are chosen with a certain probability which depends on their performance in the past iterations. As in other random algorithms, the acceptance of new solutions is controlled by means of a SA technique, diversifying the search in this way. The procedure ends when a certain stopping criterion is satisfied.

3.2 Outline of ALNS

In order to define the ALNS algorithm, we include the following elements:

- A set of heuristic methods (operators). The ALNS considers two kinds of operators, namely, repair and destroy operators. The repair operators build a new solution from a given solution while maintaining the feasibility whereas the destroy operators remove part of the solution.
- A set of variables in order to keep information about the best solution, the solution accepted by the SA and the current solution.
- A set of parameters whose values define the algorithm behaviour.
- A procedure to compute the quality of each solution.
- An initialization phase in which an initial solution and a set of initial values of the parameters are set in order to start the algorithm.

The key ideas of the ALNS can be described as follows:

Step 1: Initialization phase Construct a feasible solution and set the initial value of the parameters;
All operators have the same probability of being selected;
The stopping and the acceptance criteria are defined;
Step 2: Select an operator according to the roulette wheel mechanism;
Generate a new feasible solution and compute its profit;
Step 3: Compare the new profit against the stored profits. Apply the SA acceptance mechanism;
Keep information on the operator performance;
If a determined number of iterations are performed, update the probability of selecting each operator according to the SA results;
Step 4: Inspect the stopping criteria;
If the stop criteria is not met, go to Step 2, otherwise, the ALNS is finished.

Algorithm 1. Steps in the ALNS implementation.

3.3 The ALNS Components

The main components of our ALNS implementation are the following:

1. Neighborhood size
   Two networks G_{RN} and G'_{RN} are considered as neighbors if they have at most two different lines. Hence, the number of nodes and edges in the underlying network determines the size of the neighborhood.
2. Adaptive selection of operators
   The selection of a specific operator is made through a roulette wheel mechanism. Concretely, we associate a weight p_j with each operator j. This weight measures how well the operator has performed in the past iterations. Assuming h operators, the probability of selecting the operator m is $p_m / \sum_{j=1}^{h} p_j$.

3. Adaptive weight adjustment
   The weights p_j are updated at each iteration according to the quality of solutions. At the beginning of the ALNS execution, all weights are fixed to one and as a consequence, all operators have the same probability. At each iteration, once an operator j has been selected and applied, its score σ_j may be increased using three parameters θ_1, θ_2 or θ_3 with the next meanings:
   θ_1 : the new solution G_{RN} is better than the best global solution G_{best};
   θ_2 : the new solution G_{RN} is better than the incumbent solution G_{curr};
   θ_3 : G_{RN} is worse than the incumbent solution but it is still accepted.
   Obviously, the better the solution is, the higher the score is, i.e., $\theta_1 \geq \theta_2 \geq \theta_3$.

   After each block of s iterations, the performance of each operator j is observed and its weight is updated using the expression

$$p_j := \begin{cases} p_j & \text{if } o_j = 0 \\ (1-\varepsilon)p_j + \varepsilon\sigma_j/\nu_j o_j & \text{if } o_j \neq 0, \end{cases}$$

   where $\varepsilon \in [0,1]$ is a parameter called the reaction factor which allows controlling how quickly the weight adjustment algorithm reacts to changes in the scores. The parameter o_j is used for controlling the number of times operator j is used in the incumbent s iterations. The factor $\nu_j \geq 1$ represents the computational effort required by the operator.

   Finally, once all weights have been updated, all scores are reset to zero in order to store the performance information for the next block of iterations.
4. Acceptance and stopping criteria
   Our acceptance criterion is based on SA. We consider a standard exponential function to describe the SA, which uses two parameters: the current temperature $T_{start} > 0$ and the cooling rate $0 < \phi < 1$. At the beginning of the ALNS implementation, the temperature starts with T_{start} and after certain number of iterations, it is decreased (cooled) using the cooling rate ϕ ($T = T_{start} \cdot \phi$). The parameter T_{start} may be computed by inspecting the initial solution. In [37] T_{start} is defined in the way that a solution 5% worse than the initial solution has 50% probability of being accepted. Let G_{RN} be the current solution and $obj_{NET}(G_{RN})$ be the corresponding objective value, then, a new neighbor solution G'_{RN} is accepted if $obj_{NET}(G'_{RN}) > obj_{NET}(G_{RN})$ and is accepted with probability $e^{(obj_{NET}(G'_{RN}) - obj_{NET}(G_{RN}))/T}$ otherwise. In our case, if the difference between $obj_{NET}(G'_{RN}) - obj_{NET}(G_{RN})$ is less than ϱ% of $obj_{NET}(G_{RN})$, the acceptance probability is 0.5, i.e.,

$$e^{(obj_{NET}(G'_{RN}) - obj_{NET}(G_{RN}))/T_{start}} = 0.5$$

   or, equivalently,

$$T_{start} = (obj_{NET}(G'_{RN}) - obj_{NET}(G_{RN}))/\ln(0.5).$$

   The parameter ϱ is selected by the user.

With respect to the stopping criteria, we distinguish three different methods to control the end of the execution: the maximum number of iterations φ is reached, the final temperature T_{final} is reached or the running time exceeds a user-controlled threshold.

3.4 The ALNS Operators

The proposed ALNS considers six operators, defined as follows:

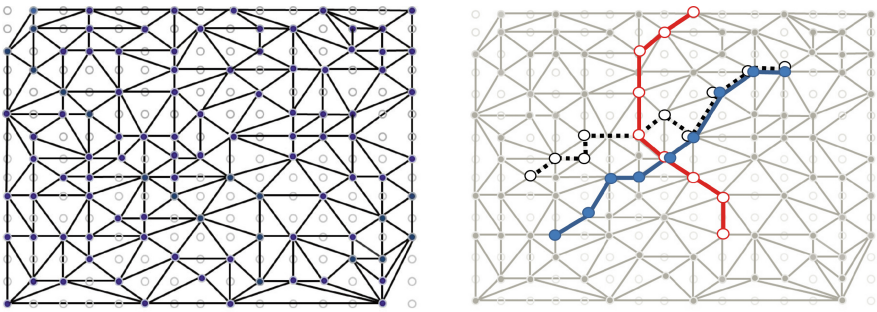
- Inserting-line operator. This operator aims at inserting a new line in the RN. To this end, two nodes representing the terminal stations, are randomly selected from the underlying network. Then, a shortest path connecting these stations in the underlying network is computed. This path configures the itinerary of the new line if the lower and upper bounds on the number of nodes of a line is fulfilled. Otherwise, a different couple of nodes are selected and the procedure above described is repeated. Once the itinerary have been set, the construction cost as well as the fixed operating cost can be computed. Since each operator has to ensure feasibility, a line is not inserted if
  - there exists an edge in the itinerary exceeding the upper bound imposed on the number of possible lines connecting each pair of adjacent nodes,
  - the itinerary is part of other existing line in the current RN solution,
  - the itinerary contains an existing line,
  - it is not connected with the existing lines in the RN.

  If finally the line is inserted, the corresponding profit is computed by means of a local search heuristic (see [18]).
- Extending-line operator. This heuristic randomly extends an existing line. First, the operator randomly choose a line and then, randomly selects the position (at the beginning or at the end in the itinerary) in order to extend the selected line. Once the line and the terminal station have been selected, a node (not belonging to the selected line) of the underlying network is randomly selected. A shortest path between this node and the terminal station of the line is computed. The itinerary of ℓ is extended according to the resulting path. A line is not extended if it reaches the maximum number of permitted nodes or there exists an edge in the itinerary exceeding the upper bound imposed on the number of possible lines connecting each pair of adjacent nodes. Finally, if the line ℓ is extended, its profit is computed using the local search heuristic.
- Delete-line operator. The delete-line operator randomly removes a line from the current RN provided the network is connected.
- Delete-part-line operator. This operator randomly selects a line ℓ to be partially removed. To do this, an intermediate node of the itinerary of ℓ and a terminal station of ℓ are randomly selected. The sub-path between both nodes is removed from ℓ. The lower bound on the number of nodes of ℓ is inspected. If the line is contained in an existing line or the network becomes unconnected after eliminating the sub-path, the removal is not considered.

- Delete-part-line and Extending-line operator. This method applies a delete-part-line and an extending-line operators in the same iteration. In case that the delete-part-line can be applied, the extending-line is later used. As can be observed, both operators work independently and, therefore, the selected line can be different for each one.
- Delete-line and Inserting-line operator. The idea of this operator consist on removing a line by means of the delete-line operator and then, if possible, apply the insert-line operator with the aim of adding a new line. As the reader can note, this method replaces a line with another line.

4 Computational Experiments

In this section, in order to show the performance of the proposed ALNS algorithm, we conduct a set of computational experiments on a medium-sized artificially generated network. The network contains 100 nodes (potential stations) and 275 links (edges). The node set was randomly selected from a 15×15 square grid with 225 nodes covering a surface of 14^2 km^2. The coordinates of each node were randomly chosen by considering a uniform distribution $U(-0.5, 0.5)$ around each coordinate (x,y) of the selected nodes in the grid, that is, by using the intervals $(x - 0.5, x + 0.5)$ and $(y - 0.5, y + 0.5)$. The edge set was then defined by using the Voronoi diagram, linking adjacent nodes and avoiding edge crossings (see Fig. 1). The length of each edge was computed as the Euclidean distance between its adjacent nodes. In order carry out the bed of experiments, an arbitrary initial solution was defined as depicted in Fig. 1.



(a) Artificial generated network (b) Initial solution

Fig. 1. Experimental scenario and initial solution

With respect to the passengers' demand, for each one of the 9900 OD pairs (100×99) among the different potential stations, the expected number of trips was obtained following a discrete uniform distribution $U(0, 43)$, given rise to an hourly OD demand matrix with a total of $107,269$ trips.

As explained previously, the modal split is described through a logit function. To this end, we need to introduce the utility associated to the competing mode (private car) as well as the associated to the railway system. Basically, the utility U_w^{RN} of using the RN, is compared pair by pair with the utility of the private car, U_w^{ALT}. As the reader may note in the following expression, U_w^{RN} is expressed in terms of monetary costs:

$$U_w^{RN} = \tau + \beta_{tt} \cdot u_w^{RN,tt} + \beta_{tr} \cdot u_w^{RN,tr} + \beta_{wt} \cdot u_w^{RN,wt},\ w \in W.$$

All parts of the trip (access, waiting, riding, transfer) were estimated in terms of time assuming a commercial speed of 40 km/h, and converted into monetary values using parameters β_{tt}, β_{tr} and β_{wt}. As a consequence, U_w^{RN} is a function of the fare and the set of edges and lines selected to perform the trip as well as the line frequencies. In this illustration, the values of time are taken from the work of [21], as shown in Table 1.

The utility of the alternative mode, U_w^{ALT} for each pair $w \in W$ is described as follows:

$$U_w^{ALT} = fc + d'_{o^w,d^w} \cdot vc + d'_{o^w,d^w} \cdot \frac{60}{v} \cdot tvc + Pt \cdot tvp + pc,\ w \in W.$$

where:

- fc, denotes the fixed cost per trip corresponding to the private car mode.
- v, represents the average speed.
- vc, defines the variable cost per km.
- tvc, denotes the unitary value of travel time.
- Pt, corresponds to the average time needed to park at destination.
- tvp, is the unitary value of parking time.
- pc, is the average cost of parking per trip.
- d'_{o^w,d^w}, corresponds to the modified Euclidean distance between the origin and destination of pair w.

In particular, each Euclidean OD distance has been multiplied by a factor 1.2, representing the impossibility of follow straight lines connecting the origin and destination of the pair. The specific values used in our experiments are included in Table 1. We want to remark that the goal of these experiments is to show the performance of the ALNS rather than the proper estimation of all these scenario-dependent parameters. Interested readers in modelling private car and public transport preferences can consult the books [20,40] and the work of [2,16,21,36,41]. [41] shows a general analysis of the value of time and [16,21] present the specific case of modelling the value of time in railway systems.

Costs concerning operation and rolling stock acquisition are taken from [22], considering the family of trains "Civia", recently used by the National Spanish Railways Service Operator (RENFE) for commuter/regional railway passengers transportation in Spain. Concretely, we have considered three different models:

Civia-463, Civia-464 and Civia-465, with 607, 832 and 997 passenger capacities (seating and standing) respectively. The remaining model parameters are shown in Table 1.

Concerning the ALNS execution, in order to select the most appropriate values for the algorithm parameters, we run consecutively the same instance three times until the coefficient of variation (the sample average divided by the sample standard deviation) reached a value of 0.1, which indicates a strong stability

Table 1. Input parameters for the computational experiments.

Parameters		
Name	Description	Value
t_{final}	Years to recover the purchase	20
h_{year}	Number of operative hours per year	6935
$t_{initial}$	Number of years spent to build the network	10
m	Model of train	463, 464, 465
$cost^{m}_{train}$	Costs for operating one train model m per kilometer [€/km]	$3, 3.1, 3.2$
$cost_{crew}$	Per crew and year for each train [€/ year]	$75 \cdot 10^3$
Inv^{m}_{train}	Purchase cost of one train Civia in €	$4.4 \cdot 10^6, 5.2 \cdot 10^6, 5.9 \cdot 10^6$
K^{m}_{train}	Capacity of each type of train [Passengers]	$607, 832, 997$
$\mathcal{H}$	Possible headways [min]	{3, 4, 5, 6, 10, 12, 15, 20}
N_{min}	Minimum number of stations for each line	5
N_{max}	Maximum number of stations for each line	16
β_{tr}	Perceived value of time spent transferring in €/min	0.25
β_{wt}	Perceived value of time for waiting at the origin station in [€/min]	0.25
β_{tt}	Perceived value of time for riding in train in [€/min]	0.083
fc	Fixed cost per trip corresponding to the private car mode [€]	1.75
v	Average speed of the private car mode [Km/h]	60.0
vc	Variable cost per km corresponding to the alternative mode [€/km]	0.12
tvc	Unitary value of time travelling in the alternative mode [€/min]	0.05
Pt	Average time needed to park at destination [min]	10.0
tvp	Unitary value of time corresponding to the parking time [€/min]	0.25
$\mathcal{L}^{\max}$	The maximum number allowed in the network	6
pc	Average cost of parking per trip [€]	Table 2

of the algorithm. Then, using the tuned ALNS, a set of 24 experiments were generated by varying fare, subsidy and the ticket price for the alternative mode, as described in Table 2. The first twelve experiments (first block) correspond to a constant value of the sum $fare + subsidy$ equal to 3. For the last twelve experiments (second block) $fare + subsidy = 2$.

Table 2. Definition of computational experiments.

Summary of experiments				
Exp. number	Fare+subsidy	Fare	subsidy	pc
1	3	0.5	2.5	0.5
2	3	0.75	2.25	0.5
3	3	1	2	0.5
4	3	1.5	1.5	0.5
5	3	0.5	2.5	0.75
6	3	0.75	2.25	0.75
7	3	1	2	0.75
8	3	1.5	1.5	0.75
9	3	0.5	2.5	1
10	3	0.75	2.25	1
11	3	1	2	1
12	3	1.5	1.5	1
13	2	0.5	1.5	0.5
14	2	0.75	1.25	0.5
15	2	1	1	0.5
16	2	1.5	0.5	0.5
17	2	0.5	1.5	0.75
18	2	0.75	1.25	0.75
19	2	1	1	0.75
20	2	1.5	0.5	0.75
21	2	0.5	1.5	1
22	2	0.75	1.25	1
23	2	1	1	1
24	2	1.5	0.5	1

The ALNS has been coded using Java and run in a computer with 8 Gb of RAM memory and a 2.8 Ghz CPU. In all cases a value of $\gamma = 3$ (line multiplicity) was considered. The cooling factor was fixed to 0.9997 and the reaction factor to 0.7. The final temperature T_{final} was set to 0.01. The score parameters θ_1, θ_2 and θ_3 were set to 10, 5 and 2, respectively. The acceptance neighborhood

parameter ϱ was set to 33%, as in [14]. The rest of model parameters are set as in [11]. In all cases, the time limit (3600s) acts as stopping criterion.

Detailed information on these solutions are collected in Tables 3 and 4. In Table 3 column "Exp. number" refers to the tested experiment. The second column shows the number of captured passengers; the third one represents the number of lines; the fourth column reports (in order) the headway corresponding to each line and the last column contains the model of train selected for each line. As in Table 3, the first column of Table 4 shows the experiment number. Afterwards, column by column, we report the computation time needed to obtain the best solution, the net profit, the network revenue, the operating cost, the building cost (including the fixed operating cost), the crew cost and finally the fleet acquisition cost.

Table 3. Computational results for the ALNS metaheuristic.

Exp. number	Demand	No. of lines	Headway	Train models
1	39493	6	[10, 10, 10, 10, 10, 10]	[463, 465, 463, 464, 464, 464]
2	36377	6	[5, 5, 5, 5, 5, 5]	[463, 465, 465, 465, 465, 465]
3	44682	6	[6, 6, 6, 6, 6, 6]	[463, 463, 463, 463, 463, 463]
4	28812	6	[15, 15, 15, 15, 15, 15]	[465, 464, 463, 463, 464, 463]
5	29547	6	[6, 10, 10, 10, 10, 10]	[464, 464, 463, 465, 464, 464]
6	28185	6	[10, 10, 10, 10, 10, 10]	[465, 463, 463, 463, 463, 464]
7	41726	6	[5, 5, 5, 5, 5, 5]	[465, 464, 463, 463, 463, 463]
8	30851	6	[12, 12, 12, 12, 12, 12]	[465, 465, 463, 464, 463, 465]
9	30851	6	[10, 10, 10, 10, 10, 10]	[464, 464, 463, 465, 463, 463]
10	31208	6	[10, 10, 10, 10, 10, 10]	[464, 465, 463, 463, 464, 464]
11	34047	6	[10, 10, 10, 10, 10, 10]	[465, 463, 463, 463, 464, 463]
12	36690	6	[10, 10, 10, 10, 10, 10]	[464, 464, 465, 465, 464, 464]
13	37201	6	[10, 10, 10, 10, 10, 10]	[465, 464, 464, 464, 464, 463]
14	37373	6	[6, 6, 6, 6, 6, 6]	[464, 463, 463, 463, 463, 463]
15	32668	6	[12, 12, 12, 12, 12, 12]	[465, 464, 464, 463, 464, 465]
16	31603	6	[12, 12, 12, 12, 12, 12]	[465, 463, 463, 464, 463, 464]
17	36520	6	[10, 10, 10, 10, 10, 10]	[465, 465, 464, 464, 463, 463]
18	40219	6	[10, 10, 10, 10, 10, 10]	[464, 463, 465, 463, 464, 463]
19	39949	6	[10, 10, 10, 10, 10, 10]	[465, 464, 463, 463, 465, 464]
20	30408	6	[12, 12, 12, 12, 12, 12]	[465, 465, 464, 463, 463, 464]
21	43361	6	[6, 6, 6, 6, 6, 6]	[463, 463, 463, 463, 464, 463]
22	28108	6	[10, 10, 10, 10, 10, 10]	[463, 463, 464, 465, 463, 463]
23	27698	6	[15, 15, 15, 15, 15, 15]	[464, 463, 464, 463, 464, 465]
24	36070	6	[6, 6, 6, 6, 6, 6]	[464, 464, 463, 463, 463, 463]

Table 4. Computational results for the ALNS metaheuristic.

Exp. number	Time	Profit	Revenue	Operating cost	Building cost	Crew cost	Acquisition cost
1	3192.90	3.70E9	7.11E9	2.91E8	2.86E9	7.65E7	1.82E8
2	3437.87	3.29E9	6.55E9	4.71E8	2.38E9	1.26E8	2.75E8
3	3481.97	3.85E9	8.04E9	5.08E8	3.26E9	1.37E8	2.82E8
4	3206.59	2.74E9	5.19E9	1.54E8	2.16E9	4.05E7	9.51E7
5	3600.82	2.34E9	5.32E9	2.82E8	2.44E9	7.36E7	1.79E8
6	2017.74	2.48E9	5.07E9	2.05E8	2.20E9	5.44E7	1.25E8
7	3561.55	3.70E9	7.51E9	5.13E8	2.85E9	1.36E8	3.09E8
8	3212.30	3.49E9	6.26E9	1.92E9	2.39E9	4.98E7	1.24E8
9	3543.65	2.69E9	5.55E9	2.32E8	2.43E9	6.10E7	1.43E8
10	2342.42	2.88E9	5.62E9	2.23E8	2.31E9	5.87E7	1.37E8
11	2428.17	3.25E9	6.13E9	2.28E8	2.45E9	6.07E7	1.35E8
12	3519.16	3.57E9	6.61E9	2.49E8	2.56E9	6.43E7	1.64E8
13	2091.95	1.27E9	4.46E9	2.80E8	2.66E9	7.32E7	1.79E8
14	3529.52	1.09E9	4.49E9	3.92E8	2.67E9	1.05E8	2.24E8
15	3167.72	1.09E9	3.92E9	2.02E8	2.44E9	5.24E7	1.33E8
16	3467.02	1.19E9	3.79E9	1.78E8	2.27E9	4.69E7	1.11E8
17	2713.09	1.19E9	4.38E9	2.49E8	2.72E9	6.50E7	1.60E8
18	3227.82	1.75E9	4.83E9	2.48E8	2.60E9	6.55E7	1.52E8
19	3484.67	1.74E9	4.80E9	2.50E8	2.58E9	6.55E7	1.59E8
20	964.50	7.67E8	3.65E9	1.98E8	2.51E9	5.21E7	1.25E8
21	3651.22	1.81E9	5.20E9	4.19E8	2.62E9	1.13E8	2.39E8
22	3552.36	8.59E8	3.37E9	2.08E8	2.13E9	5.50E7	1.25E8
23	2571.30	9.21E8	3.32E9	1.32E8	2.15E9	3.46E7	8.19E7
24	3631.94	1.32E9	4.33E9	3.73E8	2.32E9	9.98E7	2.17E8

Figure 2 resumes the different terms concerning the objective function for all the experiments. As expected, incomes are bigger for the first twelve instances. Experiments number 3 and 21 produce the best results in terms of profit for the two different blocks of experiments respectively. The structure of costs depicted in the figure denotes the complexity of the RNDLP problem. Specifically, it could be expected a higher captured demand for experiment 1 in comparison with experiment 3 as consequence of a minor fare (the RN becomes more attractive for passengers). However, as revenue is lower, a lower income gives rise to a less extensive network, given service to a minor number of O-D pairs, which translates in a minor captured demand (see Fig. 3). Moreover, the best solution (experiment 3) also corresponds to the higher construction costs and to one of the experiments with high operation costs (with exception of experiment 7).

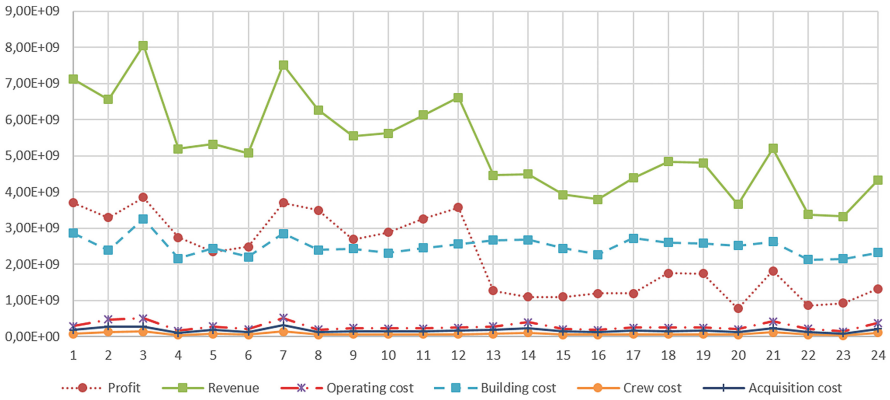


Fig. 2. Representation of costs for the set of experiments.

In order to obtain a deeper insight in the best solutions, experiments 3 and 21, Fig. 4, representing Network Profit versus time, illustrate the convergence process of the ALNS in both cases. As the reader may note, the algorithm attain a 70% of the best profit in approximately 1000 s for the experiment 3 and a 50% in 2400 s for the experiment 21. In general, the convergence in the second block of experiments (with lower sum of $fare + subsidy$) is slower than in the first one. Figure 5 shows the topology of both solutions.

Figure 6 presents a comparison between the topology of both networks, giving an idea about the coverage of the two solutions. This picture also illustrates the complexity of making the most appropriate design. As depicted, Experiment 3 covers a greater number of nodes which translates in a higher number of served O-D pairs.

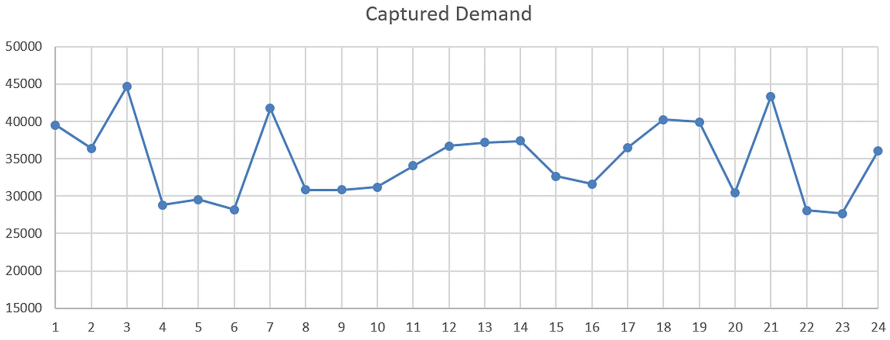
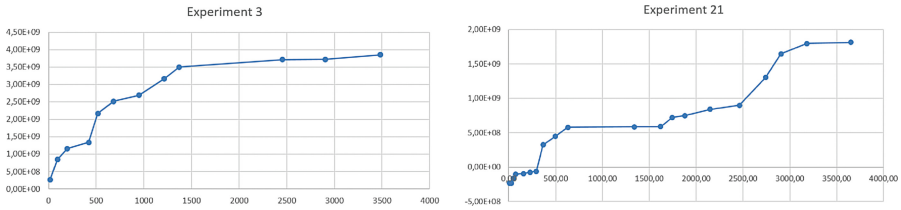
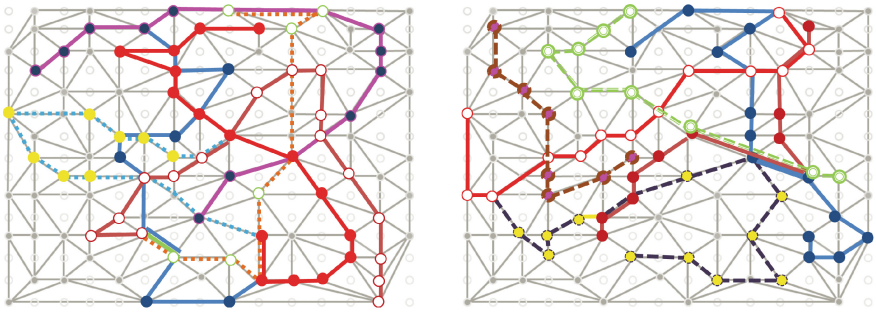


Fig. 3. Captured demand for the set of experiments.



(a) Profit evolution for Exp. n. 3 (b) Profit evolution for Exp. n. 21

Fig. 4. Improvement of the ALNS with time



(a) Network of Exp. n. 3 (b) Network of Exp. n. 3

Fig. 5. The RN solutions

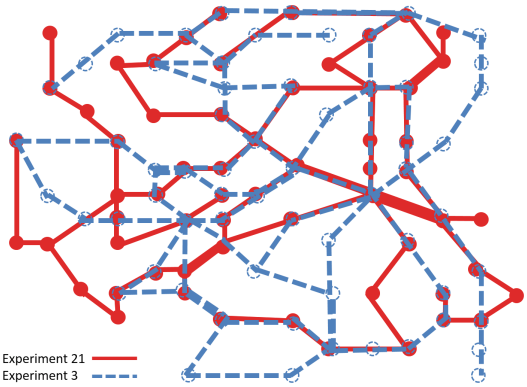


Fig. 6. Comparison of coverage Exp. 1 and 23.

5 Conclusions

In this chapter, we have presented a MINLP formulation for the integrated Railway Rapid Transit Network Design and Line Planning (RNDLP) problem, considering simultaneously the two first stages of the railway planning process. The model incorporates all relevant costs, including the network building, fleet acquisition, train operation, rolling stock and crew costs, taking into account the temporal nature of the network deployment. The proposed model reflects both, the service provider and the user points of view. The line planning phase includes line frequency and train model selection decisions, determining the final level of service for each line. The proposed formulation considers the existence of an alternative mode (private car) providing service to all origin-destination pairs. Users select the railway or the alternative mode by comparing its respective utilities by means of a logit probabilistic sub-model that include RN fare as a decision parameter. In order to solve realistic size instances of the RNDLP problem, an Adaptive Large Neighborhood Search (ALNS) metaheuristic is proposed. The ALNS performance was assessed by means of a parametric analysis in a medium-size artificial generated network. As reported in Sect. 4, the ability of the ALNS in obtaining good solutions within short computation times is demonstrated. Computational experiments provide yield useful insights into the different interactions among all the aspects related to this long-term and complex decision problem.

Nowadays, the integration of different stages of the railway planning process continues being a challenge and further research is needed in order to model and solve real and big instances. Although several attempts have been made to solve rolling stock and scheduling problems jointly, there still exists a gap in integrating strategic and tactical problems. This chapter contributes to this end.

Acknowledgements. This research work was partially supported Ministerio de Economía y Competitividad (Spain)/FEDER under grant MTM2015-67706-P and by the Canadian Natural Sciences and Engineering Research Council under grant 2015-06189.

References

1. Ahuja, R.K., Ergun, Ö., Orlin, J.B., Punnen, A.P.: A survey of very large-scale neighborhood search techniques. Discret. Appl. Math. **123**(1), 75–102 (2002)
2. Beirão, G., Sarsfield-Cabral, J.A.: Understanding attitudes towards public transport and private car: a qualitative study. Transp. Policy **14**(6), 478–489 (2007)
3. Blanco, V., Puerto, J., Ramos, A.: Expanding the Spanish high-speed railway network. Omega **39**(2), 138–150 (2011)
4. Borndörfer, R., Grötschel, M., Pfetsch, M.E.: Models for line planning in public transport. In: Hickman, M., Mirchandani, P., Voß, S. (eds.) Computer-Aided Systems in Public Transport. LNE, vol. 600, pp. 363–378. Springer, Heidelberg (2008)
5. Brueckner, J.K.: Transport subsidies, system choice and urban sprawl. Reg. Sci. Urban Econ. **35**(6), 715–733 (2005)

6. Bruno, G., Gendreau, M., Laporte, G.: A heuristic for the location of a rapid transit line. Comput. Oper. Res. **29**, 1–12 (2002)
7. Bruno, G., Ghiani, G., Improta, G.: A multi-modal approach to the location of a rapid transit line. Eur. J. Oper. Res. **104**(2), 321–332 (1998)
8. Bussieck, M., Kreuzer, P., Zimmermann, U.: Optimal lines for railway systems. Eur. J. Oper. Res. **96**, 54–63 (1997)
9. Bussieck, M.R., Lindner, T., Lübbecke, M.E.: A fast algorithm for near cost optimal line plans. Math. Method Oper. Res. **59**(2), 205–220 (2004)
10. Canca, D., Barrena, E., Algaba, E., Zarzo, A.: Design and analysis of demand-adapted railway timetables. J. Adv. Transp. **48**(2), 119–137 (2014)
11. Canca, D., De-Los-Santos, A., Laporte, G., Mesa, J.A.: A general rapid network design, line planning and fleet investment integrated model. Ann. Oper. Res. **246**(1), 127–144 (2016)
12. Canca, D., De-Los-Santos, A., Laporte, G., Mesa, J.A.: An adaptive neighborhood search metaheuristic for the integrated railway rapid transit network design and line planning problem. Comput. Oper. Res. **78**, 1–14 (2017)
13. Claessens, M.T., van Dijk, N.M., Zwaneveld, P.J.: Cost optimal allocation of rail passenger lines. Eur. J. Oper. Res. **110**(3), 474–489 (1998)
14. Coelho, L.C., Cordeau, J.-F., Laporte, G.: The inventory-routing problem with transshipment. Comput. Oper. Res. **39**(11), 2537–2548 (2012)
15. Cordeau, J.-F., Toth, P., Vigo, D.: A survey of optimization models for train routing and scheduling. Transp. Sci. **32**(4), 380–404 (1998)
16. Corman, F., D'Ariano, A.: Assessment of advanced dispatching measures for recovering disrupted railway traffic situations. Transp. Res. Rec. J. Transp. Res. Board **2289**, 1–9 (2012)
17. Cowie, J.: The british passenger railway privatisation: conclusions on subsidy and efficiency rom the first round of franchises. J. Transp. Econ. Policy **43**(1), 85–104 (2009)
18. De-Los-Santos, A., Laporte, G., Mesa, J., Perea, F.: Simultaneous frequency and capacity setting in uncapacitated metro lines in presence of a competing mode. Transp. Res. Procedia **3**, 289–298 (2014)
19. Dufourd, H., Gendreau, M., Laporte, G.: Locating a transit line using tabu search. Locat. Sci. **4**, 1–19 (1996)
20. Erlander, S.B.: Cost-Minimizing Choice Behavior in Transportation Planning: A Theoretical Framework for Logit Models. Advances in Spatial Science. Springer Science & Business Media, Heidelberg (2010)
21. Gallo, M., Montella, B., D'Acierno, L.: The transit network design problem with elastic demand and internalisation of external costs: an application to rail frequency optimisation. Transp. Res. Part C: Emerg. Technol. **19**(6), 1276–1305 (2011)
22. García, A., Martín, M.P.: Diseño de los vehículos ferroviarios para la mejora de su eficiencia energética. In: Monografías ElecRail, vol. 6. Fundación de los Ferrocarriles Españoles (2012)
23. García-Archilla, B., Lozano, A., Mesa, J.A., Perea, F.: GRASP algorithms for the robust railway network design problem. J. Heuristics **19**(2), 1–24 (2013)
24. Gendreau, M., Laporte, G., Mesa, J.A.: Locating rapid transit lines. J. Adv. Transp. **29**(2), 145–162 (1995)
25. Goossens, J.-W., van Hoesel, S., Kroon, L.: On solving multi-type railway line planning problems. Eur. J. Oper. Res. **168**(2), 403–424 (2006)
26. Guihaire, V., Hao, J.: Transit network design and scheduling: a global review. Transp. Res. Part A: Policy Pract. **42**(10), 1251–1273 (2008)

27. Gutiérrez-Jarpa, G., Obreque, C., Laporte, G., Marianov, V.: Rapid transit network design for optimal cost and origin-destination demand capture. Comput. Oper. Res. **40**(12), 3000–3009 (2013)
28. Kermanshahi, S., Shafahi, Y., Mollanejad, M., Zangui, M.: Rapid transit network design using simulated annealing. In: 12th World Conference on Transportation Research, pp. 1–15, Lisbon, Portugal (2010)
29. Laporte, G., Mesa, J.A., Ortega, F., Pozo, M.: Locating a metro line in a historical city centre: application to Sevilla. J. Oper. Res. Soc. **60**, 1462–1466 (2009)
30. Laporte, G., Mesa, J.A., Ortega, F., Sevillano, I.: Maximizing trip coverage in the location of a single rapid transit alignment. Ann. Oper. Res. **136**(1), 49–63 (2005)
31. Laporte, G., Pascoal, M.: Path based algorithms for metro network design. Comput. Oper. Res. **62**, 78–94 (2015)
32. Li, Z.-C., Lam, W.H.K., Wong, S.C., Sumalee, A.: Design of a rail transit line for profit maximization in a linear transportation corridor. Transp. Res. Part E: Logist. Transp. Rev. **48**(1), 50–70 (2012)
33. López-Ramos, F. Conjoint design of railway lines and frequency setting under semi-congested scenarios. Ph.D. thesis, Polytechnic University of Catalonia (2014)
34. Marín, A., García-Ródenas, R.: Location of infrastructure in urban railway networks. Comput. Oper. Res. **36**(5), 1461–1477 (2009)
35. Matisziw, T.C., Murray, A.T., Kim, C.: Strategic route extension in transit networks. Eur. J. Oper. Res. **171**(2), 661–673 (2006)
36. Paulley, N., Balcombe, R., Mackett, R., Titheridge, H., Preston, J., Wardman, M., Shires, J., White, P.: The demand for public transport: the effects of fares, quality of service, income and car ownership. Transp. Policy **13**(4), 295–306 (2006)
37. Ropke, S., Pisinger, D.: An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. Transp. Sci. **40**(4), 455–472 (2006)
38. Schöbel, A.: Line planning in public transportation: models and methods. OR Spectr. **34**(3), 491–510 (2011)
39. Shaw, P.: A new local search algorithm providing high quality solutions to vehicle routing problems. Technical report, University of Strathclyde, Glasgow (1997)
40. Tseng, Y.-Y.: Valuation of Travel Time Reliability in Passenger Transport, vol. 4390. Rozenberg Publishers, Amsterdam (2008)
41. Wardman, M.: Public transport value of time. Transp. Policy **11**, 363–377 (2004)

Diagnostic of the Balance and Equity of Public Transport for Tourists and Inhabitants

Maurici Ruiz[1(✉)] and Joana Maria Seguí-Pons[2]

[1] Servei de SIG i Teledetecció, Universitat de les Illes Balears, Palma, Spain
maurici.ruiz@uib.cat
[2] Departament de Geografia, Universitat de les Illes Balears, Palma, Spain
joana.segui-pons@uib.cat

Abstract. The significant development of tourism as an economic activity in some cities has led to the need to integrate potential demand for public transport generated by tourists into planning decisions. Public transportation is considered a critical factor of tourism competitiveness and plays a fundamental role in the promotion and maintenance of a sustainable tourist destination. Ensuring equity of public transport services requires specific methodological tools for diagnosis and optimization. There are few references regarding the need to integrate tourism demand in the planning of public transport of cities or the development of methods and techniques to facilitate this task. A method for evaluating potential public bus transportation demand including both residents and tourists is presented. The method uses geographic information systems and statistics, namely, the Gini coefficient, to diagnose imbalances between supply and demand and the degree of equity in transport services. The city of Palma (Spain) is selected as a case study. The results show a significant imbalance in the public transport of the city. There is a large concentration of tourist facilities in areas where the level of bus service is not sufficient for the demand of the resident population and tourists.

Keywords: Public transport · Equity · Tourism transport

1 Introduction

In recent years, the tourism sector has experienced extraordinary growth. The economic activity associated with tourism can reach between 5 and 9% of GDP (including its effects, direct, indirect and induced) [37]. The OECD average stands at just over 4% of GDP, but it accounts for 11.7% of the GDP in Spain and 44.8% of the GDP in the Balearic Islands [16]. Tourism is an expanding activity in the region, growing 3.3% per year. The displacement forecasts for 2030 foresee approximately 1.8 billion travellers [37]. Urban tourism (city tourism) has experienced a considerable rise in recent decades [36] and constitutes an economic activity with remarkable expectations. This fact has been encouraged by accessibility of low-cost airlines.

Evans and Shaw note that in recent decades, there has been a strong revival of tourism in the centre of cities, which presents a challenge for developers, operators and

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_10

local authorities [15]. This has been brought about by the rise in the supply of urban hotels in many cities and especially by the development of rent tourist platforms on the internet. It is easy to access a universal market of houses and apartments via the internet, eliminating intermediaries and simplifying rent procedures.

Infrastructure and transport services, shapers of territorial accessibility, are key factors in the competitiveness of cities as tourist destinations [35]. In this sense, public transport is essential to providing access to all facilities and services that urban tourists desire. Hall categorizes tourist mobility services into three groups: providing mobility and access to a destination area, providing mobility and access within an actual tourism attraction, and facilitating travel along a recreational route which is itself the tourism experience [21].

Negative externalities on the environment arising from tourism transport should be considered in planning tourist destinations, encouraging the use of lower polluting modes of travel. In this regard, encouraging the use of public transport by tourists is considered a fundamental strategic for ensuring the sustainability of tourism activities [35]. The increase in urban tourists can create certain stress, depending on the volume and seasonality, for the public transport system, which must provide services to tourists and the resident population.

Although the symbiotic relationships between transport and leisure are generally described as benign, analyses of public transport and leisure activities are scarcely treated. The scientific literature that analyses the integration of tourist transport with public transport planning is quite limited in journals of transport or tourism [1, 13, 15, 24, 26, 28–31, 33].

The planning and management of public transport in cities are usually performed through public or private companies that must reconcile the provision of services in an equitable manner with economic feasibility and benefits [7, 11, 17, 18, 21, 27, 32, 34].

The transport planning process should be based on an integrated analysis of the potential demand, considering the different needs of the population and the location of economic activities. This should be coordinated with the distribution of land uses and the urbanistic city model. Sustainable urban mobility plans abound in this regard and consider the use of public transport as a key element of planning.

In seasonal tourist destinations, such as the city of Palma (Mallorca, Balearic), some type of support for shuttle service is common during peak season (increased frequencies of certain lines, launching of new routes or creation of new stops). These are specific actions that seek to solve the problem temporarily but do not constitute structural integration strategies of transport tourism demand.

Sometimes the high profitability of some bus lines with the preferential use of tourism is a fundamental argument used by managers for its potentiation. This could involve deficiencies at other lines more oriented to responding to social demand. The main argument put forward is that the high effectiveness of touristic lines can equilibrate deficit lines and enable better performance of the service. This circumstance can generate inequity risks because the public transport needs of tourists are overvalued respect the social needs of the population of the city.

While on that subject, new residential tourism based on e-commerce and social networking is changing geographic patterns of tourist accommodation and also modifying the potential demand for public transport. We refer to the supply created by internet portals such as airbnb.com. In recent years, the deconcentration of supply and direct interweaving of tourists with resident have restructured transport demand models. Residential tourist, usually have less purchasing power than those staying at the regulated supply (hotels and apartments). They reach practically all the sites of the city and make use of public transport in a more homogeneous way. This softens the touristic profile of some bus lines, but the entire transport system is subjected to greater stress. As indicated by Dickinson, Robins and Lapko [12, 23] the car is the most common mode of transport used by tourists on the move, which leads to problems and externalities (congestion, parking stress, visual intrusion, noise and air quality).

Albalate et al. note that mobility is an essential question for tourists of large cities and that its planning and development have a direct effect on expanding the benefits of the tourism activities throughout the city. Tourism is a driving factor in the economic and social dynamism of a city and configures the transport demand. However, cities do not uniformly increase their transport service to give support to the tourism demand. In cities with traffic congestion problems, there is already high competence between tourism and residents in the use of public transport. In this case, tourism increases the stress of the service and generates externalities negatives [1]. In this regard, Dallen analyses a multipurpose model of tourism trips and leisure, which results in different patterns of demand and pressure on transport infrastructure [9].

Antoniou and Tyrinopoulos analyse the factors affecting the use of public transport in tourist areas by tourists and residents. They conclude that tourists base their choice on the basis of the frequency and the reliability of service, whereas residents value most other qualitative aspects. This also stresses the need to rethink policy analysis of transport demand and encourage efforts in marketing strategies, transparency and quality properly integrating tourist demand [3].

Le-Klähn indicates that public transport sometimes is promoted as a marketing strategy of tourist destinations, considering the preferences of tourists in destinations or in reducing fares [23–25]. However, Gronau and Kagermeier note that while the main focus of sustainable transport policy focuses on reducing motor vehicles and private vehicles and promoting an attractive public transport supply, that effort is not so clear when the problem is addressed for tourist transport. The main purpose is to build a supply adapted to the demand, considering all preferences and adopting a bottom-up approach [19]. Guiver et al. explains the need to adapt transport models to the needs of tourists and the important need to develop target group identification of transport needs (nature, family, sports-oriented, catchment areas, etc.) [20]. Lumsdon also focuses on this approach, stressing the need for bus transportation routes to incorporate the vision of tourists in their travel expectations [28].

This conceptual framework notes that the integration of tourism demand in the planning of public transport are still incipient. The public transport preferences of tourism are analysed and the negative externalities of not including tourism demand in a public transport system planning are identified.

In general, there is a need for further development of integrated methodologies that involve tourism demand in the planning of public transport and provide reasonable solutions within the framework of sustainable mobility.

In this paper, we carry out an integrated analysis of the potential demand for the public transport of residents and tourists. The relationship between the potential demand and the level of service is assessed by detecting geographic imbalances and calculating the level of equity using the Gini coefficient. Geographic information systems and statistical tools are combined in a specific diagnostic methodology. The city of Palma (Balearic Islands) is selected as a study area.

The questions we want to answer are as follows: What is the potential demand of public transport for residents and tourists? What is the level of public transport service in the city? Are there imbalances in public transport (i.e., in the supply-demand relationship)? Where are imbalances and inequalities located?

2 Case Study

The study case selected is the city of Palma. Palma is the capital of the Autonomous Community of the Balearic Islands (Spain). The planning and management of bus public transport of the city is developed by Empresa Municipal de Transports (EMT), which is a public company dependent of the city council.

The transport service of the city tries to cover the needs of 421,708 residents [22] in approximately 19,524 hectares. The municipality is divided into 88 neighbourhoods delimited by historical criteria. To simplify the territorial analysis carried out in this research, a grouping of neighbourhoods into seventeen homogenous urban areas that respond to a functional division criteria of the territory was performed (Fig. 1).

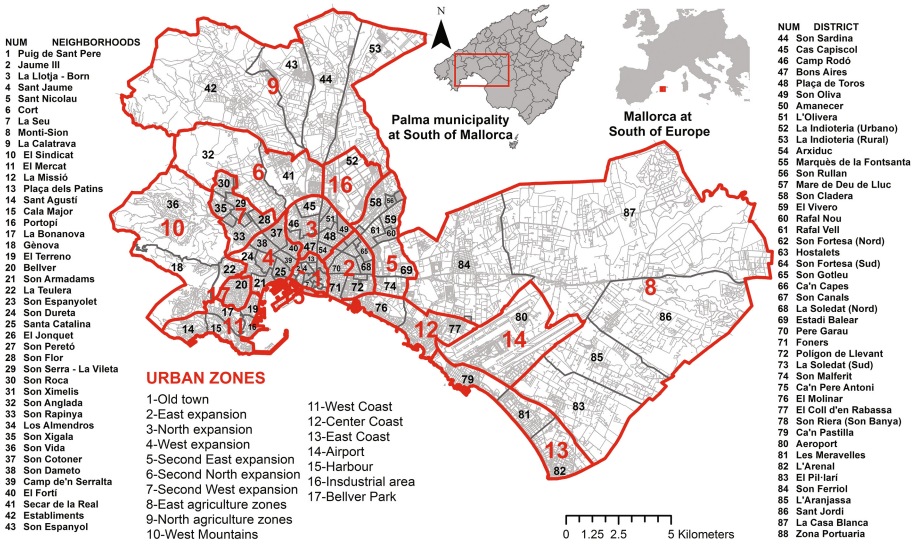


Fig. 1. Location of Palma city and division of neighbourhoods and urban zones

The supply of public transport in Palma has two subway lines: an underground city railway line (recently build, 2007) and the railway system of the island. In addition, the bus fleet cover 31 bus lines. The overall transport network is spread over 822,5 km and has 957 stops.

The road system of the city of Palma is structured in a radio network-concentric morphology, developed around the old town. The city centre is surrounded by two perimeter rings of roads. The first is formed by the roads that replaced the old walls and the other is an exterior beltway that encloses the city. This road network predetermined the development of a peripheral circulation that include radial roads to access the core of the city.

The city of Palma has had a Sustainable Urban Mobility Plan (SUMP) (developed under the project CIVITAS DYN@MO, http://www.civitas.eu/content/dynmo) since 2014. Among the main objectives of this plan is emphasizing improvements in public transport by bus and increasing the level of shuttle services to the tourist population. Measures proposed for the improvement of public transport and tourism mobility include the following: adapting the system to the public transport demand, restructuring the bus service covering the hotel industry, and improving the integration of tourism in public transport. The focus of this work is adapted to the objectives of the Palma SUMP and can provide support for decision making in this field.

3 Methodology

The method used in this investigation consists of the following phases:

- Analysis of the potential public transport bus demand of the resident population.
  The potential demand of public transport of population is represented by the resident population by geographic unit of analysis. In that respect, we used two geographic units: the urban block and the urban zone. The urban block level representation is used solely for the purpose of cartographic visualization. The source of data was the Population Census 2015 Palma (Ajuntament de Palma).
- Analysis of the potential public transport demand derived from tourism seats.
  The tourism potential demand is analysed at two levels: seats of regulated supply (hotels and apartments) and seats of non-regulated supply (individual homes and apartments) by geographical unit of analysis (urban block level and urban zones level). The representation of urban blocks serves only mapping purposes.
  The information about the regulated touristic supply was obtained from two sources: the Dirección General del Catastro (DGC) del Ministerio de Economía y Hacienda (Gobierno de España) and the Department of Tourism (DT) of the Government of the Balearic Islands.
  The data from DGC includes geographical information (shape file) and alphanumeric information of the touristic uses (table file) of each cadastral parcel. The DT has information about the total number of seats of hotels and apartments of Palma city.
  First, we made a proportional allocation of touristic seats to cadastral parcels according to its surface of touristic use. Second, using GIS tools, a cartographic

overlay was made between the cadastral parcel layer and the geographical analysis layers to obtain the number of touristic seats by urban block and by urban zone.
The non-regulated tourist seats are obtained from data provided by Airbnb.com. A shape point file represents the location of houses or apartments that have an associated attribute with the total number of touristic seats.
Another cartographic overlay was created between the layer of points of the supply and the geographical study unit layers to obtain the number of touristic seats by geographical unit (urban block and urban zone).
Finally, all the touristic seats, regulated and unregulated, were added for each geographical analysis unit (urban bloc and urban zone) to obtain the Potential Public Transport Demand of Touristic Seats.

– Integrated Demand Analysis of public transport: resident population and tourist seats.
First we proceed to normalize the potential data for each of the urban zones on a 0–1 scale according to the demand expression:

$$Normalized\ value = (Original\ Value - Minimum\ Value)/(Maximum\ Value - Minimum\ Value) \quad (1)$$

Then, the values of the potential demand (PD) of tourism seats are integrated according to expression (2):

$$Potential\ Demand\ of\ Population\ and\ Tourism = PD\ resident\ population + PD\ of\ tourism\ seats \quad (2)$$

The results of (2) are also standardized according to expression (1).

– Analysis of the level of bus public transport by urban area.
First we proceed to locate bus stops from the information provided by the Municipal Transport Company (EMT) of Palma. Then, the 959 bus stops were digitalized. Second a buffer of 400 metres was generated for each bus stop. Each buffer was assigned to a service level value based on the number of buses passing by it for a period of 12 h. This assignment takes into account the buses that pass through each bus stop considering their frequency. Finally, we calculated the bus service level by urban zone using expression (3) [8]:

$$Service\ Index_{Urban\ zone} = \sum_{1}^{n} \left(\frac{Area_{Buffer\ n}}{Area_{Urban\ Zone}} * SL_{Buffer\ n}\right) \quad (3)$$

The autocorrelation analysis of the geographic pattern of the supply was performing using the Moran index with the program GeoDa vers. 1.6. [2].

- Analysis of imbalances and equity of public transport by bus

A comparative analysis of the geographic distribution of the potential demand and supply of public transport service by geographical unit (urban zone) was performed. To contrast the different magnitudes of demand and supply, the first task was to normalize the values of each variable of the urban zones using expression (1).

The values matrix obtained is then subjected to an analysis of correlations between the different variables. In addition, a new variable "Level of Balance" is calculated for each urban area according to expression (4)

$$Level\ of\ Balance\ =\ \%\ Service\ Level\ -\ \%\ Potential\ Demand \tag{4}$$

Positive values indicate excess transportation service, whereas negative values show high values of potential demand.

– Analysis of the equity of public transport bus in relation to the population and tourist seats.

The Gini coefficient is obtained from the level of service and the resident population and tourist seats of each urban zone using expression (5):

$$\begin{gathered} \boldsymbol{GINI} = 1 - \sum_{i=0}^{n} (\delta Y_{i-1} + \overset{n}{\underset{i=0}{\delta Y_i}})(\delta X_{i-1} - \delta X_i) \\ i\ =\ row;\ n\ =\ number\ total\ of\ urban\ zones;\ Y\ =\ accumulated\ service\ level; \\ X\ =\ accumulated\ population/touristic\ seats \end{gathered} \tag{5}$$

The GINI coefficient was developed to measure the degree of inequality of a variable in a distribution of its elements [10]. The GINI coefficient ranges between 0, i.e., no concentration (perfect equality), and 1, i.e., total concentration (perfect inequality).

We used the ArcMap v.10.3 (© ESRI) software for the mapping and spatial analysis of data and Microsoft Excel and SPSS software v.22 (© IBM) for statistical analysis and graphical representation of the information.

4 Results and Discussion

4.1 Analysis of Potential Demand of Bus Public Transport

Potential demand of the resident population

The resident population of the city of Palma shows a heterogeneous pattern of geographic distribution in which areas with high population density and sparsely populated areas (Fig. 2) are combined.

The old town located at the city centre has a low population density, which shows the consequences of the progressive abandonment of residential use and its replacement by service use (preferably commercial, offices and catering establishments).

There are high levels of population in the prime concentric ring of the old town, at the widening of the city (Fig. 2). The main neighbourhoods in this area (Fig. 1) are Bons Aires, Camp Rodó, Plaça Toros, Pere Garau, etc. The population is very concentrated. The second highlighted area includes the suburbs located in the second belt of the city. They have residential uses of low and medium density maintaining the model of a dormitory town.

The Coastal urban areas also have significant populations, although lower densities; these include Arenal (east) and Cala Major (west). In these areas, tourism (hotels, apartments, complementary offers) competes with residential uses. On the west coast, the concentration of the resident population it is higher than that of tourists.

The demographic dynamics of the city in recent decades have varied greatly. Population growth continued until 2010 on the order of 10%; however because then the effects of the crisis have slowed that trend.

It is important to call attention to the fact that most of the increases of the previous decade were due to non-EU foreign immigration.

We must also mention the phenomenon of European tourists as residents that are not registered at the census but increase the permanent population of the city and generate more demand on public transport.

The transport potential demand of the resident population is oriented to pendulous travel from the periphery to the centre and directed to downtown areas with work activity (tourism, industrial and commercial). This also highlights mobility to large facilities and infrastructure (hospitals, universities, port, and airport) without neglecting mobility for leisure.

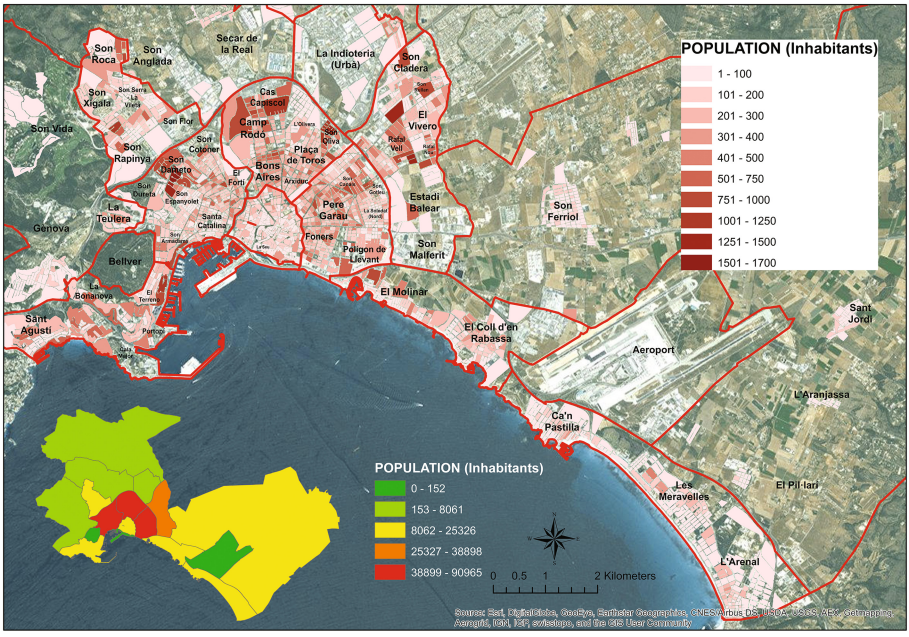


Fig. 2. Distribution of the resident population by urban block and urban zones.

Potential demand of tourism

We are going to relate the demand for transport with the location and the number of tourist seats as both regulated and non-regulated supply.

The regulated tourism supply in the city of Palma comprises a total of 237 establishments (apartments, hotels) with a total of 46,000 seats (Conselleria de Turisme, Govern de les Illes Balears, 2015) located mainly in the coastal urban areas in the east and west (Fig. 3a). They are two mature tourist destinations of the city that began their activities in the 50 s. Currently they are in a process of reforming and restructuring. They receive a majority of tourism for "Sun and Beach", which presents strong seasonality that determines its economic dynamics and transportation needs.

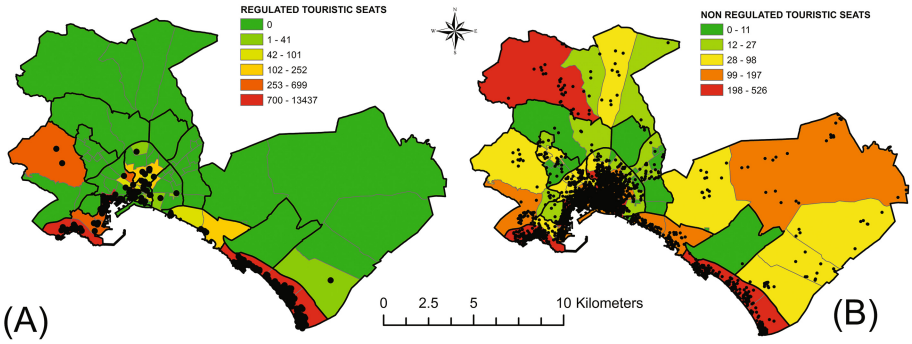


Fig. 3. Location of regulated tourist seats (a) and un-regulated seats (b)

The most central urban areas of the city, including the historic centre, have urban tourism, with high purchasing power staying in hotels, many of them newly built. This tourism sector is more oriented to shopping activities, dining and cultural activities. It is less seasonal and is more constant throughout the year. This tourism generates continuous economic activity in the city and has a moderate impact on bus transport demand.

The supply of apartments and houses for rent through the internet in Palma has experienced significant growth in recent years. This type of residential supply represents a total of 9,246 seats (only on the AirBnb platform) that are geographically much more dispersed than the regulated tourism supply (Fig. 3b). In this case, its location is not concentrated only in coastal urban areas; rather, it covers larger areas of the city. Specifically, the historical district and the first urban expansion are the preferred location. They also highlight the extensive developments located in the northwest area of the city.

Figure 4 shows the integration of the regulated and non-regulated tourism supply. A large concentration is observed in coastal areas but the supply is also widening to the West, extending into rural and natural areas of the municipality.

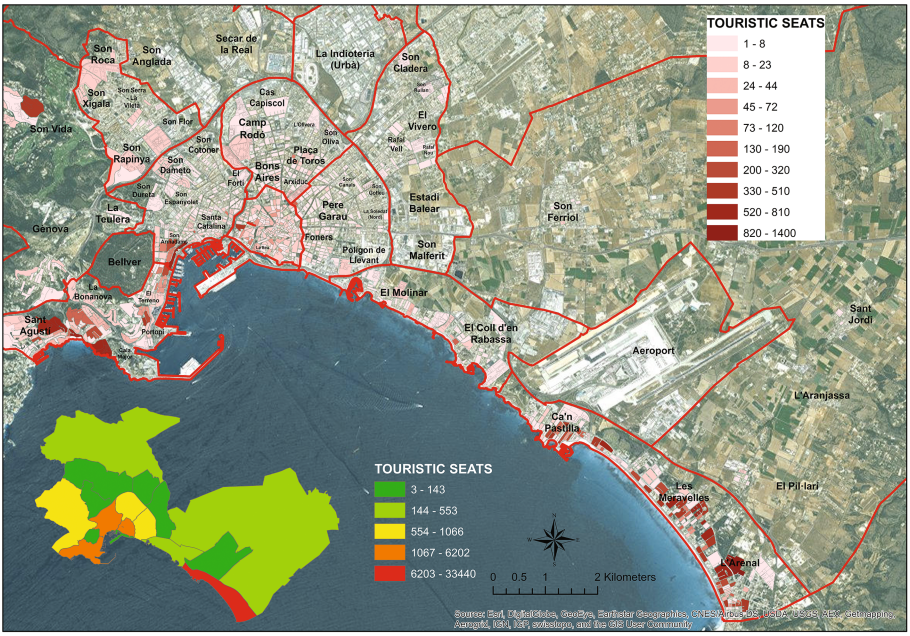


Fig. 4. Tourist seat supply, regulated and non-regulated

The occupancy rate of regulated tourism in the summer months can reach 75% capacity (Tourist occupation tables, Balearic Islands Institute of Statistics, Autonomous Community) so it is a burden of some 35,000 more people in the city who are potential users of the public transport

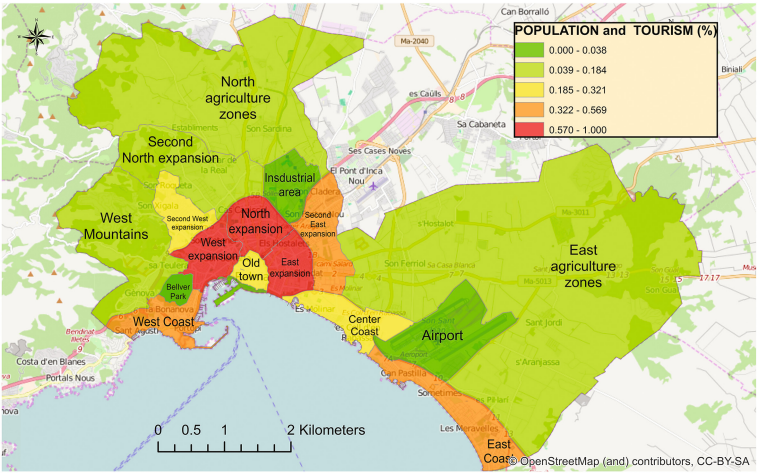


Fig. 5. Integrated potential demand for public transport: population and tourism.

Integrated potential demand: resident population and tourists seats
Figure 5 shows the potential demand for public transport integrating the resident population and tourist seats. First, the demographic burden of the neighbourhoods of the expansion appears to be high, which would represent the strong demands of the resident population. The coastal urban areas that show strong potential for tourism demand. The lowest values of potential demand are given in outlying rural areas.

4.2 Analysis of Bus Service Level Public Transport

The geographical distribution of public bus services in Palma is inhomogeneous; some areas have high levels of service where others have less. The concentric radio mode of the city configures the graph of the roads of Palma. This geographic pattern concentrates a high level of service at the city centre (Fig. 6).

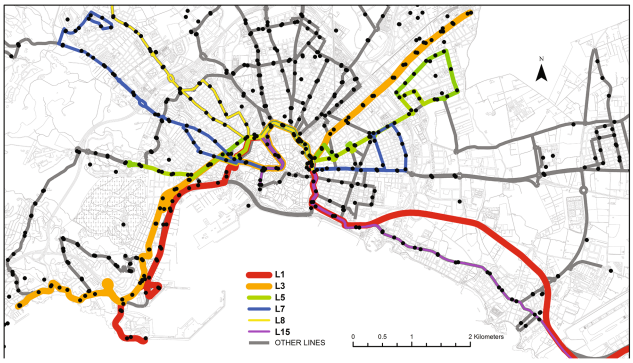


Fig. 6. Main bus lines of the city

The Moran index analyses the spatial autocorrelation of the pattern of a variable. In the case of service level, the Moran index reaches a value of 0.8, which indicates a remarkably concentrated pattern.

The low-level values of service appear at the peripheral rural areas of the city. The West Coast and the East and West zones of the second city beltway have slightly higher levels of service (Fig. 7).

4.3 Analysis of Imbalances and Level of Equity of the Bus Public Transport

Imbalances between supply and demand
Remarkably, the distribution of bus services in the city of Palma is adjusted to the potential demand of the resident population. A Pearson correlation value of 0.873 between the level of service in urban areas and the population (Table 1) is obtained. However, significant gaps are observed with excess service in the old town due to the concentric arrangement of bus lines in sparsely populated areas of the Northwest area

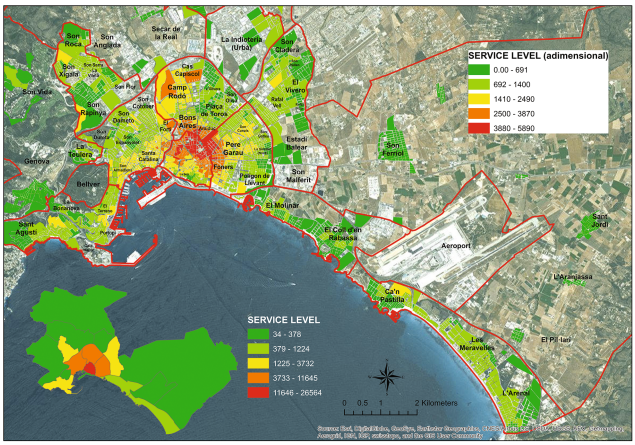


Fig. 7. Bus service level

of the city (Fig. 8A). High service needs are detected in the urban zones of the city that are more densely populated.

The balance is different for regulated tourism and non-regulated tourism. The first has less of a correlation with the level of service (r = 545), and deficits appear in the tourist coastal areas. In contrast, the service is more suited to the range of non-regulated tourism supply (r = 681). In this case, there are no significant deficits observed. This occurs because these areas largely fit with the geographical distribution of the resident population, and many of them are concentrated mainly in the old town.

The integrated potential demand (population and tourism) shows imbalances and deficits in all coastal tourist areas. The historic centre and less populated areas continue to have excess service.

The results show that the tourism potential demand is not sufficiently covered by the service supply, so an overload is generated in the system. This has special relevance during peak season.

Table 1. Pearson correlation coefficients for urban areas: analysis of variables.

	Service level	Potential demand population	Potential demand regulated touristic seats	Potential demand no regulated touristic seats	Potential demand touristic seats	Integrated potential demand population and touristic seats
Service level	1.000	0.873	0.545	0.681	0.654	0.825
Population	0.873	1.000	0.488	0.694	0.632	0.917
Tourist_Seats	0.545	0.488	1.000	0.860	0.934	0.575
AirBnb_Seats	0.681	0.694	P0.860	1.000	0.956	0.730
Tot_Tourism	0.654	0.632	0.934	0.956	1.000	0.671
Tot_Pop	0.825	0.917	0.575	0.730	0.671	1.000
Commercial	0.855	0.902	0.637	0.747	0.708	0.855
Sum_Commercial	0.830	0.840	0.489	0.607	0.543	0.790

The urban tourist areas account for nearly 85% of the tourist destinations. However, it is found that bus services for all these areas do not reach 7% (Table 2). The east and west coasts account for 73.77% of the regulated tourist destinations and have a service level of 1.82%. The resident population in tourist areas (10 of the municipality) does not reach the levels of service in other urban areas in more central urban areas.

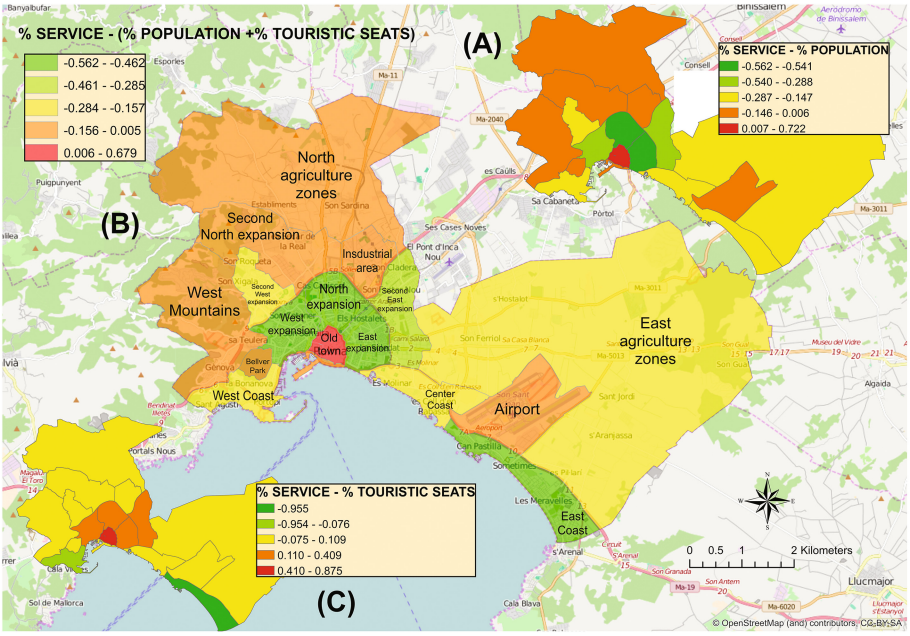


Fig. 8. Distribution of imbalances between service level and potential bus demand

Equity in the level of service

The Gini coefficient results show that the transport service for the resident population has a moderate level of equity (0.399) because bus services are concentrated in the centre of the city but the population has heterogeneous distribution patterns. Greater equity is observed in the relation between service levels and the pattern of unregulated tourism (0.253). The largest imbalance is shown for regulated tourism and level of service (0.692) because the supply is concentrated mainly in coastal areas. This confirms that the concentration of hotels and apartments in coastal urban zones unbalances the public transport model of the city. Conversely, non-regulated tourism gives stability to the transport service of the city. Its dispersed pattern is more adapted to the current transport supply. The integration of both variables (resident and tourism) does not yield significant equity results (0.424), but instead confirms the above results.

Table 2. Level service and potential demand in urban zones (%)

	Service level	Potential demand population	Potential demand regulated touristic seats	Potential demand no regulated touristic seats	Potential demand all touristic seats	Integrated potential demand population and tourist seats
Old town	39.44	6.01	3.41	28.73	7.81	20.48
East expansion	17.29	21.60	0.44	8.37	1.82	5.64
North expansion	10.73	17.53	0.91	7.18	2.00	24.63
West expansion	12.60	15.85	8.51	13.65	9.40	22.80
2nd east expansion	5.54	9.24	0.00	0.82	0.14	6.06
2nd north expansion	0.35	1.25	0.00	0.28	0.05	0.00
2nd west expansion	4.50	5.85	0.00	1.55	0.27	2.49
East agriculture zones	0.14	3.90	0.09	5.11	0.96	0.59
North agriculture zones	0.26	1.91	0.00	3.14	0.55	0.21
West mountains	0.56	1.11	1.18	2.39	1.39	0.14
West coast	4.37	5.55	11.39	12.75	11.63	2.49
Centre coast	1.64	4.83	0.59	3.19	1.04	12.92
East coast	1.82	4.48	73.47	11.41	62.70	1.52
Airport	0.05	0.04	0.00	0.03	0.01	0.00
Harbour	0.29	0.01	0.00	1.15	0.20	0.00
Industrial area	0.26	0.83	0.00	0.06	0.01	0.03
Bellver Park	0.17	0.00	0.00	0.18	0.03	0.00

5 Conclusions

Public transport planning in tourist cities should include both potential demand for the resident population as well as the potential demand from tourism, both regulated and non-regulated.

Tourist activity generates additional pressure on certain lines of public transport that need to be analysed and properly managed to ensure fairness of service.

The morphological structure of the road network of cities greatly affects public transport models and can lead to structural imbalances in the supply of services that can be increased with the location of the hotels in outlying areas.

The regulated tourism tends to be concentrated at certain places, especially in coastal cities, causing specific needs of public transport. The tourist supply non-regulated tends to have a more balanced distribution model in the territory so it is better suited to the existing transport supply.

In any case, tourism demand can overload the public transport system if is not properly planned and lead to imbalances in the service.

The results obtained in this study may have implications for transport planning at a strategic and operational level. The social equity and service level indicators proposed could help decision-making in transport planning by detecting territorial imbalances in the service level for the resident population and for the tourist population (regulated and non-regulated) also taking into account the seasonality of the tourist activity. It is important to note that the integration of the demand for public transport generated by the tourism sector requires the revision and redefinition of the concepts of public service and social equity by planners and managers. It is necessary to provide transport services to tourists without any decrease of the quality of transport service for the resident population. Tourist planning and public transport planning must be properly coordinated by adapting the service to the true needs of residents and tourists [4–6, 14]. This adaptation must necessarily include a model of economic management shared between the public administration and hotel companies.

Sustainable and innovative solutions must be provided to seasonal demand increases, avoiding spatial and temporal imbalances. The dynamic design of routes, the optimization of stops, the adaptation of the size of the bus fleets, the management of frequencies and headways, the promotion of intermodal transport, etc. are actions to be considered in order to provide flexibility to the service of public transport and improve the integration of tourist transport. The main objective is to guarantee the social, economic and environmental sustainability of public transport through a participative governance model that optimizes the integration of the tourist activity within the city without impairing the services to the resident population.

In the case study of the city of Palma, we can point to the following conclusions:

- The radio-concentric city structure determines a radio-diametrical transport service that generates imbalances between central and peripheral areas. There is a strong correlation between the level of public transport and the potential demand resulting from the resident population; nevertheless, significant differences exist in terms of excess service (the Old Town) and default service (in the beltway zones and the largest area in terms of tourist resort concentration in the east of the city).
- The potential demand generated by tourism shows imbalances with the current supply of transport. The location of tourist facilities in outlying areas of the city requires a higher level of service and disrupts the transport model.
  In this sense, we see imbalances generated by the supply of regulated tourism and competition between residents and tourists for the use of the public transport. The urban area that makes this problem more evident is the east coast of the city (s'Arenal), which was the first area of tourist concentration in the municipality of Palma, developed beginning in the 50 s.
- The demand for transport from unregulated tourism is distributed more evenly in the city and fits with the existing transport model. The tourist seats are located in

dwellings scattered throughout the municipality, which mitigates the effects of this seasonal tourist population on public transport.

Finally, it should be noted that geographic information systems combined with statistical analysis are a powerful methodological tool to diagnose the transport model of a city, integrate supply and demand and support decision making.

Acknowledgements. This work has been developed in the framework of the Project CIVITAS DYN@MO. "DYNamic citizens @active for sustainable MObility" no: 296057. Seventh Framework Programme. UE (2013-2016). The paper counts with the support of the Mobility Department of the City of Palma and the Empresa Municipal de Transports de Palma (EMT). We also express special thanks to Marina Ruiz for her help in the linguistic and the grammatical correctness of the text.

References

1. Albalate, D., Bel, G.G.: Tourism and urban public transport: holding demand pressure under supply constraints. Tourism Manage. **31**, 425–433 (2010). doi:10.1016/j.tourman.2009.04.011
2. Anselin, L., Syabri, I., Kho, Y.: GeoDa: an introduction to spatial data analysis. Geogr. Anal. **38**(1), 5–22 (2006)
3. Antoniou, C., Tyrinopoulos, Y.: Factors affecting public transport use in touristic areas **1**, 91–112 (2013)
4. Baggio, R., Scott, N., Cooper, C.: Improving tourism destination governance: a complexity science approach. Tourism Rev. **65**(4), 51–60 (2010)
5. Beaumont, N., Dredge, D.: Local tourism governance: a comparison of three network approaches. J. Sustain. Tourism **18**(1), 7–28 (2010)
6. Bramwell, B., Sharman, A.: Collaboration in local tourism policymaking. Ann. Tourism Res. **26**(2), 392–415 (1999)
7. Cha, H.: Enhancing Equity in Public Transportation Using Geographic Information Systems and Spatial. The Ohio State University (2008)
8. Currie, G.: Quantifying spatial gaps in public transport supply based on social needs. J. Transp. Geogr. **18**, 31–41 (2010). doi:10.1016/j.jtrangeo.2008.12.002
9. Dallen, J.: The challenges of diverse visitor perceptions: rail policy and sustainable transport at the resort destination. J. Transp. Geogr. **15**, 104–115 (2007). doi:10.1016/j.jtrangeo.2006.11.001
10. Delbosc, A., Currie, G.: Using Lorenz curves to assess public transport equity. J. Transp. Geogr. **19**, 1252–1259 (2011). doi:10.1016/j.jtrangeo.2011.02.008
11. Dickinson, J.E., Dickinson, J.A.: Local transport and social representations: challenging the assumptions for sustainable tourism. J. Sustain. Tourism **14**(2), 192–208 (2006)
12. Dickinson, J.E., Robbins, D.: Representations of tourism transport problems in a rural destination. Tourism Manage. **29**, 1110–1121 (2008). doi:10.1016/j.tourman.2008.02.003
13. Duval, D.T.: Tourism and Transport: Modes, Networks and Flows. Channel View Publications, Clevedon (2007)
14. Emerson, K., Nabatchi, T., Balogh, S.: An integrative framework for collaborative governance. J. Public Adm. Res. Theor. **22**(1), 1–30 (2012)
15. Evans, G., Shaw, S.: Urban leisure and transport: regeneration effects. J. Retail Leis Prop. **1**, 350–372 (2001). doi:10.1057/palgrave.rlp.5090132

16. Exceltur: Valoración turística empresarial de 2015 y perspectivas para 2016. Exceltur 81 (2016)
17. Ferguson, E.M., Duthie, J., Unnikrishnan, A., Waller, S.T.: Incorporating equity into the transit frequency-setting problem. Transp. Res. Part A Policy Pract. **46**, 190–199 (2012). doi:10.1016/j.tra.2011.06.002
18. Garrett, M., Taylor, B.: Reconsidering social equity in public transit. Berkeley Plan. J. **13**, 6–27 (1999)
19. Gronau, W., Kagermeier, A.: Key factors for successful leisure and tourism public transport provision. J. Transp. Geogr. **15**, 127–135 (2007). doi:10.1016/j.jtrangeo.2006.12.008
20. Guiver, J., Lumsdon, L., Weston, R., Ferguson, M.: Do buses help meet tourism objectives? The contribution and potential of scheduled buses in rural destination areas. Transp. Policy **14**, 275–282 (2007). doi:10.1016/j.tranpol.2007.02.006
21. Hall, D.R.: Conceptualising tourism transport: inequality and externality issues. J. Transp. Geogr. **7**, 181–188 (1999). doi:10.1016/S0966-6923(99)00001-0
22. INE. Instituto Nacional de Estadística, Padrón de Población (2015). http://www.ine.es. Accessed
23. Łapko, A.: Urban tourism in Szczecin and its impact on the functioning of the urban transport system. Procedia Soc. Behav. Sci. **151**, 207–214 (2014). doi:10.1016/j.sbspro.2014.10.020
24. Le-Klähn, D.T., Gerike, R., Michael Hall, C.: Visitor users vs. non-users of public transport: the case of Munich. Ger. J. Destin. Mark. Manag. **3**, 152–161 (2014). doi:10.1016/j.jdmm.2013.12.005
25. Le-Klähn, D.-T., Hall, C.M.: Tourist use of public transport at destinations–a review. Curr. Issues Tourism **18**, 785–803 (2015)
26. Lew, A.A., Hall, C.M., Williams, A.M.: The Wiley Blackwell Companion to Tourism. Wiley, Chichester (2014)
27. Litman, T.: Evaluating Transportation Equity: Guidance for Incorporating Distributional Impacts in Transportation Planning, vol. 8, pp. 50–65. Victoria Transport Policy Institute, Victoria (2005)
28. Lumsdon, L.M.: Factors affecting the design of tourism bus services. Ann. Tourism Res. **33**, 748–766 (2006). doi:10.1016/j.annals.2006.03.019
29. Page, S., Connell, J.: Transport and tourism. In: Lew, A., Hall, C.M., Williams, A. (eds.) The Wiley Blackwell Companion to Tourism, pp. 155–167. Wiley, Chichester (2014)
30. Page, S., Ge, Y.G.: Transportation and tourism: a symbiotic relationship? In: Jamal, T., Robinson, M. (eds.) The SAGE Handbook of Tourism Studies, pp. 371–395. SAGE Publications, London (2009)
31. Page, S.J.: Transport and Tourism: Global Perspectives. Pearson Education Limited, Essex (2009)
32. Parks, R., Rights, C.: Environmental justice and transit equity (2010)
33. Pitot, M., Yigitcanlar, T., Sipe, N.N., Evans, R.: Land use & public transport accessibility index (LUPTAI) tool: the development and pilot application of LUPTAI for the Gold Coast. In: ATRF06 Forum Papers (CD-ROM online) 2005 (2006)
34. Thompson, K., Schofield, P.: An investigation of the relationship between public transport performance and destination satisfaction. J. Transp. Geogr. **15**(2), 136–144 (2007)
35. UNEP-WTO: Making tourism more sustainable. A guide for policy makers. Environment **54**, 222 (2005). ISBN: 92-807-2507-6
36. UNWTO: Handbook on E-marketing for Tourism Destinations (2014)
37. UNWTO: How Tourism can contribute to the Sustainable Development Goals (SDGs) (2015)

On the Design of Leisure Devoted Cycling Networks

Alessandro Giovannini[1], Federico Malucelli[2], and Maddalena Nonato[3(✉)]

[1] Dipartimento di Matematica, Università degli Studi di Milano, 20133 Milano, Italy
[2] D.E.I.B., Politecnico di Milano, 20133 Milano, Italy
[3] Dipartimento di Ingegneria, Università di Ferrara, 44121 Ferrara, Italy
maddalena.nonato@unife.it

Abstract. Due to the several benefits associated to cycling, the demand for bicycle devoted infrastructures has considerably raised in the last few years. Public resource investments into infrastructure provision increased to meet demand, despite of the lack of mature design and planning methodologies. Leisure intended cycling networks in particular, which are the topic of this work, are even less addressed. We revise the critical issues in cycling network design, present a design methodology for cycle tourist networks that integrates the specific features of the problem into a combinatorial optimization framework, and test the approach on realistic data comparing with previous contributions on this topic.

Keywords: Cycling network design · Cycle tourism · Combinatorial optimization

1 Designing Cycling Networks: Facts and Challenges

Benefits associated with cycling are many, ranging from transport sustainability and public health improvement to cycle tourism development. In the last few decades, the car-centered dominating culture has been shaping urban development all over the word, yielding cycle-unfriendly cities. Nevertheless, a recent trend rediscovered the advantages related to cycling, which in turn ignited scientific publications on this topic. Functional cycling in particular, i.e., when the bicycle is intended as a mode of transport, has been widely investigated in the transportation literature, and its promotion is central to proposals willing to increase the share of soft mobility, reduce traffic congestion, and lower carbon emissions, as testified by the many European projects on this topic such as BICY [33]. As demand for cycling increases, the debate on assessed and reliable methodologies for infrastructure provision and planning becomes crucial. Positive correlation has been proven between miles of bicycle pathways per residents and percentage of cycling commuters [29]. However, infrastructure provision alone does not guarantee usage will increase as expected. Moreover, options are so many - from bike lanes, either raised or on road, to cycle ways and cycle

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_11

tracks - that selecting the highest return infrastructure investment is a complicated task. Guidelines are often sought ex-post by comparing the outcomes of different interventions: see, for example, [3] for relations between cycling facilities and levels of cycling commuting in 90 large American cities, or [16,18,20,38] for travel behavior. Nevertheless, we still miss an abstract representation of the system sufficiently comprehensive to model cyclists behavior and mode choice, so that the provision of functional cycling devoted infrastructures can be guided by a reliable prevision of future usage, thus assessing the impact of potential interventions prior to deployment, as many advocate [36].

The field of infrastructure planning for leisure cycling is even less mature. This study aims at giving a contribution in this direction: first, the key factors leading functional cycling infrastructure design and planning are revised, and similarities and differences with leisure cycling are analyzed; on these bases, we formulate the leisure cycling network design problem as a combinatorial optimization model, propose a heuristic solution approach, present results for a realistic case, discuss them, draw conclusions, and sketch future work.

1.1 Criteria for Planning Cycling Infrastructures: Leisure Versus Functional Cycling

The Safety Concern. Among the several factors influencing the decision of whether or not commuting by bike, safety stands out as the main concern and many are the proposals in the literature aimed at hazard reduction and safety level improvement of cycling facilities to make cyclists feel safe en route [1]. In particular, a highly debated issue concerns the integration of cycling infrastructures into the vehicular road network (a common practice in the U.S., U.K. and Southern Europe) as opposed to its segregation from motorized traffic (a highly valued option in Northern Europe), each having strengths and weaknesses. While a segregated cycle track would guarantee the highest safety, it can be seldom deployed along the shortest route to destination in highly urbanized settings. As a result, the detour on the safest bike track is often too long to be a practical option for commuters who have to choose between longer travel times and higher safety (see [20]). Safety perception, thought, is itself an issue: indeed, the most critical challenge in modeling cyclist behavior is the understanding of risk perception, since perceived safety often substantially differs from objective safety [30]. Moreover, the cycling culture of each country adds to the built environment in affecting cyclist behavior and risk perception (see [17,20,24]). For example, [36] reports a case where even the provision of segregated cycling infrastructure would marginally reduce the car commuting share whereas a monetary reward on cycling could induce a consistent modal shift. In conclusion, while general guidelines can be shared, investigations focused on cyclists' perception and drivers' attitude towards cyclists should be carried out on site, for each specific project, to provide tailored guidance to policy makers for infrastructure provisioning devoted to functional cycling.

Several synergies exist between leisure and functional cycling, and planning methodologies share common challenges. Indeed, the two cycling communities are often mixed, and cycling infrastructures originally designed for leisure often

get used for functional journeys by experienced cyclists as well [29]. Nevertheless, functional cyclists and leisure cyclists have specific user needs. On the one hand, commuters prefer direct routes to destination, allowing constant travel speed and no stop at junctions. On the other hand, since cycling is part of the experience when cycling for leisure, riding the fastest route to destination is not a cycle tourist priority and traveling time is gladly traded for route quality. Moreover, different classes of users exist among leisure cyclists, as later discussed in Sect. 2. Sport oriented recreational cyclists are moved by fitness improvements and value challenging itineraries embedded in healthy environments. On the contrary, family cyclists prefer less engaging trips and appreciate attractions along the way as well as easily accessible service facilities [8,21].

Safety en route is a common concern for both communities, although differently perceived. As motorized traffic represents the main source of hazard, both value traffic free or low traffic routes and clear signage. However, a 14.000 people survey in the UK [12] shows a different attitude towards traffic volume, which appears to have greater impact on leisure cyclists than on commuter cyclists, the former being generally less keen mixing with motor traffic. Such behavior is partially explained by the higher average level of experience utility bikers have, with respect to leisure bikers. Actually, there is some positive feedback in providing safe, leisure devoted, cycle routes: since inexperienced cyclists tend to overstate risk perception, offering a friendly and protected environment where they can ride and have fun in complete safety increases their confidence with this mean of transport and decreases their hazard perception, potentially enlarging the community of utility cyclists in the near future. [8] reports such a positive effect regarding the Great Western Greenway in Ireland, a cycle route created to attract visitors and cater to recreational cycling. A few years from its completion, the percentage of sustainable travel patterns of residents, school commuting by bike in particular, remarkably increased.

Safety can be evaluated based on quality indexes such as Bicycle Level of Service (BLOS) and Bicycle Compatibility Index (BCI), thus providing a quantitative methodology to set safety standards. Additional, less measurable, design criteria should be also considered when designing a cycle route, such as continuity, attractiveness, comfort, directness, societal and economical impact.

Additional Criteria. The lack of **continuity**, i.e., the possibility of reaching destination along a seamless itinerary, is often questioned by urban cyclists and it strongly affects the perceived suitability of cycling as a transport mode of choice for commuters. This concept is closely related to network **connectivity**, that is continuity for all origin - destination pairs. In a similar way, a leisure cycling route is appealing only provided that it ensures a certain quality standard throughout its way, from origin to destination; likewise, a leisure cycling network is expected to connect the local attractions in a seamless manner.

Comfort is related to surface quality and number of junctions and stops. For recreational cyclists it is also affected by accessibility of service points en route. Gradient impacts on comfort, although specific challenging itineraries can be a valid alternative to the regular track for sport oriented practitioners.

When commuting by bike, additional **travel duration** with respect to the shortest path must be limited: several studies suggest that only itineraries not longer that 33% of the shortest path are considered as viable alternatives [13]. When cycling for leisure, though, journey ambiance is favored with respect to **directness** to destination, and different classes of users should be considered, each one with its own time limit ranging from few to several hours.

In case of recreational cycling, **attractiveness** refers not only to the particular features of the pathway and to its aesthetic, but it concerns the appeal of the points of interest reachable along the pathways and to the scenery of traversed landscapes, so it varies according to the user profile. Pathway attractiveness, thought, started to be recognized as an asset by academics and city planners also for cycleways catered for commuters within a urban environment [22]. Indeed, attractiveness is among the five criteria suggested in [7] for cycle infrastructure design, and [19] mentions attractiveness as one of the attributes for a functional cycling infrastructures quality index.

Promoting cycling as a mode of transport provides direct and indirect **economic benefits** since it decreases motorized traffic and improves population health and well being. Promoting cycle tourism provides opportunities for sustainable development to many areas featuring natural and cultural attractions concentrated in a small region [23]. Cycle tourists are a particular kind of eco-tourists, who appreciate the journey as well as the destination; segregated infrastructures, scenic landscapes, and Points of Interests (PoIs) en route are a must; quite often, though, areas that would qualify in terms of attractions miss an appropriate infrastructure. This fault affects the satisfaction level, since the cycling activity is an integral part of the tourist experience. Therefore, providing adequate riding facilities is mandatory for cycle tourism development, whose economic benefits are unquestioned: the value of cycle tourism in 2012 was estimated at about 44€ billions, with significant potential to increase [37]. The Great Western Greenway is a representative success case: its creation spurred several tourism related business activities supporting the local economy and the estimated annual economic impact suggests a very short payback period for public investments of this kind.

The development of cycle networks is becoming more and more part of the public agenda at the regional, national, and even international level [33]. Since political and governmental policies shape the development of cycling infrastructures, providing guiding principles and well assessed planning methodologies to support decision makers is fundamental. Consider for example Euro Velo, the European cycle velo network (www.eurovelo.org) now extending for almost 45.000 km, which is a challenging project intended to connect all countries in Europe by an ever growing network of high-quality cycling routes [37]. This planning effort, however, often acts a posteriori, setting quality thresholds and other standards for existing bikeways to be included in the project that marginally influence planning decisions on the territory. Indeed, the planning phase of each single route is often conducted independently by each local administration as a once-off project, being managed as a resource led activity driven by

cost-benefit analysis and responding to local stakeholders expectations [6]. As a result, cycle networks tend to be discontinuously developed, potentially lacking a comprehensive vision and missing the chance of achieving long term targets [28]. On the contrary, the scientific literature advocates user needs evaluation prior to deployment as the key to fulfilling cyclists expectations and exploiting cycling facilities at their best, besides consistency and integration of all decision maker actions [9]. For example, Manton [26] provides guidelines for greenways design, emphasizing their multiobjective purpose (such as natural corridors, commuter cyclists links, and recreational traffic-free sites) while ensuring path quality and providing connection with points of interest and other networks; Weston [37] envisions a network of greenways featuring safe and continuous routes, with point of interests on route through scenic landscapes. On our side, we stress the need for embedding the system representation into an optimization model as the only mean of implicitly considering all the feasible network alternatives and supporting decision makers with reliable quantitative methodologies, rather than narrowing the choices to a limited set of precomputed options [5]. Some recent proposals in this direction are listed in the following.

1.2 Optimization Models for Cycling Network Design

A limited though significant body of literature concerns the application of optimization techniques to the solution of mathematical models for cycling network design, building on the methodologies developed for network design in other fields. This allows to potentially optimize over all the network configurations that correspond to feasible solutions to the mathematical model instead of comparing a limited number of previously designed alternatives. The resulting models, often multi-objective ones, generalize the fixed charge, multicommodity, min cost flow problem, but realistic instances are not solved to optimality.

All but one of the following are devoted to functional cycling.

In [22] a multiobjective optimization model is proposed to design a urban cycling network in Taipei. The objectives are minimizing risk and impact on traffic, maximize comfort and residents service coverage. The constraints involve bikeway type, monetary budgets, path continuities, and value ranges of decision variables. Six non dominated solutions are computed and analyzed.

In [21] the same methodology is extended to tackle the design of a recreational cycling network in the north coastal area of the Sanzhi district in Taiwan, currently featuring a fragmented offer of bikeways. The authors consider four types of bikeway retrofitting or construction (reserved or shared use bike paths, bike lanes, and bike routes) and two types of service facilities (rest areas and rental shops); constraints concern (i) connectivity among gates (existing transit stations providing access to the scenic area) and from each network node to a gate, (ii) maximum distance required for service coverage, (iii) limited budget; service demand is assumed known, given by the number of tourists currently using the access transit stations; four criteria (served demand, service station coverage, safety, attractiveness) were optimized in a 0/1 multiobjective linear programming model, solved by the ϵ-method yielding six non dominated alternatives.

In [28] cycle-network planning (CNP) at the local level is solved targeting several objectives including accident reduction, modal shift in favor of cycling, health benefits and strategic network expansion, exploiting a spatio-temporal model embedded within a GIS-based decision-support system; the development of a cycle network in a small town in the UK is used as a case study.

In [10] an optimization model is proposed that designs a minimum cost cycling network interconnecting a set of origin destination (OD) points by retrofitting existing roads such that a given quality is guaranteed (in terms of BCI on segments and LOS on intersections) and each OD path duration is bounded. The aim is to support seamless bicycle trips between all OD pairs, mending fragmented sets of existing cycle paths. Bicycle demand is not considered, as latent demand will be revealed only afterward. The methodology is applied to the county of Austin, Texas, and favorably compares to separately designing individual itineraries. However, only toys instance are solvable. In [31] the authors solve a similar problem aiming at minimizing trips length and maximizing the LOS of selected links. The terms in the objective functions are weighted according to present OD demand, disregarding latent demand.

In [27] the authors propose an optimization framework for the design of a network of bike lanes in a urban road network, aimed at identifying on which links a bike lane should be realized. The lower level of a bilevel model embeds a traffic assignment problem, while the upper level optimizes total travel time for car users and total distance that can be ridden on bike lanes subject to a budget constraint and assuming given bicycle demand. The model is heuristically solved by Genetic Algorithms and tested on an artificial grid network.

Building on top of these experiences, we propose a mathematical optimization model that either integrates as constraints or fulfills by construction the above mentioned criteria and seeks the highest utility for several classes of users. This introduces the issue of *routing for profit*.

1.3 Routing for Profit

Many applications in tourism planning are based on generalizations of the Orienteering Problem (OP) [11,34]: given a graph with travel time on edges and profit on nodes, OP consist of computing the maximum profit tour subject to maximum travel time. In tourism applications nodes represent PoIs and edges model the shortest path between PoIs [14]. They are usually embedded in decision support tools designed for tourists visiting a city and often allow to integrate user preferences [35]. A version devoted to cycle trips planning is described in [32], while the issue of thematic routes building has been recently addressed in the eco-tourism framework [2]. Again, infrastructure provision is not contemplated. In all these applications the physical (road) network providing connection between PoIs is always taken for granted, so that no design cost is taken into account. Moreover, traversed PoIs provide reward only at the first visit, so no optimal solution has multiple traversals.

The problem addressed in this work, i.e., cycle tourist network design, can not be solved by the aforementioned approaches for at least two reasons:

(1) infrastructure provision costs must be handled; (2) since biking is part of the tourist experience, successive traversals of the same edge impact total reward and must be accounted for. Both features were first addressed in [4] where the Most Attractive Cycle Tourist Path Problem (MACTPP) was introduced: it consists in computing the most rewarding itinerary from origin to destination, subject to duration and budget constraints, with decreasing attractiveness at successive traversals of the same edge and node. Due to the strong ties with OP, MACTPP was modeled by a Mixed Integer Linear Programming (MILP) model generalizing the one in [11], and easily solved by commercial solvers for realistic instances. In a later study [25], a multi-commodity version of MACTPP, named MOP-ND, was proposed to jointly handle several classes of users with different attractiveness functions and to design potentially different itineraries going from the same origin to the same destination which must share the same budget. The study experimentally shows that using more information in the planning phase pays off in terms of potential attractiveness of the resulting infrastructure. Both studies designed itineraries connecting a single origin-destination pair. In [15] the design of the whole network is finally tackled. Given a set of network access points (gates), the tools developed for the single origin-destination itineraries were exploited to build an interconnected infrastructure, linking all the gates and providing access to most PoIs. In that work, high quality paths from gate to gate were first built and then a subset was selected and the network was built out of the edges of those paths. In this work we totally revolve the decision flow: we start from selecting a set of budget compliant edges and then build the itineraries from gate to gate based on these edges, as detailed in Sect. 2.

2 Solving the Cycle Tourist Network Design Problem

2.1 Problem Statement

As discussed, the problem consists of selecting the tracks to be retrofitted, taking into account budget constraints and user preferences, in order to design a network of bike trails spanning over a restricted but compact area where several visitor attractions are located, which the network should provide access to and mutually connect along itineraries that are fit for casual, recreational, day cyclists. In our proposal, as in [21], the access to the cycling network is given through a set of gates that provide connection to other means of transport, i.e., gates are locations on the main road network or stops of the local public transport system, whether bus or train, and are equipped with parking lots and renting facilities. The cycling network should connect gates to each other and to the most attractive spots by way of scenic pathways providing maximum satisfaction to potential users, according to their preferences. While exploiting existing tracks already fit for cyclists, our planning methodology tries to consider each of them as an additional asset rather than taking all of them for granted from the beginning as *Cost for Benefit* approaches would suggest.

Note that we comply with the design criteria and recommendations discussed above. Indeed, we guarantee continuity and completeness by construction, since the resulting network connects all gates and selected routes have limited duration. Safety and comfort are ensured thanks to either selecting links already fit or retrofitting those unfit, so that all routes, from origin to destination, meet quality standards. User needs are taken into account by modeling the utility function of different user classes, and thematic itineraries will be designed between the given network gates. Directness intended as providing routes with different maximum travel durations, is guaranteed as well. The required investment is an input parameter, so that public stakeholders can set it and vary it according to availability. The model can be used as well to compute the lowest amount required to build a complete, compliant network.

More formally, we face the following combinatorial optimization problem.

We assume to be given: (i) the most sought after locations (PoIs) present in the area and the edges that connect them: the latter are the potential links of the cycle network; (ii) the locations acting as gates, giving access to the cycle network from the outside; (iii) the utility functions of a predefined set of user classes, quantifying the reward experienced by the user when biking along a link or visiting a PoI; (iv) the time required to traverse each link, in both directions; (v) the cost of upgrading each potential link in order to reach a minimum BCI threshold (0 if the link is already fit).

The problem consists of selecting the optimal set of links to be retrofitted to form a network so that: (a) each origin-destination pair of gates is connected along at least one itinerary not longer than a given duration, (b) the total budget available for upgrading is not exceeded, (c) the sum of the attractiveness of the thematic itineraries over the different user classes is maximized.

Building on the experience gained in modeling and solving MACTPP and its multi-commodity version MOP-ND, in [15] we proposed a network design methodology structured into five steps: (1) compute a set of promising paths by solving MACTPP for different time durations and total budget shares; (2) select one path for each pair of gates and user class; (3) build the network made of all the edges in the selected paths; (4) solve a MACTPP on that set of edges for each pair of gates and user class, and eventually (5) get rid of redundant edges (those never used by the computed itineraries). We call this approach *Path-First-Network-Second* (PFNS).

Despite PFNS can solve the problem, several MACTPP instances for the same origin destination pair - differing for time limit and budget share - have to be solved at step (1) to produce a sufficient variety of paths for step (2). The major drawback, however, is that PFNS is not robust with respect to budget. First, it provides no clue about the minimum budget required for a feasible solution. Second, it may fail to find a solution for a given budget even when such a solution exists, due to the fact that some paths part of that solution have not been generated at step (1). Therefore, a trial and error approach is necessary with respect to budget calibration, which is a serious pitfall for practical applications. Finally, the resulting networks are rather dense, which is not desirable. In fact,

beside being intrinsically a bi-objective problem, i.e., minimizing budget and maximizing attractiveness, a third network quality criterion concerning edge number is also present. These objectives are all conflicting, and we aim at finding Pareto (sub-)Optimal solutions; in particular we aim at being able to compute solutions with low budget whenever they exist.

The drawbacks of the PFNS approach come at no surprise since PFNS starts from individual paths, each of which is selected while unaware of its share of budget and of the set of edges used to connect the other gates. In the present work we try to overcome this weaknesses and propose an approach that reverses the previous decision flow: instead of starting from the paths and building the network out of them, here we start from a set of budget compliant edges and build the itineraries on them. We shall name this approach *Network-First-Path-Second* (NFPS) as opposed to PFNS. Since budget and total attractiveness are conflicting objectives, we do not expect NFPS to increase attractiveness at lower cost with respect to PFNS; our aim is rather to set up a methodology able to explicitly manage the budget level from the start, as an exogenous parameter. The NFPS solution algorithm is detailed in Subsect. 2.2.

2.2 NFPS: An Iterative Algorithm for Cycle Network Design

NFPS stems from an observation: the cycle tourist network design problem can be modeled by generalizing the MOP-ND model [25] so as to encompass all origin destination pairs. In detail, the resulting *complete model* consists of two parts, regarding design and routing. Design, modeled by binary variables z, concerns edge selection subject to a budget constraint; the routing part, modeled by integer variables x, is decomposable by commodity since it concerns computing the most attractive duration-constrained route for each origin-destination pair, user class, and duration. Design and routing are linked by logical constraints $z = 0 \Rightarrow x = 0$. The objective is to maximize the thematic itineraries attractiveness and depends on flow variables x. However, the number of commodities typical of realistic instances makes the automatic solution of the complete model out of reach of state of the art solvers.

This evidence leads to addressing the problem heuristically, exploiting the ability to solve a core problem, that we call the *Simplified Model* (SM) as opposed to the complete one. SM retains budget, connectivity, and the strictest time constraints requirements; the objective is to maximize first traversal attractiveness of selected edges. Its formulation is reported in (1–8). Given a graph G, SM either returns a subset of edges $E^*(G)$ admitting a feasible solution, or certifies G's infeasibility. In the first case, the value of $E^*(G)$ is computed by solving the routing part, which is now decomposable by commodity. The core procedure made of **SM + routing** is the building block of our algorithm, and it is iteratively called on different subgraphs $G' \subseteq G$ obtained by pruning different edge sets.

Let us introduce some notations to formally describe NFPS. Given the graph $G^0 = (N, A \cup E)$ where N includes PoIs, gates and junctions, E is the set of all potential links and A represents their traversal in one or the other direction,

let $\Gamma = \{1, .., n_\Gamma\} \subseteq N$ denote the set of gates and let Γ^2, indexed by $k \in \{1, .., n_\Gamma(n_\Gamma - 1)\}$, be the set of ordered pairs (γ, ν) s.t. $\gamma, \nu \in \Gamma$, $\gamma \neq \nu$. Denote by $u \in U = \{1, .., n_u\}$ a user class for which a utility function on first and second traversal on nodes and edges of G is known. For each $u \in U$ and for each pair $k = (\gamma, \nu) \in \Gamma^2$ we consider the pair $c = (u, k)$ as a commodity in the set $C = U \times \Gamma^2$. To provide the chance to choose among routes of different durations, for each commodity c we consider a set of values $T(c) = \{\tau\}$ to be used as maximum travel time, which are function of the duration of the shortest path from γ to ν.

Any set of budget compliant edges $E^* \subseteq E$ such that for each c and $\tau \in T(c)$ there is a time feasible route defined on E^* is a solution to our problem. We search for sub-optimal solutions by running the core procedure on several subgraphs $G^i \subseteq G^0$ obtained by edge pruning. The aim is to iteratively search the feasible region of the complete model laying in the vicinity of the current solution by locally modifying the most characterizing attributes. The search can be described in terms of a layered graph, whose nodes at the same layer are subgraphs generated during the same iteration by setting to 0 one edge in the solution of the father node at the previous layer. Graph G^0 is at the root of the search graph; the edges to be pruned are positive cost edges in $E^*(G^0)$, i.e., the SM solution on G^0. Indeed, since budget is the only global constraint of the complete model, what fully characterizes a solution are its unfit edges. In fact, any 0-cost edge can be added to the solution for routing purposes without affecting feasibility, as opposed to unfit edges that require retrofitting spending: in particular, the optimal solution usually corresponds to some maximal (with respect to budget) set of unfit edges. In the specific case, given $E^*(G^0)$, the n_e unfit edges with lowest positive flow are selected and as many new problems are generated by pruning (setting to 0) each one of those n_e edges, each yielding a child node at level two of the search graph together with a subgraph G' on which the core procedure will be run again to produce a different solution from the father node. In the second iteration the procedure is repeated on each of the n_e subgraphs. For each subgraph G' two cases are possible: either G' admits no feasible solutions and the search node is discarded, or SM returns a feasible set of edges $E^*(G')$ based on which, after the evaluation-routing step, n_e additional problems will be generated by the edge-pruning procedure.

From iteration three onwards, instances with the same set of pruned edges may occur; as they coincide, a single occurrence is kept, that is, more than one path from the root may lead to the same node of the search graph. Moreover, a cache mechanism records the routing solution for each subnetwork so that whenever SM returns a set of edges which has been already evaluated its reward is retrieved at no cost. To keep bearable the number of problems to be solved at each iteration, two filters are applied at different steps of the procedure. The first filter applies right after the SM solution and keeps the best n_{f1} problems based on the SM objective function value. The second filter applies once the routing part is solved, to keep the n_{f2} nodes with the best actual solution cost.

This yields at most $n_e \cdot n_{f2}$ nodes at each layer of the search graph. The values of n_e, n_{f1} and n_{f2} should be calibrated according to the specific instance size.

Basically, SM picks heuristically a maximal set of unfit edges that is feasible for routing and compliant with the current pruning. As pruning advances, the number of feasible maximal sets lowers and the process converges. However, since search is filtered, optimality is not guaranteed.

Note that SM is considerably easier than the complete model and it is quickly solved. With respect to the complete model, the following differences exist: the objective function is defined on the design variable and not on the flow variables, and the flow part reduces to a pseudo-polynomial problem for each commodity. Even the MATCP problems solved for solution evaluation purpose on the selected edges E^* are computationally less demanding than the instances solved in [4] - actually they are MACTP relaxations since design variables are set. Finally, the cache mechanism prevents solving the routing problem twice on the same subnetwork. Together with filtering, all this allows to keep the computational burden reasonable. The number of iterations depends on the average cardinality of maximal edge sets with respect to the total number of unfit edges. Stopping conditions are reached when each active node yields to the next level only nodes whose simplified model is infeasible.

The MACTP and MOP-ND formulations can be found in [4,25], respectively. For the sake of completeness, in (1–8) we report the SM formulation for a generic time limit $\tau \in T(c)$ at iteration 1, where no edge is pruned.

$$\max \sum_{[i,j]\in E} w_{ij}^h z_{ij} \qquad \text{subject to:} \tag{1}$$

$$\sum_{(i,j)\in FS(i)} x_{ij}^c - \sum_{(h,i)\in BS(i)} x_{hi}^c = b_i \qquad \forall i \in N, \forall c \in C \tag{2}$$

$$\sum_{(i,j)\in A} t_{ij} x_{cij}^c \leq \tau \qquad \forall c \in C \tag{3}$$

$$\sum_{[i,j]\in E} c_{ij} z_{ij} \leq B \tag{4}$$

$$x_{ij}^c \leq z_{ij} \qquad \forall [i,j] \in E \tag{5}$$

$$x_{ji}^c \leq z_{ij} \qquad \forall [i,j] \in E \tag{6}$$

$$z_{ij} \in \{0,1\} \qquad \forall [i,j] \in E \tag{7}$$

$$x_{ij}^c \in \{0,1\} \qquad \forall (i,j) \in A, \forall c \in C \tag{8}$$

The objective function (1) maximizes the weight of the selected edges. Selected edges must be budget compliant (4) and allow flow in both directions (5–6). For each commodity $c = (u,k)$, for $k = (\gamma,\nu)$, the flow goes from γ to ν (2) where $b_\gamma = 1$, $b_\nu = -1$ and $b_i = 0\,\forall i \in N$, along a path taking no more than τ time units (3). Note that SM is not a relaxation of the complete model so its optimal solution value does not provide a bound and cannot be used for pruning search nodes.

3 Computational Results and Conclusions

We apply the described methodology to the benchmark data used in our previous experiments. In the Trebon region, located in Southern Bohemia, Czech Republic, local authorities aim to promote sustainable economic development by facilitating the Trebon region as a cycle tourist destination. While Euro Velo route number 7 traverses Czech Republic, it does not cross the Trebon region but skims its west southern border, as route 13 does on the southern border. Both provide great hooks for visiting the Trebon region, which so far remains untouched and out of the beaten trails. The opportunities offered by cycle tourism in terms of sustainable economic growth motivated this line of research, the current step in particular where emphasis is on budget.

Hereafter we briefly recall the data description provided in [15]. We work on an abstract network, i.e., graph $G^0 = (N, A \cup E)$, made of the set of candidate links: some require an investment to be retrofitted and reconditioned into cycle tracks while others are already fitting and can be used almost at zero cost. The former are unpaved roads, canal towpath, or natural trails that are presently being used for off road cycling or hiking. The latter are minor rural paved roads with low vehicular traffic and fit for cycling. Former military roads once used to patrol the Austrian border add to this set. Indeed, the experience reported in [6] supports the choice of retrofitting and exploiting for tourist use the vast supply of rural roads currently underutilized rather than constructing new facilities, thus substantially reducing both infrastructure provision investment and environmental impact. Nodes N are PoIs or cross-roads, where usually service facilities are present. The resulting graph has 84 nodes and 146 edges, 86 of which are 0-cost edges. Arc traveling time is computed with respect to an average speed of 18 km/h and adjusted according to the gradient. The edge cost for reconditioning depends on present condition and edge length: the estimated cost of a 3 m wide path is 115 € per meter to turn it into an asphalt surface if starting from dirt road, 75 € per meter from gravel one. When an edge is made of several sections in different conditions, the total retrofitting cost is the sum of reconditioning each section to asphalt. The cost of retrofitting all edges is about $13700 \cdot 10^3$ €.

Based on the opinion of local bikers, 8 locations have been selected to operate as gates; since link gradients are low (the area is basically plain), travel direction does not affect travel time and we considered only pairs $k = (\gamma, \nu) \in \Gamma^2 : \gamma < \nu$, thus yielding 28 origin and destination pairs. The map on the left of Fig. 1 depicts the abstract graph overlapping the physical map of the region [15]. Nodes $37, 46, 49, 53, 68, 75$ are railway stops, nodes $1, 18, 49, 57, 60$ are bus stops: the former are blue circles, the latter orange ones. Both bus and train provide for bike transport. Gateways are marked as green squares and correspond to nodes $1, 18, 49, 57, 60, 70, 75, 80$. Nodes 70 and 80, shown in red on the map on the right on Fig. 1, have been chosen as gateways since they are main locations on the road network. Node 70 lays on EuroVelo route 7 and 60 is along the Iron Curtain Trail which is part of EuroVelo 13. Edges distance (km) and travel time (minutes) are reported.

The corresponding abstract graph G is depicted in the middle of Fig. 1, unfit edges marked in red. Note that on the subgraph induced by discarding unfit edges - the 0-budget scenario - some gateways are not connected while others are only by routes circuitous and long, violating the planning requirement regarding directness and continuity.

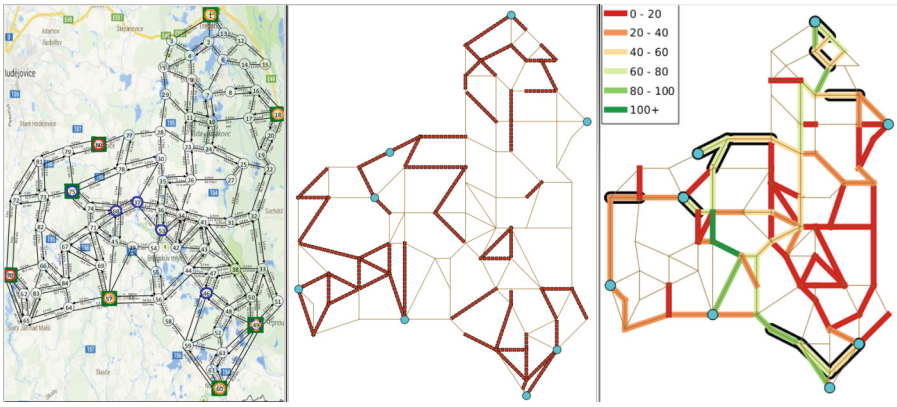


Fig. 1. Graph G overlapping the physical map of Trebon (on the left): gates are marked with colored circles. In the center, the abstract graph induced by the input data (unfit edges in red). The solution network is depicted on the right.

Concerning attractiveness, the three user classes $U = \{1..3\}$ of the previous studies (users fond of culture, of local food and traditions, and of nature and sports) have been modeled, yielding $n_C = 84$ commodities (u, k) if disregarding path orientation. In this work we are not concerned about the process according to which attractiveness has been computed, as we take these values as parametric inputs to our approach. In brief, users' point of view has been taken into account by listing all PoIs in the area that are relevant for each user class, and then by collecting and averaging the evaluation of teams of local bikers. In this way, scores for first and second traversal of each PoI have been assigned, and the attractiveness of edges and nodes was computed based on the PoIs located there and on the category of interest. The objective function is the sum of the attractiveness of the thematic itineraries selected for the three user classes. The different solutions computed along the search can be analyzed to reveal non dominated points in the three dimensional space of each user class attractiveness and provide alternative solutions to the decision maker.

The parameters and settings used in this experimental campaign are the following. The maximum number of iterations depends on n_e, n_{f2}, and on the number of edges in budget maximal sets (about 10) and on the number of unfit edges (about 60): n_e, i.e., the number of brother nodes from the same father in the search graph, is set to 5, while $n_{f1} = 20$ and $n_{f2} = 10$, which keeps the number of nodes at each iteration within $5 \cdot 10 = 50$. Figure 2 depicts a sketch

of the search graph for the first 9 levels, to give a taste of how its size evolves. In particular, the maximum size is immediately reached, that is 5 nodes at level 2, 25 at level 3 and 50 onwards to termination. Stopping conditions are reached when each active node yields to the next level only nodes whose simplified model is infeasible. On the other hand, infeasible nodes are found as early as level 3 (red nodes), meaning that there are very small subsets of edges which can not be altogether discarded. Each such set either identifies a cut separating two gateways in the topological term, or they cut all time feasible routes between origin and destination. While the first kind of sets could be identified by looking at the graph (think of gateways 60, 70 or 80), the second is not as it involves the routing part. Figure 2 also reports the number of duplicated subproblems in the search graph (blue nodes), i.e. subproblems with the same set of pruned edges. This number decreases while the solutions in the same level diversify from each other. It also shows the incidence of filtering, before and after routing, which is crucial to keep the size manageable. Note that the search does not deploy along a decision tree as a standard *branch and bound*; indeed, edge pruning may lead to duplicate nodes that are spotted and discarded.

In the simplified model the maximum path duration for commodity $c = (u, k)$ is set to $\tau = 1.5 \cdot \tau_k$ where $k = (\gamma, \nu)$ and τ_k denotes the length of the shortest path from γ to ν. When solving MACTPP to evaluate the quality of the network, the maximum duration of the itineraries is set to $\tau_1 = 1.5 \cdot \tau_k$, $\tau_2 = 1.6 \cdot \tau_k$, and $\tau_3 = 1.7 \cdot \tau_k$, to cater for different options. This yields a total of 252 MACTPP instances to be solved when evaluating the quality of the network made of the edges selected by the simplified model. However, as discussed, such MACTPP instances tend to be *easy* and each takes few milliseconds. Finally, budget level is set to $2000 \cdot 10^3$ €.

We report the results obtained by applying NFPS to the Trebon data. The simplified model and the MILP model of MACTPP have been coded in AMPL and solved by ILOG Cplex 12.5 on a quad core laptop with i7 processor.

The shape of the resulting network (reward 59849.5) is depicted on the right side on Fig. 1: it is rather sparse and itineraries tend to share the same edges, according to the colors in the legend, in particular in the middle of the network (gates are at the borders). The simplified model selects 97 edges which go down to 91 after the routing phase. Of these, 11 are unfit and use most of the available budget (1982 over 2000) as expected for a budget-maximal edge set.

Now consider the reward of the best each solution of each iteration as compared to the reward of the father node, as sketched in Fig. 3. The positive trend of the best solution reward shows that the search strategy is able to find better solutions in the proximity of good quality available ones. At the same time though, it can be observed that the father of a starring solution node is not the highest reward node of its own iteration at most layers. This fact is a point in favor of keeping a large number of active nodes and setting pretty loose thresholds for filtering (parameters n_{f1} and n_{f2}). Diversity enhances the search.

The cheapest network. NFPS was motivated by the need for keeping budget under control. Indeed, SM provides a feasible solution whenever it exists for a

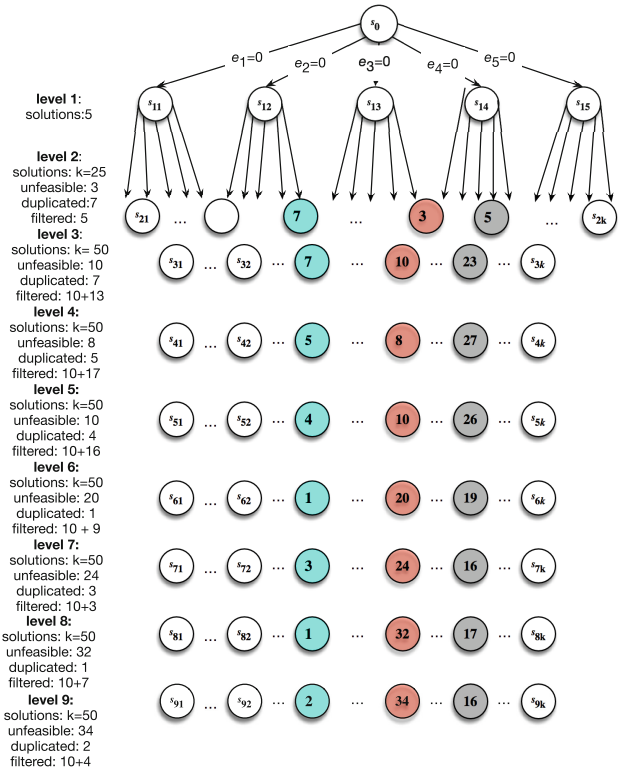


Fig. 2. The subproblems induced by edge pruning.

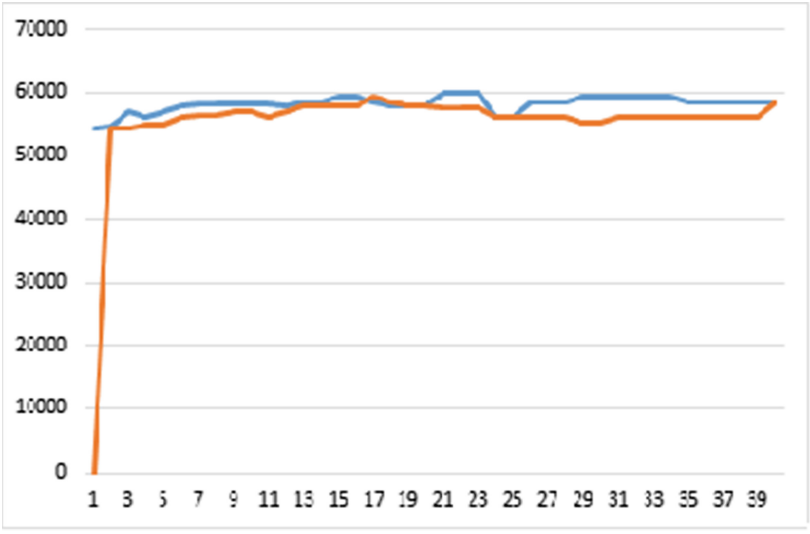


Fig. 3. Best solution reward and reward of the father node at each iteration.

given budget. By a slight change SM can be used to compute the minimum required budget, i.e., by using the retrofitting edge costs instead of edge attractiveness as objective function coefficients and getting rid of constraint 4. For our benchmark the minimum budget is $1.474 \cdot 10^3$ and reward is 54.353. Note that for this budget PFNS fails to provide a feasible solution.

Non dominated solutions. The proposed approach may also provide decision makers with appealing alternatives made of the non dominated solutions with respect to users attractiveness which are collected along the search. Since attractiveness is computed out of subjective ratings, it may be worth to consider alternative high quality solutions. As an example, consider the results for budget $1.600 \cdot 10^3$ (again, PFNS fails for this budget). NFPS takes 7 iterations and returns a reward of 55370.8.

Table 1 reports the data for the 5 best solutions, ranked according to total reward. If we disaggregate the reward components and show the contribution of each user class, it can be seen that these solutions are not dominated, and provide different thematic itineraries. Therefore, our approach can also be used to retrieve alternative solutions which are very close to the best one concerning total reward, but offer a better choice for at least one user class with respect to the other networks.

Table 1. Disaggregated and total reward for the best 5 solutions (budget 1600)

Sol	U1	U2	U3	Network reward
1	16001,63	21278,43	18090,75	55370,8
2	16264,12	21557,39	17536,42	55357,9
3	16264,52	21256,76	17774,85	55296,1
4	16249,55	21669,89	17322,96	55242,4
5	16171,97	21831,21	17187,17	55190,3

Conclusions. This paper addresses the design of cycle tourists devoted cycling networks. It provides a short review of the recent literature on cycling network design in order to retrieve the most common planning criteria and compare common practice in the design of infrastructures devoted to functional cycling versus leisure cycling. Based on these premises, we propose a mathematical optimization model which incorporates such planning criteria. These are embedded either as model constraints or in the objective function formulation, or by meeting planning requirements by construction. In particular, the model's objective function is the maximization of the network attractiveness for several classes of users, while the budget available for retrofitting or building the links of the network is formulated as a global constraint. Connection among the network access points is ensured by the flow reformulation, in which users are seen as flow units traveling from an origin gate to a destination gate, for each pair of gates. Several thresholds for route maximum duration are also formalized in the model.

Since the complete model is computationally demanding, it can not be solved by commercial solvers and a heuristic approach is proposed. The method exploits the solution of a simplified model yielding a feasible network, if any, for a given input graph. The simplified model acts as the core of an iterative algorithm due to explore (part of) the feasible region of the problem. It is run on different graphs obtained by increasingly pruning different sets of edges as long as feasibility holds. Optimality can not be guaranteed but computational results show the search capability of improving the initial solution to a large degree. The same simplified model can be used to determine the minimum budget necessary to build a network while guaranteeing all the required network quality standards that have been previously discussed. The approach has been tested for a real network, located in the Trebon region, Southern Bohemia, Czech Republic. Data regarding distances, retrofitting costs according to links status at present, and travel duration are real. Data modeling users utility functions that were previously calibrated with the contribution of local bikers, have been retrieved from previous works. The resulting case study is challenging and quite representative. Indeed, several links are already fit for cycling but either two gateways are not connected on the current infrastructure or the existing routes are too long, whereas retrofitting few links provides a much greater and better choice of itineraries. Since several mending options are possible (a combinatorial number), the only chance to scrutinize them is to implicitly consider them as the feasible solutions of a network design model, as we propose. The computational time amounts to a few hours on a regular personal computer, which is adequate when addressing a design problem. Data analysis shows that the utility function of several classes of users can be handled, providing high quality solutions that meet users needs. Indeed, the thematic routes selected for each user class and connecting the same pair of gates are based on the same network but can be rather different from each other. At the same time, our approach allows to collect the non dominated solutions computed during the search and to make them available to the decision makers to enlarge their set of options. This option can be useful when designing corridors in case of conflicting user preferences. Finally, this method allows for an incremental deployment of the network infrastructure, building some routes before others and prioritizing according to estimated demand. However, even if the implementation is carried out a step at a time, each route would still be part of a comprehensive project that guarantees the best use of the financial investment. We thus believe that this approach could lay at the heart of quantitative based tools to support local administrators in decision making.

References

1. Arup, O., and Partners Ltd., Department for Transport: Cycle Infrastructure Design. Local Transport Note 2/2008, TSO (2008)
2. Barrena, E., Laporte, G., Ortega, F., Pozo, M.A.: Planning ecotourism routes in nature parks. In: Ortegón Gallego, F., Redondo Neble, M.V., Rodríguez Galván, J.R. (eds.) Trends in Differential Equations and Applications, pp. 189–202. Springer International Publishing, Cham (2016)
3. Buehler, R., Pucher, J.: Cycling to work in 90 large American cities: new evidence on the role of bike paths and lanes. Transportation **39**, 409–432 (2012)
4. Cernà, A., Cernỳ, J., Malucelli, F., Nonato, M., Polena, L., Giovannini, A.: Designing optimal routes for cycle-tourists. TR Pro. **3**, 856–865 (2014)
5. Chang, H.W., Hsieh, H.N.: Recreational cycling routes investment selection - Hsinchu Technopolis case by applying ZOGP. J. Eastern Asia Soc. Transp. Stud. **10**, 1227–1242 (2013)
6. Cox, P.: Strategies promoting cycle tourism in Belgium: practices and implications. Tour. Plan. Dev. **9**(1), 25–39 (2012)
7. CROW: Design manual for bicycle traffic, Ede, Netherlands (2007)
8. Deenihan, G., Caulfield, B., O'Dwyer, D.: Measuring the success of the Great Western Greenway in Ireland. Tourism Manage. Perspect. **7**, 73–82 (2013)
9. Downward, P., Lumdsom, L.: The development of recreational cycle routes: an evaluation of user needs. Managing Leisure **6**, 50–60 (2001)
10. Duthie, J., Unnikrishnan, A.: Optimization framework for bicycle network design. J. Transp. Eng. ASCE **140**(7) (2014). doi:10.1061/(ASCE)TE.1943-5436.0000690
11. Fischetti, M., Salazar-Gonzalez, J., Toth, P.: The generalized traveling salesman and orienteering problems. In: Gutin, G., Punnen, A.P. (eds.) The Traveling Salesman Problem and Its Variations, pp. 609–662. Springer, New York (2007)
12. Foster, C.E., Panter, J.R., Wareham, N.J.: Assessing the impact of road traffic on cycling for leisure and cycling to work. IJBNPA **8**, 61 (2011)
13. Furth, P.G., Noursalehi, P.: Evaluating the connectivity of a bicycling network. In: TRB 94th Annual Meeting (15–5612) (2015)
14. Gavalas, D., Konstantopoulos, C., Mastakas, K., Pantziou, G.: A survey on algorithmic approaches for solving tourist trip design problems. J. Heuristics **20**(3), 291–328 (2014)
15. Giovannini, A., Malucelli, F., Nonato, M.: Cycle tourist network design. In: Proceedings of EWGT 2016, 19th EURO Working Group on Transportation Meeting, Istanbul, 5–7 September 2016 (2016, to appear on TR Pro)
16. Handy, S., van Wee, B., Kroesen, M.: Promoting cycling for transport: research needs and challenges. Transp. Rev. **34**(1), 4–24 (2014)
17. Heinen, E., Maat, K., van Wee, B.: The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances. Transp. Res. D-TR E **16**(2), 102–109 (2011)
18. Heinen, E., van Wee, B., Maat, K.: Commuting by bicycle: an overview of the literature. Transp. Rev. **30**(1), 59–96 (2010)
19. Hulla, A., O'Holleranb, C.: Bicycle infrastructure: can good design encourage cycling? Urban Plann. Transp. Res. **2**(1), 369–406 (2014)
20. Larsen, J., El-Geneidy, A.: A travel behavior analysis of urban cycling facilities in Montréal. Canada. Transp. Res. D-TR E **16**(2), 172–177 (2011)
21. Lin, J.J., Liao, R.J.: Sustainability SI: bikeway network design model for a recreational bicycling in scenic areas. Netw. Spat. Econ. **16**(1), 9–31 (2016)

22. Lin, J., Yu, C.: A bikeway network design model for urban areas. Transportation **40**, 45–68 (2013)
23. Lumsdon, L.: Transport and tourism: cycle tourism-a model for sustainable development? J. Sustain. Tourism **8**(5), 361–377 (2000)
24. Ma, L., Dill, J., Mohr, C.: The objective versus the perceived environment: what matters for bicycling? Transportation **41**(6), 1135–1152 (2014)
25. Malucelli, F., Giovannini, A., Nonato, M.: Designing single origin destination itineraries for several classes of cycle-tourists. TR Pro. **10**, 413–422 (2015)
26. Manton, R.: Novel methods for planning and design of greeways for cycling (2016)
27. Mesbah, M., Thompson, R., Moridpour, S.: Bilevel optimization approach to design of network of bike lanes. Transp. Res. Rec. **2284**, 21–28 (2014)
28. Nash, E., Cope, A., James, P., Parker, D.: Cycle network planning: towards a holistic approach using temporal topology. Transp. Plann. Techn. **28**(4), 251–271 (2005)
29. Nelson, A.C., Allen, D.: If you build them, commuters will use them: association with bicycle facilities and bicycle commuting. Transp. Res. Rec. **1578**, 79–83 (1997)
30. Parkin, J., Wardman, M., Page, M.: Models of perceived cycling risk and route acceptability. Accid. Anal. Prev. **39**(2), 364–371 (2007)
31. Smith, H.L., Haghani, A.: A mathematical optimization model for a bicycle network design considering bicycle level of service. Paper 12–3307, TRB 91st Annual Meeting, Washington DC, USA (2012)
32. Souffriau, W., Vansteenwegen, P., Vanden Berge, G., Van Oudheusden, D.: The planning of cycle trips in the province of east Flanders. Omega **39**, 209–213 (2011)
33. Trendy cycling. http://bicy.it/docs/86/Trendy_cycling_EN_web.pdf
34. Vansteenwegen, P., Souffriau, W., Van Oudheusden, D.: The orienteering problem: a survey. Eur. J. Oper. Res. **209**(1), 1–10 (2011)
35. Vansteenwegen, P., Van Oudheusden, D.: The mobile tourist guide: an or opportunity. Oper. Res. Insight **20**(3), 21–27 (2007)
36. Wardman, M., Tight, M., Page, M.: Factors influencing the propensity to cycle to work. Transp. Res. A-Pol **41**(4), 339–350 (2007)
37. Weston, R., Davies, N., Peeters, P., Eijgelaar, E.: The European Cycle Route EuroVelo: challenges and opportunities for sustainable tourism. European Parliament Report (2012)
38. Yang, C., Mesbah, M.: Route choice behaviour of cyclists by stated preference and revealed preference. In: Proceedings of Australasian Transport Research Forum (2013)

The Influence of Transportation Service Level on a Municipal Service Center's Costs: A Numerical Study Based on Supply Chain Management Models

Matan Shnaiderman(✉)

Department of Management, Bar-Ilan University, Ramat-Gan, Israel
shnidem@biu.ac.il

Abstract. This paper deals with scheduling of appointments between providers and customers (with reservations or walk-in ones) in municipal service centers. In order to improve the service level and reduce the uncertainty of the number of customers' demand, a free transportation service from the customers' locations to the service center and back is operated. An optimal transportation service level (TSL) is set in order to minimize the provider's total idle time and overtime on the one hand, and the transportation service's operation cost on the other hand. We show how the optimal number of customers to book in advance depends, analogically to inventory management models, on the ratio between the provider's idle time ("surplus") and overtime ("shortage") unit costs. The lower impact of the TSL on the demand, the lower optimal TSL and expected cost, especially if the surplus cost is higher than the shortage cost. Furthermore, we add a safety constraint, according to which, the TSL level must be high enough such that the probability for non-arrival of at-risk customers is small. We numerically find that high percentage of at-risk customers in the population, may significantly increase the TSL, and consequently, lead to meaningful jump of the expected cost (up to 26%).

Keywords: Transportation service level · Provider's idle time/overtime · Walk-in customers safety constraint

1 Introduction

Public transport service quality is an important factor in influencing travelers' behavior [3] to shift from the private car, as well as affording excluded groups, such as the elderly to increase mobility. Increased trip length or travel distance lowers activity participation rates, and as a result, the quality of life decreases, especially in case of old people [6, 7].

This paper deals with scheduling of appointments between providers and customers in service centers (medical clinics, equipment repair centers, maintenance etc.). In order to improve the service level and reduce the uncertainty of the number of customers that actually arrive at the center, a free transportation service from the customers' flexible locations to the service center and back is operated, based on the demand-responsive

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_12

transportation (DRT), or transportation-on-demand, approach, which refers to transportation of passengers between specific origins and destinations according to users' request. According to Cordeau et al. [2], transportation on demand systems have become increasingly popular in recent years, and in particular, the case of dynamical request receiving, in which vehicle routes are adjusted in real-time to meet demand. The authors mention that there are three often conflicting objectives: maximizing the number of requests served, minimizing operating costs and minimizing user inconvenience. Under DRT, several types of vehicles operating in response to calls from passengers to the transit operator, who dispatches a vehicle to pick up the passengers and transport them to their destinations vehicles. The vehicles do not operate over fixed routes or schedules, as broadly described in [16]. The vehicles may operate over variety of routes or schedules, and may pick up passengers at different points before taking them to their destinations. The authors claim that in contrast to previous researches, who only deal with optimization, the DRT approach also takes into account real-time traffic, incidents and personal behavior.

According to Quadrifoglio et al. [14], the passage of the Americans with Disabilities has led to a tremendous growth of the demand for the DRT services (more than 5000 vans and 4200 cubs provide service, generating 8 million trips per year in Los-Angeles County). Also, the average cost per passenger trip was 20.8$ in 2002, while under fixed-route lines, it was only 2.4$. The productivity and cost of the DRT is measured in their research in terms of fleet size, total miles or deadhead miles (i.e. empty trip miles driven by the vehicle). Moreover, the time-window length (which is the time range in which the provider must pick-up the customer) is an important factor impacting productivity and cost.

High importance of transportation service level (TSL) and transportation quality is expressed in transportation (access) to healthcare locations. As mentioned in [21], equal access to healthcare is a guiding principle in countries where health care is mainly provided within public health care systems. The authors consider how distance influences caseload (i.e. the number of patients using services) and service utilization of psychiatric services in acute inpatient wards, outpatient clinics and community mental health centers in Verona, Italy. They find that the caseload decreased with increasing distance. For example, at a distance of 10 km, there was decrease of 80, 60 and 85 in the three types of facilities respectively. According to Grant et al. [5], availability of transportation is essential for timely access to healthcare services. In particular, in each year, 4% of children nationwide, including 9% of children in low income households, miss an appointment because of transportation unavailability. The authors present a comprehensive literature review regarding the relationship between transportation and healthcare access, in terms of transportation disadvantages, geographical access, child health access, chronic health conditions and nonemergency medical transportation. Tsuji et al. [18] investigate the relationship between geographical accessibility (time and distance) and the utilization of outpatient services in Tokushima prefecture, Japan. Their results show negative correlations (between −0.4 to −0.6) both in distance and time.

On the other hand, there are cases in which the distance does not significantly affect the arrival rate of patients to health care locations, while high-quality transportation services do exist. Whetten et al. [19] examine how HIV-positive persons utilize

transportation services (buses and taxis, which are free of charge, as well as mileage reimbursement for private transportation), and in particular, how that utilization is influenced by distance. The authors find that 74% of participants utilized the transportation services, and that increased distance to care did not decrease utilization of the treatment program when transportation was provided to the customer when necessary. On the other hand, long distances to care and lack of adequate transportation may significantly reduce receiving services (specifically mental health care).

In this paper we consider customers for whom a reservation has been booked in advance as well as "open-access" or "walk-in" customers. All customers necessarily meet the provider, regardless of the reason for their visit (see e.g. [11] for the case in which the clinic does not accept all walk-in patients). This approach is common in planning of periodic schedules in healthcare facilities. The reason is that, as mentioned in [13], the provider aspires to match his/her daily capacity to the demand, taking into account the no-show rates, on the one hand, and to see customers when they want to be seen (similar to the idea of just-in-time models in manufacturing industry) on the other hand. The authors calculate the optimal percentage of open-access appointments in order to match daily provider capacity to the demand and maximize the expected daily number of appointments. They find that optimal percentage is significantly affected by distribution of demand for open-access appointments as well as the ratio of the show-up rates of the two customer types. The variability in customer arriving rate, and in particular, the no-show rate, may negatively impact the service level, as stated in [12]. The authors mention that high variability in the number of customers (patients) may result in provider overtime (which is the total time that the provider spends to see customers after working hours.) or, alternatively, in provider long idle time (i.e. the total time when the provider does not see patients working hours). Motivated by these issues, they consider a mean-variance model, such that the expected number of customers consulted by a provider is maximized, while its variance is minimized.

Taking into account the demand's variance, and in particular, minimizing the provider's overtime and idle time, is an approach which is based on supply chain models. Overtime and idle time are analogous to inventory shortage and surplus respectively. The lower variance of demand, the lower expected inventory costs [15]. In [10] the authors investigate how to combine appointments for pre-booked, open-access and walk-in patients, such that the expected value of a weighted summation of waiting time within the patients and the provider's idle time and overtime is minimized. They show that the optimal scheduling templates are significantly affected by, in addition to the factors mentioned above, on the coefficients of provider idle time and overtime (analogous to supply chain models). Another factor that influences the optimal scheduling is the patients' waiting time cost coefficient. One of the reasons is that long waiting time may lead to high no-show and cancelation rates (see also [1]). In [8], a model for public transportation planning, which is based on supply chain models, is presented. The public vehicles' headways and capacities are set such that the expected vehicles' empty-seat and overload, which are analogous to inventory surplus and shortage respectively, are minimized.

In contrast to the researches mentioned above, this paper combines: (a) scheduling provider and customers' appointments, and (b) transportation service operations from customers' location to the service center (and vice versa). The optimal TSL is set such that the sum of the provider's total idle time and overtime, on the one hand, and the transportation service's operation cost, on the other hand, is minimized. This level may be expressed in the transportation service zones and times as well as in the vehicles' capacity and size.

The rest of this paper is organized as follows. In Sect. 2 we formulate the problem. In Sect. 3 we calculate the optimal number of reservations that the service center should book in advance. Section 4 deals with setting of the optimal TSL, taking into account the uncertainty of the demand as well as the transportation cost. Moreover, we consider how rigorous safety approach affects the results. Section 5 concludes.

2 Problem Formulation

We consider a municipal service center which serves two types of customers, who meet a provider there. The first type refers to customers for whom a reservation has been booked in advance ("type-1"). Customers of the other type ("type-2") are "walk-in" customers, who arrive with no reservations. The quantity of customers of type-1 on a given day, denoted by q, is set in advance. The capacity of customers that can be served during the service center's daily regular time is Q. Let D be the random daily demand for service (namely, the actual number of customers that arrive to the service center during the daily regular time) and let d be its realization. This demand is a sum of two demands, D_1 and D_2 (with realizations d_1 and d_2), which denote the number of customers of type-1 and type-2 respectively. These two demands are assumed to be independent. In order to improve the service level, the service center holds vehicles that may transport customers to the service center back and forth. Let x denote the TSL, then we assume that $0 \le x \le 1$, where $x = 0$ means no transportation service and $x = 1$ means maximal service. This level affects the number of customers (of both types) that actually arrive to the service center and meet the provider. The probability that each type-1 customer arrives to the service center is $\alpha_1(x)$. The function $\alpha_1(x)$ is non-decreasing, such that $\alpha_1(0) = r_1$ (for some $0 \le r_1 \le 1$) and $\alpha_1(1) = 1$. Therefore, given the values of x and q, the demand D_1 is binomially distributed such that

$$D_1 \sim \mathrm{B}\big(q,\, \alpha_1(x)\big). \tag{1}$$

The demand D_2 satisfies a Poisson process, such that

$$D_2 \sim \mathrm{Poi}(\lambda(x)), \tag{2}$$

where $\lambda(x)$ is non-decreasing, $\lambda(1)$ is equal to a maximal arrival rate Λ and $\lambda(0) = r_2\Lambda$ for some $0 \le r_2 \le 1$. The maximal value Λ denotes the arrival rate while there are no customers who refrain from arriving to the service center due to difficulties of transport (none having of vehicle, long distance travel, difficulty of driving etc.), see [1]. The parameters r_1 and r_2 refer to the impact level of the TSL on the capacity of customers

that arrive to the service center. The lower values of those parameters the higher effect. Low values may refer to poor or old population, who is strongly supported by the transportation service. High values, on the other hand, correspond to durable population. Similarly to models deal with supply chain and inventory management [9, 20], the service center pays unit surplus cost c^+ for every unexploited service time (i.e. provider's idle time), as well as unit shortage cost c^- for every extra time (provider's overtime). In addition, there exists a transportation cost $T(x)$. This function, which is increasing in x, includes holding and operation costs of the vehicles on that day. These costs depend on the number of vehicles used, on total route duration and on total distance traveled by the vehicles [2]. Consequently, the total cost, including regular time efficiency cost (analogous to inventory cost) as well as the transportation cost, is

$$C(x) = c^+(Q-D)^+ + c^-(D-Q)^+ + T(x),$$

and its expected value is

$$\mathrm{E}C(x,q) = c^+ \sum_{d=0}^{Q-1} (Q-d)\Pr(D=d|x,q) + c^- \sum_{d=Q+1}^{\infty} (d-Q)\Pr(D=d|x,q) + T(x), \tag{3}$$

where $\Pr(\cdot)$ denotes the probability of an event. Given a TSL x as well as number of reservations q, the probabilities in the expectation (3) are calculated according to the following proposition.

Proposition 1: Assuming (1) and (2), the probability of the demand D to be equal to an integer $d \geq 0$ is

$$\Pr(D=d|x,q) = e^{-\lambda(x)} \sum_{n=0}^{\min(q,d)} \binom{q}{n} \alpha_1^n(x)\big(1-\alpha_1(x)\big)^{q-n} \frac{\lambda^{d-n}(x)}{(d-n)!}. \tag{4}$$

Proof: According to the Law of Total Probability,

$$\Pr(D=d|x,q) = \Pr(D_1+D_2=d|x,q) = \sum_{d_1=0}^{\min(q,d)} \Pr(D_1=d_1|x,q)\Pr(D_2=d-d_1|x), \tag{5}$$

and from (1) and (2) we respectively obtain

$$\Pr(D_1=d_1|x,q) = \binom{q}{d_1} \alpha_1^{d_1}(x)\big(1-\alpha_1(x)\big)^{q-d_1} \tag{6}$$

for $0 \leq d_1 \leq q$, as well as

$$\Pr(D_2=d-d_1|x) = \frac{\lambda^{d-d_1}(x)}{(d-d_1)!} e^{-\lambda(x)}, \tag{7}$$

for $0 \leq d_1 \leq d$. By substituting (6) and (7) in (5), as well as denoting n instead of d_1, Eq. (4) is received. □

Using two-stage backward dynamic programming, we first calculate the optimal number of reservations, q, as function of a value of TSL.

3 The Optimal Number of Reservations

Given the TSL x, the service center has to set the optimal value of q (denoted by q^*), namely, the number of customers reserved in advance to meet the provider on a specific day. Fix $x = x_0$, then a new objective function, Ec_2, is obtain from (3) as follows

$$Ec_2(q|x_0) = EC(x_0, q),$$

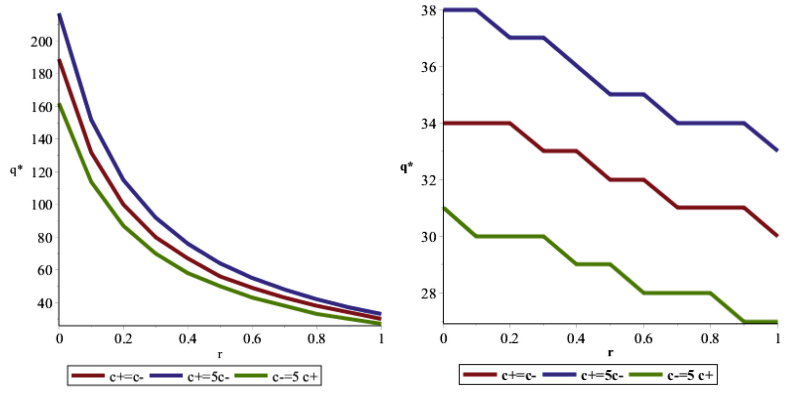
and the goal is to minimize it subject to (3) and (4).

We now show numerical results. We mention how the results (i.e. the optimal values of the decision variables as well as the minimal costs) are affected, as takes place in inventory management models, by the ratio between the unit surplus and shortage costs. In particular, we consider three cases: $c^+ = c^- = 2$; $c^+ = 5$ and $c^- = 1$; $c^+ = 1$ and $c^- = 5$.

Let $Q = 40$, $\Lambda = 10$ and assume that $\alpha_1(x)$ as well as $\lambda(x)$ is linearly increasing in [0,1].

First, we consider how the optimal number of reservations depends on the parameters r_1, r_2 and Λ.

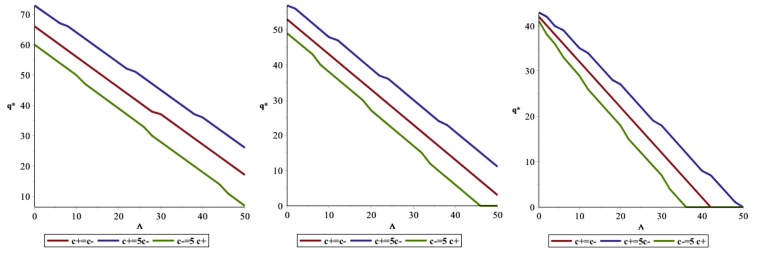
Figure 1 shows how the impact level of the TSL on customers' arriving (where "r" presents both r_1 and r_2) affects the optimal number of reservations. While the TSL is small, as illustrated in Fig. 1a for $x = 0.2$, the optimal number of reservations significantly depends on r. This optimal value is decreasing in r. If r is low, q^* may become greater than 200, as very few customers come to the service center, and the value of q^* exceeds 120 even if $c^- = 5c^+$. When r grows and tends to 1, the value of q^* may be reduced in more than 80%. On the other hand, when the TSL becomes high, then the expected quantity of customers (of both types) is significantly reduced. Moreover the impact of r on q^* becomes low, as shown in Fig. 1b for $x = 0.9$.



a. x=0.2 b. x=0.9

Fig. 1. Optimal capacities of reservations as functions of the TSL's impact level

As the average arriving rate of type-2 customers, λ, and in particular, Λ, grows, the optimal capacity of reservations of type-1 customers is supposed to decrease. Figure 2 presents those optimal values for $0 \leq \Lambda \leq 50$ as well as $r_1 = r_2 = 0.5$. These values are approximately linearly decreasing in Λ (while they are positive).



a. x=0.2 b. x =0.5 c. x=0.9

Fig. 2. Optimal capacities of reservations as functions of type-2's arriving rate

If the TSL is small and is equal to 0.2 (Fig. 2a), then it is preferable to book reservations for a few customers even if the shortage cost is five times higher than the surplus cost and Λ becomes equal to 50. When the TSL increases to 0.5, it may be optimal not to book any reservations if Λ exceeds 45 and c^- is greater than c^+, as illustrated in Fig. 2b. When the TSL is high and is close to 1, then q^* may be equal to 0, even if c^+ is 5 times greater than c^-, as shown in Fig. 2c.

Next, we calculate the optimal TSL in the service center.

4 Setting of the Optimal Transportation Service Level

Based on the previous section, let

$$\mathrm{E}c_1(x) = \mathrm{E}C(x, q^*(x)) \tag{8}$$

be the objective function while the TSL x is set under the constraint $0 \leq x \leq 1$. First, we consider how x affects the level of uncertainty.

4.1 The Influence of TSL on the Demands' Uncertainty

In the current subsection we assume that the transportation cost $T(x)$ is equal to zero, that is, the expected cost (3) contains only the surplus and shortage costs. As the variance of the total demand D becomes higher, these costs are expected to increase [15]. Since D_1 and D_2 are independent, the variance of D is

$$V_D(x) = V_{D_1}(x) + V_{D_2}(x), \tag{9}$$

and from (1) and (2) we respectively obtain

$$V_{D_1} = q^*(x)\alpha_1(x)(1 - \alpha_1(x)) \tag{10}$$

and

$$V_{D_2}(x) = \lambda(x). \tag{11}$$

As $q^*(x)$ is decreasing in x and the multiplication $\alpha_1(1 - \alpha_1)$ is decreasing in the interval [1/2, 1], we obtain from (10) the following remark.

Remark 1: Let x_0 satisfy $\alpha_1(x_0) \geq 1/2$, then the variance V_{D1} is decreasing in the interval $[x_0, 1]$. In particular, if $\alpha_1(0) \geq 1/2$, then V_{D1} is decreasing in [0, 1]. □

According to Remark 1, the uncertainty of the number of type-1 customers is supposed to decrease as the TSL grows. High TSL reduces the number of daily reservations and increases the probability of the booked customers to arrive. Thus, it usually becomes easier to estimate the exact daily quantity of type-1 customers in the service center. Next, as $\lambda(x)$ is increasing, we obtain from (11) the following.

Remark 2: The variance V_{D2} is increasing in x. □

According to Remark 2, high TSL leads to high arrival rate of type-2 customers. As a result, the uncertainty regarding the daily number of customers grows. We find that there is a conflict in the tendencies of the variances of D_1 and D_2 as x grows. Thus, taking into account only the surplus/shortage costs (namely, the uncertainty of the total daily demands), setting of the optimal TSL, denoted by x^*, is not trivial (see (9)). Under a small arrival rate Λ, the optimal TSL is supposed to be high (and close to 1). On the other hand, as Λ grows, the type-2 customers become more dominant in the daily demand, and the optimal TSL may decrease. We now numerically demonstrate these results.

Let $Q = 40$ and $c^+ = c^- = 2$. Figure 3 illustrates how the optimal TSL depends on r_1 and r_2 ("r"), for three values of Λ.

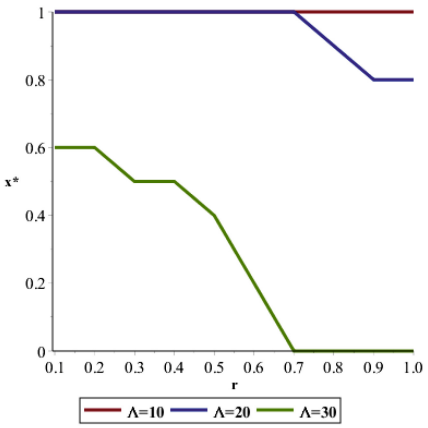


Fig. 3. Optimal TSLs under no transportation cost

While Λ is low and is equal to 10, the optimal TSL is always equal to 1, as the demand of type-1 customers is more dominant than that of type-2 customers. When Λ increases and is equal to 20, the value of x^* may be lower than 1. If r is lower than or equal to 0.7, then x^* remains 1, but if it exceeds that value, x^* decreases and becomes 0.8. As Λ becomes 30, the optimal TSL is always lower than 1. If $r = 0$, then $x^* = 0.6$, and when r exceeds 0.7, x^* vanishes, that is, in order to prevent high arrival rate $\lambda(x)$, it is preferable for the service center not to operate any transportation service. Furthermore, we can see in Fig. 3 that x^* is decreasing in r. This result is expected while the transportation cost does exist, i.e. $T(x) > 0$, since it is not preferable to pay high cost for transportation if its impact on the demand is low (see Subsect. 4.2 below). However, we receive this result while $T(x) = 0$ as well, and it can be explained as follows. If r is low, the variance V_{D1} is significantly decreasing in x, and therefore it is preferable to set a higher TSL. On the other hand, when r grows, the variance of V_{D1} is negligibly decreasing in x, while the increasing of the variance of V_{D2} remains more meaningful (recall that Λ is high). Hence, it is better to reduce the value of x in order to prevent high arrival rate of type-2 customers.

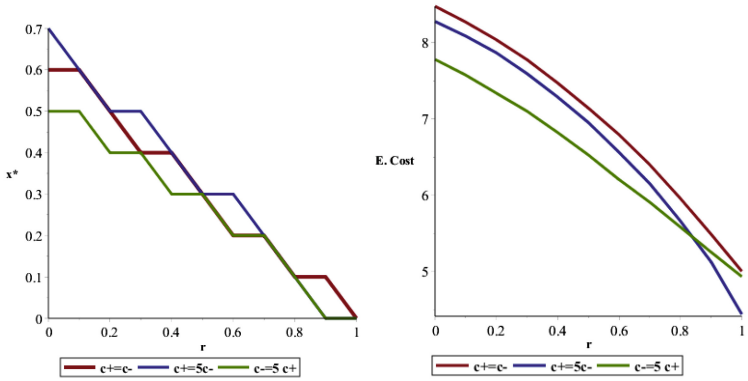
In the following subsection, the transportation cost is taken into account.

4.2 Optimal TSL Under Positive Transportation Cost

We now assume that $T(x) > 0$, that is, setting of TSL involves the transportation cost. Similarly to Shi and Xiao [17], this cost is assumed to be convex in x, namely, when x is low and is close to 0, the increasing rate of that cost is small, that is, adding of a little TSL does not significantly increase the daily expected cost. On the other hand, when the transportation cost is high and is close to 1, the increasing rate of that cost becomes greater. The reason is that increasing the TSL up to absolutely 1 (i.e. maximal transportation service) requires a lot of resources and efforts, many more than under a little lower level. As a result, the optimal TSL x^* is usually supposed to have intermediate values between 0 and 1, as we numerically illustrate in the following example.

Let $Q = 40$, and assume that the transportation cost is quadratic in the TSL, that is, $T(x) = \eta x^2/2$ (see [17]). First, we mention in Fig. 4 how the optimal TSL, as well as the expected cost, is affected by the impact level of TSL while $\eta = 8$.

The optimal TSLs decrease as r increases (Fig. 4a), since its influence on the quantity of customers that arrive to the service center becomes low, that is, paying transportation costs becomes unjustified. As a result, the expected cost is significantly reduced, especially in the case of high surplus cost, in which the TSL has the most meaningful reduction. Figure 5 presents the optimal TLS values (a) and the corresponding expected costs (b) as functions of the transportation cost coefficient η, for $r_1 = r_2 = 0.5$. Increasing of that parameter reduces of course the optimal TSL, since high investment in transportation becomes nonremunerative. The total expected costs increase in η (although the TSL is reduced). Reduction of the TSL leads to decrease of the demand, and consequently to increase of idle time (surpluses). Therefore, the most significant cost growth is in case of high surplus cost, while the least growth is corresponding to low surplus cost. In particular, the total expected cost in the former case is lower than that of the latter case while $\eta < 3$, and higher otherwise.



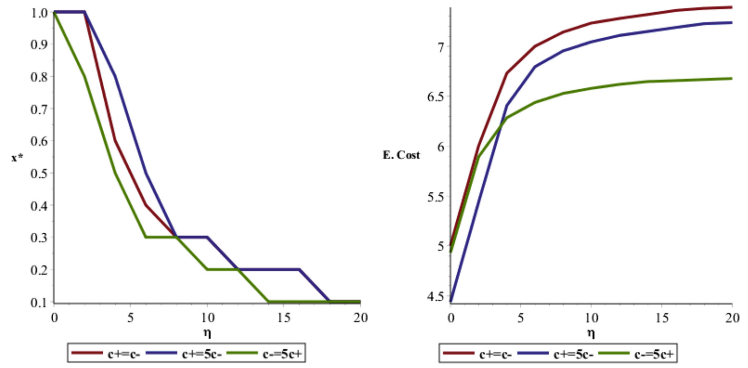
a. Optimal TSL b. Expected cost ($\times 10^3$)

Fig. 4. Optimal TSLs (a) and expected costs (b) as functions of the TLS's impact level under transportation cost

Analogous results are obtained when we consider how the optimal TSL and the corresponding expected costs are affected by the maximal possible arriving rate of type-2 customers, as illustrated in Fig. 6.

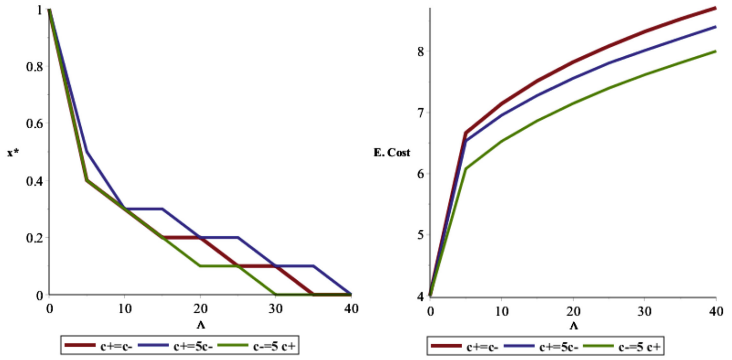
As before, the optimal TSL are reduced (Fig. 6a) whereas the total expected costs increase (Fig. 6b) in Λ. The reason for the latter result is as follows. As the optimal TSL is decreasing in Λ, then the variance of D_1 becomes higher (see Remark 1 above). Moreover, as Λ grows, the actual arriving rate $\lambda(x^*|\Lambda)$ increases (although the value of x^* decreases), namely, the variance of D_2 increases as well (Remark 2). Thus, the uncertainty of the total demand increases, and as a result, the surplus/shortage cost grows.

Finally, we note that according to Figs. 4a, 5a and 6a, the expected cost under equal unit surplus and shortage costs are always the highest in our examples, even though both of them are not high (2 rather than 5). However, while the total demand D does not exactly match the capacity Q, then a unit cost of 2 is paid. While $c^+ = 5$ and $c^- = 1$, the



a. Optimal TSL b. Expected cost ($\times 10^3$)

Fig. 5. Optimal TSLs (a) and expected costs (b) as functions of the transportation costs



a. Optimal TSL b. Expected cost ($\times 10^3$)

Fig. 6. Optimal TSLs (a) and expected costs (b) as functions of the maximal arrival rate Λ

service center expensively pays for surplus of time (namely, low demand), and therefore, the decisions are set such that high demand is almost ensured, that is, a unit cost of 1 is supposed to be paid. Similarly, while $c^+ = 1$ and $c^- = 5$, the service center expensively pays for shortage of time (high demand). Hence, low demand is almost ensured.

In the next section we consider a more rigorous approach, according to which, the TSL must be high enough such that all customers in danger (emergency cases) almost surely arrive to the service center. We call this condition a "safety constraint".

4.3 The Effect of a Safety Constraint on the Optimal TSL

Type-2 customers may need the service center service urgently due to emergency cases (we call them "at-risk-customers"). However, low TSL may prevent arriving of such customers to the service center and lead to high risk. Consequently, we formulate a constraint such that the TSL must be high enough to almost surely prevent non arrival of at-risk-customers to the service center. Our paper does not deal with customers' waiting time, but it can be assumed that at-risk customers meet the provider as immediate as possible (see e.g. [4], where waiting times are taken into account and rescheduling of elective customers due to arrival of emergency customers is considered). Let $D_2^{(\Lambda)}$ denote the demand of type-2 customers under the maximal arrival rate Λ (corresponding to the case of $x = 1$), then

$$D_2^{(\Lambda)} \sim \text{Poi}(\Lambda). \tag{12}$$

Let β be the probability of every type-2 customer to be at-risk, and let Y be the number of at-risk customers among $D_2^{(\Lambda)}$. Given a realized value $d_2^{(\Lambda)}$, the conditional distribution of Y is

$$Y_{\left| d_2^{(\Lambda)} \right.} \sim \text{B}(d_2^{(\Lambda)}, \beta). \tag{13}$$

Next, let $\alpha_2(x)$ be the probability of an at-risk customer to arrive to the service center while the TSL is x (in particular, we assume that $\alpha_2(1) = 1$). This probability is assumed to be greater than $\alpha_1(x)$, the probability among the whole type-2 customers. Also, let Z denote the number of at-risk customers who do not come to the service center, then, given the value of Y,

$$Z_{|Y} \sim \mathrm{B}\big(Y,\ 1 - \alpha_2(x)\big). \tag{14}$$

Let ε be a small positive number which measures the safety rigorousness level, then the safety constraint that we add to our problem is

$$\Pr(Z = 0\,|x) \geq 1 - \varepsilon. \tag{15}$$

Proposition 2: Under (12)–(14), the probability of Z to be zero is

$$\Pr(Z = 0\ |x) = e^{-\Lambda} \sum_{n=0}^{\infty} \left(\sum_{k=n}^{\infty} \frac{\Lambda^k}{k!} \binom{k}{n} (1-\beta)^{k-n} \right) \big(\alpha_2(x)\beta\big)^n. \tag{16}$$

Proof: Using the Law of Total Probability twice, we have

$$\begin{aligned} \Pr(Z = 0\ |x) &= \sum_{y=0}^{\infty} \Pr(Y = y) \cdot \Pr(Z = 0 | Y = y,\ x) \\ &= \sum_{y=0}^{\infty} \left(\sum_{d_2^{(\Lambda)}=y}^{\infty} \Pr(D_2^{(\Lambda)} = d_2^{(\Lambda)}) \cdot \Pr\big(Y = y \,\big|\, D_2^{(\Lambda)} = d_2^{(\Lambda)}\big) \right) \Pr(Z = 0 | Y = y,\ x). \end{aligned} \tag{17}$$

From (12), (13) and (14) we respectively obtain

$$\Pr(D_2^{(\Lambda)} = d_2^{(\Lambda)}) = \frac{\Lambda^{d_2^{(\Lambda)}}}{(d_2^{(\Lambda)})!} e^{-\Lambda}, \tag{18}$$

$$\Pr\big(Y = y \,\big|\, D_2^{(\Lambda)} = d_2^{(\Lambda)}\big) = \binom{d_2^{(\Lambda)}}{y} \beta^y (1-\beta)^{d_2^{(\Lambda)} - y} \tag{19}$$

and

$$\Pr(Z = 0 | Y = y,\ x) = \big(\alpha_2(x)\big)^y. \tag{20}$$

By substituting (18)–(20) in (17), as well as denoting n and k instead of y and $d_2^{(\Lambda)}$ respectively, we obtain (16). □

According to (16), the probability of Z to be zero is increasing in x such that

$$\Pr(Z = 0\ |x = 1) = e^{-\Lambda} \sum_{n=0}^{\infty} \sum_{k=n}^{\infty} \frac{\Lambda^k}{k!} \binom{k}{n} (1-\beta)^{k-n} \beta^n = e^{-\Lambda} \sum_{k=0}^{\infty} \left[\frac{\Lambda^k}{k!} \sum_{n=0}^{k} \binom{k}{n} (1-\beta)^{k-n} \beta^n \right] = e^{-\Lambda} \sum_{k=0}^{\infty} \frac{\Lambda^k}{k!} = 1.$$

Thus, there exists a value $x_0 \leq 1$, such that if $x \geq x_0$, then constraint (15) is satisfied. That is, the problem remains solvable under constraint (15). Let

$$x_0(\varepsilon) = \min\{0 \leq x \leq 1 | \Pr(Z = 0|x) \geq 1 - \varepsilon\}$$

be the minimal value which satisfies (15) with respect to ε, and let x^{**} be the optimal TSL subject to (15). While ε is not very small, such that $x^* \geq x_0(\varepsilon)$ (where x^* denotes the optimal TSL ignoring (15)), then $x^{**} = x^*$, namely, constraint (15) is passive and does not affect the optimal solution. On the other hand, if the rigorous level is high and ε is small such that $x^* < x_0(\varepsilon)$, then $x^{**} = x_0(\varepsilon)$, that is, (15) becomes active and increases the TSL. In particular, if the impact level of the TSL on the arrival rate of at-risk customers is low, such that

$$e^{-\Lambda} \sum_{n=0}^{\infty} \left(\sum_{k=n}^{\infty} \frac{\Lambda^k}{k!} \binom{k}{n} (1-\beta)^{k-n} \right) (r_3 \beta)^n \geq 1 - \varepsilon,$$

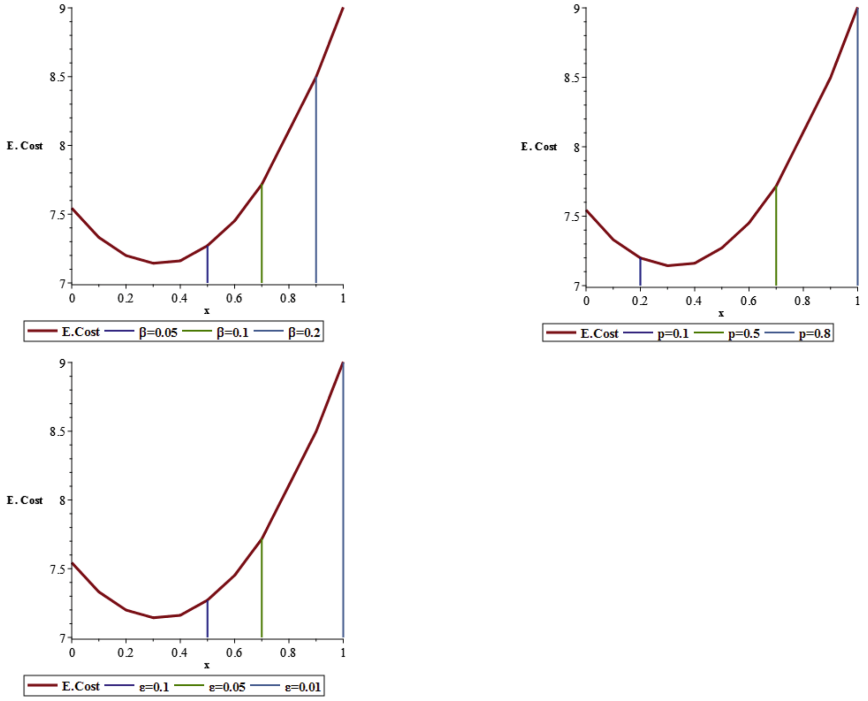
where $r_3 = \alpha_2(0)$, then $x_0(\varepsilon) = 0$, namely, constraint (15) is necessarily passive.

The impact of the safety constraint is now numerically illustrated. The fixed data from the previous subsection remain valid. Assume that $r_1 = r_2 = 0.5$, and

$$\alpha_3(x) = r_3 + (1 - r_3)x^p \tag{21}$$

for some $0 < p \leq 1$ (namely, α_3 is concavely increasing in x). Small value of p means that once x becomes little positive, then α_3 significantly grows. In other words, minimal transportation service ensures that almost all at-risk customers actually arrive to the service center. As a result, the safety constraint does not meaningfully affect the optimal solution. On the other hand, under high value of p, x must be significantly positive in order to increase α_3, that is, the safety constraint is supposed to significantly increase the optimal TSL and the expected cost. Let $r_3 = 0.7$, $\beta = 0.1$, $p = 0.5$ and $\varepsilon = 0.05$. Figure 7 illustrates how the optimal TSL as well as the expected cost is affected by the safety constraint. While this constraint is not taken into account, the optimal TSL is $x^* = 0.3$ and the minimal expected cost is 7143. The vertical lines denote the minimal TSL (i.e. x_0) corresponding to the different values of β, p and ε.

According to Fig. 7a, the safety constraint is active even if at-risk customers constitute only 5% of the type-2 customers. In this case (i.e. $\beta = 0.05$), the optimal value x^{**} becomes 0.5, and the expected cost is 7161. This growth is negligible (0.25%), but when β grows to 0.1 and 0.2, the optimal TSL increases to 0.7 and 0.9 respectively. Consequently, the expected cost significantly grows to 7716 (8%) and 8497 (19%) respectively. Next, Fig. 7b refers to the influence of the parameter p (see (21)) on the optimal TSL. While p is small (0.1), the optimal value remains 0.4 (namely, the safety constraint is passive). If $p = 0.5$, then $x^{**} = 0.7$ as before. When p increases to 0.8, then x^{**} becomes 0.8 and the expected cost grows to 8106 (13.5%). Finally, the influence of the safety rigorousness level ε is illustrated in Fig. 7c. Even if this value is not too small (0.1), constraint (15) affects the results (insignificantly) such that $x^{**} = 0.5$. When $\varepsilon = 0.05$, then $x^{**} = 0.7$. If there is a very rigorous approach regarding safety and $\varepsilon = 0.01$, then



a. Percentage of at-risk customers (β) b. Concavity of α_3 in x (p) c. Safety rigorousness level (ε)

Fig. 7. Optimal TSLs and corresponding expected cost ($\times 10^3$) under the safety constraint a function of β (a), p (b) and ε (c)

the TSL must be maximal and equal to 1. Consequently, the expected cost grows in 26% and becomes 9004.

5 Conclusions

This paper deals with combining of transportation service, as part of customer service, in a municipal service center. The transportation service level (TSL) is set in order to influence stochastic demand of customers, such that, analogically to supply chain models, expected surplus or shortage cost, caused by provider's idle time or overtime respectively, is minimized. There are customers whose visit has been booked in advance (type-1) as well as walk-in customers, who arrive to the service center without a reservation (type-2). An analytical model is numerically solved and illustrated. We show how the optimal number of type-1 customers to book in advance depends, analogically to inventory management models, on the ratio between the surplus and shortage unit costs. As the total demand's dependence on the TSL becomes higher, the optimal number of reservation is reduced (sometimes in more than 80%). As occurs in supply chain models, the surplus and shortage costs are significantly affected by the demand's uncertainty.

The variance of demand of type-1 customers is reduced as the TSL increases, while the variance of the type-2 customers' demand grows. As a result, as the arrival rate of type-2 customers grows, namely, these customers become more dominant in the service center, the optimal TLS becomes lower, such that the total demand's uncertainty as well as the transportation cost is reduced. The lower impact of the TSL on the demand, the lower optimal TSL and expected cost, especially if the surplus cost is higher than the shortage cost. In order to examine how rigorous policy in terms of safety affects the result, we add a safety constraint, according to which, the TSL level must be high enough such that the probability for non-arrival of at-risk customers is small. We numerically find that high percentage of at-risk customers in the population or, alternatively, high rigorousness policy may significantly increase the TSL, and consequently, lead to meaningful jump of the expected cost (up to 26%).

In this paper, the customers' waiting time is not taken into account, and rejecting of walk-in customers' appointments is not considered. In addition, the scheduling and TSL level are set for a single day, although it may be interesting to investigate how to plan the transportation services policy in advance (total number and sizes of vehicles, service areas etc.) in order to minimize total costs over longer periods. We leave these issues for future research.

References

1. Aboolian, R., Berman, O., Krass, D.: Profit maximizing distributed service system design with congestion and elastic demand. Transp. Sci. **46**(2), 247–261 (2012)
2. Cordeau, J.F., Laporte, G., Potvin, J.Y., Savelsbergh, M.W.P.: Transportation on demand. In: Barnhart, C., Laporte, G. (eds.) Handbook in OR & MS, vol. 14. Elsevier B.V. (2007)
3. de Ona, J., de Ona, R., Eboli, L., Mazzulla, G.: Index numbers for monitoring transit service quality. Transp. Res. Part A **84**, 18–30 (2016)
4. Erdem, E., Qu, X., Shi, J.: Rescheduling of elective patients upon the arrival of emergency patients. Decis. Support Syst. **54**, 551–563 (2012)
5. Grant, R., Goldsmith, G., Gracy, D., Johnson, D.: Better transportation to health care will improve child health and lower costs. Adv. Pediatr. **63**, 389–401 (2016)
6. Habib, K.M.N., Sasic, A., Weis, C., Axhausen, K.: Investigating the nonlinear relationship between travel activity scheduling and transportation system performances. Transp. Res. Part A **49**, 342–3357 (2013)
7. Habib, K.N.: An investigation on mode choice and travel distance demand of older people in the National Capital Region (NCG) of Canada: application of a utility theoretic joint econometric model. Transportation **42**, 143–161 (2015)
8. Hadas, Y., Shnaiderman, M.: Public-transit frequency setting using minimum-cost approach with stochastic demand and travel time. Transp. Res. Part B **46**, 1068–1084 (2012)
9. Kogan, K., Shnaiderman, M.: Continuous-time replenishment under intermittent observability. IEEE Trans. Autom. Control **55**(6), 1460–1465 (2010)
10. Peng, Y., Qu, X., Shi, J.: A hybrid simulation and genetic algorithm approach to determine the optimal scheduling templates for open access clinics admitting walk-in patients. Comput. Ind. Eng. **72**, 282–296 (2014)
11. Qu, X., Peng, Y., Shi, J., LaGands, L.: An MDP model for walk-in patient admission management in primary care clinics. Int. J. Prod. Econ. **168**, 303–320 (2015)

12. Qu, X., Rardin, R.L., Williams, J.A.S.: A mean-variance model to optimal the fixed versus open appointment percentages in open access scheduling systems. Decis. Support Syst. **53**, 554–564 (2012)
13. Qu, X., Rardin, R.L., Williams, J.A.S., Willis, D.R.: Matching daily healthcare provider capacity to demand in advanced access scheduling systems. Eur. J. Oper. Res. **183**, 812–826 (2007)
14. Quadrifoglio, L., Dessouky, M.M., Ordonez, F.: A simulation study of demand responsive transit system design. Transp. Res. Part A **42**, 718–737 (2008)
15. Raghunathan, S.: Information sharing in a supply chain: A note on its value when demand nonstationary. Manage. Sci. **47**(4), 605–610 (2001)
16. Ronald, N., Thompson, R., Winter, S.: Simulating demand-responsive transportation: a review of agent-based approaches. Transp. Rev. **35**(4), 404–421 (2015)
17. Shi, J., Xiau, T.: Service investment and consumer returns policy in a vendor-managed inventory supply chain. J. Ind. Manage. Optim. **11**(2), 439–459 (2015)
18. Tsuji, Y., Hirao, T., Fujikawa, A., Hoshikawa, Y., Yoshikoa, A., Yoda, T., Suzue, T.: Disease-wide accessibility of the elderly in primary care setting. The relationship between geographic accessibility and utilization of outpatient services in Tokushima prefecture, Japan. Health **4**(6), 320–326 (2012)
19. Whetten, R., Whetten, K., Pence, B.W., Reif, S., Conover, C., Bouis, S.: Does distance affect utilization of substance abuse and mental health services in the presence of transportation services? AIDS Care **18**(1), 27–34 (2006)
20. Zipkin, P.: Foundations of Inventory Management. McGrow-Hill Companies, Incorporated, Boston (2000)
21. Zulian, G., Donisi, V., Secco, G., Pertile, R., Tansella, M., Ammadeo, F.: How are caseload and service utilization of psychiatric services influenced by distance? A geographical approach to the study of community-based mental health services. Soc. Psychiatry (Psychiatr Epidemiol) **46**, 881–891 (2011)

A Stochastic Version of the Strategy-Based Congested Transit Assignment Model and a Technique by Smoothing Approximations

Esteve Codina(✉) and Francisca Rosell

Statistics and Operations Research Department,
Universitat Politècnica de Catalunya, Campus Nord,
Building C5, Office 216, 08034 Barcelona, Spain
esteve.codina@upc.edu

Abstract. This paper develops a stochastic version for the strategy-based congested transit assignment problem stated by Cominetti and Correa (Trans. Sci. 35(3):250–267, 2001). As a distinctive approach, this stochastic version takes into account stochastic mean waiting times of passengers at stops and in-vehicle travel times. The model is formulated as a stochastic variational inequality derived from the formulation of the deterministic version of the problem, also stated as a variational inequality problem, for which only a single solution method is known uptodate. Closely related with the stochastic model, and as a special case of it, a consistent smoothing approximation to the deterministic model is developed and it is shown that this approximation provides an alternative way of solving the deterministic model. It is also shown that both, the stochastic model and the smoothed approximation, can be solved by means of an adaptation of a path based method for the asymmetric traffic assignment problem. Computational tests have been carried out on several medium-large scale networks showing the viability of the method and its applicability to large scale transit models.

Keywords: Congested transit assignment · Stochastic variational inequalities · Strategy-based transit equilibrium · Smoothing approximation

1 Introduction

The concentration in large urban areas of a good fraction of the global population has made that their public transportation systems experience increasing levels of demand. Because of that, the development and use of congested transit assignment models has received increasing attention by researchers and practitioners of the transportation planning community. Additionally to congestion modeling, transit assignment models must take into account specific and complex aspects of these urban transportation systems, such as the common lines

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_13

problem and the route choice made by passengers. The early models took into consideration these two aspects but the effects of congestion were not taken into account. The initial paper by Chriqui and Robillard [6] introduced the notion of attractive lines. Passengers were assumed to board the first vehicle arriving at a stop of any line within the attractive set. As extensions of this seminal paper, there followed the work of Spiess [25] introducing the notion of strategy on a general multi-destination transit network and the work of Nguyen and Pallotino [22], formalizing the concept of hyperpath. The works by Cominetti and Correa [9] and Cepeda et al. [5], extended the previous concepts to congested transit networks, where due to capacity limitations in the lines, passengers had possibly to board not on the first arriving unit, but on subsequent ones. This approach, referred to in this paper as the C3F model (C3F ≡ Cepeda, Correa Cominetti, Florian), was formulated in [7] as a variational inequality problem. These aspects have been taken into account also in other contributions but, as this paper elaborates mainly on the previous references, the interested reader is directed to the monograph of Gentile and Noekel [15] or the reviews of Fu et al. [13] and Liu et al. [20]. Congestion effects on transit assignment models has led to consider queueing aspects at stops and different headway distributions. In [14] an extensive description is given of the work in this direction. Because these models are intended for planning applications, several sources of uncertainty have been addressed (see, for instance [21]). From the different stochastic models that have been proposed for the transit assignment problem, special emphasis has been done on route choice aspects (see, for instance the review in [20]), as well as on travel time variability due to daily fluctuations [21]. Although planning applications usually reproduce scenarios under a finite time horizon (e.g., peak morning period in a regular working day), aspects related to queueing models are taken from the steady state results and not for a limited or transitory time span (several hours). In this paper, the contributions may be summarized as follows: a stochastic version of the C3F model, formulated as a stochastic variational inequality, is developed which takes into account randomness in travel times and mean waiting time at stops, *with the only assumption* that mean waiting times at stops present a homoscedastic randomness. The model is not based on any route perception costs as is the case of models based on random utility maximization, such as for instance the works in [19] or in reliability-based models (e.g. [16,17]). This approach differs also from the stochastic model of Cortés et al. [10], which also elaborates on the deterministic C3F model, in that they take into account stochastic passenger's boarding decisions. Previously to the development of the C3F stochastic model, an approximation to the C3F model in its variational inequality formulation is carried out using smoothing functions. This approximation is consistent with the C3F model and it is also a special instance of the stochastic model. Both, the approximate model and the stochastic C3F model can be solved using a simplicial decomposition path-based algorithm, which for the deterministic case performs more efficiently than the MSA algorithm in [5].

The paper is structured as follows: Sect. 2 introduces initially the basic notation and gives a description of the deterministic version of the C3F model, as it

is stated in [7]. In Sect. 3 a smoothing approximation to the deterministic model based on logsum functions is described. In Sect. 4 a stochastic version of the C3F is developed and in Sect. 5 an algorithm is described for solving both the approximate and the stochastic formulations. Finally, in Sect. 6 the computational viability of the proposed models is shown on two medium-to-large scale transit networks.

2 Description and Summary of the Model's Formulation

2.1 Notation and Network Model

The notation and basic definitions are borrowed basically from [7] and, for completeness, will be repeated here almost literally. The transit network is considered in this paper in its expanded form, very similar to the one used in [25]. It will be represented by means of a directed graph $G = (N, A)$, where N is the set of nodes and A is the set of links. For a node $i \in N$, the sets of emerging and incoming links will be denoted by $E(i)$ and $I(i)$, respectively. A link will usually be represented by the index a, as a shorthand for an ordered pair of nodes (i, j), both in N.

The demand is assumed inelastic and the number of trips from i to d will be denoted by g_i^d. The set of active origin-destination (O-D) pairs $\omega = (i, d)$ on the network will be denoted by W, i.e., $W \triangleq \{(i, d) \in N \times N \mid g_i^d > 0\}$. The set of destinations in the network will be denoted by D. Also, $N_d = N \setminus \{d\}$ will denote the set of nodes excluding destination $d \in D$. The subset of nodes representing transit stops will be designated by $\hat{N}$. Also, $\hat{N}_d = \hat{N} \setminus \{d\}$.

Figure 1 schematically depicts the configuration of the expanded transit network. In this representation, transit stops are associated with a node and some of their outgoing links in the expanded network will be boarding links to a transit line. On the other hand, each transit line with vehicles halting at the stop will have a single boarding link a from the stop, a single alighting link and a single dwell link $x(a)$ to the stop. Non-boarding or non-alighting links at transit stops will model movements by walking or connections to other transit stops.

For describing passenger volumes on the transit expanded network, vector flows on links will be expressed in bold. The following notation will be used:

- v_a^d: passenger's flow at link $a \in A$ with destination $d \in D$.
- $v_i^d = \sum_{a \in E(i)} v_a^d$, i.e., the total inflow through node $i \in N$ with destination $d \in D$ and $\mathbf{v}_i^d = (..., v_a^d, ...; a \in E(i))$ is a vector of per-destination flows on boarding links at stop i.
- Also, $\mathbf{v}^d = (..., \mathbf{v}_i^d, ...;\ i \in N) \in \mathbb{R}_+^{|A|}$, $d \in D$ and the vector of link flows will be denoted by $\mathbf{v} = (..., \mathbf{v}^d, ...;\ d \in D) \in \mathbb{R}_+^{|A|\,|D|}$. As an exception, the vector of total flows on links will not be expressed in bold: $v = \sum_{d \in D} \mathbf{v}^d \in \mathbb{R}_+^{|A|}$. Its components will be $v_a = \sum_{d \in D} v_a^d,\ a \in A$.

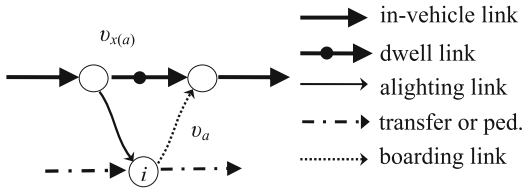


Fig. 1. Representation in the expanded network of a line serving a transit stop.

Flow balance constraints at nodes $i \in N$ will be verified by the per-destination flow vector $\mathbf{v}^d$, defining the set V^d of feasible flows for destination $d \in D$:

$$\mathrm{V}^d \triangleq \{\, \mathbf{v}^d \in \mathbb{R}_+^{|A|} \mid \sum\nolimits_{a \in E(i)} v_a^d - \sum\nolimits_{a \in I(i)} v_a^d = g_i^d,\ i \in N \},\ d \in D \tag{1}$$

The feasible set $\mathcal{V}$ of total link flows v will be then defined as: $\mathcal{V} \triangleq \{\, v \in \mathbb{R}_+^{|A|} \mid v_a = \sum_{d \in D} v_a^d,\ \mathbf{v}^d \in \mathrm{V}^d \}$ and the feasibility set for the congested transit equilibrium problem can be formulated as:

$$\mathrm{V} \triangleq \bigotimes\nolimits_{d \in D} \mathrm{V}^d \tag{2}$$

The set of boarding links a from stop i will be denoted by $\hat{E}(i)$. These links are associated with a finite effective frequency function $f_a(v)$ for that line (efff in the following), depending on the total flows v. Effective frequencies are the reciprocal of the mean waiting times at the stop until boarding onto a vehicle of the line, $\sigma_a(v) = 1/f_a(v)$. Travel times on links will be modeled by functions $t_a(v)$, $a \in A$, assumed finite on $\mathcal{V}$, i.e., $t_a(v) < +\infty, \forall\, v \in \mathcal{V}, a \in A$ and continuous. Any link different from a boarding link will be assigned an infinite frequency.

$$\hat{E}(i) \triangleq \begin{cases} \{\, a \in E(i) \mid f_a(\cdot) < +\infty \,\}, & i \in \hat{N} \\ \emptyset, & i \in N \setminus \hat{N} \end{cases} \tag{3}$$

The following notation will be used when formulating variational inequalities and optimization problems in this paper. If $C \subseteq \mathbb{R}^n$ is a convex set, its normal cone at $x \in C$ shall be denoted by $\mathsf{N}_C(x)$ and for a function $\Phi : \mathbb{R}^n \to \mathbb{R}^n$, a V.I. will be stated in the Hartmann-Stampacchia form, i.e., find $x \in C$ so that $\Phi(x)^\top (x' - x) \geq 0,\ \forall\, x' \in C$ and also in its variational condition form as $0 \in \Phi(x) + \mathsf{N}_C(x)$. This VI will also be referred to as $\mathsf{VI}(\Phi, C)$. In expressions describing optimization problems, immediately after a constraint, dual variables or multipliers for that constraint may appear after a bar "|".

2.2 A Summary of the Model's Formulation as a Variational Inequality Problem

Let S_i^d be the simplex sets $S_i^d = \{\alpha \in \mathbb{R}_+^{|\hat{E}(i)|} \mid \sum_{a \in \hat{E}(i)} \alpha_a = 1\}$ associated with a stop node $i \in \hat{N}_d$ and destination $d \in D$ and let $S = \bigotimes_{d \in D} \bigotimes_{i \in \hat{N}_d} S_i^d$. Let m

be defined as $m = |D| \cdot |A| + \sum_{d \in D} \sum_{i \in \hat{N}_d} |\hat{E}(i)|$. Then, the congested transit assignment problem can be defined as in [7] by means of a variational inequality $\mathsf{VI}(F, \mathrm{V} \times S)$, where the functional $F(\cdot, \cdot) : \mathrm{V} \times S \to \mathbb{R}^m$ is defined as:

$$F(\mathbf{v}, \zeta) = \begin{pmatrix} F_v(\mathbf{v}, \zeta) \\ F_\zeta(\mathbf{v}) \end{pmatrix} = \begin{pmatrix} T_a^d(v, \zeta_a^d) \; ; \; a \in E(i), \; i \in N_d, \; d \in D \\ -x_a^d(\mathbf{v}) \; ; \; a \in \hat{E}(i), \; i \in \hat{N}_d, \; d \in D \end{pmatrix} \tag{4}$$

where, for convenience of exposition, $x_a^d(\mathbf{v}) = v_a^d / f_a(v)$. Although $x_a^d(\mathbf{v})$ is actually a shorthand for $v_a^d / f_a(v)$ or $v_a^d \sigma_a(v)$, it can be interpreted as the total waiting time that the v_a^d passengers with destination d would spend if they choose boarding *only* on link a. In a strategy-based model, at the solution, let $E^{d*}(i)$ the set of attractive boardings at stop i. There holds then that $v_a^d \sigma_a(v) = v_{a'}^d \sigma_{a'}(v)$ for boarding links $a, a' \in E^{d*}(i)$ and thus $x_a^d(\mathbf{v})$ can be interpreted as the total waiting time for all passengers with destination d at stop i.

Then, functions T_a^d in (4) are defined as:

$$T_a^d(v, \zeta_a^d) = \begin{cases} t_a(v) + \zeta_a^d / f_a(v), & a \in \hat{E}(i) \\ t_a(v) & a \in E(i) \setminus \hat{E}(i) \end{cases} \quad d \in D, \; i \in N_d \tag{5}$$

Using vector functions $\mathrm{T}^d(v, \zeta^d) = (..., T_a^d(v, \zeta_a^d), ...; \; a \in A)$ for each destination $d \in D$, and vector functions $x_i^d(\mathbf{v}) = (..., x_a^d(\mathbf{v})...; \; a \in \hat{E}(i))$, for each stop $i \in \hat{N}_d, \; d \in D$, the previous variational inequality $\mathsf{VI}(F, \mathrm{V} \times S)$ can be written in variational condition form as:

$$(\mathtt{VI}) \quad \left[\begin{array}{l} \mathsf{Find} \; (\mathbf{v}, \zeta) \in \mathrm{V} \times S \; \mathsf{such \;\; that}, \\ \quad 0 \in \mathrm{T}^d(v, \zeta^d) + \mathsf{N}_{\mathrm{V}^d}(\mathbf{v}^d), \quad d \in D \\ \quad 0 \in -x_i^d(\mathbf{v}) + \mathsf{N}_{S_i^d}(\zeta_i^d), \quad d \in D, \; i \in \hat{N}_d \end{array} \right] \tag{6}$$

Also, variational inequality (6) can be expressed in its Hartmann-Stampacchia's form:

$$(\mathtt{VI}) \quad \left[\begin{array}{l} \mathsf{Find} \; (\mathbf{v}, \zeta) \in \mathrm{V} \times S \; \mathsf{such \;\; that}, \forall \, (\mathbf{u}, \mathrm{z}) \in \mathrm{V} \times S: \\ \quad \displaystyle\sum_{d \in D} \sum_{i \in N_d} \left\{ \sum_{a \in E(i)} T_a^d(v, \zeta_a^d)(u_a^d - v_a^d) - \sum_{a \in \hat{E}(i)} x_a^d(\mathbf{v})(z_a^d - \zeta_a^d) \right\} \geq 0 \end{array} \right] \tag{7}$$

Previous variational inequality (VI) considers congested network models without sharp capacity bounds on the flows. These are models with effective frequencies that decrease with increasing flows but do not vanish and with travel times that may increase but remain finite at any point in the feasible set of flows V expressed by (2). For this case, in [7] (Theorem 5.1), the existence of acyclic solutions is guaranteed with the only condition on the travel costs t_a and effective mean boarding times $\sigma_a = 1/f_a$ to remain positive and continuous on the feasible set of flows V.

A proper gap function for variational inequality $\mathsf{VI}(F, \mathrm{V} \times S)$ must be the following gap function $G_{\mathtt{C3F}} : \mathrm{V} \to \mathbb{R}$, in (8) and defined in [5] in order to solve

heuristically the C3F model as they proved that zeros of this gap function are solutions of it. This gap function can be written as:

$$G_{\texttt{C3F}}(\mathbf{v}) \triangleq G^{(0}_{\texttt{C3F}}(\mathbf{v}) - G^{(1}_{\texttt{C3F}}(v) \tag{8}$$

where $G^{(0}_{\texttt{C3F}}(\mathbf{v}) = \varphi(\mathbf{v}, v)$, $G^{(1}_{\texttt{C3F}}(v) = \mathsf{Min}_{\,\mathbf{u}\in \mathrm{V}}\ \varphi(\mathbf{u}, v)$ and function $\varphi(\cdot,\cdot)$ is defined as:

$$\varphi(\mathbf{u}, v) \triangleq \sum_{d\in D}\left[\sum_{a\in A} t_a(v)u^d_a + \sum_{i\in \hat{N}_d} \mathsf{Max}_{a\in \hat{E}(i)}\left\{\frac{u^d_a}{f_a(v)}\right\}\right] \tag{9}$$

The statement of the C3F model by means of (VI) in (6) or (7) in [7] (see Proposition 4.10), allows to view the algorithmic procedure stated in [5] as the only known (up-to-date) heuristic method to solve the following fix-point inclusion problem by means of an MSA procedure:

$$\mathbf{v}^* \in \mathsf{Sol}\ \mathsf{Min}_{\,\mathbf{u}\in \mathrm{V}}\ \varphi(\mathbf{u}, v^*) \tag{10}$$

Also, as stated in [7], Theorem 4.9, the previous gap function $G_{\texttt{C3F}}(\cdot)$ can be derived from the Auslender or primal gap function [1], $G_P(\mathbf{v}, \zeta)$, for variational inequality $\mathsf{VI}(F, \mathrm{V}\times S)$.

3 An Approximated C3F Model Obtained by Smoothing

In this section a VI is derived whose solutions approximate those of VI in (7), which defines the congested strategy-based transit assignment model. We will refer to that VI as the "approximate VI". The theoretical background that substantiates this derivation is left to an appendix in which a broad class of variational inequalities is considered.

For any destination $d \in D$ and stop $i \in \hat{N}_d$, with $|\hat{E}(i)|$ lines, the function $H_i : \mathbb{R}^{|\hat{E}(i)|} \to \mathbb{R}$ is defined as $H_i(\lambda) = \mathsf{max}_{a\in\hat{E}(i)}\{..., \lambda_a, ...\}$. Then the function w^d_i defined as:

$$w^d_i(\mathbf{v}) \triangleq \mathsf{Max}_{a\in\hat{E}(i)}\{..., x^d_a(\mathbf{v}), ...\} = \mathsf{Max}_{\substack{\zeta^d\geq 0 \\ s.t.}} \sum_{a\in\hat{E}(i)} x^d_a(\mathbf{v})\,\zeta^d_a \quad \sum_{a\in\hat{E}(i)} \zeta^d_a = 1 \tag{11}$$

can be approximated by the following function $\psi^d_i(\cdot,\cdot)$ for large values of the parameter θ_i:

$$w^d_i(\mathbf{v}) \approx \psi^d_i(x^d_i(\mathbf{v}), \theta_i) \triangleq \frac{1}{\theta_i}\ \log\left(\sum\nolimits_{a\in\hat{E}(i)} \exp\left(\theta_i\, x^d_a(\mathbf{v})\right)\right) \tag{12}$$

Now consider the set of solutions of linear program in (11) for some vector flow $\mathbf{v}$. Then, the gradient $\nabla_x \psi^d_i(x^d_i(\mathbf{v}), \theta_i)$ must be very close to an element

ζ_i^{d*} within the solution set of linear program (11) (see Appendix A, for a concise statement of this fact)

$$\zeta_a^d \approx (\,\nabla_x \psi_i^d(x_i^d(\mathbf{v}),\theta_i)\,)_a = \frac{\exp(\theta_i\, x_a^d(\mathbf{v}))}{\sum_{a'\in\hat{E}(i)}\exp(\theta_i\, x_{a'}^d(\mathbf{v}))},\ a\in\hat{E}(i),\ i\in\hat{N} \tag{13}$$

and the link cost functions $T_a^d(\cdot,\cdot)$ in (5), which define the variational inequality $\mathsf{VI}(F,\mathrm{V}\times S)$ can be approximated by the link cost functions $\breve{T}_a^d(\cdot,\theta_i)$ defined as:

$$\breve{T}_a^d(\mathbf{v},\theta_i)=\begin{cases} t_a(v)+\dfrac{\sigma_a(v)\exp(\theta_i\, x_a^d(\mathbf{v}))}{\sum_{a'\in\hat{E}(i)}\exp(\theta_i\, x_{a'}^d(\mathbf{v}))}, & a\in\hat{E}(i) \\ t_a(v), & a\in E(i)\setminus\hat{E}(i) \end{cases} \quad d\in D,\ i\in N_d \tag{14}$$

Then, as a direct consequence of Lemma A.1 in the Appendix, the variational inequality $\mathsf{VI}(F,\mathrm{V}\times S)$ in (7), can be approximated consistently by the following one (the approximate variational inequality)

$$(\texttt{VIapprox})\quad \left[\,\mathsf{Find}:\ \mathbf{v}\in\mathrm{V},\ \mathsf{such\ that}\ 0\in\breve{T}^d(\mathbf{v},\theta)+\mathsf{N}_{\mathrm{V}^d}(\mathbf{v}^d),\ d\in D\,\right] \tag{15}$$

in the sense that if $\theta_i\to\infty$, $i\in\hat{N}$, then solutions of (15) are also solutions of the V.I. in (6). Finally, the Auslender gap function $\breve{G}_P(\mathbf{v},\theta)$ for previous variational inequality `(VIapprox)` will then be given by:

$$\breve{G}_P(\mathbf{v},\theta)=\max_{\mathbf{u}\in\mathrm{V}}\sum_{d\in D}\sum_{a\in A}\breve{T}_a^d(\mathbf{v},\theta)(v_a^d-u_a^d) \tag{16}$$

4 A Stochastic Version for the C3F Model

In this section a stochastic version of the C3F model is considered in which mean waiting times at stops σ are considered as random variables (r.v.). Although daily fluctuations of many factors may contribute to this randomness, the operation of transit systems during a finite period is also a source of randomness. Usually congested transit assignment models under a deterministic approach use mean waiting times, either coming from delay results in queueing theory or some empirical formulas (see, for instance [11]). In a strict sense, these mean waiting times are valid on the long run. However, in practice, the number of daily arrivals of a given line at a stop is not greater than a few hundred times. In these conditions the mean waiting time may present strong deviations from the asymptotic values given by traditional queueing models and can be considered a random variable. Additionally the variability of the line headway τ may also have a relevant impact. It is well known that if $d_0=\mathbb{E}[\tau]$ is the headway's average value and C_τ is its coefficient of deviation, then the mean waiting time of a single arriving passenger is given by $\mathbb{E}[\tau](1+C_\tau^2)/2$. Then, if the average waiting time for a given non-null flow of arriving passengers is d, the ratio d/d_0 can be considered as a normalized waiting time. As an example, Fig. 2 depicts the results obtained by simulation for a bulk service queue on a limited period of

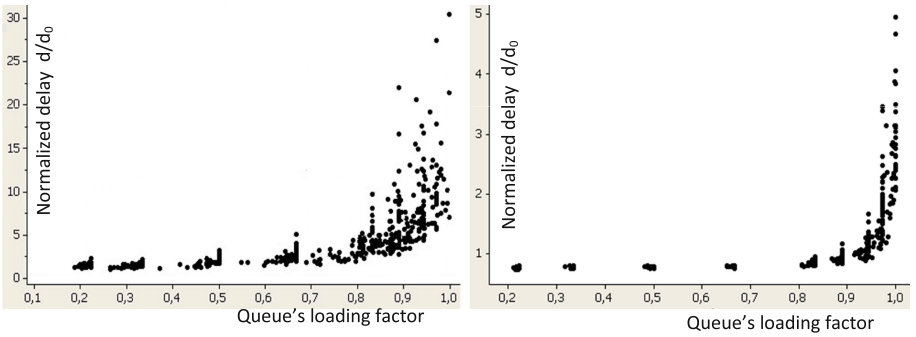


Fig. 2. Mean waiting times of clients in a finite time period with 120 server arrivals for a $M/M^{[C]}/1$ queue (left) and in a $M/D^{[C]}/1$ queue (right). Server's capacity C is constant (30 clients). The queue's loading factor is on x-axis and the normalized delay of clients d/d_0 is on y-axis.

time, showing the normalized waiting time d/d_0 versus the queue's loading factor. Each dot corresponds to a simulation with 120 server arrivals. The arrival of clients is Poissonian, whereas inter-arrival times of servers follow an exponencial distribution on the left and a degenerate distribution on the right.

One way in which stochastic variational inequalities can be considered is by means of the expected residual minimization (ERM) (see, for instance, [4]). Basically, if for scenario ω, the V.I. under consideration is $0 \in F(x,\omega) + \mathsf{N}_X(x)$, for which a gap function is $g(x,\omega)$, one seeks to solve the stochastic problem $Min_{x\in X}\mathbb{E}[g(x,\omega)]$. In our case, the gap function $G_{\texttt{C3F}}(\mathbf{v},\omega)$ will be taken into account and, considered as a r.v., will be denoted by $\tilde{G}_{\texttt{C3F}}(\mathbf{v})$. (any r.v. will wear a tilde on top in the rest of this section).

In this paper it will be assumed that in-vehicle travel times t_a are stochastic and follow a model of the form: $\tilde{t}_a(v) = \bar{t}_a(v) + \tilde{\epsilon}_a(v)$, where $\bar{t}_a(v)$ is the mean value of the r.v. $\tilde{t}_a(v)$ and the r.v. $\tilde{\epsilon}_a(v)$ is assumed to have null mean and distribution parameters which may depend on the total flows v. In these conditions $\mathbb{E}[v_a\tilde{t}_a(v)] = v_a\bar{t}_a(v)$. In the case of the effective waiting times at stops the following model will be assumed:

$$\tilde{\sigma}_a(v) = \bar{\sigma}_a(v) + \tilde{\epsilon}_a \tag{17}$$

This is a homoscedastic model where $\bar{\sigma}_a(v)$ is the mean of the r.v. $\tilde{\sigma}_a(v)$ and $\tilde{\epsilon}_a$ is a r.v. with $\mathbb{E}[\tilde{\epsilon}_a] = 0$ and with distribution parameters not dependent on the total flows v.

At this point recall the following property. Let $\tilde{\mathbf{p}}_i = (...,\tilde{p}_a,...; a \in \hat{E}(i))$ be a vector of random variables given by $\tilde{p}_a = p_a + \tilde{\epsilon}_a$ with p_a a constant and $\tilde{\epsilon}_a$ a r.v., such that $\mathbb{E}[\tilde{\epsilon}_a] = 0$ and such that $\mathbb{C}\text{ov}[\tilde{\epsilon}_a, \tilde{\epsilon}_{a'}] = m_{a,a'}$ does not depend on $\mathbf{p}_i = (...,p_a,...; a \in \hat{E}(i))$. Let then $\mathcal{H}_i(\cdot)$ be defined as:

$$\mathcal{H}_i(\mathbf{p}_i) = \mathbb{E}[H_i(\tilde{\mathbf{p}}_i)] = \mathbb{E}[\max_{a\in\hat{E}(i)}\{...,\tilde{p}_a,...\}] \tag{18}$$

It is known that (see, for instance [3], Chap. 3), in the previous conditions, $P_r(\tilde{p}_a \geq \tilde{p}_{a'},\, \forall a' \in \hat{E}(i)) = \partial \mathcal{H}_i / \partial p_a(\mathbf{p}_i) = (\nabla \mathcal{H}_i(\mathbf{p}_i))_a$.

Consider now the following VI:

$$(\texttt{stVI}) \quad \left[\, \mathsf{Find:}\ \mathbf{v} \in \mathrm{V},\ \ \mathsf{such\ \ that}\ 0 \in \bar{T}^d(\mathbf{v}) + \mathsf{N}_{\mathrm{V}^d}(\mathbf{v}^d),\ d \in D \,\right] \tag{19}$$

where functional $\bar{T}^d$ is given by:

$$\bar{T}_a^d(\mathbf{v}) = \begin{cases} \bar{t}_a(v) + \bar{\sigma}_a(v)\left(\Phi_i^d(\mathbf{v})\right)_a, & a \in \hat{E}(i) \\ \bar{t}_a(v), & a \in E(i) \setminus \hat{E}(i) \end{cases} d \in D,\ i \in N_d \tag{20}$$

$$\Phi_i^d(\mathbf{v}) = \nabla \mathcal{H}_i(..., v_a^d \bar{\sigma}_a(v), ...), \quad i \in \hat{N}_d \tag{21}$$

The following theorem shows that the Auslender gap (or primal gap) for V.I. (stVI) in (19) is an upper bound of the ERM gap function $\gamma(\mathbf{v}) = \mathbb{E}[\tilde{G}_{\texttt{C3F}}(\mathbf{v})]$. This implies that: (a) solving V.I. (stVI) is equivalent to solve the ERM problem, or equivalently, a stochastic version of (VI) defined in (6) with stochastic in-vehicle travel times and mean effective waiting times modeled as in (17), (b) if $\mathbf{v}^*$ is a solution of (stVI), then $\gamma(\mathbf{v}^*) = 0$. Because of that it can be claimed that V.I. (stVI) is the stochastic version of the deterministic congested strategy-based transit assignment model (VI) stated in (6) or (7).

Theorem 1. *If $\bar{G}_P(\boldsymbol{v})$ is the Auslender's gap for (stVI) in (19) and $\gamma(\mathbf{v}) \triangleq \mathbb{E}[\tilde{G}_{\texttt{C3F}}(\mathbf{v})]$ is an ERM gap for the stochastic version of the C3F model formulated as a VI in (6), then $\gamma(\mathbf{v}) \leq \bar{G}_P(\mathbf{v})$.*

Proof. If $\mathbf{u}, \mathbf{v}$ are vectors of link flows, it will be convenient to consider the vector of r.v. $\tilde{\mathrm{x}}_i^d(\mathbf{u}_i^d, v) = (..., u_a^d \cdot \tilde{\sigma}_a(v), ...; a \in \hat{E}(i))$, its mean $\bar{\mathrm{x}}_i^d(\mathbf{u}_i^d, v) = \mathbb{E}[\tilde{\mathrm{x}}_i^d(\mathbf{u}_i^d, v)]$ and also, $\bar{x}_i^d(\mathbf{v}) = \bar{\mathrm{x}}_i^d(\mathbf{v}_i^d, v)$.

$$\begin{aligned}
&\mathbb{E}[\tilde{G}_{\texttt{C3F}}(\mathbf{v})] \overset{(1)}{\leq} \sum_{d \in D} \left[\sum_{a \in A} \bar{t}_a(v) v_a^d + \sum_{i \in \hat{N}_d} \mathbb{E}\big[H_i(\tilde{\mathrm{x}}_i^d(\mathbf{v}_i^d, v))\big] \right] - \\
&- \min_{\mathbf{u} \in \mathrm{V}} \sum_{d \in D} \left[\sum_{a \in A} \bar{t}_a(v) u_a^d + \sum_{i \in \hat{N}_d} \mathbb{E}\big[H_i(\tilde{\mathrm{x}}_i^d(\mathbf{u}_i^d, v))\big] \right] = \\
&= \max_{\mathbf{u} \in \mathrm{V}} (v - u)^\top \bar{t}(v) + \sum_{d \in D} \sum_{i \in \hat{N}_d} \mathbb{E}\big[H_i(\tilde{\mathrm{x}}_i^d(\mathbf{u}_i^d, v)) - H_i(\tilde{\mathrm{x}}_i^d(\mathbf{v}_i^d, v))\big] \overset{(2)}{\leq} \\
&\leq \max_{\mathbf{u} \in \mathrm{V}} (v - u)^\top \bar{t}(v) + \sum_{d \in D} \sum_{i \in \hat{N}_d} \sum_{a \in \hat{E}(i)} (v_a^d - u_a^d) \mathbb{E}[\tilde{\eta}_a^x(v)\, \tilde{\sigma}_a(v)] \overset{(3)}{=} \\
&= \max_{\mathbf{u} \in \mathrm{V}} (v - u)^\top \bar{t}(v) + \sum_{d \in D} \sum_{i \in \hat{N}_d} (\bar{x}_i^d(\mathbf{v}_i^d) - \bar{\mathrm{x}}_i^d(\mathbf{u}_i^d, v))^\top \Phi_i^d(\mathbf{v}) = \bar{G}_P(\mathbf{v})
\end{aligned} \tag{22}$$

where $\tilde{\eta}_a^x(v) \triangleq (\partial H_i(\tilde{\mathrm{x}}_i^d(\mathbf{v}_i^d, v)))_a = \mathcal{I}\,[v_a^d \tilde{\sigma}_a(v) \geq v_{a'}^d \tilde{\sigma}_{a'}(v); \forall\, a' \in \hat{E}(i)]$ and $\mathcal{I}\,[\cdot]$ is the indicator function of an event. Inequality (1) follows from the fact that $\mathbb{E}[Min_{x \in X} f(x, \omega)] \leq Min_{x \in X} \mathbb{E}[f(x, \omega)]$ (See, for instance [2], Chap. 4).

Inequality (2) is because of the convexity of $H(\cdot)$ and that $\tilde{\eta}_a^x(v) = \mathcal{I}\,[v_a^d \tilde{\sigma}_a(v) \geq v_{a'}^d \tilde{\sigma}_{a'}(v); \forall\, a' \in \hat{E}(i)]$. Equality (3) is implied because $\mathbb{E}[\tilde{\eta}_a^x(v)] = P_r(v_a^d\, \tilde{\sigma}_a(v) \geq v_{a'}^d\, \tilde{\sigma}_{a'}(v); \forall\, a' \in \hat{E}(i))$. Additionally, because of mean effective waiting times are modeled as in (17),

$$\mathbb{E}[\eta_a^x(v)\, \tilde{\sigma}_a(v)] = \bar{\sigma}_a(v)\, P_r(v_a^d\, \tilde{\sigma}_a(v) \geq v_{a'}^d\, \tilde{\sigma}_{a'}(v),\ \forall\, a' \in \hat{E}(i)) = \bar{\sigma}_a(v) \frac{\partial \mathcal{H}_i}{\partial p_a}(\bar{x}_i^d(\mathbf{v})) \quad \square$$

It must be noticed that, if instead of the model in (17), it is assumed a (less realistic) model of the type $\tilde{\mathrm{x}}_i^d(\mathbf{v}_i^d, v) = (..., v_a^d \bar{\sigma}_a(v) + \tilde{\epsilon}_a, ...; a \in \hat{E}(i))$, with $\tilde{\epsilon}_a$ i.i.d. as Gumbel r.v.'s with zero mean and variance $\pi^2/(6\theta_i)$, $\forall\, a \in \hat{E}(i)$, then $\Phi_i^d(\cdot) \equiv \nabla_x \psi_i^d(x_i^d(\cdot), \theta_i)$ as defined in (13) and then the approximate V.I. defined in (15) is a special case of the stochastic V.I. defined in (19).

In practice, it is only possible to obtain values for the function $\Phi_i^d(\cdot)$ defined in (21) by using simulation. Because of these difficulties, an approximate and practical way to proceed would be based on the choice of proper parameters θ_i^d, so that the V.I. in (15) could reasonably adjust to the stochastic model given by V.I. in (19). Also, a convenient way to express model (17) for mean waiting times would be, for instance, $\tilde{\sigma}_a(v) = \sigma_{a,0}(\xi_a(\rho_a(v)) + k_1 \tilde{\epsilon}_a)$, where ρ_a is the loading factor for the boarding queue represented by link $a \in \hat{E}(i)$. These parameters, as it will be shown, would only depend on the stop i and not on the destination d. To this end, consider that,

$$\begin{aligned} \mathbb{E}[H_i(..., v_a^d \tilde{\sigma}_a(v), ...)] &= \hat{\sigma}_i v_i^d\, \mathbb{E}[\max_{a \in \hat{E}(i)} \{\frac{v_a^d}{v_i^d} \frac{\sigma_{a,0}}{\hat{\sigma}_i}(\xi_a(v) + k_1 \tilde{\epsilon}_a)\}] \\ &= \hat{\sigma}_i v_i^d\, \mathbb{E}[\max_{a \in \hat{E}(i)} \{r_a(1 + k_1 \tilde{\epsilon}'_a)\}] = \hat{\sigma}_i v_i^d\, U(\mathbf{r}_i; k_1) \end{aligned} \tag{23}$$

where $\hat{\sigma}_i = \max_{a \in \hat{E}(i)}\{\sigma_{a,0}\}$ and $\mathbf{r}_i = (..., r_a, ...)$. The arguments r_a for the function U are taken as $r_a = (v_a^d / v_i^d)(\sigma_{a,0}/\hat{\sigma}_i)\xi_a(\rho_a(v))$. Notice that, the ratio v_a^d/v_i^d is equivalent to the ratio f_a/f_i which does not depend on the destination and that the maximum value that r_a may achieve is mainly due to the level of congestion of the boarding queue given by the value of $\xi_a(\cdot)$. The following Fig. 3 shows the values of the optimal parameter θ^* for which function U is best fitted by the logsum function $\psi(\mathbf{r}_i; \theta) = \theta^{-1} \log \sum_{a \in \hat{E}(i)} \exp(\theta r_a)$ on the interval $[0, \hat{r}]^{|\hat{E}(i)|}$, for $|\hat{E}(i)| = 2, 3, 4, 5$, $k_1 \in [0, 1]$ and $\hat{r} = 10$. The probability distributions for the r.v. $\tilde{\epsilon}'_a$ are (a) $\tilde{\epsilon}'_a \sim$ unif[−1,1] and (b) a shifted k-Erlang distribution so that $\tilde{\epsilon}'_a + 1 \sim k$-Erlang and $\mathbb{E}[\tilde{\epsilon}'_a + 1] = 1$. For other values of $\hat{r}$, it has been observed that the relationship $\theta^* \hat{r} = \theta^{*'} \hat{r}'$ holds very accurately.

5 Solving the Approximate Variational Inequality

Both, the approximate variational inequality in (15) and the stochastic model in (19) have a simpler structure than (VI) in (6) or (7), as variables ζ have been eliminated by the approximation in (13). Now, link costs $\breve{T}^d$ or $\bar{T}^d$, are destination-dependent and are functions only on the vector of per-destination

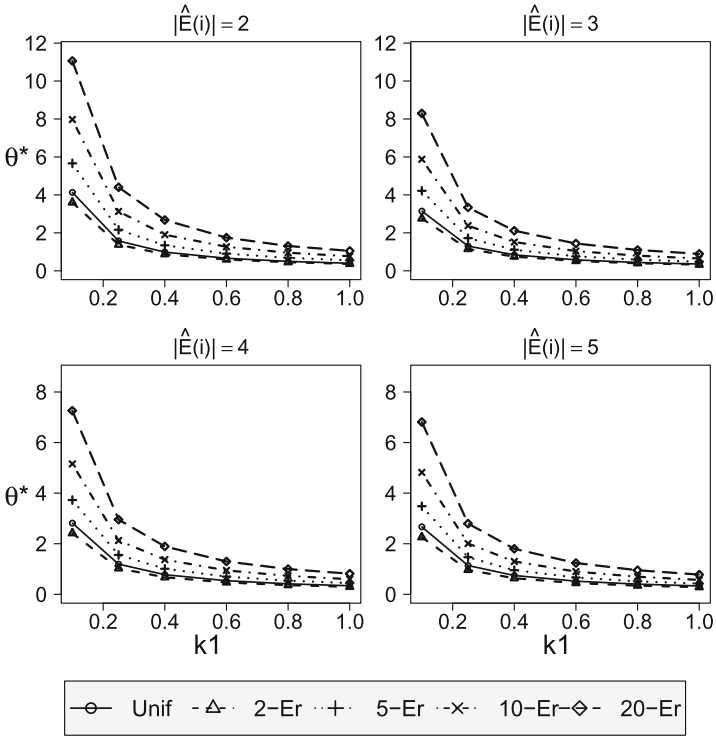


Fig. 3. Optimal values of θ depending on the dispersion parameter k_1, so that the log-sum function $\psi(\mathbf{r}_i;\theta) = \theta^{-1}\log\sum_{a\in\hat{E}(i)}\exp(\theta r_a)$ fits the U function in (23), assuming that $\tilde{\epsilon}'_a$ is distributed following a uniform distribution on [−1,1] and a shifted k-Erlang with null mean and $k = 2, 5, 10, 20$ stages.

flows $\mathbf{v}$. It will be possible to solve this approximate VI using an asymmetric traffic assignment method adapted to take into account this special structure of path costs. Because of that, the Simplicial Decomposition (SD) path-based algorithm for the fixed demand asymmetric traffic assignment problem developed in [8] has been adapted to solve the approximate model (15).

The notation used in this section is the following:

- $\Gamma^w = \{\, r \mid \textit{path joining the O-D pair } w \in W\,\}; \Gamma = \otimes_{w\in W}\Gamma^w$
- h_r^w: flow through path $r \in \Gamma^w$, $w \in W$. Also, a vector of flows on paths will be denoted by $\mathbf{h}^w$ for O-D pair $w \in W$. The polytope of these path flows will be denoted by $J^w = \left\{\mathbf{h}^w \in \mathbb{R}_+^{(n^w)} \,\middle|\, \sum_{r\in\Gamma^w} h_r^w = g_w\right\}$, and $J = \otimes_{w\in W} J^w$.
- $\mathcal{T}_r(\mathbf{h},\theta)$: cost on path r; the vector of path costs is denoted by $\mathcal{T}(\cdot,\theta) : J \to \mathbb{R}^{n_W}$, where $n_W = \Sigma_{w\in W} n^w$. The cost $\mathcal{T}_r(\mathbf{h},\theta)$ on path $r \in \Gamma^w$ will be obtained by summation of the link costs $\breve{T}_a^d$ in (14) for links on path r, for the corresponding destination d of O-D pair w. The vector of path costs for $w \in W$ will be designated by $\mathcal{T}^w(\mathbf{h},\theta)$.

On the path-flow space the approximate variational inequality in (15) will be expressed as:

$$0 \in \mathcal{T}^w(\mathbf{h}, \theta) + \mathsf{N}_{J^w}(\mathbf{h}^w),\ w \in W. \tag{24}$$

The SD algorithm specified below uses at each iteration ℓ, a subset of acyclic paths $\Gamma^{w,(\ell}$ for each O-D pair $w \in W$ and, in the master problem step, solves VI (24) restricted to the polytope $J^{(\ell} = \otimes_{w \in W} J^{w,(\ell}$ defined by these subsets of paths. The SD path-based algorithm for solving the approximate VI in (24) can be stated as follows:

Initialization: Let $\Gamma^{w,(0}$ be an initial set of acyclic paths for each O-D pair w and use them to obtain a feasible path flow vector $\mathbf{h}^{(0} \in J^{(0}$; set counter $\ell = 1$.

At iteration $\ell - th$

1. For each O-D pair $w \in W$ drop from $\Gamma^{w,(\ell-1}$ those paths with null flow in the path-flow vector $\mathbf{h}^{w,(\ell-1}$.
2. *Subproblem Step.* Increase the working set of paths by finding the shortest paths $\breve{r} \in \Gamma^w, w \in W$, using costs $\mathcal{T}^w(\mathbf{h}^{(\ell-1}, \theta)$. $\Gamma^{w,(\ell} = \Gamma^{w,(\ell-1} \bigcup \{\breve{r}\}$, $w \in W$.
   Let $\hat{\mathbf{h}}_{SP}^{(\ell}$ be the path flow vector on the shortest paths at ℓ-th iteration.
3. If $G_{rel}(\mathbf{h}^{(\ell-1}) = \mathcal{T}(\mathbf{h}^{(\ell-1}, \boldsymbol{\theta})^\top (\mathbf{h}^{(\ell-1} - \hat{\mathbf{h}}_{SP}^{(\ell}) \Big/ \mathcal{T}(\mathbf{h}^{(\ell-1}, \boldsymbol{\theta})^\top \mathbf{h}^{(\ell-1} \leq \varepsilon$ then STOP. ($\mathbf{h}^{(\ell-1}$ is an approximate solution.)
4. *Equilibration Step/Master Problem.* Obtain a new vector of path flows $\mathbf{h}^{(\ell}$ by equilibrating the costs on paths within the working sets $\Gamma^{w,(\ell}$, $w \in W$. This is equivalent to solving VI (24) with $J^{w,(\ell}$ replacing J^w. In this implementation, this V.I. has been solved by using the double projection algorithm of Khobotov [18].

6 Computational Tests

In order to prove the efficiency of the approach adopted to solve the C3F model, computational tests are shown on two medium-to-large test networks on which the VI in (15) is solved for very large values of the parameters $\theta_i = \theta$, $\forall\, i \in \hat{N}$. The dimensions of the transit networks are reported in Table 1 below.

The travel costs t_a on network links are constant. Effective frequency functions on boarding links for both networks have been adopted as $f_a(v) = \kappa_a(1 - \rho_a(v)^\beta)$, with $\beta = 0.5$, $\kappa_a = 1/6$. The functional form of the loading factor $\rho_a(v)$ for the boarding queue at boarding links $a \in \hat{E}(i)$, $i \in \hat{N}$, is given by:

$$\rho_a(v) = \frac{v_a}{c_e - v_{x(a)}} \tag{25}$$

The computational experiments have been performed on a PC with a Quad 3.20 GHz Intel Core, 650 with 4 Gb RAM and Windows 7 Service Pack 1. 40 iterations of the SD algorithm described in previous Sect. 5 were performed on

Table 1. Dimensions of the test networks. a = boarding links, y = alighting links, x = dwelling links and e = in-vehicle links.

Network	Arcs					Nodes	O-D Pairs	Dest.	Lines	Stops
	a	y	x	e	Total					
Winnipeg	950	950	817	950	6,642	2,957	5,533	126	66	327
Zaragoza	3,864	3,864	3,802	3,864	19,585	9,966	4,156	129	31	1,176

each network adopting a single value of the θ parameter for all the stops. These 40 iterations took 481 s and 354 s of CPU time for the Winnipeg's network for $\theta = 10^3$ and $\theta = 10^4$, respectively. For the Zaragoza network the CPU time required for the 40 iterations were 610 s and 607 s, also for $\theta = 10^3$ and $\theta = 10^4$, respectively. In the implementation of the computational code, the evaluation of costs in (14) were carried out carefully for large values of θ to prevent overflows and to avoid cancellation errors as much as possible. This permitted to evaluate the costs even for such large values of θ. As shown in Fig. 4, the proposed method permits to solve very accurately these instances, specially in the case of the Winnipeg network, where the relative gap G_{rel} calculated at step 3 of the SD algorithm reaches extremely low values. Taking into account Theorem 1, the gap $G_{\texttt{C3F}}(\mathbf{v}^{(k)})$ for the solutions obtained $\mathbf{v}^{(k}$, will be even smaller.

The test network of Winnipeg has been used compare the performance of the SD-algorithm and the MSA algorithm in [5] for solving the deterministic problem. The comparison has been done using the relative gap $\hat{g}_{\texttt{C3F}}(\mathbf{v}) = (G^{(0}_{\texttt{C3F}}(\mathbf{v}) - \hat{G}^{(1}_{\texttt{C3F}}(v))/\hat{G}^{(1}_{\texttt{C3F}}(v)$ for the MSA algorithm running during 1500s. The evolution of $\hat{g}_{\texttt{C3F}}(\mathbf{v})$ versus the CPU time in seconds is shown in Fig. 5. As result, although the monitored relative gaps in Figs. 4 and 5 are of different kind, it seems clear that the SD algorithm outperforms the MSA method in [5]. In terms of the computed link flows, the following Fig. 6 shows the scatterplots of total link flows v calculated by the MSA method of Cepeda et al. [5] versus the total link flows v' obtained by solving the stochastic V.I. in (19) for different values of the parameter θ_i, set to a common value θ for all the stops i in the network. Two low values, $\theta = 0.4275$ and $\theta = 2.1375$, plus two large values, $\theta = 10^3$ and $\theta = 10^4$ have been chosen for the tests. The two low values of θ have been set using the approximation to the function U in (23), *as if* $|\hat{E}(i)| = 2$ for any stop in the network and to fit a dispersion for waiting times corresponding to a) a 2-Erlang scaled by a factor $k_1 = 0.7$ ($\theta = 0.4275$) and b) a 20-Erlang scaled by $k_1 = 0.5$ ($\theta = 2.1375$). The scatterplots show clearly the matching between v and v' for large values of θ. Table 2 shows the estimation of the coefficients of the simple linear regression model $v' = \beta_0 + \beta_1 v + e$, the values of the deviation σ_{SSE} of the regression and the tests on the values of the coefficient β_0. For large values of θ, β_0 can be accepted as null. Also, the estimator $\hat{\beta}_1$ of β_1 is very close to 1 as θ increases. For the case $\theta = 10^4$, the deviation σ_{SSE} of the regression, is $\sigma_{SSE} = 0.2939$, showing also the proximity between v' and v and that the link

flows have been compared probably up to the level of accuracy obtained solving the deterministic model using the MSA algorithm.

This allows us to conclude that the SD algorithm is an efficient algorithm to solve the deterministic C3F model for the congested strategy-based transit assignment as well as the stochastic version presented in this paper.

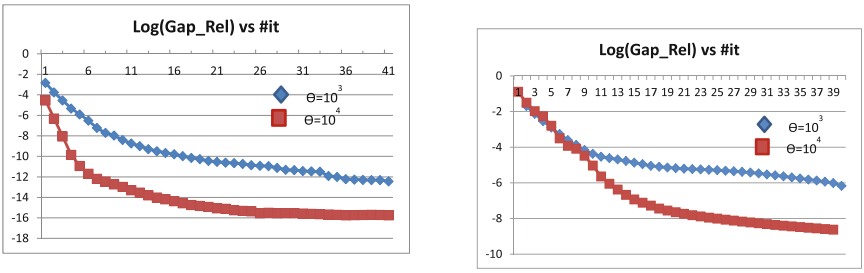


Fig. 4. Computational performance of the SD algorithm for the test networks of Winnipeg (left) and Zaragoza (right) for different values of θ.

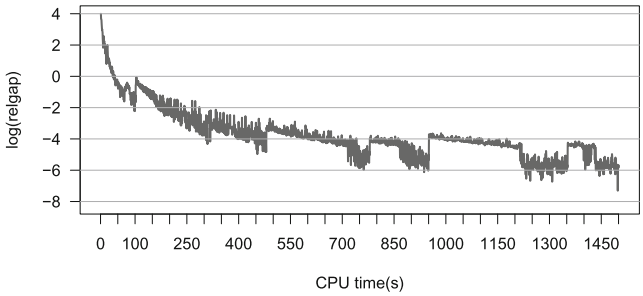


Fig. 5. Computational performance of the MSA algorithm in [5] for the test network of Winnipeg.

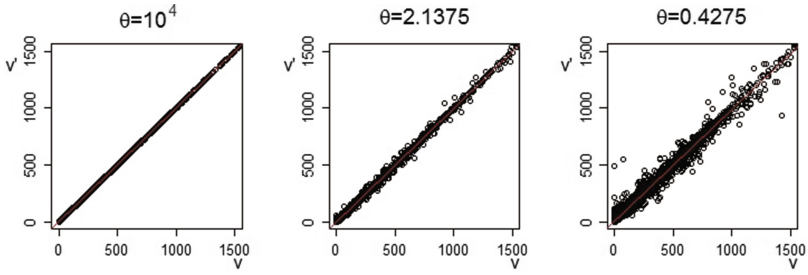


Fig. 6. Scatterplots showing the correspondence between link flows v obtained by the solving the Winnipeg's test network with the MSA algorithm in [5] and link flows v′ obtained by solving the approximate V.I. (`VIapprox`) in (15) for different values of θ.

Table 2. Regression analysis for scatterplots in Fig. 6, $v' = \beta_0 + \beta_1 v$.

θ	R^2	$\hat{\beta}_1$	$\hat{\beta}_0$	$\hat{\sigma}_{SSE}$	p-value $\beta_0 = 0$
10^4	0,9999996	0,999976	0,005367	0,2939	0,17167
10^3	0,9983678	0,998154	0,412039	18,93	0,10341
2,1375	0,9994829	0,997553	0,517923	10,64	0,00027
0,4275	0,9939224	0,984119	3,739802	36,10	0,00

7 Conclusions and Further Research

In the case that passengers follow the concept of strategies defined in [25], Cominetti and Correa [9] defined a model that takes into account the effect of congested queues for boarding the lines and in [5] the only known method up-to-date to solve that model was developed. In this paper two questions are addressed regarding this model. The first one is to state a stochastic version of it, which considers the effect of stochastic mean waiting times at boarding queues and stochastic travel times. This stochastic model is formulated as a stochastic variational inequality in the link flow space. The second question addressed is the formulation of an approximated model to the deterministic version using a smoothing approximation, which in the limit is consistent with the deterministic model. It is proved that this approximate model is, in fact a particular case of the stochastic transit assignment model, stated under very general conditions. Thus, a remarkable and distinctive approach of this stochastic model, is that it is not based on any explicitly stated assumption regarding the perception made by travelers of route costs. These assumptions are required in approaches based on random utility maximization theory, as it is the case in many current models. Also, a remarkable characteristic of the stochastic strategy-based transit assignment model presented in this paper is that it is not based on any particular probability distribution for travel costs or queueing delays at stops. The only requirement of the stated model is the homoscedasticity assumption for the randomness in queueing delays at stops. Moreover, in the paper it is presented a method for using the approximate model with a finite dispersion parameter θ to tackle networks with random queueing delays if they obey to a uniform or a k-Erlang probability distribution. Both, the stochastic model and the approximation by smoothing to the deterministic model can be solved advantageously using algorithms for the asymmetric traffic assignment problem in which path travel costs are destination-dependent. In particular, the use of a path-based simplicial decomposition algorithm for solving the models is illustrated and computational tests are shown proving the computational efficiency of the approach in obtaining very accurate solutions. Finally, although the models in this paper consider inelastic and deterministic demand, because of its compact formulation, its extension to elastic and stochastic demand can be considered in a future work.

Acknowledgments. Research supported under Spanish Research Projects TRA2008-06782-C02-02, TRA2014-52530-C3-3-P.

A Appendix: Smoothing Approximations for a Class of V.I.

Let $F(x)$ be a Lipschitz continuous function on X a polihedron on $\mathbb{R}^n$, let $f : \mathbb{R}^{p+n} \to \mathbb{R}^n$, $y : X \to \mathbb{R}^p$ be continuous functions and $H(\cdot)$ a proper, l.s.c. and sublinear function on $y(X)$ with $\partial^\infty H = \emptyset$. Consider now the following variational inequality

$$(\mathtt{VI}_1) \qquad \begin{bmatrix} 0 \in F(x) + f(\eta, x) + \mathsf{N}_X(x) \\ \eta \in \partial_y H(y(x)) \end{bmatrix} \tag{26}$$

Because of properties of H, it must be the support function of a convex set $E \subset \mathbb{R}^p$ (see, for instance [24], Proposition 3.17). Then

- $H(\mathsf{y}) = \mathsf{Max}_{\eta \in E}\, \mathsf{y}^\top \eta$ or equivalently $0 \in -\mathsf{y} + \mathsf{N}_E(\eta)$
- $\partial H(\mathsf{y}) = \{\eta \in E \mid \mathsf{y} \in \mathsf{N}_E(\eta)\}$

Thus, variational inequality $(\mathtt{VI}_1)$ can be restated as

$$(\mathtt{VI}_0) \qquad \begin{bmatrix} 0 \in F(x) + f(\eta, x) + \mathsf{N}_X(x) \\ 0 \in -y(x) + \mathsf{N}_E(\eta) \end{bmatrix} \tag{27}$$

Let S_0^* be its solution set and let $\varphi(\mathsf{y}, t)$ be a smoothing function verifying:

1. $\lim_{t \to 0+} \varphi(\mathsf{y}, t) = H(\mathsf{y})$,
2. $\varphi(\cdot, \cdot)$ is differentiable in y for $t \in]0, t_0[$, i.e. $\nabla_{\mathsf{y}} \varphi(\cdot, \cdot)$ is continuous in y for $t \in]0, t_0[$.
3. The gradient $\nabla_{\mathsf{y}} \varphi(\cdot, \cdot)$ is close to some elements in $\partial H(\cdot)$ as stated in Lemma 2.1 in [23]. More specifically, if $\{\mathsf{y}^{(k)}\}$ and $\{t^{(k)}\}$ are convergent sequences such that, $\{\mathsf{y}^{(k)}\} \to \bar{\mathsf{y}}$ and $\{t^{(k)}\} \to 0+$, then:

$$\lim_{k \to \infty} d(\nabla_{\mathsf{y}} \varphi(\mathsf{y}^{(k}, t^{(k)}),\ \partial H(\bar{\mathsf{y}})\ ) = 0. \tag{28}$$

Then Lemma A.1 below shows that previous V.I. $(\mathtt{VI}_0)$ in (27) can be approximated by the following V.I. using the smoothing function φ.

$$(\mathtt{VI}(t)) \qquad 0 \in F(x) + f(\nabla_y \varphi(y(x), t), x) + \mathsf{N}_X(x) \tag{29}$$

Let $S^*(t)$ be its solution set. Assume that it is possible to perform an ample parametrization of $(\mathtt{VI}(t))$, in (29) in the sense of Dontchev in [12]. A necessary and sufficient condition for it would be:

$$\exists\, k,\ 1 \le k \le n \quad \text{so that } \frac{df_k}{dt} = \sum_{\ell=1}^{n} \frac{\partial f_k}{\partial \eta_\ell} \frac{\partial^2 \varphi}{\partial y_\ell \partial t} \neq 0 \ \text{ at } t \in [0, t_0[\tag{30}$$

Then, the application of Theorem 7.1 in [12] results in:

$$d(x_t,\ S^*(t)) \le \kappa |t - t'| \tag{31}$$

where $\kappa < +\infty$ if variational inequality (27) has bounded solutions.

Lemma A.1. Assume that condition (*31*) applies for V.I. ($\mathtt{VI}(t)$) in (*29*) and that $\kappa < +\infty$. *Let* $S_0^* \neq \emptyset$ *be the solution set of* ($\mathtt{VI}_0$). *Let* S_{0+} *be defined as* $S_{0+} = \liminf_{t\to 0+} S^*(t)$ and assume that $S_{0+} \neq \emptyset$. Then $S_{0+} \subseteq (S_0^*)_x$

Proof. First notice that, because of (31), if $S^*(t) \neq \emptyset,\ \forall t > 0,\ t \in]0, t_0[$, then $S_{0+} \neq \emptyset$. Let $x_{0+} \in S_{0+}$. Then, it is the limit of a sequence $x_k \to x_{0+}\ \forall\{t_k\},\ t_k \to 0+$. Then, it is clear that

$$\eta_{0+} = \lim_{k\to\infty} \nabla_y \varphi(y(x_k), t_k) \in \partial H(y(x_{0+})) \tag{32}$$

and that $(x_{0+}, \eta_{0+}) \in S_0^*$, i.e. are solutions of V.I. $\mathtt{VI}_1$, (26). i.e., $x_{0+} \in (S_0^*)_x$, where $(\cdot)_x$ stands for projection on the x space. □.

References

1. Auslender, A.: Optimisation: Méthodes Numériques. Masson, Paris (1976)
2. Birge, J.R., Louveaux, F.: Introduction to Stochastic Programming. Springer Series in Operations Research. Springer, New York (1997)
3. Cascetta, E.: Transportation Systems Analysis. Models and Applications. Springer Optimization and Its Applications Series, 2nd edn. Springer, New York (2009)
4. Chen, X., Wets, R., Zhang, Y.: Stochastic variational inequalities: residual minimization smoothing sample average approximations. SIAM J. Optim. **22**(2), 649–673 (2012)
5. Cepeda, M., Cominetti, R., Florian, M.: A frequency-based assignment model for congested transit networks with strict capacity constraints: characterization and computation of equilibria. Transp. Res. B **40**, 437–459 (2006)
6. Chriqui, C., Robillard, P.: Common bus lines. Trans. Sci. **9**, 115–121 (1975)
7. Codina, E.: A variational inequality reformulation of a congested transit assignment model by Cominetti, Correa. Cepeda and Florian. Trans. Sci. **47**(2), 231–246 (2013)
8. Codina, E., Ibáñez, G., Barceló, J.: Applying projection-based methods to the asymmetric traffic assignment problem. Comput. Aided Civ. Infrastruct. Eng. **30**(2), 103–119 (2015)
9. Cominetti, R., Correa, J.: Common-lines and passenger assignment in congested transit networks. Trans. Sci. **35**(3), 250–267 (2001)
10. Cortés, C.E., Jara-Moroni, P., Moreno, E., Pineda, C.: Stochastic transit equilibrium. Transp. Res. Part B **51**, 29–44 (2013)
11. De Cea, J., Fernández, E.: Transit assignment for congested public transport systems: an equilibrium model. Trans. Sci. **27**(2), 133–147 (1993)
12. Dontchev, A.L., Rockafellar, R.T.: Ample parametrization of variational inclusions. SIAM J. Optim. **12**(1), 170–187 (2001)
13. Fu, Q., Liu, R., Hess, S.: A review on transit assignment modelling approaches to congested networks: a new perspective. Procedia Soc. Behav. Sci. **54**, 1145–1155 (2012)
14. Gentile, G., Florian, M., Hamdouch, Y., Cats, O., Nuzzolo, A.: The theory of transit assignment: basic modelling frameworks. In: Gentile, G., Noekel, K. (eds.) Modelling Public Transport Passenger Flows in the Era of Intelligent Transport Systems. Springer Tracts on Transportation and Traffic, vol. 10, pp. 287–386. Springer, Cham (2016). doi:10.1007/978-3-319-25082-3_6

15. Gentile, G., Noekel, K. (eds.): Modelling Public Transport Passenger Flows in the Era of Intelligent Transportation Systems. Springer Tracts on Transportation and Traffic. Springer, Cham (2016)
16. Jiang, Y., Szeto, W.Y.: Reliability-based stochastic transit assignment: formulations and capacity paradox. Transp. Res. Part B **93**, 181–206 (2016)
17. Jiang, Y., Szeto, W.Y., Ng, T.M., Ho, S.C.: The reliability-based stochastic transit assignment problem with elastic demand. J. Eastern Asia Soc. Transp. Stud. **10**, 831–850 (2013)
18. Khobotov, E.N.: Modification of the extragradient method for solving variational inequalities and certain optimization problems. USSR Comput. Math. Math. Phys. **27**, 120–127 (1987)
19. Lam, W.H.K., Gao, Z.Y., Chan, K.S., Yang, N.: A stochastic user equilibrium assignment model for congested transit networks. Transp. Res. **33B**(5), 351–368 (1999)
20. Liu, Y., Bunker, J., Ferreira, L.: Transit user's route-choice modelling in transit assignment: a review. Transp. Rev. **30**(6), 753–769 (2010)
21. Nielsen, O.A.: A stochastic transit assignment model considering differences in passenger utility functions. Transp. Res. Part B **34**, 377–402 (2000)
22. Nguyen, S., Pallotino, S.: Equilibrium traffic assignment in large scale transit networks. Eur. J. Oper. Res. **37**(2), 176–186 (1988)
23. Peng, J.-M.: A smoothing function and its applications. In: Fukushima, M., Qi, L. (eds.) Reformulation: Nonsmooth, Piecewise Smooth and Smoothing Methods. Applied Optimization Series, pp. 293–316. Kluwer Academic Publishers, Dordrecht (1999)
24. Rockafellar, T., Wets, R.J.-B.: Variational Analysis. Grundlehren der mathematischen Wissenschaften, vol. 317. Springer, Heidelberg (1998). doi:10.1007/978-3-642-02431-3
25. Spiess, H.: Contribution à la théorie et aux outils de planification des réseaux de transport urbains. Ph D thesis, Département d'Informatique et Récherche Opérationnelle, Publication 382, CRT, U. de Montréal (1984)

Evaluation of CO_2 Emission Reduction from Vehicles by Information Provision Using Driving Simulator

Yukimasa Matsumoto(✉) and Shogo Ishiguro

Department of Civil Engineering, Meijo University, Nagoya City, Japan
matumoto@meijo-u.ac.jp

Abstract. Reduction of CO_2 emissions has become an important social issue all over the world. Especially, the transportation sector is expected to play a sufficient role for it. Therefore, the information provision system to a vehicle for passing through the upcoming signalized intersection is proposed in this paper to decrease unnecessary vehicle movements and also to reduce the amount of CO_2 emissions from a vehicle approaching the signalized intersection. The system provides recommended speed information and accelerator-off indication, in which the vehicle could pass through the intersection or shorten the idling time if it would follow the provided information. Driving experiments were conducted with a driving simulator where the proposed information system was introduced. Multiple regression analysis showed that providing the recommended speed information and the accelerator-off indication reduced the amount of CO_2 emissions significantly. However, as subjective evaluations of the system by participants were dispersed, the information provision system is expected to be tailored to each driving characteristics.

Keywords: CO_2 · Signalized intersection · Recommended speed · Accelerator · Driving simulator · Eco-driving

1 Introduction

A vehicle is recognized to be a convenient transportation mode for moving whenever and wherever we want to go. Therefore, vehicles penetrate our lives widely and deeply, and the number of vehicles has been increasing around the world. In Japan, the number of passenger vehicles was 42,956,339 in 1995 and has been increasing to 60,517,249 in 2015, which is about 1.4 times. Consequently, our lives depend on vehicles strongly, especially around suburban area.

On the other hand, it is well known that external diseconomies arise from vehicles such as traffic congestion, traffic accidents and traffic pollution. Various measures to tackle these problems have been implemented even now. In addition, reduction of CO_2 emissions from vehicles has become an important social issue all over the world. In fact, at the COP21 held in December 2015, 195 countries adopted the Paris Agreement to address climate change in which the countries will aim to keep global temperatures from rising more than 2°C by 2100. The amount of CO_2 emissions from vehicles is therefore

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_14

required to be reduced. In fact, the amount of CO_2 emissions from a transportation sector accounts for 17.2% of the total amount from all sectors at 2014 in Japan. The ratio of CO_2 emissions from vehicles is 86.0% of the transportation-oriented emissions and the ratio from passenger vehicles is 50.8% [5]. Although an electrical vehicle, a plug-in hybrid vehicle and a hybrid vehicle generate lower CO_2 emissions, such vehicles account for less than 8% of all owned passenger vehicles in Japan at 2014.

In terms of CO_2 emissions from a vehicle, a field experiment conducted by Kadoyu et al. [9] on a street at Tsukuba city shows that the ratios of CO_2 emissions from a passenger vehicle in idling, accelerating, cruising and decelerating are 13%, 35%, 48% and 4% respectively while the ratios of travel times in the corresponding running states are 18%, 19%, 48% and 15%. This means that CO_2 emissions per unit time is the largest when a vehicle accelerates and the third largest when a vehicle idles on a street at Tsukuba city. Barth and Boriboonsomsin [2] evaluated CO_2 emissions based on the average speed of the trip from a vehicle activity database collected as a post-census travel survey in 2001 by the Southern California Association of Governments and show that grams per mile CO_2 emission rates become quite high when average speeds are very low. This also shows a distance-normalized CO_2 emission rate reaches infinity when a vehicle idles.

Idling, acceleration and deceleration of a vehicle occur frequently around a signalized intersection so that the amount of CO_2 emissions from a vehicle around a signalized intersection becomes larger. Decreasing such unnecessary vehicle movements therefore leads to reducing CO_2 emissions from vehicles. Information provision on traffic signal change to help a driver pass through a signalized intersection or shorten an idling time at a red signal is expected for decreasing such unnecessary vehicle movements.

Information provision based on upcoming traffic signal information to a driver has attracted attention from many researchers. Widodo et al. [19] estimated fuel consumption and emission rates of environment-adaptive driving with inter-vehicle communications by an autonomous running traffic flow simulator. Asadi and Vahidi [1] propose an algorithm to schedule an optimal speed trajectory for reducing idling time at stop line using upcoming traffic signal information. Tielert et al. [17] evaluated traffic-light-to-vehicle communication for reducing fuel consumption and emissions by large-scale simulations. Rakha and Kamalanathsharma [14] present a framework to enhance vehicle fuel consumption efficiency while approaching a signalized intersection through the provision of signal phase and timing information. Barth et al. [3] propose real-time dynamic advice to drivers to reduce fuel consumption and CO_2 emissions. Deceleration Support System (DSS) was designed to reduce fuel consumption by encouraging drivers to release the accelerator pedal earlier at a red signal and to stop safely without hesitation at a yellow signal [6]. Matsumoto et al. [11] simulated vehicle movements approaching a signalized intersection under information provision using a microscopic traffic simulation based on observed actual vehicle movements. Ubiergo and Jin [18] propose control algorithms for calculation of Advisory Speed Limit indications and demonstrated savings of around 15% in travel delays and around 8% in fuel consumption.

Almost all of the above studies implemented a microscopic traffic flow simulation to evaluate the proposed system and showed that the information provision to a driver was effective in improving driving behavior and also reducing the amount of CO_2

emissions from vehicles. However, the most proposed systems were evaluated under rather ideal traffic conditions in which all vehicles follow the provided information. In real world, a response of a driver to the information provision might be dispersed by each driver so that the response by each driver is expected to be grasped on real road traffic conditions.

Although field experiments are strongly desired for evaluating the proposed information system, there would be a risk of traffic accidents or any regulations not allowing these field experiments. Therefore, experiments using a driving simulator are recognized to be an effective method to evaluate a newly proposed system.

Effects to indicate the advisory speed for eco-driving to a driver were grasped by using a driving simulator. Niu and Sun [12] evaluated the effectiveness of the green wave guidance strategy and the eco-driving speed guidance strategy with a multi-vehicle driving simulator. Kircher et al. [10] investigated the difference of effects between intermittent and continuous eco-driving speed information with a truck driving simulator. Brouwer et al. [4] tailored different types of the display to present eco-driving speed feedback with a truck driving simulator. These studies indicate that providing eco-driving speed information is significantly efficient to reduce fuel consumption and also the amount of CO_2 emissions. However, these studies evaluated only effects of providing eco-driving speed information and do not compare the effect of providing other types of the information.

Indication of the accelerator pedal maneuver to a driver seems to be also efficient to reduce the amount of CO_2 emissions. Staubach et al. [16] evaluated an eco-driving support system with a driving simulator, which gave recommendation to a driver concerning gear and accelerator pedal maneuvers for efficient fuel consumption. Jamson et al. [8] and Jamson et al. [7] investigated the effectiveness of a visual or haptic eco-driving system with a driving simulator, which advised a driver of the most fuel efficient accelerator pedal angle. These studies show the indication of accelerator pedal maneuver to a driver with a visual or haptic device is effective to improve driving behavior without negative effects on safety. As mentioned above, only releasing the accelerator pedal is also shown to be effective for reducing the amount of CO_2 emissions. It seems to be easier for a driver to follow the instruction of releasing the accelerator pedal than the other maneuver on the accelerator pedal.

As many studies used a driving simulator effectively to evaluate newly proposed information system and a driving simulator makes it possible to conduct experiments without a risk of a traffic accident [15], we use a 3D driving simulator for exploring the information provision system proposed in this study. The effectiveness in reducing the amount of CO_2 emissions from vehicles approaching a signalized intersection is evaluated through a driving experiment with a driving simulator under various conditions such as different patterns of information provision and various timing to provide the information. Moreover, driving behavior and compliant response by each driver are observed through the driving experiment when the proposed information is provided. The proposed system provides two types of information to a driver, accelerator-off indication and recommended speed information, in order to pass through the upcoming signalized intersection. If a driver would follow such information, the driver could pass through the upcoming signalized intersection or shorten the idling time at a red signal.

Effects of the information provision on the driving behavior and also reducing the amount of CO_2 emissions are evaluated statistically and quantitatively through the driving experiments with a 3D driving simulator. Furthermore, the proposed information provision system is evaluated from a viewpoint of a driver by the questionnaire to the participants after the driving experiments.

2 Information Provision System

2.1 Outline of Information Provision System

For decreasing unnecessary vehicle movements such as rapid deceleration and long idling, two types of information are designed in this study. One is recommended speed information. A vehicle can pass through the upcoming signalized intersection as long as the vehicle runs at the recommended speed provided by the system. Another is accelerator-off indication, which informs a driver of releasing the accelerator pedal for passing through the upcoming signalized intersection or shortening idling time during a red signal at the intersection. The information is provided to a driver in the vehicle.

2.2 Relationship Between Information Provision and Vehicle Movement

Figure 1 shows a time-space diagram of vehicles approaching a signalized intersection, in which the horizontal axis of this figure denotes the time elapse and the vertical axis denotes the distance from the upcoming signalized intersection. The color bar on the horizontal axis indicates the signal status. In this study, for convenience, a yellow phase is regarded as the same as a red phase so that vehicles stop when the traffic signal is yellow.

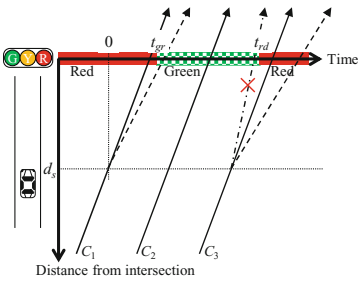


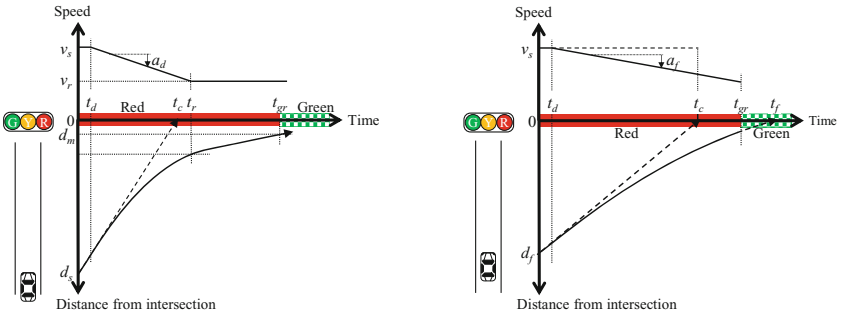
Fig. 1. Time-space diagram of vehicle approaching signalized intersection

In this figure, the vehicle C_1 stops at the intersection if the vehicle runs at a constant speed. For this movement, the information to reduce the current speed is provided at distance d_s from the intersection. A green signal starts after t_{gr} from the time when the vehicle passes d_s and the next red signal starts after t_{rd}. If the vehicle C_1 would follow the information, it could pass through the intersection like a broken line. The vehicle C_2 approaches the intersection at a constant speed and will pass through the intersection

during green time. For this movement, no information is provided. On the other hand, the vehicle C_3 approaches the intersection during green time and will stop at the intersection due to a red signal. For this movement, the vehicle C_3 could pass through the intersection like a dashed line if it would accelerate. However, this movement may raise a risk of a traffic accident. Therefore, the vehicle is informed of reducing the current speed in advance by the proposed system like a broken line so that it will shorten the idling time at the intersection during a red signal.

2.3 Method of Information Provision

The left figure of Fig. 2 shows a time-space diagram when the recommended speed is provided to a vehicle. The upper part of this figure shows the vehicle speed corresponding to the time elapse. The vehicle approaching the signalized intersection receives the recommended speed at the distance d_s from the intersection at time 0 and starts to decelerate from the current speed v_s after the reaction time t_d. The deceleration rate is assumed to be a_d. The vehicle continues to decelerate and then runs at the recommended speed v_r after time t_r. At the moment the vehicle reaches the distance d_m from the intersection, the traffic signal turns to green from red. The d_m mainly includes the queue length of the vehicles during red time and a marginal distance which corresponds to a time lag between the start of a green time and the arrival time at the intersection of the vehicle. This marginal distance helps the vehicle pass through the intersection safely because the traffic signal turns to green slightly prior to the vehicle arrival at the intersection.



(a) Recommended speed (b) Accelerator-off indication

Fig. 2. Two types of information provision

Form this figure, the recommended speed v_r can be calculated as:

$$v_r = v_s + a_d\left(t_{gr} - t_d\right) + \sqrt{a_d^2(t_{gr} - t_d)^2 - 2a_d\left(d_s - d_m - t_{gr}v_s\right)}. \tag{1}$$

For the case when the recommended speed is provided, a driver is supposed to decelerate by braking until the recommended speed. A driver must therefore glance at the speed panel to confirm whether the current speed reaches the recommended speed.

This maneuver may be a considerable load for a driver. In order to lighten such load, accelerator-off indication, which indicates a driver to release the accelerator pedal and seems to be easier for a driver to follow, is also provided in the proposed system.

The right of Fig. 2 shows a time-space diagram when the accelerator-off indication is provided to a vehicle. In this figure, d_f, t_f and t_c denote the distance of the vehicle from the upcoming signalized intersection, forecasted arrival time at the intersection with releasing the accelerator pedal and forecasted arrival time with the current running speed v_s, respectively. The deceleration rate of during releasing the accelerator pedal is assumed to be a constant of a_f and t_d is the reaction time.

From this figure, t_f is calculated as:

$$t_f = \frac{a_f t_d - v_s + \sqrt{v_s^2 + 2a_f(d_f - d_m - t_d v_s)}}{a_f}. \tag{2}$$

This figure also shows that the vehicle reaches the upcoming intersection during red time without decelerating like a broken line. However, if the vehicle would decelerate by releasing the accelerator pedal, it could pass through the intersection during green time. When this condition is satisfied, the accelerator-off indication is provided by the proposed system. As t_{gr} means the remaining red time until the signal will turn to green, the following relation can be obtained:

$$t_c < t_{gr} < t_f. \tag{3}$$

Provision of Recommended Speed Based on Fixed Location. According to the current speed v_s at the fixed location d_s and the remaining red time t_{gr} or green time t_{rd}, one of information from three patterns is provided in the type of the provision of the recommended speed. These are no information, the recommended speed information and also the accelerator-off indication. If the vehicle can pass through the upcoming signalized intersection with the current speed v_s, the current speed is provided as the recommended speed as long as the current speed is less than the speed limit or no information is provided otherwise.

If the vehicle cannot pass through the intersection with the current speed, the recommended speed v_r is calculated by Eq. (1) and then judged to be provided or not from Fig. 3 showing that patterns of the information provided by the system based on the current speed v_s at the fixed location d_s and the calculated recommended speed v_r. In this figure, the upper speed limit is set to be 60 km/h and the lower limit of the recommended speed is 15 km/h. In practice, these speed limits should be given based on the traffic conditions and the road environment. If the calculated recommended speed to pass through the intersection is higher than the current speed or lower than the lower limit, no information is provided for safety reason. If the calculated recommended speed is lower than the current speed and higher than the upper speed limit, the accelerator-off indication is provided for just shortening the idling time during a red signal. Furthermore, if the calculated recommended speed is 20 km/h lower than the current speed, in which rapid deceleration would be required, the accelerator-off indication is also provided for shortening the idling time. This allowable range of 20 km/h between the current and the

recommended speeds should be also given based on the conditions. For other cases, the recommended speed is provided to a driver. Although the accelerator-off indication is provided in this information provision type, a vehicle cannot pass through the intersection but shorten the idling time.

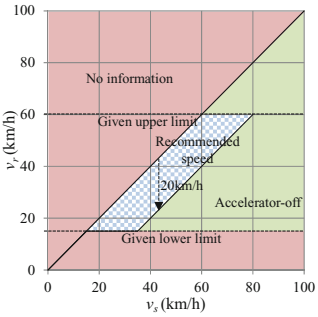


Fig. 3. Pattern of information provision

Provision of Accelerator-off Indication Based on Current Speed. The accelerator-off indication is provided with the comment of "Upcoming intersection is passable at a green signal" if the vehicle can pass through the upcoming signalized intersection by just releasing the accelerator pedal. Figure 4 represents the relationship between the current running speed v_f, the remaining red time t_{gr} and the location for the information provision d_f where the accelerator-off indication with the comment is provided. When the remaining red time is a certain time, the location for the information provision becomes farther in proportion to the current speed. When the current speed is a certain speed, the location also becomes farther in proportion to the red remaining time.

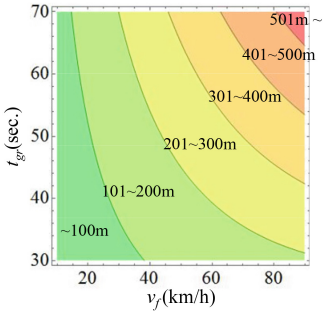


Fig. 4. Relationship between v_f, t_{gr} and d_f

For the vehicle which cannot pass through the intersection even by releasing the accelerator pedal, the accelerator-off indication is also provided at 250 m from the intersection for only reducing the idling time at a red signal.

3 Experiments Using 3D Driving Simulator

3.1 Outline of 3D Driving Simulator and Experimental Course

The information provision system proposed in this study is introduced in a fixed-base 3D driving simulator which is composed of a seat, a steering wheel, pedals, speakers and three 42" monitors having 120° field of view including side view of mirror images. The 3D virtual driving environment is generated by UC-Win/Road.

The experimental course of 3,500 m shown in Fig. 5 consists of consecutive 4 signalized intersections, where one gentle curve section is located at the middle of the course to make driving unmonotonous. The course has two lanes for both directions in which the lane width is set to be 3.25 m. A speed limit of the course is set to be 40 km/h. The parameter settings of each traffic signal are different from each other due to generating various running states of a vehicle among signalized intersections.

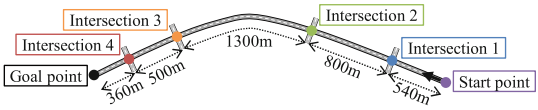


Fig. 5. Experimental course

The information provision system displays the information on the screen of the driving simulator based on the location and the speed of the current vehicle, and the remaining time of the current signal phase. The recommended speed is displayed with the comment of "Upcoming intersection is passable at a green signal" which is shown in the left of the Fig. 6. The displayed recommended speed is rounded off in unit of 5 km/h by the system considering driver's recognition time and safety. The accelerator-off indication with the comment "Idling time at upcoming intersection is shortened" shown in the middle of Fig. 6 is displayed when the vehicle cannot pass through the intersection even by safely decreasing the running speed. These patterns of the information are provided for the system of providing the recommended speed.

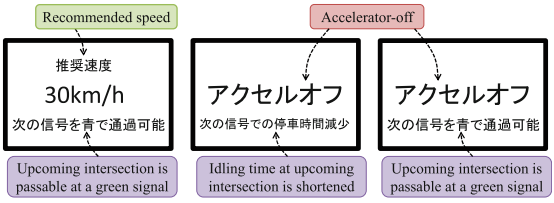


Fig. 6. Information displayed on the screen

In the type of the accelerator-off indication, the information is displayed with the comment of "Upcoming intersection is passable at a green signal" shown in the right of Fig. 6 when the vehicle can pass through the upcoming signalized intersection by just releasing the accelerator pedal. Moreover, the accelerator-off indication with the

comment "Idling time at upcoming intersection is shortened" is displayed at 250 m from the intersection for shortening the idling time even if the vehicle cannot pass through the intersection by releasing the accelerator pedal.

3.2 Design of Information Provision on Driving Simulator

In order to investigate the appropriate distance from the upcoming intersection for providing the recommended speed, the distance d_s is changed to be 250 m, 300 m or 350 m in the driving experiments. Consequently, there are 4 patterns of the providing information to pass through the intersection, the recommended speed at 250 m (RS250), at 300 m (RS300), at 350 m (RS350) and the accelerator-off indication (AO). Providing no information is denoted as NI hereafter.

The provided information is processed by an API for programming and can be displayed on the screen of the driving simulator in real time. Necessary variables like the current vehicle position, the current vehicle speed and the remaining time of the current signal phase can be collected at a rate of 60 Hz within the driving simulator.

Figure 7 shows the actual screen of the driving experiment in this study. In this experiment, it is assumed that the information is displayed on a display of a car navigation appearing the lower left part of the middle monitor with black characters on white background.



Fig. 7. Screen images of driving simulator

At every time when any information is provided, a notice tone of "dingdong" is sounded and makes a driver noticed that any information is provided on the screen.

3.3 Participants and Outline of Experiment

The participants consisted of 30 males and 3 females who were from 20 to 60 years old and with a driving license, in which one person was withdrawn due to simulator sickness. Driving experiences and primal purposes to drive a car were different considerably from person to person.

Before the start of the driving experiment, each participant took part in pre-drive interview focusing on the personal attributes. After that, the participant drove on the experimental course 2 times after a trial driving on another simple course without any information for 3 min to get used to the simulator. Then, the participant drove on the experimental course 2 times with the provision of any type of the information, such as NI, RS250, RS300, RS350 or AO. The pattern of the provided information was assigned randomly. Consequently, this driving experiment ensured every participant to

experience the driving randomly with NI, with the accelerator-off indication or with the recommended speed information. After each driving experiment, the participant was asked to complete the questionnaire about the driving experiments with the information provision.

Participants were instructed to drive as if they normally drove an actual road with a speed limit of 40 km/h. Before the start of the experiment, "You might decrease unnecessary vehicle movement if you would follow the information provided while you are driving" was told to every participant. No instruction to necessarily follow the information or to drive economically was told to them. Therefore, there were some drivers who didn't follow the provided information at all or drove at higher speed over the upper speed limit in order to finish the experiment as soon as possible. These data are excluded from the evaluation of the system.

3.4 Performance Evaluation

Experimental results are evaluated based on running speeds, maneuvers of accelerating/braking and the amounts of CO_2 emissions from the vehicle for entire experimental course.

Average Running Speed. Running speeds during the driving experiment were influenced by the provided information. Figure 8 shows the average running speed of all participants for the entire course. In this figure, the horizontal axis denotes the distance from the last signalized intersection (m) and the vertical axis denotes the average vehicle running speed (km/h). It is obvious that the average running speeds with any type of information provision starts to reduce before each intersection, because the vehicle receiving the information decelerated by following the information. Regarding the speed at each intersection, the average speed with NI tends to be the lowest. This means that the number of vehicles idling at intersection due to a red signal was the most when no information was provided.

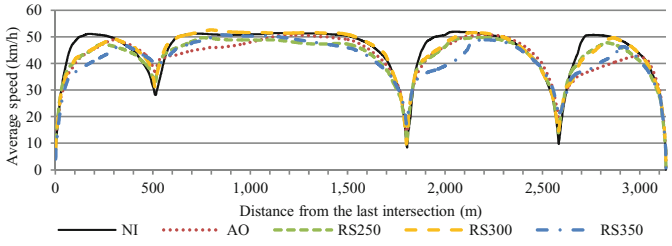


Fig. 8. Average running speed by type of information

Motion of Accelerating/Braking. Rates of accelerating and braking during the driving experiment were recorded. The rate ranges from −1.0 to 1.0 in which the states of maximum acceleration, no-step-on and maximum braking are 1.0, 0.0 and −1.0, respectively. Figure 9 shows the average rates of accelerating/braking of all participants approaching the each upcoming signalized intersection by each information provision

type. In this figure, the red color means a participant accelerates by pressing the accelerator pedal and the green color means a participant brakes the vehicle by pressing the brake pedal. No step-on state is represented by the yellow color in this figure.

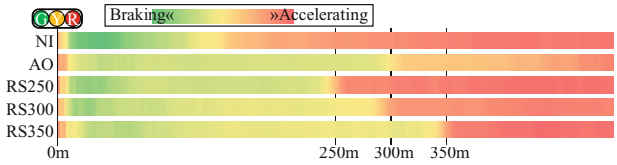


Fig. 9. Accelerating/braking rate by each information provision

From this figure, it is obvious that vehicles with NI continue to accelerate closer to the upcoming intersection than that with any information. Moreover, vehicles with NI decelerate stronger near the intersection than that with any information, which means that many stop-and-go movements occur. The vehicles with AO release the accelerator pedal faster than that with other information.

Regarding the recommended speed, it can be seen that vehicles start to decelerate after the information provision. This is grounded statistically by the Welch's *t*-test whether a significant difference of the accelerating/braking rate for each 10 m section exists from the rate with NI by a distance from the intersection, shown in Table 1. These results show that the information provision has influence on the motions of accelerating and braking significantly.

Table 1. Statistical test of average rate of accelerating/braking for each 10 m compared with NI

	191–200 m	241–250 m	291–300 m	341–350 m	391–400 m
AO	-1.575^{**}	-1.823^{**}	-1.301^{**}	-1.009^{**}	-0.945^{**}
RS250	-1.844^{**}	-0.775^{**}	−0.207	–	–
RS300	–	-1.814^{**}	-0.343^{**}	−0.144	–
RS350	–	–	-1.883^{**}	-0.833^{**}	−0.311

*5% significant, **1% significant

Amount of CO_2 Emission. Among several methods to calculate the amount of CO_2 emissions based on vehicle movements, the method proposed by Oguchi et al. [13] is used in this study. The amount of CO_2 emissions is calculated from the data acquired in the driving simulator as follows:

$$E = k_c \cdot \left\{ 0.3T + 0.028D + 0.056 \sum\nolimits_k \delta_k \left(v_k^2 - v_{k-1}^2 \right) \right\}, \tag{4}$$

where, E: the amount of CO_2 emissions (kg-C), k_c: emission factor as 0.00231 (kg-C/petrol-litter), T: travel time (sec.), D: travel distance (m), δ_k: dummy variable of accelerating (1 if accelerating or 0 otherwise), v_k: vehicle speed at k (m/sec.) and k: speed measurement point.

Figure 10 shows the average amounts of CO_2 emissions from vehicles during each driving experiment, which is calculated by the Eq. (4). Results of the Welch's *t*-test whether a significant difference of the average amount is exists from NI are also shown in this figure. From this figure, it is obvious that providing any type of information can reduce the amount of CO_2 emissions. Especially, the accelerator-off indication seems to reduce CO_2 emissions most effectively, which is about 6% reductions compared with the average amount with NI. However, when the recommended speed is provided at 300 m from the upcoming intersection, the effect to reduce CO_2 emissions becomes smaller due to dispersion of driving behavior of each participant.

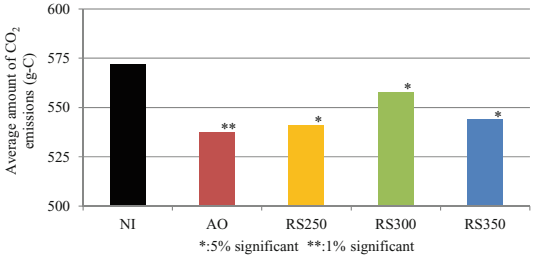


Fig. 10. Average amount of CO_2 emissions by each information pattern

3.5 Quantitative Analysis

Multiple regression analysis is conducted to grasp the information provision effects quantitatively. The objective variable is the amount of CO_2 emissions of each driving experiment and the explanatory variables are the location of the information provision, a dummy variable whether the accelerator-off indication is provided, a dummy variable of male, a period of driving experience and driving frequency per week.

Table 2 shows the result of the multiple regression analysis. The multiple correlation coefficient and *F*-value show that this analysis is significant statistically. The multiple regression coefficients indicate that the factors to reduce the amount of CO_2 emissions are any information provision and less driving frequency, and the factors to increase it are male, the period of driving experience and frequent driving. Especially, providing the accelerator-off indication can reduce 35 g of CO_2 emissions for this experimental course and providing the recommended speed at 350 m can also reduce 32 g of CO_2 emissions which is almost same reduction with providing the accelerator-off indication. From the standardized multiple regression coefficients, it can be seen that providing the recommended speed 100 m earlier has same effect as providing the accelerator-off indication. However, it is noted that this effect might be limited for the distance to provide the recommended speed information between 250 m and 350 m from the upcoming intersection.

Table 2. Result of multiple regression analysis

		Coefficient	Standardized coefficient	*t*-value
Constant		524.955		32.514**
Recommended speed distance (100 m)		−9.198	−0.414	−2.791**
Accelerator-off indication		−34.597	−0.429	−3.037**
Male		32.661	0.307	2.645*
Driving experience	More then 2 years and less than 5 years	12.063	0.140	0.921
	5 years and more	18.005	0.265	1.832
Driving frequency per week	More than 2 times and less than 6 times	−2.928	−0.0041	−0.293
	6 times and more	23.791	0.311	2.376
Multiple correlation coefficient		0.660		
F value		4.847**		

*5% significant, **1% significant

3.6 Evaluation by Participants

Each participant was asked to answer a questionnaire before the driving experiment and also after each run. The questionnaire consisted of personal attributes and evaluation of each run with any information provision.

Figure 11 shows the results of evaluations of each information provision by the participants. From this figure, it can be seen that around half of participants would intend to use this system and evaluated it useful. However, 40% and more of participants

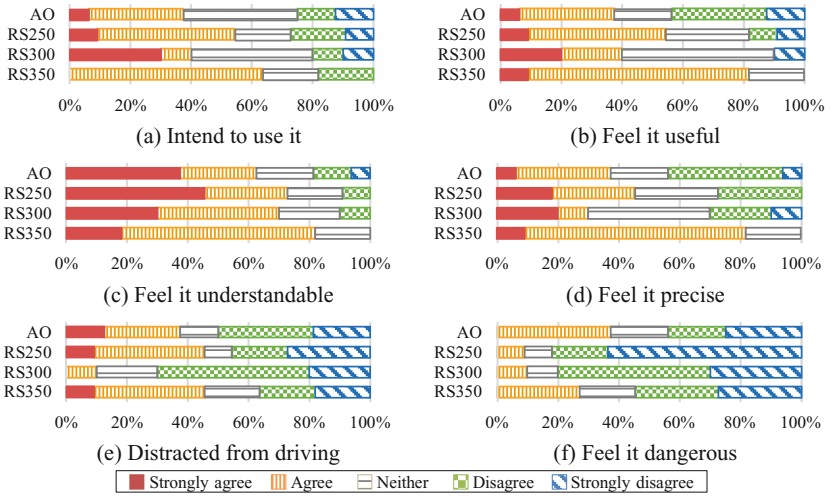


Fig. 11. Evaluation result by participant

evaluated the AO as not useful. Around 70% of participants felt the provided information understandable and positive evaluations of the recommended speed are higher than that of the AO. The evaluations on the preciseness are dispersed except the RS350 which was evaluated positively. Around half of participants felt not to be distracted from driving by the information provision. Especially, in case of the RS300, around 70% of participants evaluated it positively. Moreover, although 40% and more participants didn't feel it dangerous, around 30% of participants evaluated the AO and the RS350 dangerous which provided the information farther from the intersection. The reason behind these negative evaluations seems to be that the upcoming traffic signal changed to green from red just before the vehicle arrived the intersection due to a short marginal time set in this experience. On the other hand, around 40% of participants perceived the information provision timing precise, longer marginal time might be required based on each driving characteristic in actual situation.

4 Conclusion

For reducing the amount of CO_2 emissions from vehicles approaching a signalized intersection, the method to provide information to a driver in a vehicle was proposed in this paper. By the proposed information provision system, unnecessary vehicle movements such as rapid deceleration and long idling can be decreased. The system provides recommended speed information and accelerator-off indication, in which the vehicle could pass through the intersection or shorten the idling time if it would follow the provided information. The proposed system was introduced into 3D driving simulator, and evaluated its effectiveness for reducing CO_2 emissions. 33 participants aged from 20 to 60 years old took part in driving experiments. In this experiment, the recommended speed or the accelerator-off indication was provided to the driver based on the vehicle running condition and the upcoming traffic signal state.

As a result of multiple regression analysis, it was shown that the recommended speed and the accelerator-off indication were able to reduce CO_2 emissions significantly. Moreover, providing the recommended speed at 350 m has same effect as providing the accelerator-off indication theoretically.

While the information provision had significant effect to reduce CO_2 emissions, the evaluations of the information provision by the participants were dispersed from positive to negative. Especially, some participants felt to be distracted from driving by the information provision and around 30% of participants felt it dangerous. The reason is that the marginal time was set small in this experiment and following the recommended speed needed the participants to concentrate the maneuver of the accelerator pedal for adjusting the recommended speed, so that the participants felt stress. For future works, it is necessary to explore a suitable marginal time and a type of information provision to alleviate stress of a driver to follow the provided information.

Acknowlegement. This study has been supported by JSPS KAKENHI Grant Number 15K06262.

References

1. Asadi, B., Vahidi, A.: Predictive cruise control: Utilizing upcoming traffic signal information for improving fuel economy and reducing trip time. IEEE Trans. Control Syst. Technol. **19**(3), 707–714 (2010)
2. Barth, M., Boriboonsomsin, K.: Real-world CO_2 impacts of traffic congestion. Transp. Res. Rec. **2058**, 163–171 (2008)
3. Barth, M., Boriboonsomsin, K.: Energy and emissions impacts of a freeway-based dynamic eco-driving system. Transp. Res. Part D **14**, 400–410 (2009)
4. Brouwer, R.F.T., Stuiver, A., Hof, T., Kroon, L., Pauwelussen, J., Holleman, B.: Personalised feedback and eco-driving: An explorative study. Transp. Res. Part C **58**, 760–771 (2015)
5. Center for Global Environmental Research: National Greenhouse Gas Inventory Report of Japan. Greenhouse Gas Inventory Office of Japan, Ministry of the Environment, Japan (2016)
6. Iwata, Y., Otake, H., Takagi, M.: Results from simulation evaluation of green wave advisory system. In: Proceedings of 19th ITS World Congress, p. 9 (2012)
7. Jamson, S.L., Hibberd, D.L., Jamson, A.H.: Drivers' ability to learn eco-driving skills; effects on fuel efficient and safe driving behavior. Transp. Res. Part C **58**, 657–668 (2015)
8. Jamson, A.H., Hibberd, D.L., Merat, N.: Interface design considerations for an in-vehicle eco-driving assistance system. Transp. Res. Part C **58**, 642–656 (2015)
9. Kadoyu, K., Doi, M., Kanbayashi, K.: Grasping Property of Freight Vehicle Emissions on Real Road. NILIM 2013, National Institute for Land and Infrastructure Management, pp. 106 (2013). in Japanese
10. Kircher, K., Fors, C., Ahlstrom, C.: Continuous versus intermittent presentation of visual eco-driving advice. Transp. Res. Part F **24**, 27–38 (2014)
11. Matsumoto, Y., Oshima, T., Iwamoto, R.: Effect of information provision around signalized intersection on reduction of CO_2 emission from vehicles. Procedia Soc. Behav. Sci. **111**, 1015–1024 (2014)
12. Niu, D., Sun, J.: Eco-driving versus green wave speed guidance for signalized highway traffic: A multi-vehicle driving simulator study. Procedia Soc. Behav. Sci. **96**, 1079–1090 (2013)
13. Oguchi, T., Katakura, M., Taniguchi, M.: Carbon dioxide emission model in actual urban road vehicular traffic conditions. J. Jpn. Soc. Civ. Eng. **695**, 125–136 (2002). in Japanese
14. Rakha, H., Kamalanathsharma, R.K.: Eco-driving at signalized intersections using V2I communication. In: 14th International IEEE Conference on Intelligent Transportation Systems, pp. 341–346 (2011)
15. Rossi, R., Gastaldi, M., Biondi, F., Mulatti, C.: Warning sound to affect perceived speed in approaching roundabouts: Experiments with a driving simulation. Procedia Soc. Behav. Sci. **87**, 269–278 (2013)
16. Staubach, M., Schebitz, N., Köster, F., Kuck, D.: Evaluation of an eco-driving support system. Transp. Res. Part F **27**, 11–21 (2014)
17. Tielert, T., Killat, M., Hartenstein, H., Luz, R., Hausberger, S., Benz, T.: The impact of traffic-light-to-vehicle communication on fuel consumption and emissions. In: Internet of Things (IOT), pp. 1–8. IEEE (2010)
18. Ubiergo, G.A., Jin, W.: Mobility and environment improvement of signalized networks through vehicle-to-infrastructure (V2I) communications. Transp. Res. Part C **68**, 70–82 (2016)
19. Widodo, S., Hasegawa, T., Tsugawa, S.: Vehicle fuel consumption and emission estimation in environment-adaptive driving with or without inter-vehicle communications. In: Proceedings of the IEEE 23 Intelligent Vehicles Symposiums, vol. 2001, pp. 382–386 (2000)

Road Safety

Road Safety

Massimiliano Gastaldi

Department of Civil, Environmental and Architectural Engineering,
University of Padova, via Marzolo, 3, 35131 Padova, Italy
massimiliano.gastaldi@unipd.it

The research world continues to pay serious attention to transportation safety and especially road safety [2]. This is clear from the numbers of journals related to the topic and of papers covering several aspects of road safety: human factors, design of traffic control systems, strategies aimed at optimizing traffic and safety indicators, analysis of accidents aimed at improving the description and explanation of safety phenomena, their relationship with traffic, environmental and weather conditions, and more accurate estimations and forecasts of safety measures for all road users, drivers, cyclists, pedestrians [1, 5].

Although some successes have been achieved in the recent past, notably in developed countries around the world [4], researchers, practitioners and authorities must still face many challenges to bring transportation safety problems under control [3]. Research fosters better road safety strategies for safer mobility.

This special session focuses on some of the main issues and recent advances in road safety research. The good quality standards of the session were ensured by a selective reviewing process, which took place between June and December 2016. Each paper was reviewed by at least two critics and the editorial committee. At the end of the process, of all the contributions selected for publication by Springer, 'Advanced concepts, methodologies and technologies for transportation and logistics' (Springer series of Advances in Intelligent Systems and Computing), four were accepted for the 'Road Safety' session. These papers cover advances in the state-of-the-art related to the following road safety issues.

The chapter by Aron M., Billot R., Bhouri N., El Faouzi N.-E., Seidowsky R. entitled "**Power and Exponential Functions Relating Accidents to Traffic and Rain. Calibration on a French Network**", highlights the importance of estimating the effects of the future values of traffic variables in accidents, in order to assess a new system of traffic management, before its implementation. The main aim of the paper is to establish the relationships which quantify these effects, particularly the link between the occurrence of injury in road accidents, the prevailing traffic conditions, and the occurrence of rain. Based on data for traffic, road accidents and rain collected over a period of one year on a French urban motorway network, a set of safety performance functions were estimated: each provides the accident risk per vehicle-kilometer for a certain type of accident, according to the variable *occurrence of rain* and to the *level of a traffic variable* (average speed, occupancy, etc.). Analyses were carried out separately by lane and for two types of accidents: single-vehicle and multiple-vehicle.

The chapter by Haj-Salem H., Farhi N., Lebacque J.P., Bhouri N. entitled "**Development of Coordinated Ramp-Metering based on Multi-Objective Non-linear Optimization Functions: Traffic and Safety**", focuses on the extension of the OASIS (Optimal Advanced System for Integrated Strategies), to take into account the multi-objective non-linear optimization technique for coordinated ramp metering. This function includes two cost functions: classical traffic performance indices and a new safety index (Risk Model). A brief description of Risk model construction is provided. In order to include the risk index, OASIS was revisited and off-line simulation studies were conducted on a real test site corresponding to the A6W France motorway in the southern part of the 'Ile de France' motorway network. Five consecutive on-ramps were examined as controls. The results are very promising.

Bella F. and Silvestri M. in the chapter "**Drivers' Behavior at Bicycle Crossroads**" report the outcomes of a driving simulator study to evaluate the effectiveness of two countermeasures at bicycle crossroads in order to improve cyclist safety: colored pavement markings and raised islands. The mean speed profiles and statistical analysis of the driver's speed along the sections approaching the bicycle crossings showed that driver behavior was positively affected by the peculiarities of the two countermeasures, which determined better yielding behavior. The subjective assessment of participants revealed that colored pavement markings were considered the most effective in terms of driving aids.

Starting from the consideration that automatic incident detection methods for basic freeway segments are widely reported in the literature, but their application to freeway ramp merging zones is limited, the final chapter entitled "**Automatic Incident Detection On Freeway Ramp Junctions. A Fuzzy Logic-Based System Using Loop Detector Data**" by Rossi R., Gastaldi M., Gecchele G., introduces a control system which can identify incidents from vehicle loop detector data on such zones. The system was developed with fuzzy logic concepts and calibrated with data from micro-simulation experiments. The main finding of this study is that the detection system shows excellent False Alarm Rates (FAR) and satisfactory Detection Rates (DR) and Mean Time To Detection (MTTD), generally better than those obtained with the traditional California#7 comparative algorithm.

I would like to thank all the authors for submitting good-quality papers, all the reviewers for their careful work, and the Guest Editors of the book - Jacek Żak, Yuval Hadas and Riccardo Rossi - for inviting me to be coordinator for the 'Road Safety' session.

References

1. Elvik, R., Høye, A., Vaa, T., Sørensen, M.: The Handbook of Road Safety Measures, 2nd ed. Emerald Group Publishing Limited, Bingley (2009)
2. Hagenzieker, M.P., Commandeur J.J.F., Bijleveld F.D.: The history of road safety research: a quantitative approach. Transp. Res. Part F: Traffic Psychol. Behav. **25**,150–162 (2014)

3. Hakkert, A.S., Gitelman, V.: Thinking about the history of road safety research: past achievements and future challenges. Transp. Res. Part F: Traffic Psychol. Behav. **25**, 137–149 (2014)
4. World Health Organization: Global Status Report on Road Safety (2015)
5. Young, W., Sobhani, A., Lenn, M.G., Sarvi, M.: Simulation of safety: A review of the state of the art in road safety simulation modeling. Accid. Anal. Prev. **66**, 89–103 (2014)

Power and Exponential Functions Relating Accidents to Traffic and Rain. Calibration on a French Network

Maurice Aron[1(✉)], Romain Billot[2], Neïla Bhouri[1], Nour-Eddin El Faouzi[3], and Régine Seidowsky[1]

[1] Université Paris-Est, IFSTTAR, COSYS-GRETTIA, Marne-la-Vallée, France
mauraron@gmail.com,
{neila.bhouri,regine.seidowsky}@ifsttar.fr
[2] Telecom-Bretagne, Plouzané, France
romain.billot@telecom-bretagne.eu
[3] Université Lyon, ENTPE, IFSTTAR, LICIT UMR_T9401, Lyon, France
nour-eddin.el-faouzi@ifsttar.fr

Abstract. Relations between the occurrence of road accidents, traffic and rainfall conditions are valuable in setting safety objectives for traffic management, and in assessing the safety impacts of new traffic management systems, prior to their implementation. Based on traffic, road accidents and rain data collected over one year, on a French urban motorway network, a set of safety performance functions were estimated; each of them provides the accident risk per vehicle-kilometer for a certain type of accident, according to the occurrence of rain, and to the level of a traffic variable (average speed, occupancy, percentage of tailgating…). Analyses were carried out separately by lane and for two types of accidents: single-vehicle accidents and multiple-vehicle accidents. The relationships, although statistically significant, have yet to be validated by the treatment of another set of accidents.

Keywords: Traffic data · Accident · Rain · Surrogate data · Traffic indicators · Urban motorway · Risk · Safety performance function · Logistic regression, power model

1 Introduction

In order to assess a new traffic management, before its implementation, it is necessary to assess the impact of the future values of the traffic variables on accidents. The aim of this paper is to establish the relationships which quantified this impact. For safety reasons, drivers adapt their speed, relative speed, time gap and lane, according to the infrastructure (bends, slopes, intersections), traffic conditions (speed of the vehicle ahead or on the adjacent lane, density) and weather conditions. It must be remembered that the performance of vehicles and drivers decreases on slopes and bends, and some danger may come from close vehicles. Despite this adjustment, the accident rate remains related to infrastructure, [17], weather [5, 6], and traffic conditions [9].

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_15

Relations between traffic conditions, infrastructure elements and accidents are discrete or continuous. In discrete relationships, the risk per vehicle-km is computed by traffic flow regime, a traffic flow regime being a homogeneous group of traffic flow conditions, for different weather and infrastructure conditions. Golob *et al.* [9] found different accident rates according to the type of crash and traffic conditions, to the temporal variations in volume and speed; Abdel-Aty *et al.* [2] used traffic conditions and rain occurrence as accident precursors; [14] highlighted the impact of speed variation on accidents.

In continuous relationships, the rate of accident is a continuous function of the traffic variable. According to Nilsson or Elvik, the risk is a continuous power function of the speed [7, 13]; Hauer and Elvik also proposed an exponential function [8, 11]. When appropriate, continuous relationships give a quick understanding of the risk and might be included in simulations or in traffic management algorithms.

The objective of this paper is to model the relationship between the risk of accidents and the traffic, according to the findings of Golob *et al.* [9] or Abdel-Aty *et al.* [2], who identified, among other things, the ways in which congestion affects road safety. Our approach is also in the continuity of Nilsonn, Elvik, Hauer [7, 8, 11, 13], focusing on continuous relationships, and aims to demonstrate how speed, density and some other traffic variables are related to the occurrence of accidents.

The role of speed in road safety has been demonstrated. “Speed” refers to different quantities: the speed limit at a national level, on a network, on a particular section; speeds of individual drivers recorded at particular points, or their distribution on a route; average speed on a spatial range; temporal average at a given point, by lane, or for all lanes. Depending on what “speed” is, the analytical pattern, the numerical values, and the relevance of models vary. In this paper, we consider the 6-minute average speed observed. Nilsonn as well as Elvik and Hauer models have been tested on some French interurban motorways; they have been adapted to take into account the different motorway lanes and rain. Other models have also been considered.

In the following, we present the data in Sect. 2, followed by some continuous models, which relate the risk to different traffic indicators, in Sect. 3. Section 4 describes the relationships which have proved to be significant, while their limits are discussed in Sect. 5. In Sect. 6, the relationships between the obtained risk models are developed further. Section 7 contains the conclusion and perspectives; main numerical results are to be found in the Appendix.

2 Traffic, Accident and Meteorological Data

Meteorological Data. The occurrence of rain, at the time and place of the accident, is recorded in the accident database; in the case of no-accident, the rain occurrence is provided, every six minutes, at the meteorological station of Marignane, Marseille airport, less than 30 km from all points of the network. This station is managed by Météo-France, the French Agency in charge of weather forecasts. During rain, the percentage of injury accidents or fatalities (13%) is greater than the percentage of vehicle-kilometers travelled (5.3%); this confirms the danger due to rain.

Traffic Data. Between June 2009 and May 2010, the French centre for studies on risk, mobility and environment, CEREMA, collected traffic data (vehicle lengths, speeds) on the "Marius" network. This network is 150 km long and is made up of the urban parts of motorways A7, A50, A51 and A55 around Marseille (Fig. 1).



Fig. 1. The Marius urban motorway network, near Marseille (France)

The sections have either two lanes per direction (here called middle and slow lanes) or three lanes (fast, middle and slow lanes); 104 available traffic stations by direction, (one station every 750 m) are available on the main carriageway and on the ramps. Data are recorded every six minutes; the whole one-year database was used for calibrating the traffic-safety relationships. Given that there is not much missing data, the traffic pattern based on available data is assumed to be representative. A weather station is located in Marignane airport, in the North-West.

The traffic counts and the distance between sensors lead to an estimated 1.5 billion vehicle-kilometers. 5.3% vehicles travel during rain, and 15% travel at night. Although defining night as being from sunset to sunrise would be a more accurate reflection of the contribution of darkness to the occurrence of accidents, here night is always defined as being from 8 pm to 6 am, which brings some homogeneity for the types of travel and driver, and for their tiredness.

Accident data. The French police collect the characteristics of all road fatalities and injury accidents. A few characteristics are used here: date, hour, minute, precise location, number of vehicles involved, rain information.

Over one year, 292 injury accidents or fatalities were recorded on the Marius network. Missing data affects 18 accidents for which the direction of the accident is missing, or its location, or the traffic data shortly before. Table 1 gives the distributions of the remaining 274 accidents, according to rain and presence of a PTW.

Relationships have been estimated by lane because averaging the traffic indicators between lanes may hide certain phenomena such as heterogeneous speeds or densities between lanes. Although the lane where an accident begins is generally unknown, the accidents occurring when inserting from or to a ramp are mentioned in the database.

Table 1. Distribution of accidents, according to the presence of rain and PTW

	Rain	No_Rain	Total	PTW	No PTW
Total	36	238	**274**	52	222
Night time	9	47	**56**	8	48
Daytime	27	191	**218**	44	174
Daytime, rain confirmed(*)	16	187	203	42	161
Daytime, rain not confirmed(*)	**11**	**4**	15	2	13

(*) The rain information considered here comes from the police report; however, the meteorological station does not confirm rain information in 11 cases by rain, and in 4 cases by no-rain (daytime).

These accidents are not included in the estimations of the relationships for the middle and fast lanes. For every accident, traffic data from the first upstream sensor, when available, are considered; when unavailable, traffic data at the first downstream sensor, or at the second upstream or downstream sensors, are considered. Traffic conditions may change at the moment of the accident; also it is mandatory, when estimating a relationship between accidents and traffic conditions, to use traffic conditions before the accident. As accident times are only estimated by the police on their arrival on site afterwards, for every accident we examined the series of speeds recorded at the upstream sensor for forty minutes until the time estimated by the police. When traffic conditions did not change, we considered the 6-minute period ending before the accident time (taking into account a time offset equal to the average travel time between the sensor and the accident location); when one single drop in speed occurred, we considered the 6-minute period ending before the drop; when several drops were recorded, we selected the period ending eighteen minutes before the accident (eighteen minutes was found to be the time-lag for which the sensitivity analysis conducted in [4] provides the highest correlation between speed and single vehicle accidents).

It is at night that 20% of all accidents occur, but for only 15% of the vehicle-kms traveled. This means an increase in risk at night.

Types of accident. Relationships between traffic and accidents have been calibrated by type of accident. Two types of accident have been considered: accidents implying a single vehicle, and accidents implying multiple vehicles.

Table 2. Number of accidents with available traffic data by lane and type of accident

	Single vehicle	Multiple vehicles	Total
Slow lane	46	147	193
Slow lane(*)	46	137	183
Middle lane(**)	44	128	172
Fast lane(***)	39	117	156

(*) Excluding accidents on ramps.
(**) There are fewer accidents on the middle lane because of missing traffic data.
(***) There are fewer accidents counted on the fast lane because there is no “fast” lane for accidents occurring on a two-lane motorway section.

25 out of the 218 daytime accidents are linked either to a breakdown of a vehicle, or to the driver (drowsiness, alcohol…) or to the presence of a pedestrian; these accidents are probably not linked to the traffic conditions, so they have been discarded from the analysis. The distribution of the 193 remaining accidents according to the type of accident is given in Table 2.

3 Statistical Models and Traffic Indicators

Safety Performance Indicators are generally based on negative binomial models, or on distributions like the Poisson-Maxwell-Conway distribution which better fit the dispersion [12]. The risk "*R*" of accident by vehicle-kilometer is related in our approach to different variables within a logistic regression. Six types of relationships have been tested on thirteen variables (traffic indicators).

3.1 Six Models

The name of a model is constituted by a part indicating its pattern ("*POW*" for the power function, "*EXP*" for the exponential function,..) and by subscripts indicating, if relevant, that the sets of accidents and traffic conditions have been restricted for the estimation of the model: The subscript "N" (for no-rain) indicates that accidents and traffic conditions during rain have been excluded. The subscript "R" (for "Ramp") indicates, when the model is estimated on the slow lane, that accidents occurring near a ramp have been included. The subscript "*" indicates that accidents implying a PTW have been excluded.

For a given weather condition (rain or no-rain), and a given type of accident, the six-minute periods are grouped by average speed intervals: the speed interval for a group "*i*" of periods is such that there is at least one accident of the given type occurring during the six-minute periods associated to this group.

1. In the power model, which is generally applied to speed only, risk is proportional to an exponent of the traffic indicator.

$$(POWER) : R_i = \alpha_V V_i^{\beta + \gamma_{Rain}.Rain_i} + \varepsilon_i. \text{ for } i = 1 \ldots n \quad (1)$$

   ε_i is the deviation, assumed to be Gaussian-distributed.
   R_i is the risk by vehicle-km, V_i is the average speed for the group i of periods; $Rain_i = 1$ if rain occurs during each period of this group (0 otherwise); the number of accidents (for a given type of accident) and the number of vehicle-kms are associated to this group.
   α_v, β and γ_{Rain} are deterministic parameters to be estimated. n is the number of groups of periods. This type of model applies also to other relationships, where V_i is replaced by the traffic indicator of another variable, with an analogous process for forming the groups of periods.
2. We introduce "logistic" power models, linking the logit of the risk $\text{Log}[R_i/(1 - R_i)]$ to the indicator (V) and to the occurrence of rain. Computing confidence intervals

on parameters requires an assumption on the distribution of deviations - here assumed to be Gaussian. The second model is written as follows, using α as the logarithm of α_v:

$$(POW): \ \mathrm{Log}[R_i/(1-R_i)] = \alpha + \beta.\mathrm{Log}(V_i) + \gamma_{Rain}.Rain_i + \varepsilon_i. \tag{2}$$

The number of accidents during rain being low, we have also proposed a simplified model, noted (POW_{N}), without the rain coefficient.

3. Hauer [11] and Elvik [8] proposed an exponential model; we add to this model a term which models the rain impact:

$$(EXPO): R_i = e^{\alpha + \beta V_i + \gamma_{Rain}.Rain_i} + \varepsilon_i \tag{3}$$

4. The associated logistic model reads:

$$(EXP): \ \mathrm{Log}[R_i/(1-R_i)] = \alpha + \beta.\, V_i + \gamma_{Rain}.Rain_i + \varepsilon_i. \tag{4}$$

5. (V_i) being positive, its square is an increasing function, and replaces (V_i) without changing the sign of β in model $\left(EXP^2\right)$:

$$\left(EXP^2\right): \ \mathrm{Log}[R_i/(1-R_i)] = \alpha + \beta.\, V_i^2 + \gamma_{Rain}.Rain_i + \varepsilon_i. \tag{5}$$

6. The function $\mathrm{Log}^2(V_i)$ has been successfully tested for some indicators (the percentages); it is a decreasing function when V_i is less than 1; this would imply a change of interpretation of the sign of β, unless considering $\left(-\mathrm{Log}^2(V_i)\right)$ as follows:

$$(LOG^2): \mathrm{Log}\left[\frac{R_i}{1-R_i}\right] = \alpha - \beta.Log^2(V_i) + \gamma_{Rain}.Rain_i + \varepsilon_i \tag{6}$$

7. Parabolic models (excluding rainy conditions): When risk is not monotonous with the traffic indicator, modeling requires one more parameter. However, due to an insufficient number of accidents, it was not possible to estimate four parameters; models are therefore estimated here on datasets excluding rain, so the rain coefficient can be removed; "γ" designs the new coefficient – the coefficient of the square of the indicator in parabolic models, which takes into account the traffic indicator and its square; the direction of variation of the risk depends on whether the traffic indicator V is below/above the value $-\beta/(2.\gamma)$.
The (EXP_{NP}) model comes from the exponential model:

$$(EXP_{\mathrm{NP}}): \ \mathrm{Log}[R_i/(1-R_i)] = \alpha + \beta.\, V_i + \gamma\, V_i^2 + \varepsilon_i \tag{7}$$

where the subscript "P" (for Parabolic) indicates a parabolic term.

8. The (MIX) model combines a part coming from a power model, and a parabolic part coming from an exponential model:

$$(MIX_{\mathrm{NP}}):\ \mathrm{Log}[R_i/(1-R_i)] = \alpha + \beta.\mathrm{Log}(V_i) + \gamma\, V_i^2 + \varepsilon_i. \tag{8}$$

3.2 Statistical Analysis

The traffic and accident databases were separated into different cases, according to the time of day (night time/daytime), lane, weather (rain or not). Independent analyses were performed for every combination of cases.

Three accident datasets were considered: the whole dataset, or all accidents except those during rainy conditions, (disregarding the traffic data when raining), or all accidents except those involving Power-Two-Wheelers (PTW).

The logistic regression Generalized Linear Model (GLM) was used with the software R ®; it processes the vectors of number of accidents $(Acc_i,)$ and number of vehicle-kilometers (N_i) by group of periods.

3.3 Thirteen Traffic Indicators

Even with individual data on upstream sensors, it is impossible to identify, among others, the driver responsible for the accident. What we wanted to know was whether the parameters of the whole set of drivers are different just before an accident. The thirteen indicators presented here are computed from the aggregation of traffic data over 6-minute periods:

1. Average speed V_i, by 6-minute period in km/h. We use here the arithmetic speed average (time mean speed).
2. Occupancy – it is the percentage obtained by summing the "occupancy times" (in seconds) of vehicles passing in a 360-second period, and then by dividing the sum by 360; the occupancy time of vehicle j of length L_j and speed V^j is equal to $3.6\left(L_j+\lambda\right)/V^j, \lambda$ being equal to 1 m, the length of the inductive loop; the unit factor is "3.6".
3. Relative speed in km/h ("RelSpeed") is the difference between the speeds of two consecutive vehicles on the same lane. The sum of relative speeds on a period is of no interest, because the speed of a vehicle generally appears twice in the sum with opposite signs, and thus disappears. Since negative relative speeds have no safety impact, they were disregarded. The indicator proposed here is the sum of relative speeds, when positive, divided by the traffic count.
4. Indicators 4, 5, 6, and 7: Time headway is here the difference between the arrival times at the sensor of the fronts of two successive vehicles. Indicator 4 is the 6-minute average time headway ("Average TH"); indicators 5–7 are the percentages of tail-gating (less than 2 s, "%TH <2"), short (less than 1 s, "%TH <1") and very short headways (less than 0.5 s, "%TH <0.5").
5. Indicators 8, 9, 10, and 11: Uno defined the "*PICUD*" (Potential Index for Collision with Urgent Deceleration) [16]. It is the estimated difference (in meters) between the

stopping locations of vehicles "j" and "$j-1$", assuming that the leader $j-1$ brakes at the very instant t^j when the follower "j" passes the sensor. The leader is then located at $V^{j-1}(t^j - t^{j-1})$ meters downstream the sensor; its rear end is located at $V^{j-1}(t^j - t^{j-1}) - L^{j-1}$ meters before, assuming that the traffic sensor records times of passage of the front of vehicles. When the *PICUD* is negative, there is a collision danger, which would have been avoided, if the follower had a space gap greater by $\{-PICUD\}$ before the braking of the leader. The follower brakes with the same deceleration (here the deceleration is $\Gamma = 6.25$ m/s^2) at time $t^j + T$, after a reaction time $T = 1$ s, when the vehicle is located at $V^j.T$ meters downstream the sensor:

$$PICUD^j = \frac{(V^{j-1})^2 - (V^j)^2}{2\Gamma} - T.V^j + \left(t^j - t^{j-1}\right).V^{j-1} - L^{j-1} \tag{9}$$

The length L^j of the second vehicle replaces the length L^{j-1} of the leader in Eq. (9) when the sensor records times of passage of the rear end of vehicles.

Indicator 8 is the absolute value of the sum of negative $PICUD^j$, over the 6-minute period, divided by the traffic count of the period.

Indicators 9, 10, and 11 are the percentages of drivers for which the *PICUD* is below a threshold (respectively 0 m ("*%PIC*<0"), minus 10 m ("*%PIC*<−10"), minus 20 m ("*%PIC*<−20").

6. Indicators 12 and13: *PICUD* becomes "*PICUDbis*", by removing the reaction time in Eq. (9).

   Indicator 12 ("*PICUDBis*") is the sum of negative $PICUDbis^j$, divided by the traffic count.

   Indicator 13 is the percentage of negative *PICUDbis* ("*%PICBis*<0").

4 Significant Relationships

The risk of single-vehicle and multi-vehicle accidents is related to average speed in Sect. 4.1, to occupancy in Sect. 4.2, to time headway in Sect. 4.3, to relative speed in Sect. 4.4 and to *PICUD* indicators in Sect. 4.5. The logistic form of the models has been preferred. Results are presented by lane, with or without the impact of rain, only for daytime accidents and traffic. Numerical values of the parameters of logistic regressions are presented in the tables given in the Appendix. The lines of these tables are numbered for easy reference: the lines whose numbers end with the suffix "−1" represent the risk of single-vehicle accidents, while suffix "−2" is related to multiple-vehicle accidents. The relationships with a P-value lower than 5% are generally considered as significant. However, as the number of accidents in our database is not very large, and in the hope of identifying useful relationships, we also considered relationships with a P-value threshold of 10%; these relationships will have to be confirmed on a larger database. In the tables, for each relationship, the first line gives the number of the relationship, the type of model, the values of the two or three

parameters, the Null deviance D_0, and the AIC criterion; the second line gives the name of the traffic indicator, the standard deviations "(σ)" of the coefficients, the P value (P %) rounded to zero when less than 0.1; the residual deviance D_R and the degree of freedom #Fre.

4.1 Risk and 6-Minute Average Speed (Daytime)

In normal weather conditions, there are relationships between single-vehicle accidents and speed (see the first subsection), between multiple-vehicle accidents and speed (the second subsection); for the whole set of weather conditions, there are relationships between single-vehicle accidents, speed and rain (the third subsection), We did not find any significant relationship between multiple-vehicle accidents, speed and rain, likely because, during rain, there is a concomitant (slight) decrease in speed and an increase in risk.

Models with a rain term, single-vehicle accidents. The risk of single-vehicle accidents (daytime) is significantly related to the 6-minute average speed of the fast lane by power and exponential relationships. Rain is a significant contributing factor. When Power Two Wheeler accidents are excluded (model *POW**), the exponent β is very high ($\beta = 7.86, \sigma_\beta = 2.2$) with 52% of the deviance explained. The rain coefficient $\gamma_{rain} = 1.88$ (its standard deviation σ_γ being 0.37); this corresponds to multiply the risk, in case of rain, by $e^{1.88} = 6.5$. When including PTW, the speed exponent decreases to $\beta = 2.66\,(\sigma_\beta = 1.3)$ with 38% of deviance explained, see Table 4 in the Appendix, lines 2-1 and 6-1; the decrease of β means that PTW accidents are not particularly correlated to a high observed average speed. We imagine that when the average speed is low or moderate, some PTW drive between the lanes with a higher speed, but this speed has no impact on the average speed because the magnetic loop of the traffic sensor is not implemented too close to the adjacent lane.

The exponential model (*EXP*) gives a speed exponent $\beta = 0.03(\sigma_\beta = 0.014)$ which passes to $0.073(\sigma_\beta = 0.020$ when PTW accidents are excluded (model *EXP**), see Table 4 in the Appendix, lines 1-1 and 7-1.

On the middle lane, no logistic model appears significant. However, the second lane (from the right) of both two-lane sections and three-lane sections are grouped in this work as the "middle" lane, which brings some heterogeneity.

On the slow lane, no significant relationship was found. Speed is less homogeneous on this lane, due to the presence of trucks and ramps. The speed average, an average of inhomogeneous quantities, is not a good indicator.

Normal weather conditions, single-vehicle accidents. In normal weather conditions, i.e. when excluding traffic and accidents during rain, the percentage of explained deviance is only due to the traffic indicator, and is generally smaller than when the rain impact is added.

On the fast lane the model, the power model remains significant only when PTW are excluded (POW_{N*}), with a percentage of explained deviance of 26%, instead of 52% in (POW_*), and a slightly smaller exponent β, (Table 4, line 3_1).

Exponential models are also significant with smaller β, as well as with and without PTW (Models (EXP_N) and (EXP_{N*}) lines 5-1 & 8-1. Since many PTW postpone their trip

when it rains, which decreases their risk exposure during rain, the comparison of risks between rainy and non-rainy conditions is fully justified only when excluding PTW.

Multi-vehicle accidents are related to the speed with negative β coefficients (Table 4, lines 9-2 to 14-2). This might mean that the presence of a high number of low-speed vehicles (trucks, entering vehicles) brings some danger. But it might just mean that, when speed decreases, accidents are more likely to result from crashes between vehicles than to be single-vehicle accidents.

Three tentative conclusions. First, the average speed is a better indicator when excluding PTW. Second, exponential models are robust. Third, speed must be estimated on the fast lane.

4.2 Risk and 6-Minute Occupancy, Daytime, Multiple-Vehicle Accidents

Positive coefficient β are obtained, which means that the higher the occupancy, the higher the rate of multiple-vehicle accidents.

Including accidents related to a ramp, on the slow lane, model (EXP^2), with a rain coefficient $\gamma_{rain} = 0.5$ explains a limited percentage (20%) of the deviance of the logit of the risk of multiple-vehicle accidents (Table 5, line 26-2 in the Appendix).

Excluding rainy conditions, model (POWN) explains 36% of the deviance on the middle and slow lanes, and 10% only on the fast lane; on the fast lane, excluding rainy conditions and PTW, model (EXPN*) explains 23% of the deviance (Table 5, lines 25-2, 27-2, 28-2) - see also the comparison of the models, Fig. 5.

However, the parabolic model EXPNP (with $\beta > 0$ and $\gamma < 0$) indicates that when occupancy is very high (greater than 32%), the risk decreases on the slow lane (relation 29-2, Table 5); it is the same for the middle lane (relation not included in Table 5).

4.3 Average Time Headway and Percentages of Short Time Headway

Average time headway (TH). Average time headway is proportional to the inverse of the traffic count; thus a high TH (a low traffic flow) occurs either when there is a low traffic demand, or when there is congestion. The first case, which is the most frequent, implies that, for a high TH, when vehicles are far from each other, the higher the TH, the lower the rate of multi-vehicle accidents. Negative β and positive γ_{rain} have effectively been obtained on the middle lane, with a very limited percentage of explained deviance (Table 6, lines 31-2 and 32-2). At the same time, the rate of single-vehicle accidents is higher. A positive β explaining the single-vehicle accidents is thus expected and obtained on the three lanes for all models. Table 6, line 30-1 gives the results of the model (*POW*) for the slow lane; the coefficients are similar for other lanes and the percentage of explained deviance varies from 39% to 50% according to the lane.

Percentage of short time headway. The higher the percentage of short headway, the lower the single-vehicle accident rate. β exponents are negative (Table 6, line 33-1); when including rain, 38% to 68% of the deviance, according to the lane, are explained.

Relationships between risk of *multiple-vehicle* accidents and the three percentages of time headway are either insignificant, or explain a very small part of the deviance, except when a parabolic term is added. Excluding rainy conditions, parabolic models such as (EXP_{NP}) or (MIX_{NP}) (with $\beta < 0$ and $\gamma > 0$) indicate, for all lanes and for the three percentages, that the risk of *multiple-vehicle* accidents is high, both for low and high percentages of short time headway, (Table 6, lines 34-2 to 42-2).

4.4 Risk and Sum of Positive Relative Speeds (Divided by the Traffic Count)

Abdel-Aty *et al.* found that the speed variance (obtained from a series of ten consecutive 30″ average speeds) was a relevant accident precursor [1, 3]. This means that changes in 30″ traffic conditions are correlated to accident occurrence. This does not mean that heterogeneous speeds (between two consecutive vehicles on the same lane) are correlated to accidents and that more homogeneous speeds (as obtained with an adaptive speed control device) would be safer; although we tried to check this point with our relative speed indicator, which measures this speed heterogeneity, on our limited dataset we found little evidence on this point.

The risk of single-vehicle accidents increases with the sum of positive relative speeds (divided by the traffic count). In all lanes the β coefficients of the power model are positive: 27% to 64% of the deviance is explained, depending on the lane. The numerical values of the parameters are given here only for the fast lane (Table 6, line 43-1). This might come from correlations between relative speed and traffic density, and between traffic density and accident type. Symmetrically, with the model (*POW*), the risk of multiple-vehicle crashes seems to decrease (β negative) when the indicator increases; 33.6% and 45% of the deviance of the *multiple-vehicle* accident risk are explained on the middle and the slow lanes (Table 6, lines 44-2 and 45-2. However, with parabolic models MIX_{NP} or EXP_{NP}, the parabolic coefficients γ are positive, implying an increase in risk on the middle and slow lanes when the relative speed indicator is sufficiently high (above 1.5 km/h) (Table 6, lines 46-2 and 47-2).

4.5 *PICUD*-Based Indicators

Single-vehicle accident occurrence is related, with positive γ_{rain} and β coefficients, to *PICUD*-based indicators:

- the absolute value of the average of negative *PICUDbis* on the fast lane (line 48-1),
- the percentage of *PICUD* less than minus 20 m (fast lane, without Power–Two-Wheelers, line 49-1),
- the percentages of negative *PICUDbis* on the fast and middle lane (lines 50-1 and 51-1).

Multi-vehicle accident occurrence is related to rain with a positive coefficient and, with a negative β, to the percentages of *PICUD* less than minus 10 or 20 m (Table 6, lines 52-2 and 53-2); however these relationships are obtained only when including accidents related to the ramp entrance/exit (which was tested on the slow lane only); no significant relationship appears between multi-vehicle accident occurrence and the averages of negative *PICUD* or *PICUDbis*.

When rain is excluded, the rate of multi-vehicle accidents is related to the percentages of *PICUD* less than minus 10 m or minus 20 m. However, these relationships are obtained:

- on the fast lane when excluding PTW (model $LOG^2_{N^*}$), with a small percentage of deviance explained (line 54-2); the risk increases with the percentage $(\beta > 0)$,
- on the slow and the middle lanes, when adding a parabolic term (EXP_{NP}, lines 55-2 and 56-2); the risk increases with high percentages $(\gamma > 0)$.

The percentage of negative *PICUD*, when excluding rainy conditions, is related to the risk of multiple-vehicle accidents by the parabolic model EXP_{NP}, with a positive coefficient γ, which shows that risk increases with high percentages of negative *PICUD*. However, the percentage of explained deviance (26%) is limited (Table 6 line 57-2).

5 Discussion

Accidents are due to an inappropriate speed, relative speed or lane change at an individual level; however, at an aggregated level, a correlation between risk and an aggregated traffic variable might come via another correlated variable. Different interpretations/misinterpretations have to be considered. They are discussed here only in the case of speed:

1. Continuity between individual and aggregated traffic values: a high average speed results from many risky drivers with high individual speeds: it is likely, but not certain, that a high average speed results in a higher risk.
2. A good correlation does not mean causality. An accident due to a high density also appears with a low speed, because high density is correlated to a low speed, but this is misleading.
3. Misinterpretation. “A low speed” can be inappropriate if speed is not sufficiently low. During rain, the speed decreases, but too little, hence the risk increases. If the presence of rain is not identified, a misleading negative correlation between risk and speed would appear. All contributing factors have to be identified, and then introduced.
4. Accounting for the many contributing factors is difficult, with only a limited number of accidents. Accidents occur due to multiple factors: presence of rain, of a curve, of an access ramp, etc. Using the speed average may be inadequate, due to the sensitivity of the extremes ends of the distribution (the risky drivers). If all but the risky drivers respect the speed limit, this will imply a negative correlation between *average* speed and risk.

5. Classifying accidents as single or multiple-vehicle accidents: even if speed had no effect on risk, speed would have a positive correlation with single-vehicle accident risk and a negative one with multiple-vehicle accident risk: this is due to the negative correlation between speed and density, combined with the correlation between density and type of accident. When traffic density is low, few vehicles are close to another vehicle, and so more accidents are in fact "single-vehicle" accidents.
6. Non monotonous relationships. Multiple-vehicle accidents increase with traffic density, but for very high density, some crashes only lead to material damage (excluded here). The pattern of the relationship linking the risk to the traffic indicator is not monotonous when several basic phenomena occur. Progress on that topic should be possible by refining the analysis on the basis of more accident data, thus allowing a more disaggregated level.

A lack of correlation may come from the absence of any relationship, from confounding various opposite effects, or from insufficient data; a lack of linear correlation may also come from a non linear relationship.

6 Deepening the Relations Between the Relationships

In order better to understand the relevance of the models, the occurrence of accidents and the quality of data and data processing, three points need to be highlighted:

- Both the power and the exponential functions model the relationships between accident risk and speed or occupancy. How is this possible, given that the functions are different?
- What are the relationships between the total accident risk (sum of risks for single- and multiple-vehicle accidents) and speed?
- Combining the relationship between speed and risk with the relationship between risk and occupancy gives a speed-occupancy relationship. How credible is this new relationship, which comes under the category of speed-density relationships? The more incredible the new relationship, the more incredible the two risk models which have been combined.

6.1 Models Relating the Accident Risk and Speed

In the exponential model, the risk "R" is linked to the speed by a negative exponential function; this implies that the logarithm of the risk is a linear function of the speed.

Here, a linear relationship between the logit of the risk, defined as $\mathrm{Log}[R/(1-R)]$, and the speed has been calibrated within a logistic regression. As $(1-R)$ is very close to 1, the logit of the risk is numerically very close to the logarithm of the risk. Also, in Figs. 2, 3, 4 and 5, the Y-axis, instead of being entitled "logit of the risk" has been entitled "Logarithm of the risk" for the sake of readability. In these figures, the exponential model appears as a straight line, whereas the power model appears as a logarithm function.

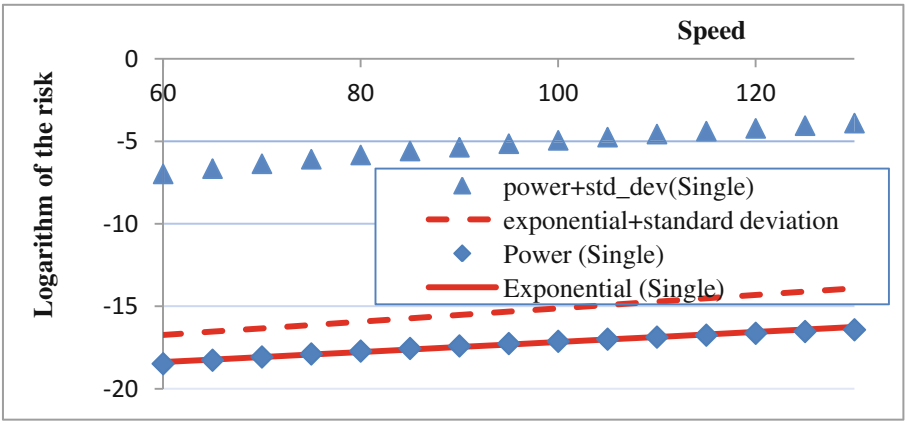


Fig. 2. Daytime risk of single-vehicle accident related to speed: Power & exponential models (fast lane)

Daytime accident risk and single-vehicle accidents. Figure 2 gives the modeled risks for single-vehicle accidents related to speed (daytime, fast lane, average 6-minute speeds between 60 and 130 km/h). The power model (represented by a diamond without any solid line, (parameters from Table 4, relation 2-1) is impressively close to the exponential model (the solid line) (parameters from Table 4, relation 1-1).

In the case of rain - not displayed here - both models remain very close together: 1.6886 or 1.69669 need to be added to the logit of the risk for the power model and for the exponential model, respectively.

The curves "exponential+Standard Deviation (Single)" and "power+Std_Dev (Single)" have been built by adding one standard deviation to parameters α and β of the models; the first of these two curves is closer to the curve displaying the risk than the curve "power+Std_Dev (Single)". This shows that the exponential model has a lower standard deviation than the power model and indicates that the exponential model should be chosen.

Accident risk and multiple-vehicle accidents. As already stated in Sect. 4.1, the relationship between the average speed and risk is closer when excluding accidents implying PTW, perhaps because the speed of some PTW, which weave in and out between two lanes, is broadly independent of the recorded average speed; in addition, as the number of accidents in our database is limited, excluding rain decreases the number of parameters of the models and in turn makes the relationship easier to identify.

In the cases where there is no rain and no Power-Two-Wheelers are involved, Fig. 3 gives the logit of the modeled risk for single-vehicle accidents and for multiple-vehicle accidents, assuming either a power model or an exponential model.

These curves are based on relations 3-1, 8-1,13-2,14-2 of Table 4*, Appendix.*

The exponential model (dashed straight line) and the power model (stars) also coincide when considering the multiple-vehicle accidents. Both risks decrease with

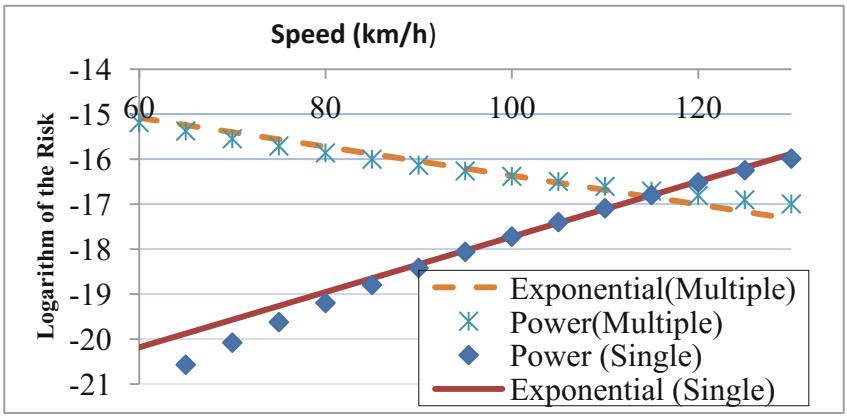


Fig. 3. Daytime risk related to the speed on the fast lane for single and multiple-vehicle accidents, power and exponential models; rain and PTW excluded.

speed. Although we have no real proof, this decrease is probably not really due to speed, but to the combination of two factors:

1. Correlations between high speed and low density,
2. With low density traffic there are few vehicles on the road; accidents generally concern a single vehicle thus decreasing the occurrence of multiple-vehicle crashes.

Total accident risk and speed. The total daytime risk of accidents (single- or multiple-vehicle) is less sensitive to speed, because of a balance between the increase and decrease described previously.

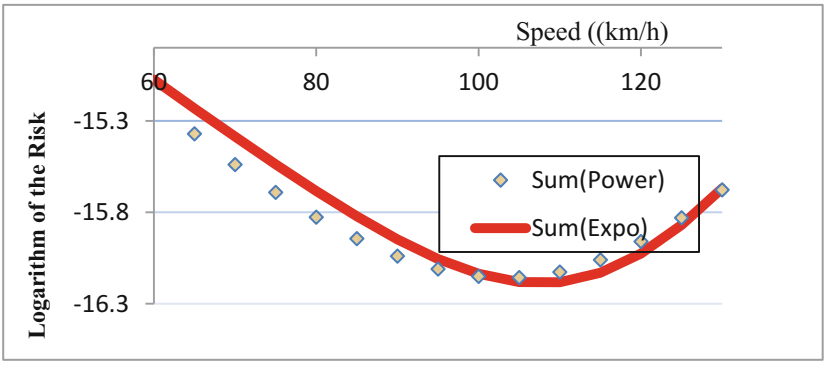


Fig. 4. Total daytime risk (single- and multiple-vehicle accident) related to the speed on the fast lane, power and exponential models; rain and PTW excluded.

The total risk at low speeds is high, probably due to congestion. The total risk decreases while speed increases up to 110 km/h and then increases for higher speeds. By definition, at least two vehicles are involved in a multiple-vehicle accident, whereas just one vehicle is involved in a single-vehicle accident[1]. It would be interesting to consider the risk by vehicle, which would be obtained by weighting each accident by the number of vehicles involved in the accident. Multiple-vehicle accidents would be weighted by a weight of at least two and would have a greater impact on the total risk.

6.2 Models Relating the Accident Risk and Occupancy

The power and exponential models have also been used to model the relationship between accident risk and occupancy.

In the case of no rain, both the power and exponential models relate significantly both risks (single- and multiple-vehicle accidents) and occupancy on every lane. Their numerical parameters are given in Table 5; these models are displayed for the fast lane on Fig. 5, from the relations given Table 5 lines 17-2, 18-2, 21-1, 23-1).

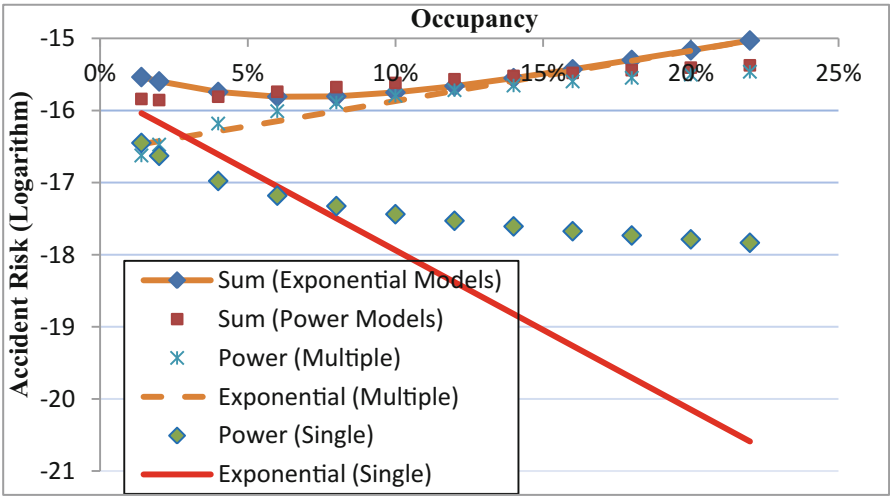


Fig. 5. Risks of single- and multiple-vehicle accidents, and total risk related to the occupancy of the fast lane, power and exponential models, daytime, no rain.

The risk of single-vehicle accidents *decreases* when occupancy increases, since with a high occupancy rate there are many vehicles on the road, and accidents generally concern multiple vehicles.

[1] However, a number of drivers involved in a single-vehicle accident claim that the accident is due to a runaway vehicle.

The power and the exponential models are close together in the case of multiple-vehicle accidents; however, they differ in the single-vehicle case. Remember that the number of single-vehicle accidents used to calibrate the models is rather low (23). For single-vehicle accidents, the exponential model must be preferred to the power model because of the lower values of two criteria (Residual Deviance and AIC-Akaike Information).

The total risk is minimal (exponential model) when occupancy is 8%; for this occupancy, there are too few vehicles on the road to cause multiple-vehicle accidents, and average speed is generally not high enough to cause single-vehicle accidents. When occupancy is higher, there is a strong increase in the total risk.

6.3 Consistency Analysis: The Case of Speed and Occupancy Risks

There is a relationship between density ρ and occupancy O, assuming that all vehicles have the same constant speed during a period. Let L be the average length of the vehicles and $\boldsymbol{l}$ be the length of the traffic sensor (in meters); $\rho = 1000.O/(L+l)$ or:

$$O = (L+\boldsymbol{l}) \cdot \rho/1000 \tag{10}$$

In France, the values of L and $\boldsymbol{l}$ are 4.2 m and 1 m.

Let us now consider a model giving the risk related to speed, and a model giving the risk related to occupancy. If these models are consistent, their combination should be credible.

As the risk model of single-vehicle accidents related to speed is monotonous (increasing), it is possible to invert it and to produce a monotonous (increasing) "speed related to risk" model; combining this new model with the (decreasing) "risk related to occupancy" model results in a new "speed related to occupancy" model which is decreasing. Then, using the occupancy-density relationship, a speed-density relationship is obtained:

Another speed-occupancy or speed-density relationship is obtained when inverting the monotonous (decreasing), model giving the risk of multiple-vehicle accidents in function of speed, with the increasing model giving the risk of multiple-vehicle accidents in function of occupancy.

Our very aim is not to identify a speed-density relationship, but to analyze the relationship produced. The risk models will not be considered as credible if the speed-density relationship. produced by their combination is not credible.

Let us consider both exponential models for risk related to speed and risk related to occupancy; then the logarithm of the risk is linear related to speed and linear related to occupancy; it leads to a linear relation between speed and occupancy (or density) which is the relation proposed by Greenshields [10].

Considering both power models for risk related to speed and risk related to occupancy leads to a power speed-density relationship:

$$V = \mu \cdot \rho^{\delta} \text{ with } \delta < 0 \tag{11}$$

Considering the power model (Eq. 2) for risk related to speed and the exponential model (Eq. 4) for risk related to occupancy leads to the Underwood speed-density relationship:

$$Log\, R/(1-R)] = \alpha_V + \beta_V \cdot Log(V) = \alpha_o + \beta_o \cdot O = \alpha_o + \beta_o(L+\boldsymbol{l}) \cdot \rho/1000) \tag{12}$$

where α_V and β_V are the parameters of the power model linking risk to speed; α_O and β_O are the parameters of the exponential model linking risk to occupancy.

Rewriting Eq. (12): $\text{Log}(V) = (\alpha_o - \alpha_V)/\beta_V + \rho.(\beta_o/\beta_V)\,(L+l).\,/1000$ or:

$$V = e^{(\alpha_o - \alpha_V)/\beta_V} . e^{\beta_o/\beta_V (L+\boldsymbol{l})\rho/1000} \tag{13}$$

This is the Underwood [15] fundamental diagram:

$$V = V_f . e^{(-\rho/\rho_{\text{cr}})/} \text{ with : } V_f = e^{(\alpha_o - \alpha_V)/\beta_V} \text{ and } 1/\rho_{\text{cr}} = -(\beta_o/\beta_V . (L+\mathbf{l})/1000) \tag{14}$$

where V and ρ are the variables (average speed and density); V_f, is the free-flow speed and ρ_{cr} is the critical density at which the traffic flow is maximum.

When a credible speed-density relationship is, as inquired, obtained, this does not prove that the risk models are good, because the risk has disappeared from Eq. (13). Let us imagine that 50% of accidents are missing; the risk would be divided by two, but this would be unnoticeable in Eq. (13).

Fundamental diagrams which link speed to density are always using the space-mean speed (i.e. the harmonic mean). It is then necessary, for passing from risks to fundamental diagrams, to express risks in function of the space-mean speed rather than the time-mean speed (i.e. the arithmetic speed). The coefficients α, β, γ_{Rain} of the power model (Eq. (2)) relating risk and speed are replaced by the coefficients α_{Space}, β_{Space}, γ'_{Rain} of model (15):

$$\text{Log}[R_i/(1-R_i)] = \alpha_{Space} + \beta_{Space}.\text{Log}\left(V_{s,i,}\right) + \gamma'_{Rain}.Rain_i + \varepsilon_i \tag{15}$$

where $V_{s,i}$ is the harmonic mean speed at period i; α_{Space}, β_{Space} and γ'_{Rain} are obtained either by a direct calibration of model (15), or from the α, β and γ_{Rain} parameters of model (2) by minimizing the sum of the square of the differences between the risk related to the time-mean and the space-mean speeds:

$$\sum\nolimits_i \left(\alpha_{Space} + \beta_{Space}\text{Log}\left(V_{s,i}\right) + \gamma'_{Rain}Rain_i - \alpha - \beta\,\text{Log}(V_i) - \gamma_{Rain}Rain_{,i}\right)^2 \tag{16}$$

This second approach has been done here. For the case of {fast lane, no rain, daytime}, three power models relating risk to space-mean speed have been calibrated:

– risk for single vehicle accidents: (PTW excluded): Relation 3-1 is modified in Relation $3\text{-}1_{\text{Space}}$

– risk for multiple vehicle accidents: Relation 10-1 is modified in Relation 10-2_{Space}
– risk for multiple vehicle accidents (PTW excluded); Relation 14-2 is modified in Relation 14-2_{Space}

The values of the parameters of these relationships are in the Appendix, Table 4. The inverse of these three power models have been combined with exponential models relating risk to occupancy (relations 24-1, 17-2, 25-2); three Underwood speed-density relationships have been obtained; there parameters are in Table 3:

Table 3. Values of the parameters of the underwood speed-density relationships obtained by power and exponential models relating risks of single- or multiple-vehicle accidents to space-mean speed and occupancy, daytime, no-rain, fast lane.

	V_f	ρ_{cr}	Capacity
Single-vehicle accident without PTW	132.6	53.1	2589
Multiple-vehicle accident	120.3	55.7	2465
Multiple-vehicle accident without PTW	119.3	58.1	2550

The free-flow speed is credible; the critical density and the capacity, although rather high, are not incredible, taking into account, first that it is a fast lane and second that a higher capacity appears when considering data aggregated on a 6-minute period rather than over one hour.

The three corresponding speed-density curves are given in Fig. 6:

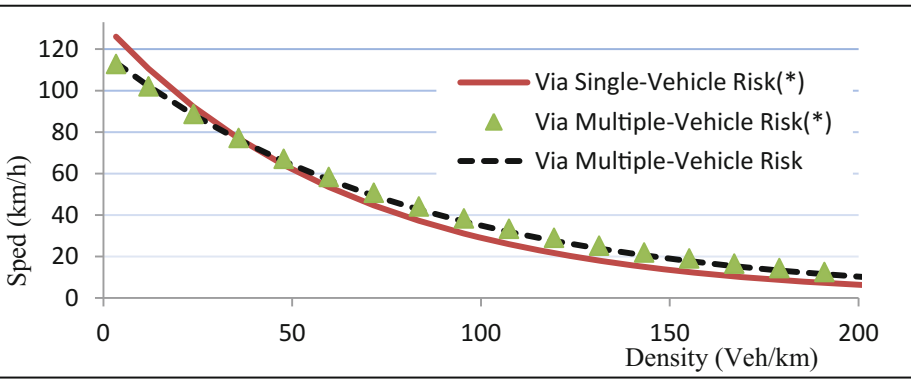


Fig. 6. Three Underwood speed-density relationships, no-rain, with and without (*) PTW, fast lane, daytime (The star(*) in the legend indicates than PTW have been excluded. The three speed-density relationships appear very close together.)

7 Conclusion and Perspectives

Using significant relationships, it has been shown that some variables are significantly linked to accidents:

- for single-vehicle accidents, the 6-minute average speed on the fast lane; and the average time headway (on every lane),
- for multiple-vehicle accidents, the percentages of short time headway (for every lane); excluding rain; the power and the exponential models relate significantly accidents and occupancy on every lane - but the correlation is low on the fast lane; moreover, for the slow and middle lanes, risk increases with occupancy until a threshold, then decreases.

It is more difficult to understand other relationships linking accidents to relative speed or to the "*PICUD*", a variable based on the collision computation. These relationships might come from a correlation between relative speed or the *PICUD* with the traffic density, followed by a mechanical effect of the level of density on the type of accident.

Moreover, the behavior of a single driver who is responsible for an accident does not systematically appear at an aggregated 6-minute level.

Such relationships, if correctly validated and integrated in traffic management tools, should be useful to anticipate the safety impact of a new traffic management scheme.

We also investigated the relationships *between* the risk relationships, and we think that this approach provides a better understanding of data, accidents and models.

The perspectives of this work are three-fold: improving the data processing, validating and assessing the predictive power of such relationships as accident precursors, and integrating them into traffic management systems.

Improving data processing is at three levels:

(1) Improving data. The number of accidents considered is not sufficient to validate the relationships; the time of the accident is not perfectly known; the distance between two successive traffic stations is too high to capture very local problems; data are static, and do not describe the beginning or the ending of a bottleneck; Ppwer-Two-Wheels are not counted.
(2) Data processing. Some analyses must be disaggregated, according to the type of section and infrastructure. Various aggregation times should be tested.
(3) Traffic indicators. Other relevant indicators should be added (Time to Collision, Post-Encroachment Time, etc.). Selected percentiles might replace the 6-minute average in the indicators. Inter-lane indicators, based on relative speed between lanes and on gap availability, and platoon indicators, should be introduced. Future research should also include the development of models that take into account the various contributing variables in the same model, and that also take into account some interactions between the variables.

The perspective also is to go on to validation and to transferability. Validating the relationships on other periods is essential, as well as studying their transferability onto other sites. The power of such relationships as "accident precursors" should be assessed first by analyzing the rates of false alarm and of "no detection" they imply (Abdel-Aty et al., 2005) and second by checking whether, beyond correlation, the traffic indicators really contribute to the risk. In the case where this power is sensitive, the final step would consist in integrating such relationships as safety criteria in traffic management algorithms.

Acknowledgements. We are very grateful to the "COMET" project (https://sites.google.com/site/orsicomet/home), launched by IFSTTAR, which brings together different participants in order to share knowledge and to develop tools for traffic management during adverse meteorological conditions. We also thank the reviewers for their relevant and fruitful remarks.

Appendix: Significant Relationships

Table 4. Parameters of significant relationships relating risk to average speed and rain (day-time)

Number	Model	Constant	β	γ	D_0	AIC
	Lane	(σ_α) (P%)	(σ_β) (P%)	(σ_γ) (P%)	(D_R)	#Fre
1-1	*EXP*	−20.205	0.030	1.697	59.9	131
Speed	Fast	(1.44) (0%)	(0.01) (2.5%)	(0.33) (0%)	(36.2)	#41
2-1	*POW*	−29.382	2.662	1.689	59.9	132
Speed	Fast	(6.13) (0%)	(1.31) (4.3%)	(0.33) (0%)	(37.0)	#41
3-1	POW_{N^*}	−48.158	6.609		26.9	74
$3\text{-}1_{Space}$		−43.282	5.562			
Speed	Fast	(11.7) (0%)	(2.51) (0.8%)		(19.7)	#23
4-1	POW_N	−27.399	2.235		28.9	98
Speed	Fast	(6.78) (0%)	(1.45) **(12%)**		(26.0)	#31
5-1	EXP_N	−19.584	0.025		28.9	97
Speed	Fast	(1.58) (0%)	(0.01) (10%)		(25.8)	#31
6-1	POW_*	−54.0	7.86	1.88	56.6	102
Speed	Fast	(10.3) (0%)	(2.21) (0%)	(0.37) (0%)	(27.2)	#33
7-1	EXP_*	−25.1	0.07	1.88	56.6	103
Speed	Fast	(2.2) (0%)	(0.02) (0%)	(0.37) (0%)	(27.3)	#33
8-1	EXP_{N*}	−23.9	0.06		26.9	75
Speed	Fast	(2.5) (0%)	(0.02) (0.8%)		(19.9)	#24
9-2	EXP_N	−12.828	−0.033		109.4	224
Speed	Fast	(0.35) (0%)	(0.003) (%)		(51.1)	#72
10-2	POW_N	−5.149	−2.392		109.4	229
$10\text{-}2_{Space}$		−6.913	−2.015			
Speed	Fast	(1.26) (0%)	(0.28) (0%)		(56.0)	#72
11-2	Log^2	−10.173	−0.283		109.4	228
Speed	Fast	(0.66) (0%)	(0.03) (0%)		(54.6)	#72
12-2	EXP_N^2	−14.01	−2.109e-04		109.4	222
Speed	Fast	(0.22) (0%)	(2.5e-05) (0%)		(49.0)	#72
13-2	EXP_{N*}	−13.1543	−0.0321		89.4	197
Speed	Fast	(0.40) (0%)	(4.4e-3) (0%)		(46.9)	#65
14-2	POW_{N*}	−5.620	−2.337		89.4	200
$14\text{-}2_{Space}$		−7.344	−1.969			
Speed	Fast	(1.43) (0%)	(0.3) (0%)		(49.5)	#65

Table 5. Parameters of significant relationships relating risk, occupancy and rain (daytime)

Number	Model	α	β	γ	D_0	AIC
	Lane	(σ_α) (P%)	(σ_α) (P%)	(σ_γ) (P%)	(D_{Res})	(#Fr)
15-1	*POW*	−18.751	−0.545	1.796	94.8	126
Occup.	Fast	(0.31) (0%)	(0.07) (0%)	(0.33) (0%)	(41.1)	#34
16-2	POW_{N*}	−15.238	0.352		75.0	224
Occup.	Fast	(0.42) (0%)	(0.15) (2.4%)		(69.3)	#68
17-2	EXP_N	−16.565	6.969		97.2	256
Occup.	Fast	(0.16) (0%)	(1.32) (0%)		(74.9)	#78
18-2	POW_N	−14.823	0.422		97.2	269
Occup.	Fast	(0.38) (0%)	(0.14) (0.3%)		(87.2)	#78
19-1	*EXP*	−15.622	−24.474	1.993	94.8	117
Occup.	Fast	(0.26) (0%)	(4.75) (0%)	(0.33) (0%)	(32.9)	#34
20-1	POW_*	−18.676	−0.466	2.033	64.2	99
Occup.	Fast	(0.39) (0%)	(0.09) (0%)	(0.36) (0%)	(27.7)	#29
21-1	POW_N	−18.595	−0.503		47.4	87
Occup.	Fast	(0.36) (0%)	(0.09) (0%)		(27.3)	#23
22-1	EXP_*	−15.762	−28.054	2.115	77.1	231
Occup.	Fast	(0.31) (0%)	(6.1) (0%)	(0.37) (0%)	(59.9)	#73
23-1	EXP_N	−15.727	−22.108		47.4	82
Occup.	Fast	(0.30) (0%)	(5.60) (0%)		(22.5)	#23
24-1	EXP_{N*}	−16.1008	−20.1865		21.581	64
Occup.	Fast	(0.37) (0%)	(7.11) (0.4%)		(13.2)	#20
25-2	EXP_{N*}	−16.7577	6.5275		75.0	213
Occup.	Fast	(0.17) (0%)	(1.48) (0.4%)		59.4	#68
26-2	EXP_R^2	−16.117	14.423	0.482	144	351
Occup.	Slow	(0.1) (0%)	(2.26) (0%)	(0.28) (9%)	(116)	#98
27-2	POW_N	−13.069	1.182		128	182
Occup.	Slow	(0.4) (0%)	(0.16) (0%)		(81.6)	#82
28-2	POW_N	−12.787	1.321		142	178
Occup.	Middle	(0.4) (0%)	(0.18) (0%)		(91)	#79
29-2	EXP_{NP}	−17.473	20.029	−31.340	127	283
Occup.	Slow	(0.3) (0%)	(4.17) (0%)	(10.5) (0%)	(75)	#84

Table 6. Significant relationships. relating risk, various traffic indicators and rain (daytime)

Number	Model	α	β	γ	D_0	AIC
Variable	Lane	(σ_α) (P%)	(σ_α) (P%)	(σ_γ) (P%)	(D_{Resi})	#Fr
30-1	*POW*	−18.6	1.15	1.25	86	154
Average TH	Slow	(0.3) (0%)	(0.22) (0%)	(0.35) (0%)	(51)	#46
31-2	*EXP**	−15.8	−0.12	0.52	63	246
Average TH	Middle	(0.2) (0%)	(0.07) (6.5%)	(0.3) (11%)	(56)	#80
32-2	MIX_{NP}	−15.2	−1.08	0.014	61	237
Average TH	Middle	(0.3) (0%)	(0.41) (0.8%)	(0.007) 6%	(54)	#75

(*continued*)

Table 6. (*continued*)

Number	Model	α	β	γ	D_0	AIC
Variable	Lane	(σ_α) (P%)	(σ_α) (P%)	(σ_γ) (P%)	(D_{Resi})	#Fr
33-1	*EXP*	−15.4	−3.58	1.67	69.3	133
%TH<2s	Middle	(0.4) (0%)	(0.85) (0%)	(0.34) (0%)	(30.5)	#46
34-2	EXP_{NP}	−14.0	−8.84	8.56	70.8	243
%TH<2s	Fast	(0.4) (0%)	(1.67) (0%)	(1.7) (0%)	(48.9)	#85
35-2	EXP_{NP}	−14.8	−7.44	8.60	70.7	253
%TH<2s	Middle	(0.5) (0%)	(2.0) (0%)	(2.2) (0%)	(57.5)	#82
36-2	EXP_{NP}	−15.4	−4.38	6.59	74.2	287
%TH<2s	Slow	(0.4) (0%)	(2.24) (5%)	(2.6) (1%)	(64.4)	#95
37-2	EXP_{NP}	−14.36	−14.47	26.21	94.8	235
%TH<1s	Fast	(0.2) (0%)	(2.5) (0%)	(5.5) (0.0)	(62.1)	#72
38-2	MIX_{NP}	−17.2	−0.47	6.62	68.4	242
%TH<1s	Middle	(0.4) (0%)	(0.14) (0.1%)	(4.1) (11%)	(59.7)	78
39-2	MIX_{NPR}	−16.9	−0.34	6.87	59.3	247
%TH<1s	Slow	(0.4) (0%)	(0.13) (0.7%)	(4.1) (9(%))	(52.8)	78
40-2	MIX_{NPR}*	−18.4	−0.47	168.5	56.9	145
%TH<0.5s	Fast	(0.5) (0%)	(0.11) (0%)	(51.4) 0.1%	(40.8)	42
41-2	EXP_{NP}	−15.7	−63.1	1565	36.1	131
%TH<0.5s	Slow	(0.2) (0%)	(37.6) (9.3%)	(969) (10%)	(33.3)	#37
42-2	EXP^2_N	−16.0	−308.5		54.6	183
%TH<0.5s	Middle	(0.1) (0%)	(136) (2.4%)		(47.7)	#56
43-1	*POW*	−17.0	1.95	1.59	57.2	111
RelSpeed	Fast	(0.16) 0%	(0.45) (0%)	(0.33) (0%)	(20.6)	#36
44-2	POW_R*	−16.2	−1.41	0.52	122	277
RelSpeed	Slow	(0.1) (0%)	(0.2) (0%)	(0.3) (9.8%)	(67)	#85
45-2	POW_N	−16.1	−1.58		106	228
RelSpeed	Slow	(0.1) (0%)	(0.2) (0%)		(50)	#71
46-2	MIX_{NP}	−16.8	−2.00	0.42	100	225
RelSpeed	Middle	(0.3) (0%)	(0.4) (0%)	(0.2) (10%)	(60)	#66
47-2	EXP_{NP}	−13.0	−4.60	1.42	106	228
RelSpeed	Slow	(0.5) (0%)	(1.2) (0%)	(0.66) (3%)	(48)	#71
48-1	*EXP*	−18.2	1.78	1.6	28	63
PicudBis	Fast	(0.3) (0%)	(0.9) (3.9%)	(0.5) (0.2%)	(18)	#18
49-1	EXP^{2*}	−17.7	60.66	1.9	40	94
%PIC<−20	Fast	(0.3) (0%)	(36.8) (9.9%)	(0.4) (0%)	(20)	#33
50-1	EXP^2	−17.3	246.5	1.6	52	113
%PiCbis<0	Middle	(0.2) (0%)	(140) (7.8%)	(0.3) (0%)	(33)	#35
51-1	EXP^2	−17.4	211.2	1.8	57	92
%PiCbis<0	Fast	(0.2) (0%)	(110) (5.5%)	(0.3) (0%)	(33)	#24
52-2	POW_R	−17.3	−0.48	0.51	97	288

(*continued*)

Table 6. (*continued*)

Number	Model	α	β	γ	D_0	AIC
Variable	Lane	(σ_α) (P%)	(σ_α) (P%)	(σ_γ) (P%)	(D_{Resi})	#Fr
%PIC<−10	Slow	(0.3) (0%)	(0.08) (0%)	(0.28) (7%)	(62)	#95
53-2	POW_R	−18.4	−0.55	0.54	78	183
%PIC<−20	Slow	(0.5) (0%)	(0.1) (0%)	(0.28) (5%)	(43)	#54
54-2	$LOG^2_{N^*}$	−17.0	0.05		63	142
%PIC<−20	Fast	(0.2) (0%)	(.01) (0%)		(35)	#44
55-2	EXP_{NP}	−15.2	−41.30	238.09	67	177
%PIC<−20	Middle	(0.2) (0%)	(8.0) (0%)	(64) (0%)	(31)	#59
56-2	EXP_{NP}	−15.2	−73.26	725.46	61	148
%PIC<−20	Slow	(0.2) (0%)	(17.4) (0%)	(250) 0%	(28)	#46
57-2	EXP_{NP}	−14.2	−11.50	16.4	71	245
%PIC<0	Fast	(0.4) (0%)	(2.6) (0%)	(4.4) (0%)	(63)	#84

References

1. Abdel-Aty, M., Uddin, N., Pande, A., Abdalla, F.M., Hsia, L.: Predicting freeway crashes from loop detector data by matched case-control logistic regression. Transp. Res. Rec. **1897** (1), 88–95 (2004)
2. Abdel-Aty, M., Uddin, N., Pande, A.: Split models for predicting multivehicle crashes during high-speed and low-speed operating conditions on freeways. Transp. Res. Rec. **1908**, 51–58 (2005)
3. Abdel-Aty, M.A., Pemmanaboina, R.: Calibrating a real-time traffic crash-prediction model using archived weather and ITS traffic data. IEEE Trans. Intell. Transp. Syst. **7**(2), 167–174 (2006)
4. Aron, M., Seidowsky, R., Ditchi, N.: Traffic indicators and accidents: the case of a motorway network in the South of France. Transportation Research Board, 92nd Annual meeting, paper#13-4638 Washington D.C. (2013)
5. Bergel-Hayat, R., Debarrh, M., Antoniou, C., Yannis, G.: Explaining the road accident risk: weather effects. Accid. Anal. Prev. **60**, 456–465 (2013)
6. Brodsky, H., Hakkert, A.S.: Risk of a road accident in rainy weather. Accid. Anal. Prev. **20** (3), 161–176 (1988)
7. Elvik, R.: A re-parameterisation of the power model of the relationship between the speed of traffic and the number of accidents and accident victims. Accid. Anal. Prev. **50**, 854–860 (2013)
8. Elvik, R.: Speed and Road Safety – New Models. TØI Report 1296/2014 Oslo, 39 p. (2014). https://www.toi.no/getfile.php/Publikasjoner/T%C3%98I%20rapporter/2014/1296-2014/1296-2014-elektronisk.pdf
9. Golob, T.F., Recker, W.W., Alvarez, V.M.: Freeway safety as a function of traffic flow. Accid. Anal. Prev. **36**(6), 933–946 (2004)
10. Greenshields, B.D., Bibbins, J.R., Channing, W.S., Miller, H.H.: A study of traffic capacity. Highw. Res. Board Proc. **14**, 448–477 (1935)
11. Hauer, E.: Speed and safety. Transp. Res. Rec. **2103**, 10–17 (2009)

12. Lord, D., Guikema, S.D.: The Conway–Maxwell–Poisson model for analyzing crash data. J. Appl. Stoch. Models Bus. Ind. **28**(2), 122–127 (2012)
13. Nilsson, G.: Traffic safety dimensions and the power model to describe the effect of speed on safety. Bulletin 221. Lund Institute of Technology, Department of Technology and Society, Traffic Engineering, Lund (2004). http://lup.lub.lu.se/record/21612
14. Park, P.Y.-J., Saccomanno, F.: Reducing treatment selection bias for estimating treatment effects using propensity score method. J. Transp. Eng. Am. Soc. Civ. Eng., Manuscript No. TE/2005/023489, February 2007
15. Underwood, R.T.: Speed, Volume, And Density Relationship: Quality and Theory of Traffic Flow Yale Bureau of Highway Traffic, New Haven, pp. 141–188 (1961)
16. Uno, N., Iida, Y., Itsubo, S., Yasuhara, S.: A microscopic analysis of traffic conflict causes by lane-changing vehicle at weaving section. In: Proceedings of the 13th Mini-Euro Conference on "Handling Uncertainty in Transportation Analysis of Traffic", pp. 143–148 (2003)
17. Venkataraman, N.S., Ulfarsson, G.F., Shankar, V., Oh, J., Park, M.: Model of relationship between interstate crash occurrence and geometrics: exploratory insights from a random parameters negative binomial approach. Transp. Res. Rec. **2236**, 41–48 (2011)

Development of Coordinated Ramp-Metering Based on Multi-objective Nonlinear Optimization Functions: Traffic and Safety

Habib Haj-Salem(✉), Nadir Farhi, Jean Patrique Lebacque, and Neïla Bhouri

COSYS/GRETTIA, 14-20 Bd Newton, 77447 Marne-La-Vallee Cedex 2, France
{habib.haj-salem,nadir.farhi,
jean-patrique.lebacque,neila.bhouri}@ifsttar.fr

Abstract. This paper is focused on the extension of the OASIS (Optimal Advanced System for Integrated Strategies) in order to take into account the multi-objective nonlinear optimization technique for coordinated ramp metering. The multi-objective function includes two costs functions: traffic and safety (Risk model) indices. OASIS is revisited and off-line simulation studies are conducted on real test site corresponding to A6W France motorway located in the south part of "Ile de France" Motorway networks. Five consecutive on-ramps are considered for the control. The obtained results are very promising.

Keywords: Risk index · Multi-objective non linear optimization · Coordinated ramp metering · Off-line evaluation

1 Introduction and Background

A number of approaches have been developed in the past for the design of control strategies that involve control measures such as route recommendation via Variable Message Signs (VMS) devices or equipped vehicles, ramp metering, motorway-to-motorway control (MTMC), automatic incident detection (AID), hard lane shoulder etc. Several approaches were investigated including expert systems, fuzzy systems, neural networks, classical feedback control and optimal control strategies based on either linear or non linear approaches.

In practice, traffic control systems within corridors or motorway networks have been developed independently for each individual control measure attempting to optimize traffic flow on the motorway or the urban road networks or on the both components. With respect to the optimal control, the problem formulation was focused on the development of the integrated strategies which simultaneously take into account several control strategies such as ramp metering, MTMC, user's guidance. This global point of view suggests that control measures within the entire network should be designed in an integrated way, or, at least, they should be suitably coordinated during operation, so as to meet

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_16

the overall objectives [9,10]. These approaches are very promising. Nevertheless, whatever considered approach, the used cost objective functions are mainly focused on the optimization, in the considered system, of the traffic indices such as the minimization of the Total Time Spent index (TTS), the travel time, the maximization of the Total Travel Distance (TTD), the mean speed etc.

However, except the Automatic Incident Detection (AID) strategies which are focused on real time accident/incident detection aiming at minimizing the detection time and the traffic impact of the accident, the safety index is not taken into account. In general, the safety aspect is considered as an external cost function and not included in the real time control strategies. Safety indices are computed during the evaluation process. The classical approaches consist in collecting first the incident/accidents traffic data during the experimented scenarios. After the experimentation phase, the safety evaluation will commence. These evaluations are based on the statistical analysis of the number of accidents before and after the implementation of the tested strategies. Therefore, the constitution of the accident database must include a minimum number of accidents in order to guaranty the statistical significance of the evaluation process. This means that the field data collection period must have a long time periods (comprising between 5 to 10 years), which is the "price to pay" for having a significant safety evaluation result.

On the other hand, the introduction of electronic devices and computerized systems in the vehicle technologies have significantly contributed to user's safety and comfort. Nevertheless, the prediction of the crash in real time is still in investigation phase and some research efforts are dedicated in this area [4,11]. During the last decade, there is increased focus on the development of real time ("*potential crash*") prediction algorithm on urban motorway traffic [12,16,20].

This paper is focused on the development of optimal control strategy based on multi-objective index function, including traffic and safety indices. The first part of this paper is dedicated to the description of the risk model building. This risk model is validated on the ringway of Paris. The second part of this paper is focalized on the description of the generalized problem formulation based on the application of the optimal control for the development of integrated control strategies. The numerical solution of a formulated large-scale nonlinear optimal control problem is effectuated by application of a non-linear optimization technique based on the optimal control theory, which is able to deal straightforwardly with non-linear features. In this paper, focus is made on the coordinated ramp metering. The third part of this paper is dedicated to the reformulation of the cost function by including the risk index model. Summary of the developed risk model is described along with several simulation tested and evaluated scenarios. Investigation of the efficiency of this multi-objective non-linear approach is performed by using a macroscopic simulation model namely METACOR [5] which is able to simulate traffic flow phenomena in corridor networks with arbitrary topology.

2 Summary Description of the Risk Index Model

The developed risk model is based on the collection of traffic measurements synchronized with incident/accidents detection. The considered site was the ringway of Paris. Incidents/accidents database were collected from the implemented real time Automatic Incident Detection system (AID) and the operator reported incident/accident files. The characteristics of the accidents include: starting time (H:min), end-time, location, weather conditions, severity. The constituted data base covers 3 year periods (2002–2004). During this period, the total number of accidents collected is around 900. For each incident/accident, the considered measurement stations are depicted in the Fig. 1. Two upstream and two downstream measurement stations were considered. For each accident, the collected traffic data (traffic volume, occupancy rates and speed) covers two hours (one hour before and one hour after the crash). The collection time interval of the measured data is equal to one minute.

Before the statistical analysis, a selection of valid accidents is performed. The applied selection criteria correspond to the availability of all data measurements: 120 in total (60 = 1 h before the accident and 60 after the accident), the same topology of the accident location (same number of lane = 4 lanes) and sunny days. Among the total number of accidents(900), only 85 accidents are selected for the statistical analysis. Consequently, the total number of collected measurements is equal to $120 * 85 = 8200$. Each measurement includes two upstream and downstream measurement stations of each lane (4 *lanes*) of traffic volume, occupancy rates and speed for each minute time slice. In the constituted data base, each observation line includes $4 * 3 * 4 = 48$ measured traffic variables.

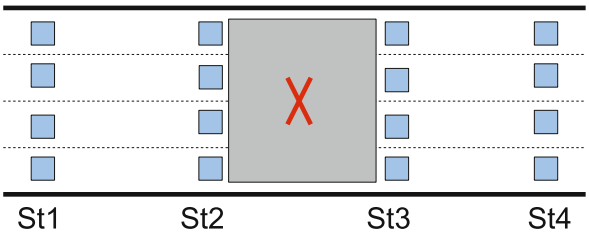


Fig. 1. Topology of the considered stretch measurements for each crash

The applied methodologies are based on statistical analysis of the traffic conditions before the accident. The main objective of this step is to analyze the traffic conditions before the occurrence of the accident and to extract the most important traffic variables to be used for the risk modeling. A series of multivariate statistical methods are used, aiming at finding the relationship between the occurrence of the accident and the traffic conditions on the stretch using the real data measurements [6,7,12].

The multivariate statistical approaches include the following steps:

1. Correlation: What are the non correlated variables to be used?
2. Principal Component Analysis (PCA): Reduction of the used variables
3. Clustering: To understand the traffic dynamics before the accidents indicated by the dynamic behavior between clusters
4. Logistic regression: Model building bases on the *logit* model application

Extensive iterative statistical computing of the four steps were performed using SAS statistical software. In particular, the following scenarios were considered:

– upstream measurements by lanes and by stations
– downstream measurements by lanes and by stations
– overall measurements by lanes and by stations

In order to alleviate the paper, reader can refer to the cited references for more details. In particular, the four steps results by scenario are well described in the reference [6] along with extended details and results analysis. In summary, the best scenario found corresponds to the use of the upstream and downstream measurement stations (means of the four lanes). In this case, the used state variables were limited to: Q1, Oc1, Oc2, Oc3 and Oc4. These 5 significant state variables to be used are provided by the statistic computation of steps (1) and (2). The clustering computations (step 3) provide 5 clusters with the following characteristics:

– **Cluster** 1**:** is the most dense one (more than 36% of all observations cases). It is characterized by quite homogeneous occupancy rates(Occ) and average flow over the 4 stations, (Occ of 11% to 12%) and a flow of 1450 to 1500 vehicles per hour and per lane) characterizing fluid traffic conditions.
– **Cluster** 2**:** presents a very high Occ on the (upstream) stations St_1 and St_2 and rather average downstream (14% to 18%). As for the flow, it is rather stable and low compared to other clusters. This cluster contains 20% of the overall observations.
– **Cluster** 3**:** presents high occupancy rates over all stations. The upstream flow is around the capacity whereas the downstream flow is less than the capacity (hight level of congestion: 40% and 37%). This cluster contains 12% of the overall observations.
– **Cluster** 4**:** has average Occ close to the usual 20% critical value, increasing from upstream to downstream (26.7%) at station St_4. The flows are higher than the other clusters, up to 1774 veh/h/lane at station St_2.
– **Cluster** 5**:** has a high average Occ (around 37%) at all stations and a lower flow (around 1230 veh/h/lane). When we consider the accidents and attribute to each time step the cluster number to which it belongs, we observe that 43 accidents out of 85 studied (51%) present cluster change during the last six minutes.

For 60 accidents (i.e. more than 70% of them), the last time step belongs to cluster 2, characterized by a rarefaction shock wave (congested upstream and fluid downstream).

The Risk model (R) is set to $R = 1$ for the observations belonging to cluster 2 and $R = 0$ elsewhere. The calibration of the Risk model is based on 55 accidents.

2.1 The Logistic Regression Model

In our case, the predictor corresponds to the occurrence of the accident ($R = 1$) or not ($R = 0$) which is a binary variable. In case of binary dependent variable, the linear regression should not be applied due to the possibility to have a probability higher to 1. Logistic regression, also called a *logit* model, is used to model dichotomous outcome variables. In the *logit* model the *logodds* of the outcome is modeled as a linear combination of the predictor variables.

The "logit" model solves these problems:

$$\ln \frac{p}{1+p} = AX + b \tag{1}$$

$$\frac{p}{(1-p)} = exp(AX + b) \tag{2}$$

where:

p is the probability that the event R occurs: $p(R = 1)$ $p/(1-p)$ is the "odds ratio" $ln[p/(1-p)]$ is the log odds ratio, or "logit"

The logistic regression model is simply a non-linear transformation of the linear regression. The "logistic" distribution is an S-shaped distribution function which is similar to the standard-normal distribution (which results in a probit regression model) but easier to work with in most applications (the probabilities are easier to calculate). The logit distribution constrains the estimated probabilities to lie between 0 and 1.

For instance, the estimated probability is:

$$p = R = \frac{1}{(1 + exp(-AX - b))} \tag{3}$$

This approach is applied for the Risk model development. The "*logit*"regression model found is the following:

$$RiskIndex = \frac{1}{1 + \exp -(\alpha + \beta Oc_1 + \gamma Oc_2 + \delta Oc_3 + \theta Oc_4 + \Phi Q_1)} \tag{4}$$

where:

Oc_i is the occupancy rates of the station (i); $i = 1..4$

Q_1 is the traffic volume of station 1.

$\alpha = -7.1677; \beta = 0.2122; \gamma = 0.1389; \delta = -0.1061;$

$\theta = -0.2052; \phi = 0.00038.$

Analyse des estimations de la vraisemblance maximum

Paramètre	DF	Estimation	Erreur std	Khi 2 de Wald	Pr > Khi 2
Intercept	1	-7.1677	0.6941	106.6361	<.0001
moytost1	1	0.2122	0.0118	320.8749	<.0001
moytost2	1	0.1383	0.00899	237.0292	<.0001
moytost3	1	-0.1061	0.00935	128.6439	<.0001
moytost4	1	-0.2052	0.0112	333.1159	<.0001
moyqst1	1	0.000385	0.000074	27.2964	<.0001

Fig. 2. Logistic regression: Table of *SAS* output statistics

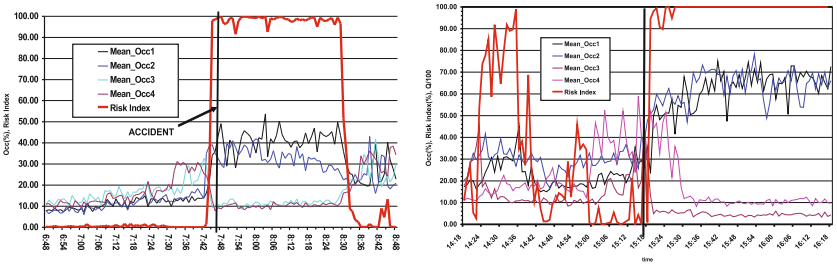
The risk index model calibration was performed on 55 accidents whereas the validation has been proceed on the rest (30 accidents) of the database (Fig. 2).

Screening the found parameters of the risk model, the following remarques can be drawn:

– the downstream parameters are negative. This means that the increase of the downstream occupancy rate measurements leads the decrease of the risk index value
– on the contrary, the upstream parameters are positive which leads, in case of increasing upstream occupancy rate measurements, the increase of the risk index.

For the risk model validation, the same parameters are used for 30 accidents which are not used for the calibration. The obtained results of the 30 accidents are very close to the two accidents selected. The risk index time evolution results are depicted in Fig. 3. The output results indicate that the risk index value is maximal just before the occurrence of the accident. We can observe that the recorded occurred accident time in both figures are not exactly synchronized when the Risk index is maximum. This remarque is related to the recording of accident times by the operator. In some cases the recorded times include some reporting error: it could be higher or less (max 6 min $(+,-)$ are observed).

In Fig. 3 (*Acc*-1), the risk index is maximum just before the occurrence of the accident. However, in case of (*Acc*-2), we can observe that before the occurrence



"Acc-1 " "Acc-2"

Fig. 3. Time evolution of the risk index on the ringway of Paris

of the accident (14 h:18–15 h:18), the risk index is very high without occurrence of the accident. This means that the traffic flow is completely instable during this period (1 h). This instability is translated by time variation of the risk index. However, at 15 h:18, the accident is occurred. When the risk index is maximum, accident can be occurred or not.

The developed risk index can be used off-line as an evaluation index during the evaluation process. In particular the computation of the cumulative Risk index can be considered as a safety external index for the performance assessment between scenarios (e.g. assessments of without and with control). In this case, computation of the risk index using only measurements will leads to the dramatically reduction of the field test periods.

In real time, the computation of the cumulative risk index on the considered motorway axis using real time measurements can be used as safety monitoring tool (e.g. safety user warning system). In our case the risk index is integrated in a multi-criterion function to be optimized in real time (safety index combined with a traffic index).

In this paper, the developed risk model will be integrated in the cost function of OASIS Kernel for coordinated ramp metering strategy development using non linear optimization.

3 Integrated Control: Mathematical Problem Formulation

The control strategy is an essential part of any traffic-responsive motorway control system. In particular, control such urban intersection control, ramp metering, motorway to motorway control ($MTMC$) and route guidance aim at improving significantly the transport efficiency (so the safety) and reducing of environmental pollution [9]. In the past, traffic control systems within corridors or motorway networks have been developed independently for each individual control measure attempting to optimize traffic flow on the motorway or the urban road network. However, a traffic corridor constitutes an entity in terms of operational objectives, user requirements, and impact of individual control devices [10]. Therefore, the integrated control approaches have an additional potential in comparison with the individual ones. This global point of view suggests that control measures within the entire network should be designed in an integrated way, or, at least, they should be suitably coordinated during operation, so as to meet the overall objectives. The need for an integration of road and motorway traffic control measures in metropolitan areas emerges from the following considerations:

- Integration of control measures is expected to improve the use of overall network capacity.
- Particular control measures, such as driver information or route guidance, are only reasonable on the basis of the integrated network traffic state.
- Particular operational objectives, such as preference of motorway paths, can only be achieved by consideration of the overall network.

Due to the lack of a general theoretical background, the non-linear nature of the traffic process, and the constraints imposed by the control measures, the extraction of the decision rules is a highly complicated and time consuming task. The optimization techniques based on the optimal control theory allow to deals straightforwardly with important non-linear features, and have been implemented and used in many fields. Moreover, this approach seems promising because the control decisions are based on the minimization of an arbitrary control criterion rather than regulating towards a certain state of the process and because it considers explicitly the control constraints.

Consider a traffic process described by the general state equation reads:

$$x(k+1) = f[x(k), u(k), d(k)] \tag{5}$$

where $x(k) \in R^n$, $u(k) \in R^m$, $d(k) \in R^n$, denote the state, the control and the disturbance vectors respectively. The disturbance vectors include the demands at the origins, accident/incident locations, weather conditions etc.

In our case, the state vector $x(k)$ correspond to the mathematical macroscopic traffic model equations. The dynamic macroscopic traffic modeling equations are based on 3 main variables: traffic volume $q(x,t)$, density $\rho(x,t)$ and speed $v(x,t)$. The variable $d(k)$ corresponds to the traffic demand vector at each origin of the considered network whereas $u(k)$ is the vector of the control variable. The control vector includes several control strategies such as ramp-metering, motorway to motorway control, speed limits, route guidance, intersection control etc. In general, the control variables are subject to some constraints expressed as: $U_{min} \leq u(k) \leq U_{max}$.

3.1 Summary Description of the Used Macroscopic Traffic Flow Model Equations

In this section, the mathematical equations of the used traffic model are described. In this study the MAGISTER multi-model Kernel is used. MAGISTER Kernel includes several macroscopic traffic models:

- First order modeling (LWR) for the urban modeling
- 4 s order models (ACL: Acceleration Limit model [14]; METACOR [5], ARZ [17], GSOM [14])

According to the macroscopic nature of these models, the state variables of the simulated traffic process are the density $\rho(vh/km)$, the mean speed $v(km/h)$ and the traffic volume (or flow) $q(vh/h)$. Among the second order models in MAGISTER, the METACOR model is used. METACOR is able to simulate urban and motorway network for arbitrary topology including non-oriented destination (no OriginDestination (OD) matrix is needed), only the global demand at origins must be provided. The oriented destination modeling is included also. In this case, the OD matrix must be provided as an input to the model.

The motorway network is represented as a directed graph whereby the links of the graph represent motorway stretches. Each motorway stretch has uniform

characteristics, i.e. no on-/off-ramps and no major changes in geometry. The nodes of the graph are placed at locations where a major change in road geometry occurs, as well as at junctions, on-ramp, and off-ramps.

The time and space arguments are discretised. The time discretisation is the same for the whole network, while the space discretisation is defined for each link separately. The discrete time step is denoted by T. A link l is divided into N_l segments of equal length. For each segment i of each link l at each time instant $t = kT, k = 0, 1, K$, where K is the time horizon, the following macroscopic variables are defined:

- Traffic density $\rho_{l,i}(k)$ (veh/km) is the number of vehicles in segment i of link l at time $k\,T$ divided by the length of the segment.
- Mean speed $v_{l,i}(k)$ (km/h) is the mean speed of the vehicles included in segment i of link l at time $k\,T$.
- Traffic volume or flow $q_{l,i}(k)$ (veh/h) is the number of vehicles leaving segment i of link l during the period $[kT, (k+1)T]$, divided by T.

Let us summarize, in discrete time, the main mathematical equations (non-oriented destination modeling) for a link (l) and a segment (i): The conservation equation on each link (l) and segment (i) reads:

$$\rho_{l,i}(k+1) = \rho_{l,i}(k) + \frac{T}{\Delta_i}[u_{l,i-1}(k)q_{l,i-1}(k) - u_{l,i}q_{l,i}(k)] \tag{6}$$

Relationship between volume, speed and density:

$$q_{l,i}(k) = \rho_{l,i}(k)v_{l,i}(k) \tag{7}$$

Speed equation:

$$\begin{aligned} v_{l,i}(k+1) = v_{l,i}(k) + \tfrac{T}{\tau}[F(\rho_{l,i}(k)) - v_{l,i}(k)] + \tfrac{T}{\Delta_i}v_{l,i}(k)[v_{l,i-1}(k) - v_{l,i}(k)] \\ - \tfrac{\nu T}{\tau \Delta_i}\tfrac{\rho_{l,i+1}(k)-\rho_{l,i}(k)}{\rho_{l,i}(k)+\kappa} \\ - \tfrac{\delta T}{\Delta_i}\tfrac{q_{ramp}(k)v_{l,i}(k)}{\rho_{l,i}(k)+\kappa} \\ - \tfrac{\phi T}{\Delta_i}\tfrac{(\lambda_l-\lambda_{l+1})}{\lambda_l}\tfrac{\rho_{l,i}(k)v_{l,i}^2}{{\rho_{cr}}^l} \end{aligned} \tag{8}$$

Fundamental diagram equation:

$$v(\rho_{l,i}(k)) = V_{f,l}exp\left[-\frac{1}{a}\left[\frac{\rho_{l,i}(k)}{\rho_{cr}^l}\right]^a\right] \tag{9}$$

where

T is the simulation time step

Δ_i = length of the segment i of the link l

F: Fundamental diagram of the link l expressed by the Eq. (9).

$u_{l,i}$: denotes the control variables at the segment i (corresponds to the capacity restriction (severity) due to an accidents/incidents at the link l and segment i)

V_f, ρ_{cr}, a: denote the free-flow speed, the critical density and the parameter of the fundamental diagram of link l respectively.

$\tau, \kappa, \nu, \delta, \phi$: are the model parameters: delay term, the anticipation term and other weight terms.

The set of parameters $P = (V_f, \rho_{cr}, \mathrm{a}, \tau, \kappa, \nu, \delta, \phi)$ reflect particular characteristics of a given traffic network and depend upon street geometry, vehicle characteristics, drivers' behavior, etc.

At the origins of the simulated network a queuing modeling is applied. This modeling permits the queue length estimation and the waiting time. These indices are very useful for the evaluation of the ramp metering strategy performance.

In general, the input traffic to the network is provided at each origin where the demand is supposed to be known.
Assume that, at each origin (o), the demand during the time slice k is $d_0(k)$. The dynamic evolution of the queue length at the origin (o) can be written as:

$$Lq_o(k) = Lq_o(k-1) + T[d_0(k) - q_o(k)] \tag{10}$$

where $q_o(k)$ is the injected traffic volume in the downstream link of the considered origin. When the demand at the origin is higher than the supply of the downstream link, queue at the origin will appear. This simple queuing modeling is very useful for the application of ramp metering strategies. This variable will be considered as a state variable.

In summary, the state variable of the simulation model are:

$$x = [\rho_i \quad v_i \quad Lq_o]^T \tag{11}$$

3.2 Definition of the Traffic Cost Function

The cost criterion may be arbitrary chosen. In general, the traffic criterion used for the optimal control could be a combination of several criterion such as: the Total Travel Distance (TTD), the Total Time Spent (TTS), the travel time, the generalized mean speed $V_{moy} = TTD/TTS$, the queue length and the waiting time at all origins, the total fuel consumption etc. The general form of the traffic cost criterion is the following:

$$J_{Traf} = T \sum_{K} \left\{ \sum_{l} \sum_{i} \rho_{l,i}(k) \Delta_{l,i} + \sum_{o} Lq_o(k) \right\} \tag{12}$$

where K is the time horizon; i: the number of segment of the link l and (o): the number of origins.

In this study, the safety index reported in Eq. (4) is added to Eq. (12) in order to constitute the global cost function to be optimized, reads:

$$J_{Glob} = J_{Traf} + \alpha J_{Saf} \tag{13}$$

where the parameter (α) is a weighted factor to be calibrated during the optimization process.

3.3 Numerical Solution of the Optimal Control Approach

The optimization algorithm is one of the key ingredients in the optimal control process we attempt to resolve. In particular, since the objective is to control online, with a quick reaction to uncertainties, the convergence towards the optimal solutions has to be a low CPU time process. According to the standard methodology based on the optimal control theory [18], we can solve the optimization problem described in the next paragraph by using an optimization algorithm based on the first derivatives.

A discrete-time optimal control problem is formulated as:

$$\min J = \psi(x(K)) + \sum_{k=0}^{K-1} \phi(x(k), u(k)) \tag{14}$$

Subject to:

$$x(k+1) = f\left[x(k), u(k), d(k), P\right]; \quad x(0) = x_0 \tag{15}$$

The variable P corresponds to the set of model parameters and the constraints:

$$U_{min} \leq u(k) \leq U_{max} \tag{16}$$

This type of algorithm is well adapted but requires the explicit derivatives calculation in order to limit the computational effort. In this context, the derivatives concerns the cost criterion with respect to $u(k)$ and to $x(k)$:

$$g(k) = \frac{\partial J}{\partial u(k)} = \left[\frac{\partial f}{\partial u(k)}\right]^T \lambda(k+1), k = 0..., K-1 \tag{17}$$

The term λ denotes the constant which can be computed backwards, using the following equations:

$$\begin{gathered} \lambda(K) = 0 \\ \lambda(k) = \frac{\partial \phi}{\partial x(k)} + \left[\frac{\partial f}{\partial x(k)}\right]^T \lambda(k+1), k = 0..., K \end{gathered} \tag{18}$$

With all these ingredients, it is possible to optimize efficiently any cost function (13), by using some appropriate non-linear iterative optimization algorithm based on a search direction. This search direction can be found using any appropriate algorithms such as "Direction Set Methods" or "Variable Metric Methods". The optimization process steps (for the control strategy) are the following:

1. Start with random admissible trajectories of the variable (u_0), and compute the associated cost function J_0 (Eq. 13).
2. For the iteration it, compute the gradients using Eq. (18).
3. Compute the search direction using an appropriate non-linear optimization algorithm, and the new trajectory (u^{it+1}).

4. Evaluate the new trajectory by computing the cost function J^{it}.
5. if $(J^{it} - J^{it-1})/J^{it} \leq \epsilon$ then stop
   else go to step 2 with iteration $it = it + 1$

The associated algorithm, (OASIS), has been revisited by including the safety index. This tool is a generic algorithm and can be applied to any corridor network with arbitrary topology. Some investigations have been conducted on two different optimization algorithms, known to be efficient on several different optimization problem:

– A recent and very simple but fast algorithm, well-known in the field of nonlinear neural networks, called RPROP [19],
– The conjugate gradient [8], comparable to the quasi-Newton method but which requires intermediate storage of order N versus $N \times N$ for the quasi-Newton method.

On the other hand, there is not, as far we know, any overwhelming advantage that the quasi-Newton-method hold over the conjugate gradient techniques. Both algorithms are reliable and provide good results in terms of optimal solutions and convergence rates. Nevertheless, the *RPROP* algorithm seems more faster to find the *best or local minima*.

As indicated in the Eq. (13), the global cost function includes both functions: traffic and safety and both functions are antagonistically. As a matter of fact, the safety index is minimal (= 0) in both traffic condition limits: (a) No traffic (no accident) and (b) Blocked traffic (maximal density on the motorway: all users are stopped and not move). When the traffic evolves in time, the safety index will automatically increase.

On the contrary, the traffic index will be optimal when the traffic volume is around the capacity of the infrastructure. In this case, the safety index will increase. This is the antagonistic behavior of the safety and traffic cost functions. This problem is well known as $Pareto-optimal\ multiobjective\ solution$ [2,13]. Let us remind the basic properties.

Basic Properties of Pareto-Optimal Solution: Without loss of generality, let us consider the following multiobjective optimization problem:

$$\max f_1(X), f_2(X), f_3(X), ..., f_N(X) \tag{19}$$

where $N =$ number of objective functions; $X = \{x_o, x_l, ..., x_{n-1}\}$ is a n-dimension variable vector.
Assuming $X_0, X_1, X_2 \in \psi$,

1. X_1 is said to be dominated by (or interior to) X_2 if $f(X_1)$ is partially less than $f(X_2)$, i.e. $f_N(X_1) \leq f_N(X_2)$ for $\forall i = 1, 2...N$ and for $\exists\ i = 1...N$ with $f_i(X_1) < f_i(X_2)$
2. X_0, is said to be Pareto-optimal (or non-dominated), if there does not exist any $X \in \psi$ such that X dominates X_0. This definition is based on the intuitive conviction that the point X is chosen as optimal if no criterion can be

improved without degrading at least one other criterion. If a solution is not dominated by any other solutions, then this solution is said to be a non-dominated solution.

Usually, multiobjective optimization problems have a group of trade-off, non-dominated optimal solutions that can satisfy the global cost function. In our case, the weighted parameter (α) must be calibrated and fine tuned in order to minimize the non dominated solutions in case of similar traffic conditions. The used criteria is the value of the cumulated TTS index and the cumulated risk index. The α Risk value do not dominate the optimal solution.

4 Simulation Studies and Hypothesis

4.1 Test Site Description

The test site is located in the south of the *Ile de France* (IDF) Motorway network (Fig. 4). This site is the most critical area of the overall motorway network in the (IDF) motorway networks. Morning and evening peak congestions extend over several hours and several kilometers whilst fairly dense traffic is regularly observed even during off-peak hours. The total length of the A6W motorway axis is approximately 20 km (only the direction towards Paris is considered). This motorway axis includes 5 consecutive on-ramps which are fully equipped with loop detectors and traffic signal lights. The carriageway is equipped with detector stations, one measurement stations each 500 m for traffic volume, occupancy and speed measurements respectively. This test site was the test bed of several ramp metering strategies: isolated and coordinated [3,6].

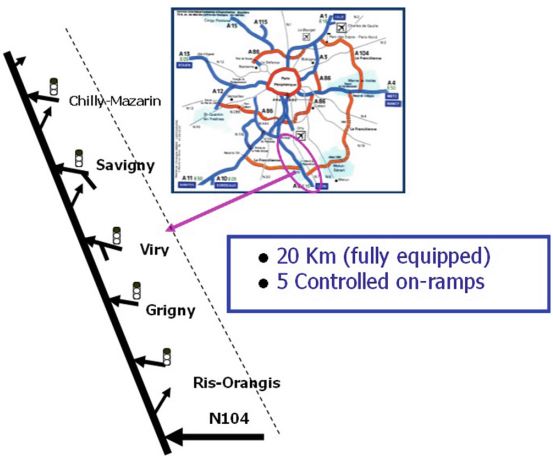


Fig. 4. A6W test site

The developed optimization algorithm (OASIS) has been implemented in the kernel of the simulation tool namely MAGISTER. The developed algorithm can be applied to any corridor network without any specific development. The first step was the METACOR model calibration and validation. This step is performed using real data measurements. The used data was split into two parts: the first part (one week from 5 h:00 to 21 h:00) is used for the model calibration and another week is used for the model validation.

The second step consists to reconstruct the demand profiles using measurements at each origin of the considered site. Several simulation runs were needed for the fine tuning of these demand profiles. They must reconstitute the congestion at the same time and the same locations than the real measurements. More description can be found in [6].

The last step is the test of the developed integrated control strategy OASIS using the multi-criteria objective function.

4.2 Risk Model Validation

Before using the safety index model in the cost function to be optimized, it is necessary to proceed to the risk model validation. A data base is constituted using the real data collection system and the existing files of the reported accidents/incidents. The used data corresponds to the collected accident synchronized with measurement traffic data. The total number of collected accidents is equal to 60. After data cleaning and accidents selection criteria (sunny day, night excluded), only 20 accidents are selected for the risk model validation. The same parameters found on the ring way of Paris are conducted on A6W motorway axis. Among the 20 accident, two accidents are chosen (see Fig. 5.). For the other accidents, the time evolutions of the risk index and traffic variables are

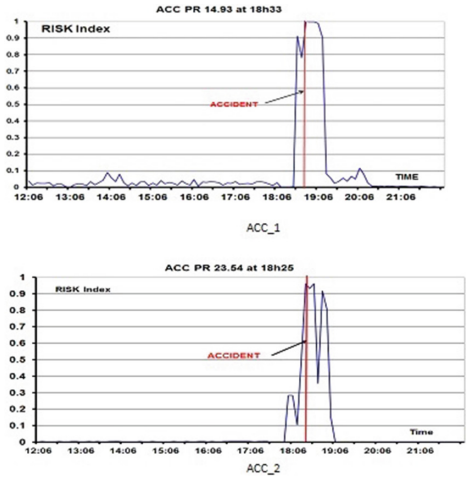


Fig. 5. Time evolution of the risk index for two accidents: Acc1 and Acc2

very similar. Figure 5 depicts the risk index time evolution of two accidents. We can underline that: using the **same parameters** found on the ring way of Paris and applying on A6W motorway axis, the obtained results are very promising. Without any parameter calibrations, the risk index value is maximal before the occurrence of the accident. Consequently, we can assume that the computation of risk index can be considered as a robust safety index and the same parameters can be used in the multi-criteria cost function optimization.

4.3 Scenario Definitions

In order to highlight the impact of the risk index function in the optimization, three main scenarios are considered:

1. Optimization of the risk index only
2. Optimization of the Traffic index only
3. Optimization of both indices: Traffic and Risk (multi-criteria)

The No control case is considered as Reference. For all considered scenarios, the same demand profiles at the origins are applied. These demand profiles are constructed from collected measurements on A6W. In particular, during the peak morning periods, the demand levels at the controlled on-ramps are very high (around 1600 to 2100 vh/h). All fixed demand profiles started and finished at fluid conditions respectively. Consequently, all vehicles are served leading to have the same value of the cumulated TTD index for each scenarios. In this case, comparative assessment will be based on the Total Time Spend (TTS) index. These demand profiles start at 5 h:00 in the morning and complete at 15 h:00. Consequently, the time Horizon of the optimization is over the duration of the simulation horizon $H = 10$ h. For each scenario, the traffic (TTS) and Risk indices are computed and reported in the output files of the simulation. The applied optimization is called Open-Loop Control because the optimization is performed on the overall time Horizon. The characteristics of the ramp metering control strategy is the following:

– The cycle duration is equal to 40 s. This value is found as the optimal cycle during the passed field trails on the same site.
– In the simulation, the control variable is the split which is defined as: *Green duration/cycle time*.
– the imposed constraints on the control variable are: $U_{min} = 0.25$ and $U_{max} = 1$. This means that the minimum of green duration is equal to 10 s whereas the maximum green = Cycle = 40 s.

Optimization Problem Dimension: The simulation time step of the simulation T is equal to 10 s. The total number of links and segments is equal to 13 and 31 respectively. The simulation time horizon (H) is equal to 10 h which equal to: $10 * 3600/10 = 3600$ time steps. The number of state variables are equal to 3 per segment (speed, density and Lq) and the number of control variables is equal to

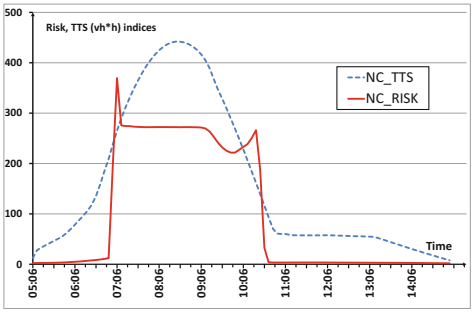


Fig. 6. No Control (NC): Time evolution of the TTS and the Risk indices

5 (5 on-ramps). Hence, the optimization problem dimension can be computed as: $(3600)(31 * 3 + 5) = 352800$ which a very-very large non linear optimization problem to be solved.

Before the results analysis for all scenarios, the prerequisite step is the calibration of the weighted parameter (α) of the global cost function reported in the Eq. (13). The used approach is the "*trial and error*" method. Several simulation runs were conducted in order to obtain a fine tuning of this parameter. The best value found is equal to ($\alpha = 2.5$). Figure 6 depicts the time evolution of no control case of the TTS and the risk index respectively.

The CPU time consuming by scenario is reported in the Table 1. The used computer is a PC, DELL precision T5610 using Windows OS with 64 GB of RAM and two CPU: 2.1 GHz. With respect to the size of the optimization problem, the CPU time consuming is largely low. The ratio between the real-time and the time consuming is roughly equal to 200. This means that the optimized control strategy is running 200 faster than the real time and could be integrated in the Traffic Management System (TMS) without any problem of computation time consuming.

Table 1. CPU Time consuming

Scenarios	Nb. Iter	CPU Time(sec)
Risk	130	60
Traffic	110	45
Risk+Traf	115	55

Figure 7 depicts the evolution of the optimized cost function J_{glob} for each iteration and during the global time horizon ($H = 10$ h) of the three optimized scenarios: Risk, Traffic and the Risk+traffic respectively. Compared to the optimization of the Risk or the Risk+traffic cost functions, we can observe that the optimization of the TTS index converges quickly to the optimal control trajectories.

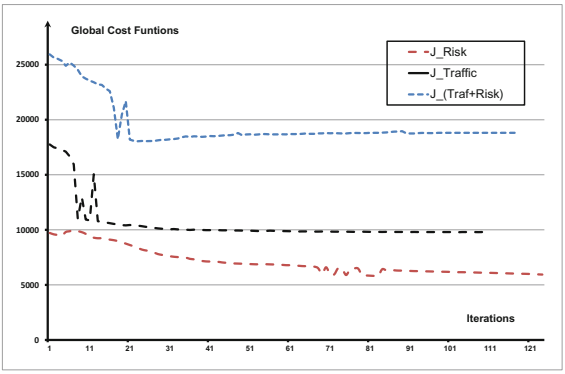


Fig. 7. CPU time consuming

5 Simulation Result Analysis

5.1 Control Output Trajectories Analysis

Before analyzing the performance in term of traffic and safety improvements by each scenarios, it is interesting to analyze first the output of the optimized control trajectories by scenarios. In our case, the optimized control variables concern 5 on-ramps. Among the 5 on-ramp, the on-ramp "*Savigny*" is selected. This on-ramp is the most critical one. During the pick morning period the observed demand is around 2200 vh/h. This level of the demand generates a high level of congestion which spill-back to the downstream on-ramps until the main origin of the A6W. Figure 8 depicts the time evolution of the ramp metering rate of the 3 scenarios.

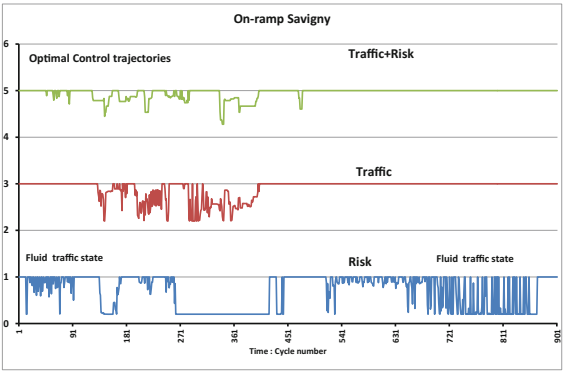


Fig. 8. Control trajectories by scenarios

Screening the control variable trajectory results, the following remarks can be drawn:

- Optimizing the Risk index only (*scenarios*1), the convergence speed of the optimization is slow compared to the other scenarios. On the other hand, the ramp metering control is always active in all traffic situations: fluid or congested. This behavior is observed also for the other on-ramps. Consequently, implementing this scenario in the Traffic Management System will not be satisfactory for the traffic managers and not acceptable by the users. In particular, controlling all the on-ramps when the traffic conditions are fluid does not have sense. Nevertheless, the safety index is dramatically improved.
- In case of scenarios (2) where the traffic index is optimized, the time evolution of control variables is correct and corresponds to real activation in the field of the ramp metering technique. When the traffic is fluid no control is applied whereas in congestion, the ramp metering is activated until the end of the congestion.
- In case of scenarios (3) where both Traffic and safety indices are optimized, the control variables are more smoothed compared to the other scenarios. However, the ramp metering is started earlier and finishing little bit late compared to the scenarios (2). In this case a trade-off is found between both indices and the *Pareto* principle is satisfied.

5.2 Global Performance Analysis

Table 2 includes all computed performance indices and the gain compared to the No Control (NC) case. In particular, the main considered performance index is the cumulated Total Time Spent (TTS) in the system over the time Horizon of the simulation. As indicated above, optimizing the Risk index only leads to a dramatic improvement of the safety index. The obtained gain is around 39% whereas the TTS index is improved by 12%. The same remark can be drawn in case of scenarios 2 where the TTS index only is optimized. The improvement of the TTS index is around 27% whereas the Risk is improved by only 10%. However, in case of multiobjective optimization scenario, the obtained optimal solutions are located in between both scenarios: (1) and (2). In this tested scenario, both indices are improved. As a matter of fact, the gain obtained are comprising between the optimization of the TTS (scenarios (2)) and the Risk (scenarios (1)). These solutions are subject to the Pareto-optimal multiobjective conditions.

Screening the time evolution of the TTS index depicted in Fig. 9, in case of the optimized multiobjective functions (traffic and safety), the time evolution of TTS is located between both scenarios: (1) and (2). At the beginning of the congestion, the optimization of the TTS index, gives better results than all scenarios. On the contrary, the optimized risk index only corresponds to worse solution with respect to the TTS index. This is due to the generation of long queue at the on-ramps leading to the increase of the waiting time and the TTS at each controlled on-ramps.

Table 2. Global performance of the 3 scenarios

	NC		OASIS		Gain(%)	
Optimized index	TTS(vh*h)	Risk	TTS(vh*h)	Risk	TTS(vh*h)	Risk
Traffic	16179	9615	11726	8688	**27**	**10**
Risk	16179	9615	14108	5910	**12**	**39**
Risk+Traf	16179	9615	12386	8191	**24**	**15**

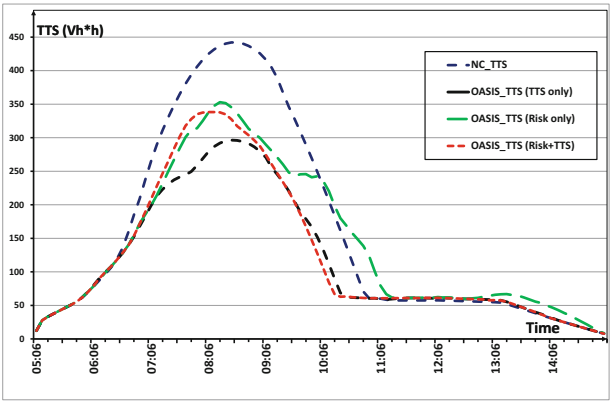


Fig. 9. TTS time evolution by scenarios

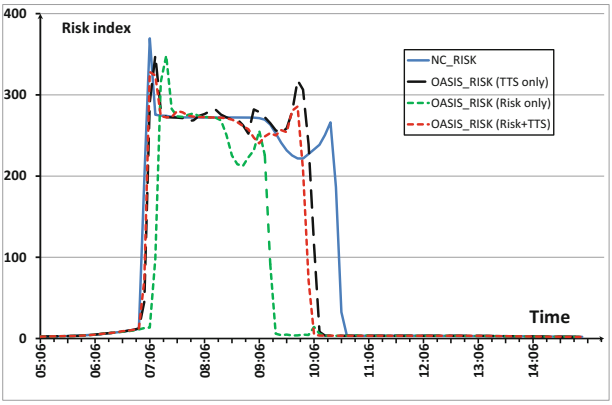


Fig. 10. Risk time evolution by scenarios

Figure 10 depicts the time evolution, on 6 min time slice, of the cumulated risk index of the considered axis for each scenario. The No-Control case is considered as a reference. Compared to the time evolution of the TTS index indicated in Fig. 9, similar time evolution of the Risk index is observed. In case of scenario (2) where the risk only is optimized, a delay is observed at the beginning of the

congestion and the reduction of the level of the risk at the end of congestion. This is due to earlier starting time of the control which leads to generation of long queue at the on-ramps. In case of optimizing of Risk and TTS, the time evolution of the Risk index is in between both scenario (1) and scenario (2).

6 Conclusion and Next Steps

During the calibration and validation steps of the risk index on the test site, the obtained results are very promising. The behavior of the risk index is satisfactory with respect to the real occurrence of the accidents. Applying the risk index model on two real occurred accidents on A6W, the risk index is maximal before the occurrence of both selected accidents. In particular, the same parameters found on the ringway of Paris are used. Consequently, the risk index model can be applied, without any calibration efforts, to any motorway networks. This is a relevant result for the real time application.

The application of the non-linear multi-objective optimization approach using "OASIS" for coordinated ramp metering is very promising in terms of performance and computational effort. The obtained results indicate that the improvement of both indices TTS and $Risk$ is largely acceptable when the multi-objective cost function is optimized. Compared to the optimization of the $Risk$ or the TTS only, the optimization of multiobjective cost function results are globally better.

In this study, the considered time horizon of the optimization is equal to the duration of the simulation run. In consequence, the applied strategy is in open loop control. In order to close the loop, the hierarchical control scheme (rolling horizon) will be applied. In OASIS, the rolling Horizon technique is already developed. The next step will consist in undertaking the results analysis in case of close loop control.

On the other hand, in this paper, only coordinated ramp metering is considered. The next step consists in expanding this approach in order to make integrated control using the same multi-objective cost function. The integration control corresponds to the optimization of the global multi-objective cost function by taking into account simultaneously several types of control strategies such as ramp metering, Motorway-to-motorway control and main-lane speed limits etc.

References

1. Abdel-Aty, M., Pande, A.: Identifying crash propensity using specific traffic speed conditions. J. Saf. Res. **36**, 97–108 (2005)
2. Arrow, K.J.: Social Choice and Individual Values. Wiley, New York (1951)
3. Bhouri, N., Haj Salem, H.: Isolated vs. coordinated ramp metering: field evaluation results of travel time reliability and traffic impact. Transp. Res. Part C (TRC) **28**, 155–167 (2011)
4. Dilmore, J.: Implementation strategies for real-time traffic safety improvement on urban freeways. MA thesis, University of Central Florida, Orlabdo (2005)

5. Elloumi, N., Haj-Salem, H., Papageorgiou, M.: A macroscopic modelling tool for urban corridor. TRISTAN II: TRIennal Symposium on Transportation Analysis, Capri-Italy, Congress Center, DJune, 23–28 1994, vol. I, pp. 135–149 (1994)
6. EURAMP Project: Deliverable D3.5, Safety Critical Issue For Ramp metering, European commission, Brussels (2006)
7. Everitt, B.: Cluster Analysis. Heinemann Educ Books, London (1974)
8. Fletcher, R.: Practical Methods of Optimization, 2nd edn. Wiley, Chichester (1987)
9. Haj Salem, H.: Summary of field results. In: Transportation Research Record (TRR) N1603, pp. 90–98. National Academic Press, Washington DC (1997)
10. Haj Salem, H., Papageorgiou, M.: Ramp metering impact on urban corridor traffic: field results. Transp. Res. A **29 A**(4), 303–319 (1995)
11. Haj-Salem, H., Farhi, N., Lebacque, J.P.: Risk index model building: application for field ramp metering safety evaluation on urban motorway traffic in Paris. In: Bulgarien Academy of Sciences, Cybernetics and Information Technologies, Sofia, vol. 13(4) (2013). Print ISSN: 1311–9702; Online ISSN, pp. 1314–4081, doi:10.2478/cait-2013-0051
12. Haj-Salem, H., Ramananjoana, C., Lebacque, J.P.: Risk index model for real-time prediction of potential traffic crash on urban motorway. In: Proccedings of the EWGT2006, Advances in Traffic and transportation Systems Analysis, BARI, Italy (2006)
13. Kaplow, L., Shavell, S.: Fairness versus Welfare. Harvard Univ. Press, Cambridge (2002)
14. Lebacque, J.P.: Two-phase bounded acceleration traffic flow model: analytical solutions and applications, TRB 82nd annual meeting. TRR **1802**, 220–230 (2003)
15. Lebacque, J.P., Khoshyaran, M.M.: A variational formulation for higher order macroscopic traffic flow models of the GSOM family. In: ISTTT-20 Conference (2013)
16. Lee, C., Hellinga, B., Saccomanno, F.: Evaluation of variable speed limits to improve traffic safety. Transp. Res. Part C **14**, 213–228 (2006)
17. Mammar, S., Lebacque, J.P., Haj-Salem, H.: Resolution of the Aw Rascle & Zhang macroscopic traffic flow model. In: Proceedings of the IMA Conference, London (2005)
18. Papageorgiou, M., Marinaki, M.: A Feasible direction algorithm for the numerical solution of optimal control problems. Internal report N1995-4, Technical University of Crete, Greece (1995)
19. Riedmiller, M., Braun, H.: The RPROP algorithm. In: Proceedings of the IEEE International Conference on Neural Networks (ICNN 93)
20. Thomas Glob, F., Recker, W.W.: Method for relating type of crash to traffic characteristics on urban freeway. Transp. Res. Part A **38**, 55–80 (2002)

Drivers' Behavior at Bicycle Crossroads

Francesco Bella(✉) and Manuel Silvestri

Department of Engineering, Roma TRE University, Rome, Italy
{francesco.bella,manuel.silvestri}@uniroma3.it

Abstract. The main objective of this study was to examine the effect on drivers' behavior of several countermeasures at bicycle crossroads during the driver–cyclist interaction (a cyclist that crosses the road) by the use of a driving simulator. Three conditions of cyclist crossroads were investigated: baseline condition (no treatment), raised island and the colored paved markings. Forty-two participants drove a suburban scenario with several bicycle crossroads having different configurations and presence/absence of cyclist. 378 speed profiles were plotted from 150 m before each crossroads. The analysis revealed that the driver's speed behavior was fully consistent with the behavioral models in literature. A further analysis was performed on driver's speed by ANOVA. Results showed that there was a main effect of the countermeasures, which determined lower speeds compared with that of the baseline condition. In addition, driver behavior was actually characterized by the specific peculiarities of the two countermeasures and both determined a better yielding behavior. However, the outcome of the questionnaire to participants showed that colored paved markings was considered as the most effective in terms of driving aid.

Keywords: Road safety · Driver's behavior · Driving simulator · Cyclist crossroad

1 Introduction

The cyclists on the roads are continuously increased in last years and this needs to be taken into account in the mobility and traffic studies, as well as the potentialities and the risks related to the presence of these road users have to be considered. One of the main problems is the lack of adequate paths, which should ensure high safety levels during the travel. Moreover, a crucial point for the safety of cyclists is the crossroad, which is the point where the interactions with the motorized vehicles are more significant. The cycle tracks in fact are more safe along road sections and more dangerous at intersections [23]. European accident statistics show that 55% of the cyclist fatalities occur in urban areas [13]. In addition, 31% of cyclist fatalities occur at intersections and about 56% of these fatalities are located at crossroads [13]. Despite the problem is relevant the efforts to make safer the driver – cyclist interaction in urban and suburban areas seem to be not enough. It is widely known that reduce the risks connected to the driver - cyclist interaction is linked to a good cycling infrastructure [e.g. 25, 29], which includes the bicycle crossroads. However, the critical issues of the driver – cyclist interaction at crossroads

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_17

are attributable both to driver's behavior and to the poor layout of the crossroad, which does not always allow the driver to yield to the cyclist. For example, a study of Wood et al. [30] found that visibility is the main factor to which cyclists and car drivers attribute crashes involving cyclists. As reported in Kim et al. [19], also the speed is the factor that contributes to serious and fatal outcomes in vehicle – bicycle accident, while Minikel [20] highlights that higher vehicles' speed, higher traffic volumes, and the presence of heavy vehicles are detrimental to cyclist safety. Concerning the driver behavior, the paper of Räsänen and Summala [24] shows the yielding decision process of vehicle with the consideration of vehicles detecting vision and time. It suggests that adequate vehicles approaching speed to intersections have a positive effect on their yielding decision to crossing cyclists. The driver yielding behavior plays an important role in the interaction with the cyclist. A theoretical framework of the decision process that occurs during driver-cyclist interaction is provided by the "Threat Avoidance Model" developed by Fuller [14]. This model of the driver's behavior proposes that within the context of the motivation for a particular journey (usually a specific destination within a specific period of time), the driver behavior is focused on the avoidance of averse or potentially adverse stimuli in the road-traffic environment [15]. This model implies that when confronted with a discriminative stimulus for a potential aversive event, what driver does depend specifically on the rewards and punishments for alternative responses. In a vehicle–cyclist interaction at a bicycle crossroad, the cyclist presence is the discriminative stimulus. Such an adverse stimulus can cause: (a) an "anticipatory avoidance response" or (b) a "non-avoidance response". In the first case, the driver considers the cyclist presence to be a "threat", and then, he slows down; in this way, the cyclist can pass before the driver. In this case, the driver is "punished" with a loss of time. In the second case, the driver maintains the same speed because he considers the cyclist presence to be a "threat" but chooses a "non–avoidance response", signaling to the cyclist that he has no intention to yield; then, two possible conditions could occur: (a) the driver passes first or (b) he is forced to brake. The first action is a "reward" for the driver because he does not stop and, thus, does not suffer delay; in the second case, the cyclist assumes a "competitive behavior", and therefore, the driver is forced to a delayed avoidance response (braking) or a collision occurs. Finally, this model suggests that the driver can experience a "no discriminative stimulus" (he does not see the cyclist), and therefore, he does not expect a "threat". In this case, two possible conditions could also occur: (a) the interaction with the cyclist does not cause a risk (the cyclist does not cross) or (b) a delayed avoidance response is required to avoid an accident. In other works, Silvano et al. [26, 27] proposed a hierarchical discrete model to explore the main factors of drivers' yielding behavior process in the vehicle – cyclist interactions. The assumption is that in order to avoid the collision due to the conflicting trajectories (driver and cyclist are arriving at the conflict zone, which is the common crossing area where collision occurs if both of them maintain their current trajectories), a "negotiation" begins upstream the conflict zone. In this situation, the driver has the responsibility to yield; however, when the driver perceives the conflict with the cyclist he can decide, at a certain point, if yield or not. On overall, the decision process of the driver supposes two levels: (a) a conflict event is perceived or (b) a yield event is observed. Several factors can

contribute to the yielding process, as the actual possibility to stop the vehicle to yield, the actual possibility to detect the cyclist, the attitudes of the driver.

According with the literature, interaction between vehicles and cyclists at bicycle crossroads are critical situations, in which the driver has to be influenced in order to adapt his speed if the cyclist is present, avoiding the occurrence of accidents and improving the cyclist safety. For this reasons, this paper aims at: (a) analyzing the driver behavior at bicycle crossroads in the presence of several countermeasures, which are aimed to allow the driver to adapt his speed and, thus, improve the ability of driver to yield to cyclist; (b) comparing the countermeasures and identifying the most effective treatment for bicycle crossroads.

Because of the importance of the speed in the driver – cyclist interaction, the countermeasures at bicycle crossroads have to highlight the presence of the crossroads and induce the driver to slow down. Two countermeasures were analyzed: colored paved markings and raised island. The first is a red painting of the bicycle crossroad and its aim is to point out the crossroads and focus the attention of the approaching driver on it, contributing at the speed reduction [18]. The second is a physical facility in the middle of the road that determines a narrowing of the lane and induces the driver to adapt his speed [10]. Accomplishing these aims was possible by the use of a driving simulator that, mainly, allows risk avoidance for the experimenters and full control of the experimental conditions.

2 Method

This study was conducted using the advanced driving simulator of the Department of Engineering – Roma Tre University. Several studies have demonstrated that driving simulators are useful tools for the evaluation of the driver's behavior as induced by the road configurations and traffic conditions [e.g. 4–7, 8, 9]. Moreover, driving simulators are ideal tools for studies whose field survey is made impossible by the implicit high risks that the experimenters would be subjected to and the difficulty of ensuring controlled experimental conditions. More specifically, concerning the study of the interaction between vehicle and cyclist, a recent driving simulator study [12] was carried out to analyze how the number of cyclists, the cyclist's approach direction, and the cyclist's action affect the speed and mental workload of drivers approaching rural intersections.

In the present study a multi – factorial experiment was designed to analyze the effect of three configurations of bicycle crossroads (a baseline condition and two countermeasures) and driver – cyclist interaction conditions on drivers' behavior while approaching the bicycle crossroads.

2.1 Road Scenario

The experimental road scenario was a two – lane suburban road about 7.6 km long in which also the cyclist that crosses the road was simulated. To ensure the same approaching condition, 9 signalized intersections were placed in advance of each bicycle crossroad. Each driver was obligated to stop at the signalized intersection, due to the red

light that turned on when the driver was at approximately 100 m from the intersection. The distance between the signalized intersection and bicycle crossroad was equal to 400 m, which allowed the drivers to reach a congruous speed for the simulated scenario. The posted speed limit was 50 km/h while the cross-section was 10 m wide, formed by two 3.00 m wide lanes, two paved shoulders 0.50 m wide and two curbs 1.50 m wide, according the Italian road design guidelines [22]. According to the Italian Highway Code [21], the bicycle crossroad was 1.50 m wide and three vertical signs were posted: one at 150 m in advance of the bicycle crossroad to signal the presence of it. One at 10 m, to advice the driver to yield if the cyclist was present and one in correspondence of the bicycle crossroad. This configuration represents the baseline condition (Fig. 1a). In addition to the baseline condition, two types of countermeasures were placed in the scenario: colored paved markings and raised island. The first was the red painting of the bicycle crossroad (Fig. 1b), while the second was a physical facility 40 m long, 0.50 m wide and 0.05 m height (Fig. 1c), implying the narrowing of the vehicle lane to 2.75 m.

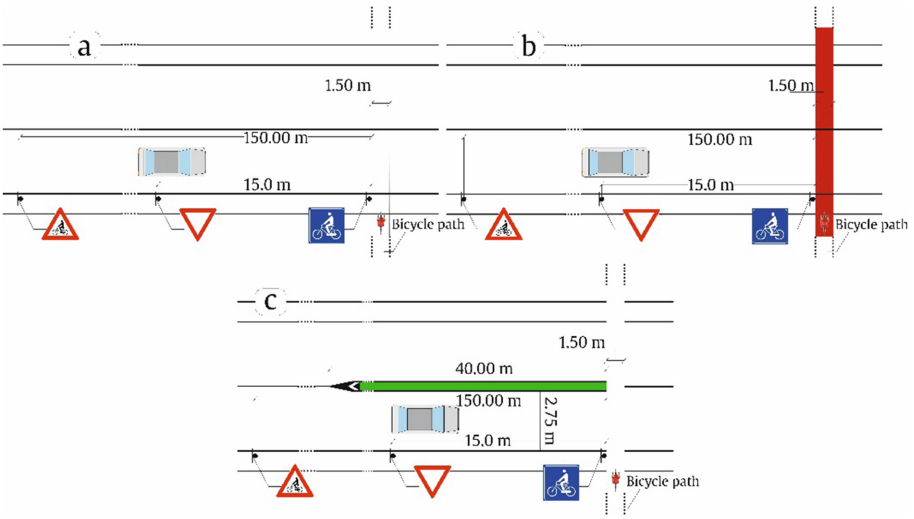


Fig. 1. (a) Baseline, (b) Colored paved markings, (c) Raised island

A cyclist coming from driver right side passed through such bicycle crossroads. The cyclist was triggered at 20 m from the collision point when the driver was at 50 m from it. According to several data in literature [11, 16, 28], the speed of the cyclist was 20 km/h. Assuming the driver speed equal to 50 km/h, in such condition the driver and cyclist have the same time to arrive (3.6 s) to the collision point.

It should be noted that this condition is representative of a theoretical driver – cyclist interaction, which occurs only if the driver adopts the hypothesized speed value.

Because of the high variability of the drivers' speed while approaching the bicycle crossroad and with the aim to ensure adequate sample sizes for the actual conditions of driver-cyclist interactions recorded during the simulated drives, the cyclist presence was repeated two times for each of the three configurations of the bicycle crossroad.

To avoid a potential effect of the order on the driver's behavior, 3 road scenarios that have a different sequence of the 9 combinations of bicycle crossroad (baseline condition, raised island and colored paved markings) x cyclist (cyclist absence and two conditions in which the cyclist was crossing the road) were implemented in the driving simulator. Each scenario was driven by one of the 3 groups into which the participants were divided (see next section on participants).

2.2 Driving Simulator

The driving simulator of the Department of Engineering – Roma Tre University is an interactive fixed-base driving simulator. It was previously tested, calibrated, and validated [2, 3] as a reliable tool for the study of the driver's speed behavior. The hardware interfaces (wheel, pedals and gear lever) are installed on a real vehicle. The driving scene is projected onto three screens: one in front of the vehicle and one on either side, which provide a 135° field of view. The resolution of the visual scene is 1024 × 768 pixels with a refresh rate of 30 to 60 Hz. The system is also equipped with a sound system that reproduces the sounds of the engine.

2.3 Procedure

The experiment was conducted with the free vehicle in its own driving lane. In the other driving lane, a slight amount of traffic was distributed to induce the driver to avoid driving into that lane. The simulated vehicle was a standard medium-class car with automatic gears. The data recording system acquired all of the parameters at spatial intervals of 2 m. The driving procedure consisted of the following steps: (a) communicating to the driver about the duration of the driving and the use of the steering wheel, pedals, and automatic gear; (b) training at the driving simulator on a specific alignment with a length of approximately 5 km; (c) filling in a form with personal data, years of driving experience, average annual distance driven; (d) driving one of the three road scenarios with a specific bicycle crossroads-cyclist sequence; (e) filling in of a questionnaire about the discomfort that is perceived during driving, to eliminate from the sample driving performed under anomalous conditions. This questionnaire consisted of 5 questions, with each question addressing a typed of discomfort: nausea, giddiness, daze, fatigue, other. Each question could be answered by a score of 1–4 in proportion to the level of discomfort experienced: null, light, medium, and high. The null and light level for all four types of discomfort is considered to be the acceptable condition for driving; (f) filling in of a questionnaire about how the driver perceived the countermeasures. For this question the participant could choose between (a) driving aid and (b) obstacle; for each of them, the driver could answer by a score from 0 to 3 in proportion to the perceived effectiveness. Drivers were instructed to drive as they normally would in the real world.

2.4 Participants

Forty-two drivers (24 men and 18 women), whose ages ranged from 24 to 59 (mean = 29.3; SD = 8.5) and who had regular European driving licenses for at least three years were selected to perform the driving in the simulator (mean = 11.0; SD = 8.0). They were chosen from students and staff of the University and volunteers from outside of the University. The drivers had no prior experience with the driving simulator and had an average annual driven distance of at least 2500 km. According to the questionnaire on perceived discomfort, all of the participants experienced null or light levels of discomfort. Thus, the sample used for the analysis consisted of all 42 drivers, which were divided into 3 groups; the 3 groups drove different scenarios, which were each characterized by a specific sequence of bicycle crossroad/cyclist.

3 Data Processing

To analyze the drivers' speed behavior while approaching the bicycle crossroads, the speed data were recorded starting from 150 m in advance of each one of the 9 bicycle crossroads. The speed profile of each driver was plotted for each selected section. A total of 378 speed profiles were plotted (42 drivers × 9 bicycle crossroads). To analyze the driver behavior in the interaction with the cyclist, the variable TT_{arrive} (in s) was determined as

$$TT_{arrive} = \frac{L_{Vi}}{V_i} \tag{1}$$

where:

- V_i: the driver's initial speed value (in m/s), identified at the moment in which the driver starts to decrease his speed, releasing the accelerator pedal;
- L_{Vi}: the distance (in m) from the bicycle crossroad where the V_i value is located.

Therefore, TT_{arrive} is the time left for the driver to arrive at the bicycle crossroad in the moment in which he started to decelerate, regardless of the cyclist position. It should be noted that TT_{arrive} implicitly determined a classification of the interactions based on the driver's characteristics. Drivers with low "availability" to yield determined low TT_{arrive} because they tended to start to slow down when they were close to the bicycle crossroad and/or from high initial speeds. Drivers with high "availability" to yield, instead, determined high values of TT_{arrive} because they tended to start to reduce the speed when they were far from bicycle crossroad and/or from low initial speeds.

Speed profiles also showed several events when drivers did not yield because they accelerated to pass the conflict point before the cyclist. However, no case of collision was recorded. Table 1 shows, for the 3 configurations of bicycle crossroad, the number of driver-cyclist interactions, the mean, maximum and minimum values of TT_{arrive}, the number of driver-cyclist interactions for several groups of values of TT_{arrive} and the number of interactions where the drivers did not yield.

Table 1. Driver-cyclist interactions recorded at the driving simulator

Counter measures	No. of driver-cyclist interaction **(a)**	TT_{arrive} [s]						No. of failed yields **(b)**
		mean [s]	max [s]	min [s]	$TT_{arrive} \leq 3$	$3s < TT_{arrive} \leq 4s$	$TT_{arrive} > 4$	
Baseline cond.	72	3.6	12.7	0.3	32	19	21	12
Raised island	79	3.8	9.4	1.3	30	25	24	5
Colored paved markings	78	4.2	11.7	1.9	21	22	35	6
Total (a + b)	**252**							

The speed values located at 60 m, 40 m and 20 m before the collision point were also used for the analysis aimed at highlighting the effectiveness of the countermeasures in terms of driver speed adaptation for the conditions of absence and presence of a cyclist.

4 Data Analysis and Results

A first analysis was based on the mean speed profiles for different groups of TT_{arrive} values. The purpose was to analyze the driver's speed adaptation while approaching the bicycle crossroad under different driver's characteristics (i.e. different values of TT_{arrive}) and how this influence occurs under different configurations of bicycle crossroad.

A further analysis was based on the speed values located along the approaching section at 60 m, 40 m and 20 m before the collision point to assess the effectiveness of the countermeasures in terms of driver speed adaptation for the conditions of absence and presence of a cyclist. More specifically, an analysis of variance (ANOVA) was performed to investigate the effects on the dependent variable (speed along the approaching section) due to the three factors: countermeasures (baseline condition, raised island, and colored paved markings), driver - cyclist interaction conditions (presence, absence of the cyclist) and point of speed recording (60 m, 40 m and 20 m before the collision point).

4.1 Mean Speed Profiles

Mean speed profiles were plotted for each countermeasure, for 3 groups of TT_{arrive} values and for the cyclist absence condition (Fig. 2). The mean speed profiles for the "no cyclist" condition were obtained from the events in which the driver, while approaching the bicycle crossroad, did not encounter the cyclist (42 data for each countermeasure).

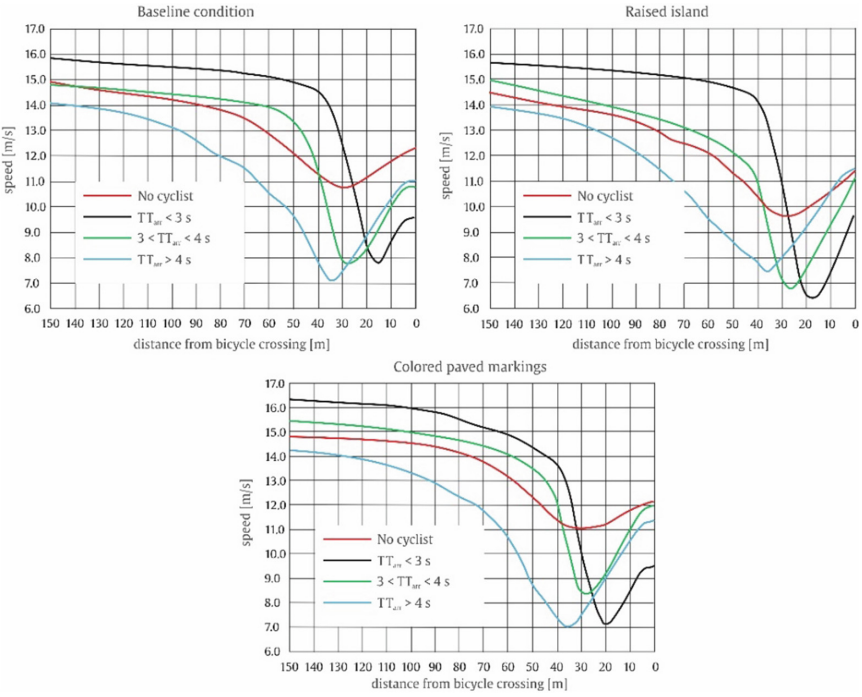


Fig. 2. Mean speed profiles (a) Baseline (b) Raised island (c) Colored paved markings

For all the countermeasures and for $TT_{arrive} < 3$ s, the speed profile was higher than those for higher values of TT_{arrive}. However, in the last section in advance of the bicycle crossroad, approximately 40 m in advance of the bicycle crossroad, the driver decreased the speed abruptly from approximately 14.5 m/s to 7 m/s because he had to yield to the cyclist that is crossing. The minimum speed value was reached at 20 m from the bicycle crossing for the raised island and the colored paved markings, while for the baseline condition it was located at 15 m from the bicycle crossing. Moreover, the minimum speed values were approximately 6.5 m/s for the raised island and the colored paved markings, while for the baseline conditions was 7.8 m/s.

For all the countermeasures and for $3 \text{ s} < TT_{arrive} < 4$ s, the speed profiles showed that the speed values were lower than those for the condition of $TT_{arrive} < 3$. The beginning point of the abrupt speed reduction was located farther from the bicycle crossroad; for the baseline condition, it was approximately at 50 m from the bicycle crossroad, while for the raised island and the colored paved markings it was located at approximately 40 m. However, the speed value at which the deceleration begins was higher for the baseline condition (approximately 13.7 m/s) than that for the raised island and the colored paved markings (approximately 12 m/s and 13 m/s, respectively). The minimum speed values were located at approximately 30 m from the bicycle crossing for the baseline condition and the colored paved markings, while it was located at approximately 25 m from the bicycle crossing for the raised island. However, the minimum

speed values were higher for the baseline condition and the colored paved markings (approximately 8 m/s) than that for the raised island (approximately 7 m/s).

For all the countermeasures and for $TT_{arrive} > 4$ s, the speed profiles were lower than those for the other TT_{arrive} values. The speed reduction occurred gradually, starting from more than 90 m away from the bicycle crossroad. The corresponding speed value was approximately 13 m/s. The minimum speed was located at approximately 35 m from the bicycle crossing and the values were approximately 7 m/s.

For the condition of no cyclist and for all the countermeasures a gradual speed reduction was observed, starting from a speed value approximately of 14 m/s until the minimum speed value. The minimum speed was reached at approximately 30 m from the colored paved markings and the baseline condition, while it was located at approximately 25 m for the raised island. However, the minimum speed value was lower for the raised island (approximately 9.6 m/s) than those for the baseline condition and colored paved markings (10.7 m/s and 10.9 m/s, approximately).

4.2 Driver's Speed Behavior

The interaction and main effects on the speed along the approaching section due to the independent factors were analyzed with ANOVA. The Bonferroni correction was used for multiple comparisons. For the analysis, SPSS (Statistical Package for Social Science) software was used.

Main Effects. ANOVA revealed that there was a statistically significant (al level of 5.2%) main effect of the countermeasures ($F_{2,1116} = 2.919$, $P = 0.052$, partial Eta squared = 0.005, observed power = 0.570). Pairwise comparison showed that the driver speeds for the colored paved markings and the raised island (10.9 m/s and 10.8 m/s, respectively) were less than that for the baseline condition (11.5 m/s). This outcome is consistent with the aim of the countermeasures, which provide a better information about the presence of the bicycle crossroad helping him to adapt the speed.

Statistical analysis showed that there was a significant main effect of the distance from the bicycle crossroad ($F_{2,1116} = 73.731$, $P = 0.000$, partial Eta squared = 0.117, observed power = 1.000). Pairwise comparison showed that the speed at 60 m from bicycle crossroad (12.9 m/s) was statistically significantly higher than that at 40 m from bicycle crossroad (mean difference = 1.6 m/s; $P = 0.000$) and at 20 m from bicycle crossroad (mean difference = 3.8 m/s). Moreover, the speed at 40 m from bicycle crossroad was statistically significantly higher than that at 20 m from bicycle crossroad (mean difference = 2.2 m/s; $P = 0.000$). This result was expected because the driver, while approaching the bicycle crossroad, reduced his speed in order to detect the presence of the cyclist.

ANOVA showed that there was a significant main effect for the cyclist conditions ($F_{1,1116} = 9.950$, $P = 0.002$, partial Eta squared = 0.009, observed power = 0.883). Moreover, the speed for the condition of cyclist presence (10.8 m/s) was lower than that for the condition of cyclist absence (mean difference = 0.8 m/s; $P = 0.002$). This result showed that the driver was affected by the presence of the cyclist and, thus, reduced the speed to give way.

Interaction Effects. A significant interaction effect was found for distance from the bicycle crossroad by driver – cyclist interaction ($F_{2,1116} = 13.038$, $P = 0.000$, partial Eta squared = 0.023, observed power = 0.997). No other significant interaction effect was found. Results showed that the driver's behavior was affected by the presence of the cyclist in the section between 40 m to 20 m from the bicycle crossroad. At 60 m and at 40 m from bicycle crossroad, in fact, the driver's speed is almost the same among the cyclist conditions (Fig. 3a). Therefore, in the section nearest the bicycle crossroad the cyclist condition influenced the driver's behavior. In the section between 40 m to 20 m from the bicycle crossroad the speed reduction was 3.0 m/s when the cyclist was present, while for the condition of cyclist absence was 0.5 m/s. Thus, when the cyclist was present the driver decreased the speed in order to yield to the cyclist. For the condition of no cyclist the driver did not perceive the presence of the cyclist and, thus, he maintained approximately the same speed.

It should be noted that, even if not statistically significant, the interaction effect distance from bicycle crossroad by countermeasure ($F_{4,1116} = 0.492$; $P = 0.742$) provided interesting results (Fig. 3b).

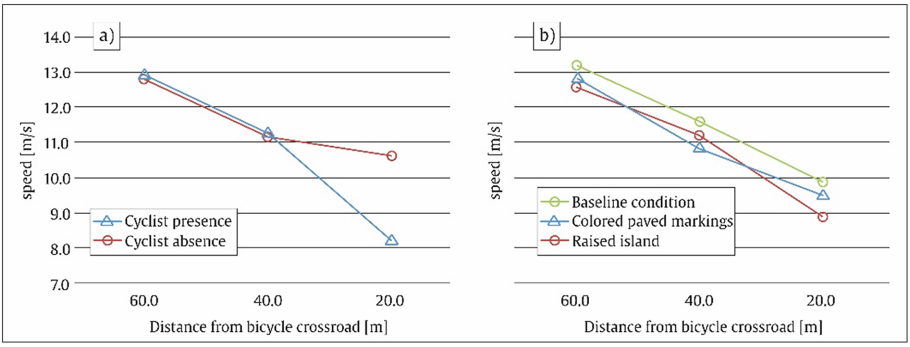


Fig. 3. Interaction effects: (a) distance from bicycle crossroad by cyclist condition (b) distance from bicycle crossroad by countermeasure

When the driver was at 60 m from bicycle crossroad, he was not influenced by the colored paved markings and raised island (mean driver speed of 12.9 and 12.6 m/s, respectively) because he was far from the bicycle crossroad. Passing from 60 m to 40 m in advance of the bicycle crossroad, the driver reduced the speed more for the colored paved markings (2.0 m/s) than the raised island (1.4 m/s). From 40 m to 20 m in advance of the bicycle crossroad, the driver continued to decrease the speed but to a lesser extent for colored paved markings. This points out that:

- the colored paved markings provides an advanced information to the driver about the presence of a bicycle crossroad, allowing him to decrease (with higher rates) the speed since 60 m in advance the bicycle crossroad;
- the raised island, instead, induces a driver speed reduction that is more pronounced in the section between 40 m and 20 m, where it is physically present; in this section,

the countermeasure determines a reduction of the lane width, which produces a greater influence in terms of speed reduction.

In other terms, it seems that the driver's behavior is actually affected by the specific peculiarities of the two countermeasures (better visibility and information to the driver for the colored paved markings and greater "physical" influence for the raised island in the section nearest to the bicycle crossing).

4.3 Outcome of the Questionnaire

The results of the questionnaire show that 71% of the drivers (30 of 42) reported the highest level of driving aid for the colored paved markings, while only 19% of the drivers (8 of 42) reported the same level of driving aid for the raised island. For the baseline condition, the highest level of driving aid was reported by 40% of the drivers (17 of 42). Consistently with this outcome, 79% (33 of 42) of the drivers reported the lower score for the obstacle driving effect of the colored paved markings, while only 31% (13 of 42) of the drivers reported the lower score for the obstacle driving effect of the raised island. For the baseline condition, the lower score for the obstacle driving effect was reported by 76% of drivers (32 of 42).

5 Discussion

5.1 Mean Speed Profiles

As reported in Table 1, the analysis of drivers' mean speed profiles showed that the yielding compliance was increasing with the raised island (in 6.3% of interactions, 5 of 79, drivers did not yield) and the colored paved markings (in 7.7% of interactions, 6 of 78, drivers did not yield) compared to the baseline condition (in 16.7% of interactions, 12 of 72, drivers did not yield), which is consistent with the aim of improving the driver – bicyclist condition through the use of countermeasures at bicycle crossroads.

As expected, the analysis of the mean speed profiles revealed that the driver's speed behavior was affected by different groups of TT_{arrive} values and therefore by different drivers' characteristics. In fact, for all the countermeasures, the mean speed profiles highlighted that:

- drivers started to decrease the speed from lower values with the increase of TT_{arrive}
- drivers adopted a less abrupt speed reduction with the increase of TT_{arrive}

More specifically, for all of the countermeasures and for $TT_{arrive} < 3$ s, the driver was approaching the bicycle crossroad with high speed values and adopted the most abrupt speed reductions. This behavior highlights a low "availability" of the driver to yield (or a certain driver's aggressiveness). The driver would have the priority at the bicycle crossroad, and thus, he maintains the same speed until he is close to the conflict point; then, he is forced to brake to avoid hitting the cyclist.

For $3 \text{ s} < TT_{arrive} < 4$ s the driver adopted lower speed and less abrupt speed reductions than those observed for $TT_{arrive} < 3$ s. This behavior reveals that the driver realizes that

he cannot pass before the cyclist and starts to decelerate farther from the bicycle crossroad. This behavior was more accentuated for $TT_{arrive} > 4$ s, where the driver adopted the lower speeds and the less abrupt speed reductions highlighting a careful behavior.

It should be noted that these drivers' behaviors are completely consistent with the "Threat Avoidance Model" developed by Fuller, the hierarchical discrete model of Silvano et al. [26, 27] and with the results of Bella and Silvestri [7]. In particular, the behavior observed for $TT_{arrive} < 3$ s can be related to the "non-avoidance response". The driver, in fact, maintains the same speed because he considers the cyclist presence to be a "threat" but chooses a non-avoidance response, signaling to the cyclist that he has no intention to yield. This "negotiation" phase, however, is always won by the cyclist (into the simulated scenario, the cyclist was set to cross regardless of the driver's behavior) and the driver is forced to a delayed avoidance response (braking) to avoid the collision.

The behavior observed for $TT_{arrive} > 4$ s (and to a lesser extent, also that for 3 s $< TT_{arrive} < 4$ s) can be related, instead, to the case of "anticipatory avoidance response". The driver considers the cyclist presence to be a "threat" and he slows down; in this way, the cyclist can pass before the arrival of the driver and a yield event is observed.

Only slight differences were observed among the countermeasures for different values of TT_{arrive}. However, the slight differences that were revealed from the analysis of the mean speed profiles did not highlight a clear trend that enables to express considerations on the induced effects by the several countermeasures.

A more evident difference between the countermeasures were observed for the cyclist absence condition. For this condition, the minimum speed reached by the driver was lower for the raised island (9.6 m/s) compared with that for baseline condition (10.7 m/s) and colored paved markings (10.9 m/s). This relationship was expected because the raised island causes a physical narrowing of the driving lane, inducing the driver to adopt a lower speed.

5.2 Drivers' Speed Behavior

Statistical analysis showed that the effect of the countermeasures on driver's speed was statistically significant (al level of 5.2%). The mean speeds for the raised island and the colored paved markings were lower than that for the baseline condition. The observed effects of the countermeasures on drivers' speed behavior were similar with those reported by Hallmark et al. [18] and Campbell et al. [10] for the colored paved markings and by Bared et al. [1] and Gross et al. [17] for the raised island. These countermeasures highlight and advise the driver of the presence of a bicycle crossroad and, thus, the possible presence of a cyclist that could cross. In this way, the driver can anticipate the maneuver to reduce the speed and make safer the interaction with the cyclist.

The effect of the distance from the bicycle crossroad was statistically significant on drivers' speed behavior. The result highlighted that the driver adapted his speed while approached the bicycle crossroad in order to yield to the cyclist. This outcome is consistent with the behavioral model developed by Fuller [15] and with the hierarchical discrete model of Silvano et al. [26, 27]. In particular, the failed "negotiation" upstream the conflict zone due to the behavior of the cyclist (which crosses regardless of the driver

behavior) arises the occurrence of the situation in which the driver has to force the braking and give way to the cyclist to avoid the conflict. Statistical analysis showed a significant effect of the driver - cyclist interaction conditions on drivers' speed behavior. In particular, the cyclist presence induced the driver to reduce the speed in order to yield to the cyclist. For the no cyclist conditions, instead, the driver did not perceive the presence of the cyclist and, thus, maintained a higher speed.

In addition, the results highlighted that there was a significant interaction effect for the distance from the bicycle crossroad by driver – cyclist interaction conditions. On overall, when the driver was farther from the bicycle crossroad (from 60 m to 40 m) he adopted the same speed both for the cyclist absence and presence. This outcome can be explained as the "no discriminative stimulus", because the driver is too far from the bicycle crossroad and, thus, he is not influenced by the possible presence of the cyclist. Near to the bicycle crossroad (from 40 m to 20 m) the driver was influenced by the presence/absence of the cyclist and changed his speed. In particular, the driver adopted higher speed for the condition of no – cyclist. This finding suggests that the "discriminative stimulus" is triggered when the driver is near the collision point. In this condition the driver reacts to the presence of the cyclist by reducing his speed, or by maintain his speed because he has not perceived the cyclist presence. The others interaction effects were not statistically significant.

6 Conclusions

The results based on the mean speed profiles showed that the driver behavior was fully consistent with the behavioral model developed by Fuller [15] according to which the driver could adopt a "non– avoidance response", warning the cyclist of his intention to not give way, or could adopt an "anticipatory avoidance response", slowing down and giving way to the cyclist. Moreover, the outcomes of the driver mean speed profile were consistent with the hierarchical model of Silvano et al. [26, 27], in which the driver starts a "negotiation" with the cyclist upstream the conflicting zone to ensure his priority. However, due to the experiment design, only the case in which the cyclist wins the "negotiation" occurred, which consistently represents the case in which the driver forces the brake to avoid the conflict with the cyclist.

The statistical analysis showed that the effect on drivers' speed due to the countermeasures was statistically significant at level of 5.2%. For the colored paved markings and the raised island lower speeds were recorded (10.9 m/s and 10.8 m/s, respectively) compared with that obtained for the baseline condition (11.5 m/s).

The results on driver speed due to distance from the bicycle crossroad showed that the driver reduced significantly his speed while he approached the bicycle crossroad. This behavior was expected; the driver, in fact, decreases the speed near the bicycle crossroad because he perceives the presence of the cyclist, which induces him to decelerate and give way.

The statistical analysis showed also that the driver – cyclist interaction conditions affected significantly the driver speed. In particular, for the condition of cyclist presence the driver speed was lower (statistically significant) than that for the no cyclist condition.

When the cyclist is not present, the driver does not perceive the possible "treat" of a potential conflict and, thus, he maintains higher speeds. The interaction effect distance from bicycle crossroad by driver – cyclist interaction conditions was statistically significant. Comparing the outcomes for the conditions of cyclist presence and absence, the results showed that the driver behaves differently only in the section near the bicycle crossroad (between 40 m to 20 m from the bicycle crossroad).

Despite the interaction effect of the distance from bicycle crossroad by countermeasure was not statistically significant, the results highlighted the drivers' behavior was actually influenced by the specific peculiarities of the two countermeasures.

In summary, the result of the present study highlights that the countermeasures are effective and influence the driver's behavior consistently with their expected functions. In addition, both countermeasures determined a better yielding behavior compared to that for the baseline condition (see Table 1); therefore they can be considered equivalent in terms of yielding behavior, although the questionnaire revealed a greater acceptance of the colored paved markings compared with that of the raised island. This result was also reasonable, because the colored paved markings is a purely perceptual cue which does not cause a physical obstruction, while the raised island is a "physical" cue and, as such, less "accepted" by the driver.

Further studies might examine other variables of the driver behavior (such as deceleration or reaction time) in order to investigate additional effects produced by the countermeasures, as well as the analysis referred to a group of cyclists that is crossing instead to the single cyclist.

References

1. Bared, J.: Two Low-Cost Safety Concepts for Two-Way STOP-Controlled, Rural Intersections on High-Speed Two-Lane, Two-Way Roadways. Federal Highway Administration. FHWA-HRT-08-063 (2008)
2. Bella, F.: Validation of a driving simulator for work zone design. Transp. Res. Rec. J. Transp. Res. Board **1937**, 136–144 (2005). http://dx.doi.org/10.3141/1937-19
3. Bella, F.: Driving simulator for speed research on two-lane rural roads. Accid. Anal. Prev. **40**, 1078–1087 (2008). http://dx.doi.org/10.1016/j.aap.2007.10.015
4. Bella, F.: New model to estimate speed differential in tangent–curve transition. Adv. Transp. Stud. Int. J. **15**, 27–36 (2008). doi:10.4399/97888548186823
5. Bella, F.: How traffic conditions affect driver behavior in passing maneuver. Adv. Transp. Stud. Int. J. (Special Issue), 113–126 (2011). http://dx.doi.org/10.4399/978885484657911
6. Bella, F.: Driver performance approaching and departing curves: Driving simulator study. Traffic Inj. Prev. **15**, 310–318 (2014). doi:10.1080/15389588.2013.813022
7. Bella, F.: Operating speeds from driving simulator tests for road safety evaluation. J. Transp. Saf. Secur. **6**(3), 220–234 (2014). http://dx.doi.org/10.1080/19439962.2013.856984
8. Bella, F., Silvestri, M.: Effects of safety measures on driver's speed behavior at pedestrian crossings. Accid. Anal. Prev. **83**, 111–124 (2015). http://dx.doi.org/10.1016/j.aap.2015.07.016
9. Bella, F., Silvestri, M.: Driver's braking behavior approaching pedestrian crossings: A parametric duration model of the speed reduction times. J. Adv. Transp. **50**, 630–646 (2016). doi:10.1002/atr.1366

10. Campbell, J.L., Richard, C.M., Graham, J.: NCHRP Report 600 B: Human Factors Guidelines for Road Systems. TRB, Washington, DC (2008)
11. Dozza, M., Schindler, R., Bianchi-Piccinini, G., Karlsson, J.: How do drivers overtake cyclists? Accid. Anal. Prev. **88**, 29–36 (2016)
12. Duivenvoorden, K., Hogema, J., Hagenzieker, M., Wegman, F.: The effects of cyclists present at rural intersections on speed behavior and workload of car drivers: A driving simulator study. Traffic Inj. Prev. **16**, 254–259 (2015). doi:10.1080/15389588.2014.937484
13. ERSO: European Road Safety Observatory. Traffic Safety Basic Facts – Cyclist (2015)
14. Fuller, R.: A conceptualization of driver behaviour as threat avoidance. Ergonomics **27**, 1139–1155 (1984)
15. Fuller, R.: Predicting what a driver will do: Implication of the threat avoidance model of driver behavior–road user behavior. Theory and Research. In: Proceedings of 2nd International Conference on Road Safety. Groningen, Netherlands (1987)
16. García, A., Domènech, A.A., Llorca, C., Gómez, F.A., Ferrer, V.: Effects of road geometry on the interaction between cyclist and vehicles on two-lane rural roads. In: 5th International Symposium on Highway Geometric Design (2015)
17. Gross, F., Jagannathan, R., Hughes, W.: Two low-cost safety concepts for two-way, stop-controlled intersections in rural areas. Transp. Res. Rec. **1280**, 11–18 (2009)
18. Hallmark, S.L., Peterson, E., Fitzsimmons, E., Hawkins, N., Resler, J., Welch, T.: Evaluation of Gateway and Low Cost Traffic Calming Measures for Major Routes in Small Rural Communities. Center for Transportation Research and Education, Iowa State University, Ames (2007)
19. Kim, J.K., Kim, S., Ulfarsson, G.F., Porrello, L.A.: Bicyclist injury severities in bicycle-motor vehicle accidents. Accid. Anal. Prev. **39**, 238–251 (2007). doi:10.1016/j.aap.2006.07.002
20. Minikel, E.: Cyclist safety on bicycle boulevards and parallel arterial routes in Berkeley, California. Accid. Anal. Prev. **45**, 241–247 (2012). doi:10.1016/j.aap.2011.07.009
21. Ministry of Infrastructures and Transports. Nuovo codice della strada, D.L. 30/4/1992 n.285, Gazzetta Ufficiale Repubblica Italiana, n. 114. (In Italian) (1992)
22. Ministry of Infrastructures and Transports. Decreto Ministeriale del 5/11/2001 Norme funzionali e geometriche per la costruzione delle strade. Roma (in Italian) (2001)
23. Räsänen, M., Summala, H.: Attention and expectation problems in bicycle–car collisions: An in–depth study. Accid. Anal. Prev. **30**, 657–666 (1998)
24. Räsänen, M., Summala, H.: Car drivers' adjustments to cyclists at roundabouts. Transp. Hum. Factors **2**, 1–17 (2000)
25. Scheepers, P., Hagenzieker, M., Methorst, R., van Wee, B.: A conceptual framework for road safety and mobility applied to cycling safety. Accid. Anal. Prev. **62**, 331–340 (2014)
26. Silvano, A.P., Ma, X., Koutsopoulos, H.N.: A hierarchical modelling framework for vehicle-bicycle interactions at roundabouts. In: International Cycling Safety Conference, vol. 3 (2014)
27. Silvano, A.P., Ma, X., Koutsopoulos, H.N.: When do drivers yield to cyclists at unsignalized roundabouts? Empirical evidence and behavioral analysis. In: Transportation Research Board 94th Annual Meeting (2015)
28. Walker, I., Garrard, I., Jowitt, F.: The influence of a bicycle commuter's appearance on drivers' overtaking proximities: An on-road test of bicyclist stereotypes, high-visibility clothing and safety aids in the United Kingdom. Accid. Anal. Prev. **64**, 69–77 (2014)
29. Wegman, F., Zhang, F., Dijkstra, A.: How to make more cycling good for road safety? Accid. Anal. Prev. **44**, 19–29 (2012)
30. Wood, J.M., Lacherez, P.F., Marszalek, R.P., King, M.J.: Drivers' and cyclists' experiences of sharing the road: Incidents, attitudes and perceptions of visibility. Accid. Anal. Prev. **41**, 772–776 (2009)

Automatic Incident Detection on Freeway Ramp Junctions. A Fuzzy Logic-Based System Using Loop Detector Data

Riccardo Rossi(✉), Massimiliano Gastaldi, and Gregorio Gecchele

Department of Civil, Environmental and Architectural Engineering, University of Padova, Via F. Marzolo 9, 35131 Padova, Italy
{riccardo.rossi,massimiliano.gastaldi,gregorio.gecchele}@unipd.it

Abstract. Vehicle loop detectors or other equipment installed on highway sections are commonly used for monitoring traffic flow conditions on road networks. For operational analysis, it is essential to be able to distinguish low levels of service due to over-saturated conditions from those caused by extraordinary events such as incidents. In the case of incidents, prompt responses are crucial for activating any required countermeasures, such as rescue activation or traffic detours. Automatic Incident Detection methods for basic freeway segments are widely reported in the literature, but their application to freeway ramp merging zones is limited. This work introduces a control system which can identify incidents from vehicle loop detector data on freeway ramp merging zones. The system was developed with fuzzy logic concepts and calibrated with data from micro-simulation experiments. The main finding of this study is that the detection system, despite its simplicity, shows excellent False Alarm Rate (FAR) and satisfactory Detection Rate (DR) and Mean Time To Detection (MTTD), generally better than those obtained with the traditional California#7 comparative algorithm.

Keywords: Automatic Incident Detection · Freeway ramp · Fuzzy logic

1 Introduction

In urban areas, recurring congestion during morning and/or evening peak periods is due to over-saturated traffic flows, which are generally easy to forecast. Non-recurring congestion takes place because of a rapid capacity drop, for unexpected reasons which are generally identified as incidents. Incidents include accidents, disabled vehicles on the road, spilled loads, temporary maintenance and construction work, signal and detector malfunctioning, and other particular and unusual events which disrupt the normal flow of traffic and cause delays. Whatever the reason, congestion increases fuel consumption and vehicle emissions due to increasing speed differentiation, decreased travel speed and increased travel time. Although it is difficult to reduce these negative effects in cases of recurring congestion, it is possible of non-recurring congestion by means of efficient traffic management systems - in particular, accurate detection of incidents.

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_18

Incidents can be detected by traffic management systems with automatic or non-automatic methods. Automatic Incident Detection (AID) methods are more common and extensively applied in practice [30]: they can analyse traffic data and detect incidents quickly by means of various types of sensors and techniques (e.g., image processing, detector/sensor-based incident detection algorithms, probe vehicle data). Although there is ample literature on the available methods and applications, it is mainly related to basic freeway segments, and AID applied to freeway ramp merging zones is limited. This paper uses a fuzzy logic-based approach to detect incidents in freeway ramp merging zones. Fuzzy logic theory is effective in dealing with complex non-deterministic problems in an accurate and interpretable way. This paper introduces fuzzy logic into the AID algorithm and uses fuzzy linguistics to describe the variation law of traffic parameters.

The paper is organized as follows. Section 2 gives a broad summary of the various kinds of incident detection algorithms. Section 3 presents the proposed Fuzzy Logic-Based Incident Detection System; Sect. 4 introduces the comparative algorithm California#7. Evaluation of both systems, tested with simulated datasets, is provided in Sects. 5 and 6. Conclusions and further developments of the research are given in Sect. 7.

2 Theoretical Background

Automatic algorithms are algorithms which automatically prompt an incident alarm when traffic data from traffic sensors satisfy certain pre-set conditions; non-automatic algorithms or procedures are based on human observer reports (i.e., driver-based "sensors"). Incidents can be automatically detected in several ways. Although some methods based on video processing have been adopted to detect traffic incidents [29], they are sensitive to outdoor environmental factors (e.g., static shadows, snow, rain, glare). Inductive loop detectors (ILD) are the most commonly used sensors in traffic surveillance and management applications, they can collect traffic data stably and do not suffer from the effects of outdoor environments.

Several algorithms based on data from fixed in-road sensors have varying data requirements, principles and structural complexity. Traditional incident detection algorithms are grouped into seven categories according to their basic principle: (1) comparative; (2) statistical; (3) time series; (4) filtering/smoothing; (5) traffic modeling; (6) artificial intelligence; and (7) image processing (1). All these algorithms use loop detector or loop-emulating data collected at points along the roadway and all are applied to freeways. A review of the literature shows that the approaches are usually defined for basic freeway segments and then also applied to other segments, including ramp merging zones [23,24], generally with worse performance.

Comparative algorithms are well-established and well-known, and compare the values of observed traffic parameters (volume, occupancy, speed) with pre-established thresholds and prompt an incident alarm when values exceed those

thresholds. Comparative algorithms include decision trees (DT), also called California algorithms [26,31,32], pattern recognition (PATREG) algorithms [10] and the APID algorithm [28].

Statistical algorithms use statistical techniques to determine whether observed detector data differ statistically from estimated or predicted traffic characteristics. Examples are the standard normal deviate (SND) algorithm [13] and the Bayesian algorithm [26,40].

Time series algorithms use time series models to predict normal traffic conditions and detect incidents when detector measurements deviate significantly from model outputs. Methods used to predict time-dependent traffic for incident detection include, among others, the autoregressive integrated moving-average (ARIMA) model [3] and the high occupancy (HIOCC) algorithm [10].

Smoothing and filtering methods are designed to remove from traffic data any short-term noise or inhomogeneities which cause false alarms, and thus allow true traffic patterns to be more visible and more readily detectable as true incidents [4]. Smoothing is a mathematical method which produces a weighted average of a given traffic variable; filtering algorithms use a linear filter which allows low-frequency components of the detector data to pass but removes their undesirable high-frequency portions. Representative algorithms include double exponential smoothing (DES) [11], low-pass filter (LPF) [6], and discrete wavelet transform and linear discriminant analysis (DWT-LDA) algorithms [2].

Traffic modeling approaches have been adopted for incident detection based on comparisons between observed traffic parameters and parameter values estimated by the models. Traffic modeling algorithms include dynamic models [41], the catastrophe theory model and its modifications [15,16,33] and the low-volume (LV) incident detection algorithm [14].

Artificial intelligence refers to a set of procedures which apply approximate reasoning and uncertainty in complex decision-making and data analysis. Artificial intelligence methods applied in automatic incident detection include neural networks [1,2,8,12,34,38,39], fuzzy logic [5,27], and a combination of the two [21,22]. Among the Artificial Intelligence methods, fuzzy logic, and fuzzy inference systems in particular, appear to be promising, thanks to their accuracy, interpretability and robustness. Fuzzy systems can not only handle linguistic concepts, incorporating human expert knowledge with its graduations, but can also express the behavior of the system in an interpretable way. In addition, they are universal approximators which can perform robust non-linear mapping between inputs and outputs.

In case of *image processing methods*, programs may be applied to extract traffic measures, such as volume, occupancy, speed, and/or queue length, from video images, using the image-processing unit (consisting of a surveillance video camera and an image processing computer program) as a loop detector.

In this paper, the incident identification system was developed by integrating data-driven procedures and expert knowledge in a common Fuzzy Logic framework [19]. A set of "if-then" rules, expressed in verbal terms, is built from the fuzzy knowledge base, properly describing the cause-effect mechanism of the incident detection process.

3 Methodology

3.1 System Structure

The basic freeway segment system developed by the authors [35] was modified to include the presence of the ramp. Figure 1 shows how the system operates in a two-lane basic freeway segment, in which two detector stations are presumed to be placed upstream and downstream of the location of the incident.

Two Fuzzy Inference Systems (FIS) simultaneously analyse traffic data collected by each detector station and identify anomalies due to incidents, according to the differences among observed traffic parameters and normal traffic conditions (i.e., without incidents or accidents) (Fig. 1(a)). The approach is similar to that of traditional comparative algorithms, which compare the value of observed traffic parameters with pre-established threshold values. However, in the system proposed here, differences are computed on the Volume-Speed Diagram (VSD) of traffic flow: if in the same traffic conditions (i.e., traffic flow) differences between observed and normal states in the VSD are "low", the system is in a "normal state"; otherwise, an "abnormal state" exists. Since the location of the incident is unknown, the identification system is applied to each detector station to evaluate the state of the road section for each time interval (Fig. 1(b)). Depending on the relative location of the incident, FIS1 or FIS2 detects an "abnormal state" in the section caused by an incident (red car), which can prompt an incident alarm if that abnormal state lasts for a set time (e.g., green cars queuing due to an accident).

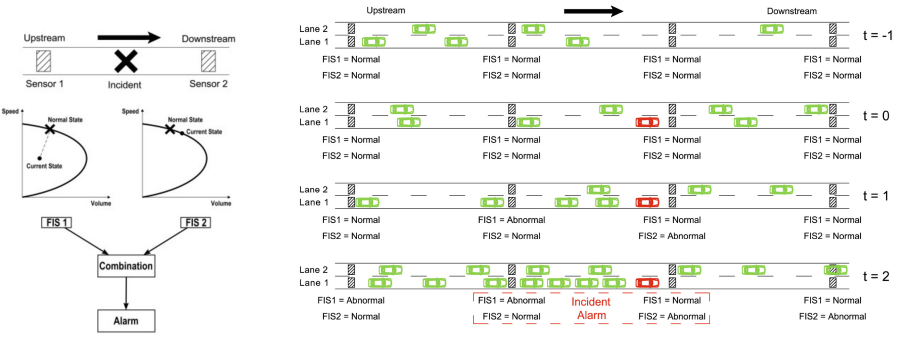
The Volume-Speed Diagram (VSD) may be obtained analysing existing traffic data or using a traffic simulation model; in this paper the second approach was adopted. Simulation runs were performed for different traffic flow values in normal conditions (without incident) for different flow rates determining 5-min experimental points as reported in Sect. 5. The Volume-Speed Diagram was obtained by interpolation of the experimental points.

The basic freeway segment system was modified by adopting a two-level structure (Fig. 2). At the first level, local FISs evaluate the presence of incidents at the borders of the merging zone by comparing current traffic conditions with normal ones: FIS1 is upstream of the major road; FIS2 is downstream of it, and FIS3 is upstream of the minor access ramp. At the second level, a global FIS combines the information provided by the local FISs and prompts an incident alarm.

3.2 Fundamentals of Fuzzy Inference Systems

Fuzzy inference systems are common fuzzy logic applications, which are able to incorporate human expert knowledge with its nuances, as well as to express the behavior of the system in an interpretable way [19]. A typical fuzzy inference system is shown in Fig. 3.

Input and output variables are partitioned in fuzzy sets (membership functions), representative of linguistic variables. The underlying relationship between



(a) System structure (b) Example of implementation

Fig. 1. Sketch of basic freeway segment system: (a) system structure; (b) example of implementation

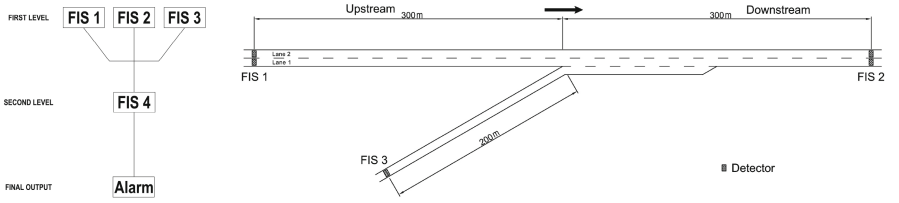


Fig. 2. Sketch of proposed incident detection system

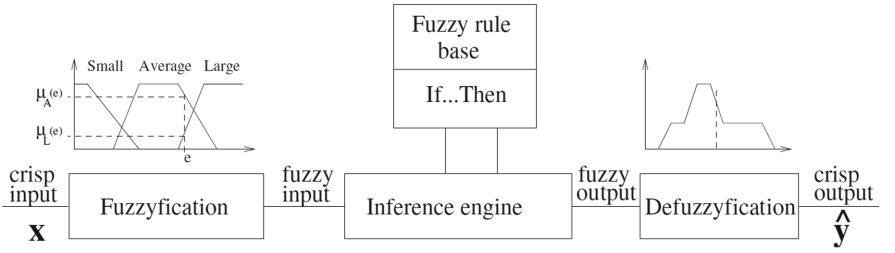


Fig. 3. Simplified structure of a fuzzy inference system (FIS). Source [19]

the input and output spaces is modelled using linguistic if-then inference rules, which represent joint sets of possible input and output values. Input elements are fuzzified to determine the grade of membership (GoM), ranging from 0 (not belonging) to 1 (full belonging), they belong to the linguistic variables. Fuzzified values are processed by the inference engine according to the fuzzy rule base. The rules are aggregated yielding a possibility distribution representing the inferred output; a defuzzification stage is required to obtain a crisp (finite) value from the output possibility distribution. Fuzzy inference systems, which represent the core of AID proposed model, have been developed according to these considerations.

3.3 Fuzzy Model Identification

The proposed system is based on traffic data such as volume, occupancy and speed, collected by commonly used detectors (e.g., ILD). As listed in Table 1, first-level FISs have the same input variables (density and space mean speed): FIS2 examines them in absolute terms; FIS1 and FIS3 examine them relatively (difference between actual and usual conditions, measured as euclidean distance). Output variables STATE1, STATE2 and STATE3 represent the state of the merging zone at a certain time, as detected by each FIS. At the second level, FIS4 combines the states of the system provided by local FISs into the global state of the merging zone (STATE4).

The FISs were developed from traffic simulation data (see Sect. 5) with FisPro, an open-source software available free on the Internet [18]. The membership functions of the premise and consequence fuzzy sets (input and output variables) are identified with the K-means algorithm [20] and the rules of inference with the so-called FPA (Fast Prototype Algorithm, [17]), according to Mamdani's product-sum inference. For the sake of brevity, Fig. 3 shows an example of the membership functions of the linguistic terms in the premise and consequence fuzzy sets for the FISs.

The fuzzy system knowledge base of each FIS is characterized by three fuzzy sets in the domain of all input variables and by two triangular membership functions in the domain of the outputs, representing normal and abnormal conditions, respectively (Fig. 4).

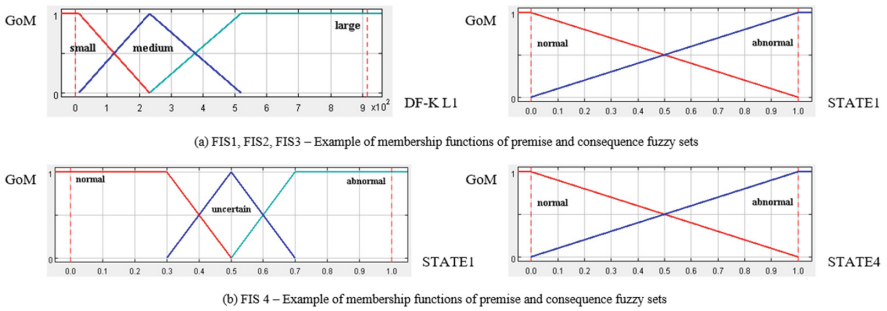


Fig. 4. Example of premise and consequence fuzzy sets for FIS1, FIS2, FIS3 and FIS4

Similar rules were identified for all simulated flow rates; reasonable combinations of input variables were covered by inference rules, integrating data-driven procedures with expert knowledge when necessary [19]. As an example, two rules for calculating traffic volumes on the major part of the road of 3,000 vehicles/h for FIS1 and FIS4 are:

Table 1. Input and output variables adopted by system FISs.

FIS	Type	Name	Description
FIS1(U)	Input	DF-K L1	Distance current/normal condition in Flow-Density plane (Major road-Lane 1)
FIS1(U)	Input	DF-SMS L1	Distance current/normal condition in Flow-Space Mean Speed plane (Major road-Lane 1)
FIS1(U)	Input	DF-K L2	Distance current/normal condition in Flow-Density plane (Major road-Lane 2)
FIS1(U)	Input	DF-SMS L2	Distance current/normal condition in Flow-Space Mean Speed plane (Major road-Lane 2)
FIS1(U)	Output	STATE1	State of the system for FIS1
FIS2(D)	Input	K L1	Density observed in Major road-Lane 1
FIS2(D)	Input	SMS L1	Spatial Mean Speed observed in Major road-Lane 1
FIS2(D)	Input	K L2	Density observed in Major road-Lane 2
FIS2(D)	Input	SMS L2	Spatial Mean Speed observed in Major road-Lane 2
FIS2(D)	Output	STATE2	State of the system for FIS2
FIS3(R)	Input	DF-K L1	Distance current/normal condition in Flow-Density plane (Major road-Lane 1)
FIS3(R)	Input	DF-SMS L1	Distance current/normal condition in Flow-Space Mean Speed plane (Major road-Lane 1)
FIS3(R)	Input	DF-K L2	Distance current/normal condition in Flow-Density plane (Major road-Lane 2)
FIS3(R)	Input	DF-SMS L2	Distance current/normal condition in Flow-Space Mean Speed plane (Major road-Lane 2)
FIS3(R)	Output	STATE3	State of the system for FIS3
FIS4(G)	Input	STATE1	State of the system for FIS1
FIS4(G)	Input	STATE2	State of the system for FIS2
FIS4(G)	Input	STATE3	State of the system for FIS3
FIS4(G)	Output	STATE4	State of the system for FIS4

Description of FIS:(U) = major road upstream FIS, (D) = major road downstream FIS, (R) = ramp FIS, (G) = global FIS

If (DF-K L1 is Medium) And (DF-K L2 is Large)
And (DF-SMS L1 is Small) And (DF-SMS L2 is Medium)
Then (STATE1 is Abnormal)

If (STATE1 is Normal) And (STATE2 is Normal) And (STATE3 is Uncertain)
Then (STATE4 is Normal)

The first-level fuzzy output variables (STATE1, STATE2, STATE3) are defuzzified with the centroid method [25], yielding three indexes, representing

the state of the merging zone as detected at each border and providing some information concerning the uncertainty of the estimate. Values of less than 0.5 are associated to normal traffic condition (absence of incident), values of more than 0.5 are associated to incident (abnormal) condition. Similarly, the final state of the road segment (STATE4) is determined by the composition of the first-level outputs by FIS4, and by the defuzzified output variable in normal or incident conditions.

4 California Algorithm #7

California Algorithm #7 (CA7) is a comparative algorithm which uses pre-set thresholds to classify current road condition, based on occupancy data from two adjacent detector stations [26]. The algorithm calculates spatial difference in occupancy (OCCDF), relative spatial difference of occupancies (OCCRDF) and occupancy values obtained from downstream detectors (DOCC) and compares these inputs with three pre-set thresholds (T1, T2, T3). T1 is the maximum value of OCCDF, T2 the maximum OCCRDF value, and T3 the maximum DOCC value, all in normal conditions. How these indicators are calculated is given in Eqs. 1 to 3.

$$OCCDF(i,t) = OCC(i,t) - OCC(i+1,t) \tag{1}$$

$$OCCRDF(i,t) = (OCC(i,t) - OCC(i+1,t))/OCC(i,t) \tag{2}$$

$$DOCC(t) = OCC(i+1,t) \tag{3}$$

where i denotes the detector station number, $i+1$ the downstream detector station for i station, and t the time period: $OCC(i+1,t)$ is the occupancy value obtained from detector station $(i+1)$ in time period t.

Comparing thresholds and inputs according to a decision tree process, the algorithm can identify the actual state of the road among the four states identified for CA7. The road is in state 0 when there are no incidents, in state 1 when there is the possibility of an incident but none has yet been detected, in state 2 when an incident is detected, and in state 3 when the incident continues.

In this case study, the CA7 algorithm was applied with data from upstream and downstream detector stations in the major road segment.

5 Comparative Evaluation

5.1 Performance

AID performance is usually evaluated according to the following measures of effectiveness: detection rate (DR), false alarm rate (FAR), and mean time to detection (MTTD) [7]. In this study, these measures were used to compare incident detection algorithms (proposed method and CA7) [9].

Detection rate DR is defined as the percentage of incident cases detected correctly by the algorithm, as shown in Eq. 4:

$$DR = \frac{\text{number of incidents detected}}{\text{total number of incidents}} \cdot 100 \tag{4}$$

False Alarm Rate FAR is an index representing the rate at which non-incidents are falsely classified as incidents, thus giving rise to a false alarm (Eq. 5):

$$FAR = \frac{\text{number of false alarm cases}}{\text{total number of algorithm applications}} \cdot 100 \tag{5}$$

Mean Time To Detection MTTD (Eq. 6) is defined as the average time taken by an algorithm to detect incidents. It is obtained by measuring the delay between the time when an incident occurs and the time when it is successfully detected by the algorithm (an alarm is given), averaged over m incidents detected during a certain interval of time.

$$MTTD = \frac{(t_i + \ldots + t_m)}{m} \tag{6}$$

The primary aim of an AID algorithm is to maximize incident DR and minimize FAR, i.e., to reduce the time delay between the occurrence of an incident and its detection by the algorithm (MTTD). Since these criteria are in conflict with each other, a trade-off must be made when the algorithm is implemented, according to priorities (e.g., whether a high DR is acceptable at the cost of a relatively high FAR, or not). In critical safety applications, a perfect DR is necessary even though it may be more expensive in terms of false alarms; for the problem of freeway incident detection, a FAR of 1.0% and below is acceptable [38].

5.2 Data Description

For calibration and testing, data were obtained from a microsimulation traffic model, developed with SIAS-Paramics®software [36,37] for traffic modeling, due to the scarcity of real incident data from Italian freeways.

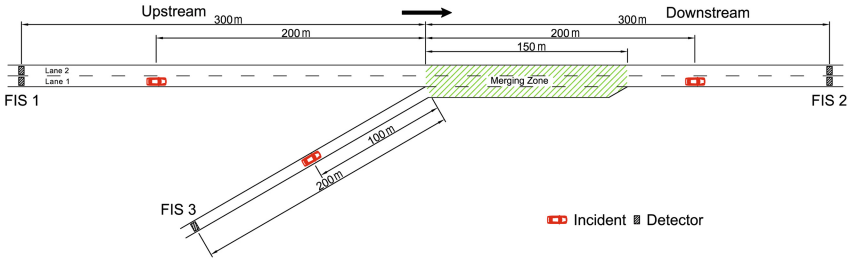


Fig. 5. Sketch of model adopted for simulation

The model includes a major road (freeway segment) and a minor road (ramp) with a merging zone of 150 m at the point of confluence (Fig. 5). Traffic flow data were collected by virtual detector stations embedded in the road at a distance of 100 m from each other; each station has one loop detector for each lane and collects data at 1-min intervals. In order to analyse the performance of the system in varying traffic conditions, the major road section was loaded with three flow rate values (i.e. 1,000, 2,000 and 3,000 vehicles/hour) and the minor road sections with 900 vehicles/hour.

Simulations lasted 180 min and included only cars, the speed limits of which were set at 90 km/h for the major road and 50 km/h for the minor road. The microsimulation model outputs depend on the random seed number used in each model run; multiple runs, with different random number of seeds, must be carried out to provide stable and robust outputs.

In the present study, 10 model runs were performed for each traffic volume value, with a duration of 180 min. Since data were aggregated on the basis of 1-min time intervals, 180 experimental points (speed, density, flow) were given by each detector station for each lane. 1-min experimental points were calculated following these steps:

1. Flow. It was obtained as the number of vehicle passing on the loop in the selected time interval;
2. Speed. Space mean speed was estimated as the harmonic average of vehicle speeds observed on the loop in the selected time interval;
3. Density. Density value K was estimated from occupancy measure O obtained from loop sensor, adopting Eq. 7:

$$\begin{gathered} O = c_K \cdot K \\ \text{where } c_K = L + d \end{gathered} \tag{7}$$

where L is the average vehicle length and d is the loop length.

Incidents were generated by the software "incidents editor", which defines the characteristics of the vehicle which caused the incident, including its duration (in our case, 20 min) and the speed of the vehicle (0 km/h). The location and lane of incidents varied according to the three case studies in the study area (Fig. 5):

– Upstream: incident on major road, lane 1, 200 m upstream of confluence point;
– Downstream: incident on major road, lane 1, 200 m downstream of confluence point;
– Ramp: incident on minor road, 100 m upstream of confluence point.

Traffic data from simulations were used for calibration (FIS identification and composition) and validation of the proposed incident recognition system. Calibration used data recorded at the borders of the merging zone (on both major and minor roads) from 7 runs for each flow rate. The remaining data were used to test the performance of the system.

6 Experimental Analysis

The performance of the proposed model and CA7 were evaluated by standard parameters (DR, FAR, MTTD), assessing the effects of the following two factors:

– traffic flow conditions (1,000, 2,000 or 3,000 vehicles/hour);
– location of incident (ramp, upstream, downstream).

Table 2 lists the results obtained from incident simulations.

Table 2. Proposed system and CA7 performance.

Incident position	Main road traffic volume [vehicles/h]	DR [%]		FAR [%]		MTTD [min]	
		Proposed	CA7	Proposed	CA7	Proposed	CA7
U	1000	100	10	0,0	0,3	0,9	0,0
U	2000	100	100	0,1	0,1	4,0	6,1
U	3000	70	100	0,1	0,4	5,9	2,9
R	1000	90	20	0,2	0,2	2,1	11,0
R	2000	100	10	0,0	0,0	3,2	8,0
R	3000	100	10	0,0	0,1	3,0	1,0
D	1000	100	0	0,0	0,3	0,9	-
D	2000	100	0	0,0	0,1	0,8	-
D	3000	100	100	0,0	0,4	0,5	10,0

Incident positions: (U) = upstream, (D) = downstream,(R) = ramp

The following observations were made according to the performance values listed in Table 2:

– The Detection Rate (DR) of the proposed system is generally high; partial exceptions are the cases of the incident located on the ramp with low traffic volume and the upstream incident with high traffic volume;
– The False Alarm Rate (FAR) is very low in all cases, an ideal condition for the problem of freeway incident detection [38];
– The lowest value of Mean Time to Detection (MTTD) is obtained for the downstream incident and increases for ramp and upstream conditions;
– Compared with CA7, the proposed method shows better performance:
  - Lower (or equal) FAR in all conditions tested;
  - Higher (or equal) DR in all conditions tested, excluding the case of upstream incident with high traffic volume;
  - Lower MTTD when incident occurs on the ramp or downstream.

7 Conclusions

This paper presents a fuzzy logic-based approach to detect incidents in freeway ramp merging zones, fuzzy linguistics being used to describe the variation law of traffic parameters. A traffic microsimulation tool (S-Paramics) was adopted to evaluate the performance of the proposed system in terms of standard indices, i.e., detection rate (DR), false alarm rate (FAR), and mean time to detection (MTTD). The results were obtained for three flow rate values for the major road and three locations of incidents in the merging zone (ramp, upstream, downstream).

Although the system is simple and its structure easily interpretable, the proposed incident detection system shows excellent FAR and satisfactory DR and MTTD. Comparison of it with the traditional comparative California#7 algorithm shows that the proposed method generally has better performance, excluding a small number of conditions.

Future research could extend the present work in several directions, including:

- extension of analysis to real incident data (e.g., I-880 database), to evaluate the effect of the detection accuracy and reliability of sensors in real-world traffic conditions;
- comparison of the proposed system with other algorithms (traditional and more recent ones) to extend its validation;
- inclusion of other types of vehicles to represent real traffic conditions better.

References

1. Abdulhai, B., Ritchie, S.G.: Enhancing the universality and transferability of freeway incident detection using a Bayesian-based neural network. Transp. Res. C-Emer. **7**, 261–280 (1999)
2. Adeli, H., Samant, A.: An adaptive conjugate gradient neural network-wavelet model for traffic incident detection. Comput. Aided Civ. Inf. **15**, 251–260 (2000)
3. Ahmed, M.S., Cook, A.R.: Application of time-series analysis techniques to freeway incident detection. Transp. Res. Rec. **841**, 19–21 (1982)
4. Balke, K.N.: An evaluation of existing incident detection algorithms. FHWA - Report No. FHWA/TX-93/1232-20, Washington D.C. (1993)
5. Chang, E.C.-P., Wang, S.-H.: Improved freeway incident detection using fuzzy set theory. Transp. Res. Rec. **1453**, 75–82 (1994)
6. Chassiakos, A.P., Stephanedes, Y.J.: Smoothing algorithms for incident detection. Transp. Res. Rec. **1394**, 8–16 (1993)
7. Chen, S., Wang, W., van Zuylen, H.: Construct support vector machine ensemble to detect traffic incident. Expert Syst. Appl. **36**, 10976–10986 (2009)
8. Cheu, R.L., Ritchie, S.G.: Automated detection of lane-blocking freeway incidents using artificial neural networks. Transp. Res. C-Emer. **3**, 371–388 (1995)
9. Cheu, R.L., Srinivasan, D., Loo, W.H.: Training neural networks to detect freeway incidents by using particle swarm optimization. Transp. Res. Rec. **1867**, 11–18 (2004)

10. Collins, J.F., Hopkins, C.M., Martin, J.A.: Automatic incident detection TRRL algorithms HIOCC and PATREG. TRRL Supplementary - Report No. 526, Crowthorne, Berkshire, U.K. (1979)
11. Cook, A.R., Cleveland, D.E.: Detection of freeway capacity-reducing incidents by traffic stream measurements. Transp. Res. Rec. **495**, 1–11 (1974)
12. Dia, H., Rose, G.: Development and evaluation of neural network freeway incident detection models using field data. Transp. Res. C-Emer. **5**, 313–331 (1997)
13. Dudek, C.L., Messer, C.J., Nuckles, N.B.: Incident detection on urban freeway. Transp. Res. Rec. **495**, 12–24 (1974)
14. Fambro, D.B., Ritch, G.P.: Evaluation of an algorithm for detecting urban freeway incidents during low-volume conditions. Transp. Res. Rec. **773**, 31–39 (1980)
15. Forbes, G.J., Hall, F.L.: The applicability of catastrophe theory in modeling freeway traffic operations. Transp. Res. A-Pol. **24**, 335–344 (1990)
16. Gall, A.I., Hall, F.L.: Distinguishing between incident congestion and recurrent congestion: a proposed logic. Transp. Res. Rec. **1232**, 1–8 (1989)
17. Glorennec, P.-Y.: Algorithmes d'apprentissage Pour Systems d'infrence Floue. Edition Hermès, Paris, France (1999)
18. Guillaume, S., Charnomordic, B.: Learning interpretable fuzzy inference systems with FisPro. Inform. Sci. **181**, 4409–4427 (2011)
19. Guillaume, S., Charnomordic, B.: Fuzzy inference systems: an integrated modeling environment for collaboration between expert knowledge and data using FisPro. Expert Syst. Appl. **39**, 8744–8755 (2012)
20. Hartigan, J.A., Wong, M.A.: A K-means clustering algorithm. Appl. Stat. **28**, 100–108 (1979)
21. Hsiao, C.-H., Lin, C.-T., Cassidy, M.: Application of fuzzy logic and neural networks to automatically detect freeway traffic incidents. Transp. Eng. J. ASCE **120**, 753–772 (1994)
22. Ishak, S.S., Al-Deek, H.M.: Fuzzy ART neural network model for automated detection of freeway incidents. Transp. Res. Rec. **1634**, 56–63 (1998)
23. Jin, J., Ran, B.: Automatic freeway incident detection based on fundamental diagrams of traffic flow. Transp. Res. Rec. **2099**, 65–75 (2009)
24. Karim, A., Adeli, H.: Comparison of fuzzy-wavelet radial basis function neural network freeway incident detection model with california algorithm. Transp. Eng. J. ASCE **128**, 21–30 (2002)
25. Klir, G., Yuan, B.: Fuzzy Sets and Fuzzy Logic. Theory and Applications. Prentice Hall PTR, Upper Saddle River (1995)
26. Levin, M., Krause, G.M.: Incident detection: a Bayesian approach. Transp. Res. Rec. **682**, 52–58 (1978)
27. Lin, C.K., Chang, G.L.: Development of a fuzzy-expert system for incident detection and classification. Math. Comput. Model. **27**, 9–25 (1998)
28. Masters, P.H., Lam, J.K., Wong., K.: Incident detection algorithms for COMPASS an advanced traffic management system (1991)
29. Michalopoulos, P.G., Jacobson, R.D., Anderson, C.A., DeBruycker, T.B.: Automatic incident detection through video image processing. Traffic Eng. Control **34**, 66–75 (1993)
30. Parkany, E., Xie, C.: A complete review of incident detection algorithms and their deployment: what works and what doesn't. The New England Transportation Consortium - Report No. NETCR37 (2005)
31. Payne, H.J.: Development and testing of incident detection algorithms, volume 1: summary of results. FHWA - Report No. FHWA-RD-76-19, Washington D.C. (1976)

32. Payne, H.J., Tignor, S.C.: Freeway incident-detection algorithms based on decision trees with states. Transp. Res. Rec. **682**, 30–37 (1978)
33. Persaud, B.N., Hall, F.L.: Catastrophe theory and patterns in 30-second freeway traffic data implications for incident detection. Transp. Res. A-Pol. **23**, 103–113 (1989)
34. Ritchie, S.G., Cheu, R.L.: Simulation of freeway incident detection using artificial neural networks. Transp. Res. C-Emer. **1**, 203–217 (1989)
35. Rossi, R., Gastaldi, M., Gecchele, G., Barbaro, V.: Fuzzy logic-based incident detection system using loop detectors data. Transp. Res. Procedia **10**, 266–275 (2015)
36. SIAS Limited. S-Paramics Reference Manual (2006)
37. SIAS Limited. S-Paramics Assignment Stability 1.1 Installation & User Guide (2012)
38. Srinivasan, D., Sanyal, S., Sharma, V.: Freeway incident detection using hybrid fuzzy neural network. IET Intell. Transp. Syst. **1**, 249–259 (2007)
39. Stephanedes, Y.J., Liu, X.: Artificial neural networks for freeway incident detection. Transp. Res. Rec. **1494**, 91–97 (1995)
40. Tsai, J., Case, E.R.: Development of freeway incident detection algorithms by using pattern recognition techniques. Transp. Res. Rec. **722**, 113–116 (1979)
41. Willsky, A.S., Chow, E.Y., Gershwin, S.B., Greene, C.S., Houpt, P., Kurkjian, A.L.: Dynamic model-based techniques for the detection of incidents on freeways. IEEE Trans. Automat. Control **25**, 347–360 (1980)

Artificial Intelligence and Soft Computing in Transportation and Logistics

Artificial Intelligence and Soft Computing in Transportation and Logistics

Michele Ottomanelli

Technical University of Bari, via Orabona, 4, 70125 Bari, Italy
michele.ottomanelli@poliba.it

Transportation and logistics systems play a crucial role in the development of an area because of their economic, societal, and environmental importance. Transportation and logistics systems are actually a very challenging industry characterized by an intrinsic high complexity due to the non-linear interactions of many actors and components. In fact, sector is made up of heterogeneous elements that are spatially and functionally distributed. These elements are referred to human actors and or artificial elements. The first are characterized by different behaviour and perception or judgment capability with individual or shared objective. Also artificial elements are qualified by different performances, functionality and operating framework [2]. In addition, many parameters and variables are characterized by uncertainty, subjectivity, imprecision and ambiguity. Human operators, traffic controllers, dispatchers, drivers, passengers, engineers, and planners use subjective knowledge, approximate values, and/or linguistic information when making decisions.

Such an environment makes transportation and logistics problem dominated by very complex dynamics and uncertainty. In addition, most of the real-life transportation and logistics problem are combinatorial by nature (i.e. NP-hard problems).

Examples of this kind of problems are the dynamic and static routing and scheduling of vehicles, crew rostering and vehicle fleet planning for airlines, railroads, truck operations and public transportation services, transportation networks design and optimization, locational problems, crowd evacuation, etc.

The treatment and solution of such complex problems need the development and use of innovative intelligent systems that fuse knowledge, techniques, and methodologies from a wide range of scientific areas [1–3, 7]. To this purpose, in the last decades, novel methods and approaches have been proposed in literature. They are based on computational tools that allow to compute with words (for example, Fuzzy Logic) [1, 6–8], learn and adapt (i.e. Artificial Neural Networks) [1, 2, 5], and perform systemic random search and optimization (for ex. Evolutionary and Swarm Intelligence Algorithms) [4] that succeed where many traffic and transportation engineering problems are combinatorial by nature and cannot be solved via exact methods (such as Linear Programming and Branch-and-Bound) either because of the large dimension or because it is difficult/not possible to split them into smaller sub-problems, as well as because of the high computational time.

This collection of computational tools can be summarized in Artificial Intelligence and Soft Computing methods that is the topic of this chapter.

Five papers were accepted for this chapter, based on a rigorous review process, that includes independent reviews by selected researchers, as well as the evaluation by the

guest editors. These papers present some advances applied to different field of transportation engineering: (1) airport operations, (2) ports operations, (3) logistics; (4) traffic signal control, (5) route choice.

Marinelli, Palmisano, Ottomanelli and Dell'Orco explored swarm intelligence methods in the paper "**Optimizing airport gate assignments through a hybrid metaheuristic approach**" that deals with air transport management. In particular, they face the problem of gate assignments of flights that is a key issue to solve in real time when service disruptions occur. They propose an innovative algorithm based on a hybrid approach called Biogeography-based Bee Colony Optimization (B-BCO). It properly combines two metaheuristics: biogeography-based (BBO) and bee colony optimization (BCO) algorithms. The proposed model is then applied to the case study of International Airport of Malpensa (Milan, Italy).

The section entitled "**Shipping liner company stowage plans: an optimization approach**" by Ambrosino, Paolucci and Sciomachena aims at determining the optimal stowage plan within a Multi-Port Master Bay Plan Problem for each port included in the circular route of a containership. The work extends the modelling capability of a previous research by considering in the formulation of the MP-MB problem also the minimization of the number of hatch movements during the unloading and loading phases. To solve the problem they propose an improved Mixed Integer Linear programming Heuristics that can be applied to containership with capacity up to 18000TEUs.

Evolutionary Intelligence is used in the paper "**A multi-depot periodic vehicle routing model for petrol station replenishment**" by Carotenuto, Giordani, Massari and Vagaggini, in which the authors face the problem of petrol stations replenishment. They propose to model the problem as a generalization of the Multi-Depot Periodic Vehicle Routing Problem (MDPVRP) and provide a mathematical formulation of it that is solved out using a heuristic based on a hybrid Genetic Algorithms because of the high number of instances considered in the problem.

Jin and Ma face the problem of traffic signal control in the paper "**A Multi-criteria Intelligent Control for Traffic Lights using Reinforcement Learning**". They propose an adaptive signal control system capable of taking multi-criteria strategies into account. A general multi-agent framework is introduced for modelling signal control operations. The behaviour of each cognitive agent is modelled by a Markov decision process (MDP). Reinforcement learning algorithms are applied to solve the MDP problem.

The paper entitled "**Analysis of imprecise perception in route choice considering fuzzy costs**" by Binetti, De Mitri and Ottomanelli is devoted to the classical route choice modelling. In order to consider uncertainty of users perception and trying to overcome some limits of well know logit based choice models (i.e. IIA property) the authors propose a path choice model in which costs are represented by fuzzy numbers. In the research, users' estimation costs relevant to choice a route over a road network is then modelled in a real uncertain, imprecise and ambiguous environment. In addition, an index for evaluating the overlapping degree of each path with the others is introduced.

References

1. Transportation Research Board: Artificial Intelligence in Transportation, Circular E-C113 (2007)
2. Teodorovic, D.: Fuzzy logic systems for transportation engineering: the state of the art. Transp. Res.: Part A **33**, 337–364 (1999)
3. Marinelli, M., Caggiani, L., Ottomanelli, M.: Managing the uncertainty of data fusion from different sources in modelling route choice behaviour. In: IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, pp. 202–207. IEEE Press (2015)
4. Zhang, S., Lee, C.K.M., Chan, H.K., Choy, K.L., Zang, W.: Swarm intelligence applied in green logistics: a literature review. Eng. Appl. Artif. Intell. **37**, 154–169 (2015)
5. Dougherty, M.: A review of neural networks applied to transport. Transp. Res. Part C, **3**, 247–260 (1995)
6. Avineri, E.: Soft computing applications in traffic and transport systems: a review. In: Hoffmann, F., Köppen, M., Klawonn, F., Roy, R. (eds.) Soft Computing: Methodologies and Applications, pp. 17–25. Springer, Berlin (2005)
7. Kikuchi, S., Chakroborty, P.: Place of possibility theory in transportation analysis. Transp. Res. Part B, **40**, 595–615 (2006)
8. Henn, V., Ottomanelli, M.: Handling uncertainty in route choice models: from probabilistic to possibilistic approaches. Eur. J. Oper. Res. **175**, 1526–1538 (2006)

Optimizing Airport Gate Assignments Through a Hybrid Metaheuristic Approach

Mario Marinelli, Gianvito Palmisano(✉), Mauro Dell'Orco, and Michele Ottomanelli

D.I.C.A.T.E.Ch., Technical University of Bari, Via E. Orabona, 4, 70126 Bari, Italy
{m.marinelli,gianvito.palmisano,dellorco,
m.ottomanelli}@poliba.it

Abstract. The gate assignment problem (GAP) is one of the most important problems that operations managers face daily. The GAP aims at determining an assignment of flights to terminal and ramp positions (gates), and an assignment of starting and ending times of the processing of a flight at its position. The objectives related to the flight gate assignment problem (FGAP) include the minimization of the number of flights assigned to remote terminals and the minimization of passengers' total walking distance. The main aim of this research is to find a novel methodology to solve the FGAP. In this paper, we propose a hybrid approach called Biogeography-based Bee Colony Optimization (B-BCO). This approach is obtained by properly combining two metaheuristics: biogeography-based (BBO) and bee colony optimization (BCO) algorithms. The proposed B-BCO model integrates the BBO migration operator into to bee's search behavior. The obtained results show the better performances of the proposed approach in solving FGAP when compared to BCO.

Keywords: Flight gate assignment · Bee Colony Optimization · Biogeography Based Optimization · Multicriteria analysis · Airport operations

1 Introduction

The complexity of airports management has increased significantly. Flight delays or accidents might happen if operations are not handled well, and the domino effect could occur affecting the whole operations of airports. In airports, the tasks related to flight gate assignment problem (FGAP) are some of the most important daily operations. Several researches have been dealt with FGAP in order to solve it in spite of its complexity. The main purpose of flight-to-gate assignments is to assign aircraft to suitable gates allowing passengers to conveniently embark and disembark. The distance a passenger has to walk in any airport to reach various key areas, including departure gates, baggage belts, and connecting flights provide an important performance measure of the quality of any airport. Certain walking distances are fixed, others are dynamic. In particular, the distances traveled from check-in counters to gates and from gate to gate (in the case of transfer or connecting passengers), change depending on how scheduled flights are assigned to gates.

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_19

Gate scheduling defines the assignment of flights to terminals or ramps, called gates. In recent years, all around the world, the demand for air transport has grown significantly. Most airlines adopt the structure of hub-and-spoke, which makes those hub airports even more crowded. Although these hubs usually have more infrastructure resources such as several runways and multiple terminals, they are still insufficient due to the participation of new airlines, more aircraft and higher frequency of flights. As a result, current schedule of the gate assignment for flights tends to be tight each other; in other words, there is less separation time between successive flights in the airports which have insufficient gates. Such schedules are usually not robust when faced with disruptions. In fact, there are many factors that affect gate assignments during real operations, including flight or gate breakdown, flight earliness or tardiness, emergency flights, severe weather conditions, errors made by staff and many others. If any of these disruptions lead to gate conflicts, knock-on-effect may affect the following schedules not only related to one gate, but also many airlines' following connections. Flight delay or cancel, transfer passenger missing connection are all among the events it may bring out. Therefore, gate reassignment is urgent especially for large disruptions such as arrival delay for a too long time, gate breakdown or airport temporary closure.

Gate assignment problem (GAP) is to assign a gate for each of the flight for a set of flights. Flight schedule defines the time period necessary to carry out operations to process a flight and the gate assigned to that flight. One of the most important objectives of flight gate assignment problem (FGAP) is to minimize the distance walked by passengers and the distance between connecting flights.

The most important information for gate scheduling refers to the arrival and departure time, the type of aircraft, and connections between pairs of successive flights served by the same aircraft. In the following sections, we describe the proposed model of the Biogeography-based Bee Colony Optimization (B-BCO), obtained by the combination of the Biogeography-Based Optimization (BBO) and Bee Colony Optimization (BCO). The model has been applied to the case study of the Italian international airport of Milan-Malpensa. The results and analysis are then reported. Finally, some concluding remarks are given.

2 A Survey of Models for the Flight-Gate Assignment Problem

There is an amount of research papers on airport gate assignment problems. Gate assignment problem (GAP) is to assign a gate for each of the flight for a set of flights. Dorndorf et al. [11, 12] and Bouras et al. [6] have written comprehensive reviews of the formulations and methodologies proposed in the literature.

Dorndorf et al. [13] consider a gate assignment problem which includes the weighting sum of three objectives: airline preferences, tows, and robustness. They transformed this problem for the special case with no shadow restrictions into an equivalent Clique Partitioning Problem and solves it using an ejection chain algorithm. Based on this work, Dorndorf et al. [14] further incorporated a forth objective that minimize the deviation from a given reference schedule and shown that it could be easily integrated into their previous existing model. Ding et al. [9, 10] focused on the over-constrained

gate assignment problem which aimed to minimize the number of ungated flights and total walking distances for transfer passengers. Tabu search and simulated annealing are used to solve their problems. Kim et al. [22, 23] considered three objectives in the gate assignment problem: passenger transit time, aircraft taxiing time, and expected gate conflict time.

Babic et al. [1] formulated the gate assignment problem as a linear problem. A branch-and-bound algorithm is used to minimize the walking distance and to find the optimal solution, but transfer passengers are not considered.

Most of the formulations are based on heuristic approaches. Xu and Bailey [33] define a tabu search algorithm for an airport with the aim of minimizing the passenger walking distances, considering connecting flights. They propose a two-stage algorithm to solve the problem. It uses both a greedy strategy to minimize the number of un-gated flights and an improved tabu search heuristic to minimize the total connection time. Drexl and Nikulin [15] consider multiple objectives such as minimizing the number of ungated flights and the total walking distances, as well as maximization of the total gate assignment preference score.

Other models try to improve the performance of static gate assignment by taking into account stochastic flight delays. Hassounah and Steuart [20] show that planned buffer times could improve schedule punctuality. Hamzwawi [19] introduces a rule-based system for simulating the assignment of gates to flights and for evaluating the effects of particular rules on gate utilization. Gosling [17] describes an expert system for gate assignment that has been applied to a major hub of Denver Stapleton airport. Srihari and Muthukrishnan [29] use a similar approach to solving the FGAP, applying additionally the sensitivity analysis. Yan and Chang [34] and Yan and Huo [35] used a fixed buffer time between two contiguous flights assigned to the same gate, in order to absorb the stochastic flight delays. Gu and Chung [18] designed a genetic algorithm, which generates effective solutions by calculating minimum extra delayed time schedules. Bolat [5] proposed mathematical models and heuristic procedures to provide solutions with the minimum dispersion of idle periods for the FGAP. Yan et al. [36] proposed a simulation framework, which can not only analyze the effects of stochastic flight delays on static gate assignments but also evaluate flexible buffer times and real-time gate assignment rules.

As for the types of approach, the FGAP has been formulated through a wide variety of models: Bard et al. [2] introduced an integral minimum cost network flow model, which reconstructs the operating programs of airlines in response to delays, transforming the routing problem in a time - based network, where the total time horizon is divided into distinct periods. Under some slight conditions, an optimum of the new model corresponds to the optimal solution of the original problem.

Cheng [8] describes the integration of mathematical programming techniques into a knowledge-based gate assignment system to provide partial parallel assignments with multiple objectives. Both optimization and rule-based approaches have been combined with simulation analysis in Baron [3]. A comparison of different metaheuristics (Genetic Algorithm, Tabu Search, Simulated Annealing) applied to the FGAP has been carried out by Cheng et al. [7]. Moreover, Hu and Di Paolo [21] have proposed an improved Genetic Algorithm applied to the FGAP considering a multi-objective function. These

metaheuristics differ from the proposed B-BCO algorithm because they are based on a solution *improvement* approach possibly not efficient with NP-hard problems subject to very strict constraints, like in the FGAP. In fact, these approaches can easily generate infeasible solutions that should be properly penalized through a carefully designed fitness function. Instead, the proposed B-BCO algorithm is based on a solution *construction* approach that always generates feasible solutions and improves them over iterations. In this paper, we extend a previous study by Marinelli et al. [24] in order to improve the proposed BCO algorithm introducing a novel hybridization with BBO.

3 Mathematical Formulation

In this paper, we discuss the FGAP which attempts to schedule a set of flights during any given planning day and assign them to different gates so as to minimize the number of flights that are not assigned to any gate; additionally, we seek to minimize walking distances between gates and hence minimize connection times. In our model, we will not stipulate secondary constraints which account for boarding times and other buffers between arrival and departure times since these are easily dealt with by extending flight arrival and departure durations.

Operation managers seek to minimize distances that passengers have to walk to reach departure gates, baggage belts, and connecting flights since this is a key quality performance measure for an airport. Aircraft stands at the terminal and off-pier stands on the apron are often simply referred to as "gates". As the gate assignment is a type of job-shop scheduling problem, its complexity increases exponentially as constraint size changes (e.g. number of flights, available gates, aircraft, flight block time, etc.). The NP-hard characteristic of the problem implies that there is no known algorithm for finding the optimal solution within a polynomial-bounded amount of time. When an aircraft arrives at the airport, it can be either assigned to the fixed terminal gates or, in particular conditions, it can be assigned to a remote terminal gate. All the fixed gates are usually equipped with passenger bridges, whereas passengers from flights assigned to remote gates can be transported to the terminal building by transfer bus shuttle services. The connection through the busses increases the connection time, so the main objective is minimizing total passenger walking distance and connection time.

Let us consider the following variables:

N number of flights
M number of gates, including remote gates
$f_{j,o}$ number of passengers from flight j to the baggage claim area
$f_{o,j}$ number of passengers from the check-in area to flight j
$f_{j,r}$ number of passengers from flight j to flight r
$w_{i,o}$ walking distance between gate i and baggage claim area
$w_{o,i}$ walking distance between check-in area and gate i
$w_{i,k}$ walking distance between gate i and gate k
$Y_{i,j}$ binary value representing the association of gate i to flight j

$Y_{k,r}$ binary value representing the association of gate k to flight r
$C_{j,i}$ binary value representing the compatibility between flight j and gate i

In this paper, the flight gate assignment problem is considered as composed of two main criteria:

- Minimization of total walking distance (TWD), including respectively the distance that a passenger walks to reach departure gates, baggage claim area, and connecting flights:

$$min\ \ TWD = \sum_{i=1}^{M}\sum_{j=1}^{N} f_{o,j} \cdot w_{o,i} \cdot Y_{i,j} + \sum_{i=1}^{M}\sum_{j=1}^{N} f_{j,o} \cdot w_{i,o} \cdot Y_{i,j} + \sum_{i=1}^{M}\sum_{k=1}^{M}\sum_{j=1}^{N}\sum_{r=1}^{N} f_{j,r} \cdot w_{i,k} \cdot Y_{i,j} \cdot Y_{k,r} \tag{1}$$

- Minimization of the number of flights assigned to remote terminal gates (*RG*), corresponding to the maximization of the number of flights assigned to fixed gates (*FG*):

$$min\ \ RG = \sum_{i \in RG}\sum_{j=1}^{N} Y_{i,j} \tag{2}$$

This optimization problem is subject to the following constraints:

1. compatibility between gate and airplane: a small aircraft can be assigned to a big gate, but a large aircraft cannot be assigned to a small gate. A large gate has the flexibility to accommodate various sizes of aircraft whereas a small gate is more limited. The compatibility is usually provided by the airport regulations:

$$C_{j,i} = 1 \qquad \text{if } Y_{i,j} = 1, \qquad 1 \le i \le M, \qquad 1 \le j \le N \tag{3}$$

2. every flight j must be assigned to exactly one gate including remote gates:

$$\sum_{i=1}^{M} Y_{i,j} = 1, \qquad 1 \le j \le N; \tag{4}$$

3. prevent schedule overlapping of two flights if they are assigned to the same gate:

$$t_{i,j}^{dep} < t_{i,z}^{arr} \quad \text{if } Y_{i,j} = 1 \text{ and } Y_{i,z} = 1, \qquad 1 \le i \le M, \qquad 1 \le j, z \le N \tag{5}$$

where $t_{i,j}^{arr}$, $t_{i,z}^{dep}$ are respectively the arrival and departure time of flights j and z associated to gate i.

The constraint (4) expresses the necessity that every flight must be assigned to one and only one gate; the constraint (5) prevents the schedule overlapping, in case two flights are assigned to the same gate.

Compatibility between stands and planes is usually provided by airport regulations: a small aircraft can be assigned to a large stand, but the vice versa is impossible. A large stand can accommodate various sizes of aircraft while the flexibility of a small stand is more limited. However, to minimize the space requirement, airport terminals are usually built with stands of different sizes, so that very large aircraft have their operations restricted to a few stands and, therefore, to a few gates.

The program (1)–(5) is evidently a multi-objective optimization problem, whose techniques of solution can vary widely, passing from outranking methods [27] to interactive programming [4, 16] or, more recently, to the Best - Worst method [26]. Since a deep study of the multicriteria analysis is out of the aim of this work, for the sake of simplicity we have used the most popular method for multiobjective optimization, the weighted sum method [30]. The method transforms multiple objectives into an aggregated objective function by multiplying each normalized objective function $f_i(x)$ by a weighting factor and summing up all weighted objective functions

$$V(x) = \sum_{i=1}^{\gamma} \delta_i \cdot f_i(x) \tag{6}$$

where the coefficient δ_i is the weighting factor for the i-th objective function and γ is the number of functions to be aggregated. If $\sum_{i=1}^{\gamma} \delta_i = 1$ and $0 \le \delta_i \le 1$, the weighted sum is said to be a convex combination of objectives.

In our case $\gamma = 2$, therefore, $\delta_2 = 1 - \delta_1$ and the combination of Eqs. (1) and (2) gives:

$$\begin{aligned} min\ F = \delta_1 \cdot \left(\sum_{i=1}^{M} \sum_{j=1}^{N} f_{o,j} \cdot w_{o,i} \cdot Y_{i,j} + \sum_{i=1}^{M} \sum_{j=1}^{N} f_{j,o} \cdot w_{i,o} \cdot Y_{i,j} + \sum_{i=1}^{M} \sum_{k=1}^{M} \sum_{j=1}^{N} \sum_{r=1}^{N} f_{j,r} \cdot w_{i,k} \cdot Y_{i,j} \cdot Y_{k,r} \right) \\ + (1 - \delta_1) \cdot \sum_{i \in RG} \sum_{j=1}^{N} Y_{i,j} \end{aligned} \tag{7}$$

The program (7) with the constraints (4) and (5) is hardly, if not at all, manageable by ordinary analytical tools; therefore, in this paper, we have proposed a metaheuristic, based on the collective intelligence concepts, called Bee Colony Optimization (BCO).

In the next section, we present the proposed methodology based on the fusion of two metaheuristics, Biogeography-Based and Bee Colony Optimization metaheuristics, to solve this problem.

4 The Proposed Optimization Model

In this paper, we present an extension of a previous work by Marinelli et al. [24] where a metaheuristic approach based on the Bee Colony Optimization (BCO) was proposed. That approach is extended to obtain a new hybrid methodology (B-BCO) obtained by combining BCO and the Biogeography-Based Optimization (BBO). The hybrid model allows to combine the convergence of individual models capacity and thus allows to determine more of feasible solutions.

4.1 Biogeography-Based Optimization (BBO)

Biogeography-based optimization (BBO), proposed by Simon [28], is a stochastic optimization technique for solving multimodal optimization problems. It is based on the concept of biogeography, which deals with the distribution of species that depend on different factors such as rainfall, diversity of topographic features, temperature, land area, etc. Biogeography describes how species migrate one island or habitat to another, how new species arise and how species become extinct. The BBO algorithm uses these concepts to improve the habitat suitability index (HIS) of all habitats, which results in evolving the initial random solutions for a particular problem. The BBO algorithm starts with a random set of habitats. Each habitat has different individuals that correspond to the number of variables of a particular problem. In addition, each habitat has its own immigration, emigration, and mutation rates. This mimics the characteristic of various geographically separated locations in nature.

A habitat is an island which is geographically isolated from other habitats. The habitat having high suitability index (HSI) means that the habitat is geographically well suited for species living. The species have more possibility to migrate into the neighboring habitat when (i) the high HSI habitats have more species (ii) habitat has low HSI. This process is named as emigration. In immigration, the species move towards the high HSI habitat having few species. The immigration and emigration of species in a habitat are known as migration. The migration of species in a single habitat is explained in Fig. 1. From this figure, it has been seen that when there are no species in a single habitat, then the immigration rate (λ) of the habitat will be maximized (I). On the other hand, when the number of species increases, then only a few can successfully survive. Therefore, the immigration rate is decreased and become zero at $S = S_{max}$. Similarly, considering the emigration curve in the figure, the emigration rate (μ) increases with the increases of the species and hence the emigration rate is higher for the more crowded

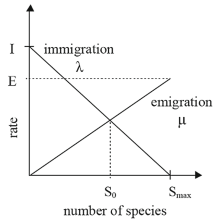


Fig. 1. Species model of a single habitat.

habitat than the less populated habitat. The maximum possible immigration rate occurs when there are zero species in the habitat. Increasing the number of species, the habitat becomes more crowded and, therefore, able to accommodate a number of lower species, thereby decreasing the rate of immigration.

However, species diversity is correlated with HSI, so when more species arrive at a low HSI habitat, the habitat's HSI will tend to increase. The immigration rate λ and the emigration rate μ are functions of the number of species in the habitat. The maximum possible immigration rate I occurs when there are zero species in the habitat. S_{max} is the largest number that the habitat may contain; not being able accommodate new species, it has a rate of immigration zero. If there are no species in the habitat, then the emigration rate is zero. With the increase in the number of species in the habitat, the habitat becomes ever more crowded, therefore of the representative species leave the habitat and increases the rate of emigration (Fig. 1).

The optimization steps of the BBO algorithm are described as follows while the flow chart of the proposed approach is described completely in Fig. 2.

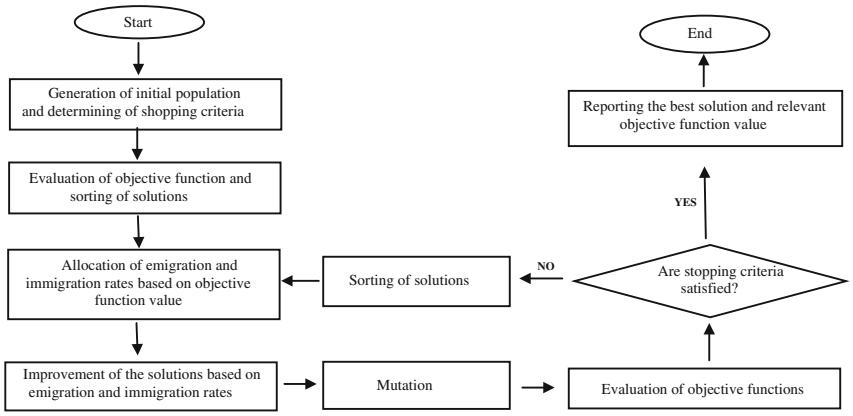


Fig. 2. General steps of the BBO algorithm.

The initialization phase of the BBO algorithm is to set the parameter of BBO algorithm. The parameters are maximum species count (S_{max}), maximum migration rate (E), immigration rate (I), the maximum mutation rate (m_{max}), habitat modification probability (p_{mod}), the number of habitats and the number of iterations. In BBO algorithm, each individual or decision variable is considered as a "habitat" with a habitat suitability index (HSI) to measure the individual where the length of the habitat is depending on the number of decision variables. This HSI measures the goodness of the solution and is same as that of fitness value in population-based optimization algorithms. The corresponding variables of individuals that characterize habitability feature are called as suitability index variables (SIVs). These initialized randomly within the maximum and minimum limits as $x_{min,i} + U_d(x_{max,i} - x_{min,i})$ where U_d is the uniformly distributed random number lies between (0, 1). Each habitat, in a population size NP, is represented by N-dimensional vector as $H = [\mathrm{SIV}_1, \mathrm{SIV}_2, \ldots, \mathrm{SIV}_N]$ where N is the number of the SIVs

(features) to be evolved for optimal HSI. HSI is the degree of acceptability that is determined by evaluating the objective function i.e. HSI = $f(H)$.

There are mainly two main operators: migration (which include both emigration and immigration) and mutation. Migration is the adaptive process which is used to change the existing habitat's SIV. Each individual has its own immigration rate λ and emigration rate μ for improving the solution and thus evolve a solution to the optimization problem and is modified based on its probability p_{mod}, that is a user-defined parameter. Based on this probability, its immigration rate (λ) of each habitat is decided whether to modify each SIV's or not. Another habitat is selected based on emigration rate μ, its SIV will migrate randomly to the selected habitat's SIV. A good solution is a habitat with relatively high μ and low λ, while the converse is true for a poor solution. High HSI habitat resists changing more the low HSI habitat. Poor solutions accept more useful information from a good solution, which improves the exploitation ability of the algorithm. At initially, the immigration and emigration rates are the functions of the number of species in the habitat and they are evaluated as:

$$\lambda_k = I\left(1 - \frac{k}{S_{max}}\right) \tag{8}$$

$$\mu_k = E\left(\frac{k}{S_{max}}\right) \tag{9}$$

where I is the maximum possible immigration rate; E is the maximum possible emigration rate; k is the number of the species of the kth individual, and S_{max} is the maximum number of species. I and E are often set equal to 1, or slightly less than 1. The pseudo code of the migration operator is depicted in Algorithm. On the other hand, in BBO, mutation is used to increase the diversity of the population to get the best solution. The low and high HSI valued habitats have less possibility to mutate while average HSI valued habitats have more possibility to mutate. To see how the proposed method works, a flow-chart of the BBO is reported in Fig. 2 and a conceptual picture of the migration between the habitats is shown in Fig. 3.

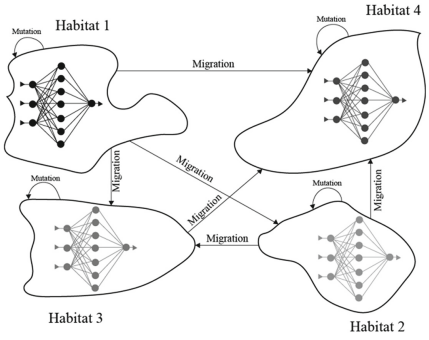


Fig. 3. Conceptual model of migration between habitats.

4.2 Bee Colony Optimization (BCO)

The Bee Colony Optimization (BCO) is a metaheuristic inspired by the foraging behavior of honeybees. It is one of the first algorithms that uses basic principles of collective bee intelligence for solving combinatorial optimization problems. BCO is a stochastic, random-search population-based technique. The BCO was first applied to transportation problems by Teodorovic et al. [31, 32]. It was motivated by the analogy found between the natural behavior of bees searching for food and the behavior of optimization algorithms searching for an optimum in combinatorial optimization problems. The main idea is to build the multi-agent system (a colony of artificial bees) able to efficiently solve hard optimization problems. Artificial bees investigate through the search space looking for feasible solutions. In order to increase the quality of produced solutions, autonomous artificial bees collaborate and exchange information. Sharing the available information and using collective knowledge, artificial bees concentrate on more promising areas and slowly abandon solutions from those less promising. Step by step, artificial bees collectively generate and/or modify their solutions. BCO performs its search in iterations until some predefined stopping criterion is satisfied. All artificial bees are located in the hive at the beginning of the search process. During the search process, artificial bees communicate directly. Each artificial bee makes a series of local moves, and in this way incrementally constructs a solution of the problem. Bees are adding solution components to the current partial solution until they create one or more feasible solutions. When flying through space, artificial bees perform forward step or backward step. During a forward step, bees create various partial solutions. They do this via a combination of individual exploration and collective experience from the past. After that, they perform a backward step, i.e. they return to the hive. In the hive, all bees participate in a decision-making process. The search process is composed of iterations. Each iteration ends when one or more feasible solutions are created (Fig. 4).

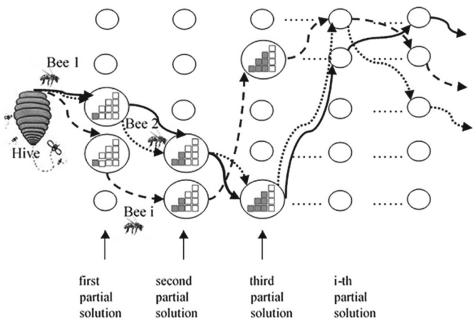


Fig. 4. Stages of the Bee Colony Optimization.

Like Dynamic Programming, the BCO also solves combinatorial optimization problems in stages. Each of the defined stages involves one optimizing variable. Let us denote by ST = $\{st_1, st_2, \ldots, st_m\}$ a finite set of pre-selected stages, where m is the number of stages. By B we denote the number of bees to participate in the search process, and by

I the total number of iterations. The set of partial solutions at stage st_j is denoted by S_j ($j = 1, 2,\ldots, m$). The following is the pseudo-code of the Bee Colony Optimization.

Initialization. Determine the number of bees *B*, and the number of iterations *I*. Select the set of stages ST = {st_1, st_2 ,..., st_m}. Find any feasible solution x of the problem. This solution is the initial best solution.
Set i = 1. Until i = I, repeat the following steps:
  Set j = 1. Until j = *m*, repeat the following steps:
    Forward step: Allow bees to fly from the hive and to choose B partial solutions from the set of partial solutions S_j at stage st_j.

    Backward step: Send all bees back to the hive. Allow bees to exchange information about the quality of the partial solutions created and to decide whether to abandon the created partial solution and become again uncommitted follower, continue to expand the same partial solution without recruiting the nestmates, or dance and thus recruit the nestmates before returning to the created partial solution.
  Set, j = j + 1.
  If the best solution x_i obtained during the ith iteration is better than the best-known solution, update the best-known solution ($x = x_i$).
Set i = i + 1

4.3 Biogeography-Based Bee Colony Optimization (B-BCO)

The previous two metaheuristics have been combined to obtain the Biogeography-Based Bee Colony Optimization (B-BCO). In this model, the bees tend always moving to find a partial solution. With the introduction of the concepts of migration and immigration of BBO, bees along optimal partial routes may move from one habitat to another (migration operator). In this work, B-BCO is used to find an optimal path through an artificial network that represents the decision space. The network is composed of different habitats and for each habitat, there are different layers (previously called 'stages') which represent the set of flights, temporally ordered according to a given schedule. Each node

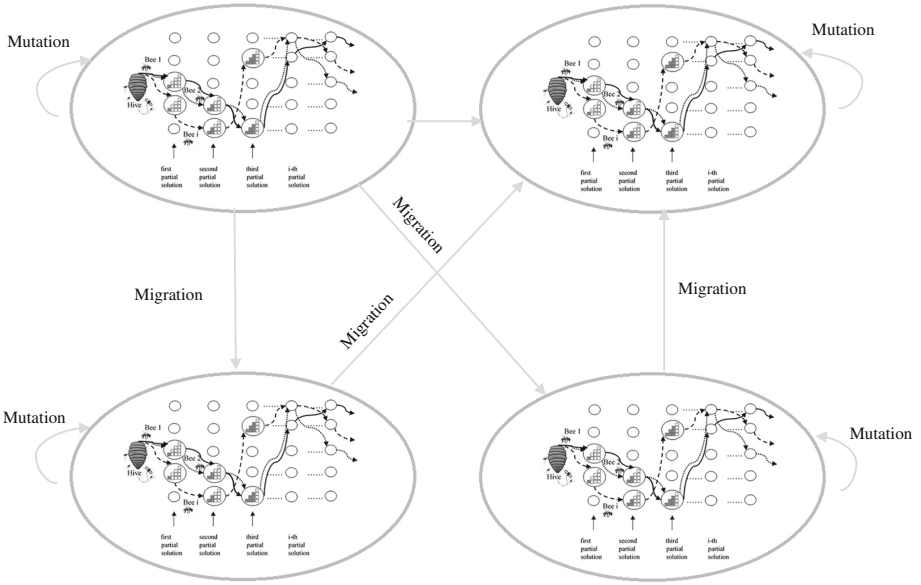


Fig. 5. The resulting artificial network for each habitat and between different habitats.

represents an association of a flight F_i to an available gate G_j in the airport, so it refers to variable $Y_{i,j}$ in the problem formulation. During a single iteration, each bee finds partial solutions and chooses the next node through a roulette wheel selection. All the partial solutions are identified observing the constraints of the optimization problem and the associated fitness value is given by the objective function. As a result, a path of the artificial network corresponds to a particular flight gate assignment found by a bee in the colony. At the end of each iteration, all the solutions found are evaluated referring to the associated fitness value and the best assignment is saved. Thus, a new iteration starts searching for new solutions until the maximum number of iterations is reached. In the following, the pseudo-code of the algorithm is reported. Figure 5 shows the networks obtained for each habitat and the emigration operations between different habitats.

Initialization. Determine the number of bees B, the number of iterations I and number of habitats H. Select the set of stages ST = {st_1, st_2 ,…, st_m}. Find any feasible solution x of the problem. This solution is the initial best solution.
Set i = 1. Until i = I, repeat the following steps:
Set j = 1. Until j = m, repeat the following steps:
For each habitat
Perform forward and backward step
Compute species number sorting bees from most fit to least fit
Calculate λ, μ of each bee in the habitat
For each bee in the habitat
If rand < λ *then*
Select an emigrating bee with probability μ
End if
End for each bee
End for each habitat
Set j = j + 1.
If the best solution x_i obtained during the ith iteration is better than the best-known solution, update the best-known solution ($x = x_i$).
Set i = i + 1.

5 Application and Results

In order to evaluate the outcomes of the proposed B-BCO, the method has been applied to the case of the Milano-Malpensa international airport (in the following called Malpensa). The outcomes have been compared to the BCO approach. Malpensa airport has two terminals, for international and domestic flights, and an area reserved for freight traffic called respectively Malpensa 1, Malpensa 2 and Malpensa Cargo. The airport is strategically important both for Italy and Europe. In 2012, the Malpensa airport was ranked second in Italy after Rome-Fiumicino airport for overall passenger traffic, with about 18.5 million passengers (on average 50 000 per day), and in the first place for freight traffic, with 414.317 tons.

We have taken into account the flight scheduling of May 2012. The database consists of 178 flights and 65 gates. The proposed approach has been applied considering the structure of the airport and, in particular, an additional constraint related to the assignment of a flight to international or domestic gates based on its origin/destination. The compatibility between gate and airplane has been determined according to Malpensa Airport Regulations [25].

Results have been carried out in terms of best objective function values obtained for different δ_1 values (Figs. 6, 7, 8 and 9). Table 1 reports in detail the obtained results for two models in terms of TWD (total walking distance), RG (flights to remote gates) and best objective function value F^* for the considered δ_1 values.

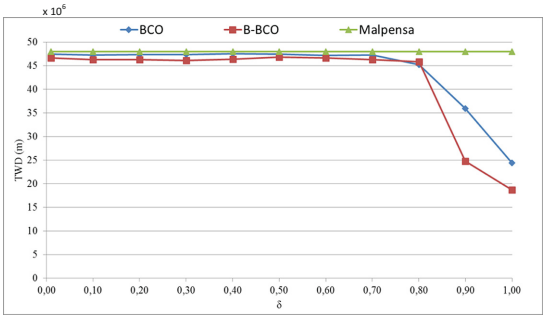


Fig. 6. Resulting total walking distance obtained by B-BCO (red line), BCO (blue line) for different δ_1 values and Malpensa schedule (green line).

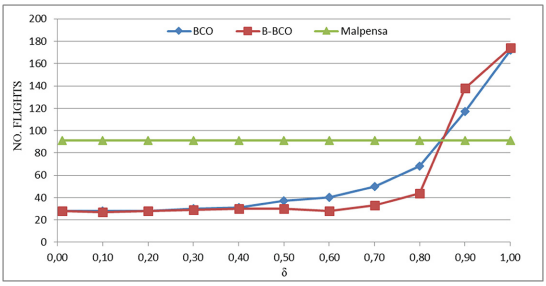


Fig. 7. The resulting number of flights assigned to remote gates (RG) obtained by B-BCO (red line), BCO (blue line) for different δ_1 values and Malpensa schedule (green line).

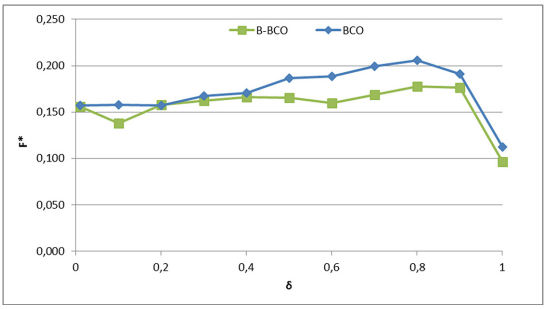


Fig. 8. Resulting best objective function value F^* found by BCO and B-BCO for different δ_1 values.

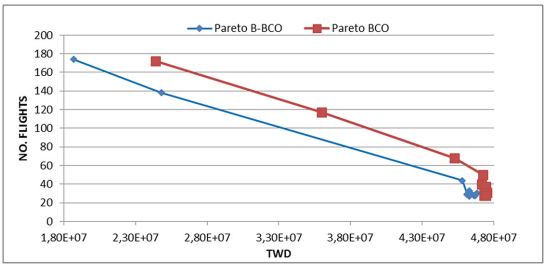


Fig. 9. The resulting Pareto front obtained by the optimal solutions found.

Table 1. Comparison of the results obtained by BCO and B-BCO.

δ_1	TWD (m)		RG (flights)		F^*	
	BCO	B-BCO	BCO	B-BCO	BCO	B-BCO
0.01	47458686	46704031	28	28	0.157	0.156
0.10	47326985	46284727	28	27	0.158	0.138
0.20	47412610	46345799	28	28	0.157	0.158
0.30	47346654	46142048	30	29	0.167	0.162
0.40	47531467	46368517	31	30	0.171	0.166
0.50	47476033	46853129	37	30	0.186	0.166
0.60	47159774	46652882	40	28	0.189	0.160
0.70	47256210	46318181	50	33	0.199	0.169
0.80	45257497	45813261	68	44	0.206	0.178
0.90	35979382	24810019	117	138	0.191	0.176
1.00	24411882	18695425	172	174	0.112	0.096
% improvement	**5%**		**6%**		**9%**	

A sensitivity analysis has been made in order to highlight the role of the variable δ_1 in the decision-making process.

We can observe that the solutions found by B-BCO are almost always better than the BCO solutions. As a matter of fact, the total walking distance is almost always lower (red line in Fig. 6) than Malpensa and BCO resulting in an improvement of 5%. The number of flights assigned to remote gates is lower than the others (improvement of 6%) except for $\delta_1 > 0.8$ where B-BCO tends to minimize the TWD. In terms of best objective function value F^*, the B-BCO outperforms the BCO approach with an improvement of 9%.

Finally, in Fig. 9, results show that all points of the obtained BBCO Pareto front are below the BCO Pareto front. Thus, they are better for both considered objectives.

6 Conclusions

In this paper, we have proposed a novel hybrid optimization algorithm to solve the flight gate assignment problem (FGAP). This hybridization has been obtained combining two promising metaheuristics: Biogeography-based (BBO) and Bee Colony Optimization

(BCO) algorithms. The problem has been defined considering two objective functions: minimization of the total walking distance and the number of flights assigned to remote gates, subject to schedule and compatibility constraints. These objectives have been reduced to a single one introducing a weighting parameter.

The proposed algorithm, called Biogeography-based Bee Colony Optimization (B-BCO), has been applied to a real schedule of Malpensa airport. Results highlight the effectiveness of the proposed B-BCO when compared to the BCO approach and Malpensa schedule.

Advantages introduced by the proposed hybrid approach are related to a better solution quality and better convergence capabilities. In particular, through B-BCO, we have obtained an improvement of 9% related to optimal objective function values F^*. Moreover, the solutions found by B-BCO are usually better than BCO in terms of both total walking distance and number of flights assigned to remote gates.

Moreover, the results show that the proposed approach offers acceptable solutions, validated on real data and that it could be applied to other airports.

Future research will focus on a more complex formulation of the problem considering more criteria and constraints related to airline companies' preferences andagreements.

References

1. Babic, O., Teodorovic, D., Tosic, V.: Aircraft stand assignment to minimize walking. J. Transp. Eng. **110**, 55–66 (1984)
2. Bard, J., Yu, G., Arguello, F.: Optimizing aircraft routings in response to groundings and delays. IIE Trans. **33**, 931–947 (2001)
3. Baron, P.: A simulation analysis of airport terminal operations. Transp. Res. **3**, 481–491 (1969)
4. Benayoun, R., de Montgolfier, J., Tergny, J., Larichev, O.: Linear programming with multiple objective functions: step-method (stem). Math. Program. **1**, 366–375 (1971)
5. Bolat, A.: Models and a genetic algorithm for static aircraft gate assignment problem. J. Oper. Res. Soc. **52**, 1107–1120 (2000)
6. Bouras, A., Ghaleb, M.A., Suryahatmaja, U.S., Salem, A.M.: The airport gate assignment problem: a survey. Sci. World J. **2014**, 1–27 (2014)
7. Cheng, C.-H., Ho, S.C., Kwan, C.-L.: The use of meta-heuristics for airport gate assignment. Expert Syst. Appl. **39**, 12430–12437 (2012)
8. Cheng, Y.: A knowledge-based airport gate assignment system integrated with mathematical programming. Comput. Ind. Eng. **32**, 837–852 (1997)
9. Ding, H., Lim, A., Rodrigues, B., Zhu, Y.: New heuristics for the over constrained airport gate assignment problem. J. Oper. Res. Soc. **55**, 760–768 (2004)
10. Ding, H., Lim, A., Rodrigues, B., Zhu, Y.: Aircraft and gate scheduling optimization at airports. In: Proceedings 37th Annual Hawaii International Conference on System Sciences, pp. 5–8 (2004)
11. Dorndorf, U., Drexl, A., Nikulin, Y., Pesch, E.: Flight gate scheduling: state-of-the-art and recent developments. Omega **35**, 326–334 (2007)
12. Dorndorf, U., Jaehn, F., Lin, C., Ma, H., Pesch, E.: Disruption management in flight gate scheduling. Stat. Neerl. **61**, 92–114 (2007)
13. Dorndorf, U., Jaehn, F., Ma, H., Pesch, E.: Modelling robust flight-gate scheduling as a clique partitioning problem. Transp. Sci. **42**, 292–301 (2008)

14. Dorndorf, U., Jaehn, F., Ma, H., Pesch, E.: Flight gate scheduling with respect to a reference schedule. Ann. Oper. Res. **194**, 177–187 (2012)
15. Drexl, A., Nikulin, Y.: Multicriteria airport gate assignment and Pareto simulated annealing. IIE Trans. **40**, 385–397 (2008)
16. Geoffrion, A., Dyer, J., Feinberg, A.: An interactive approach for multicriterion optimization with an application to the operation of an academic department. Manage. Sci. **19**(4–Part–1), 357–368 (1972)
17. Gosling, G.: Design of an expert system for aircraft gate assignment. Transp. Res. **24A**, 59–69 (1990)
18. Gu, Y., Chung, C.A.: Genetic algorithm approach to aircraft gate reassignment problem. J. Transp. Eng. **125**(5), 384–389 (1999)
19. Hamzwawi, S.: Management and planning of airport gate capacity: a microcomputer-based gate assignment simulation model. Transp. Plann. Technol. **11**, 189–202 (1986)
20. Hassounah, M., Steuart, G.: Demand for aircraft gates. Transp. Res. Rec. **1423**, 26–33 (1993)
21. Hu, X., Di Paolo, E.: An efficient genetic algorithm with uniform crossover for the multi-objective airport gate assignment problem. In: Goh, C.K., Ong, Y.S., Tan, K.C. (eds.) Multi-Objective Memetic Alg. SCI, vol. 171, pp. 71–89. Springer, Heidelberg (2009). doi: 10.1007/978-3-540-88051-6_4
22. Kim, S.H., Feron, E.: Numerical analysis of gate conflict duration and passenger transit time in airport. arXiv Preprint arXiv:1308.6217 (2013)
23. Kim, S.H., Feron, E., Clarke, J.P., Marzuoli, A., Delahaye, D.: Airport gate scheduling for passengers, aircraft, and operation. In: Tenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2013) (2013)
24. Marinelli, M., Dell'Orco, M., Sassanelli, D.: A metaheuristic approach to solve the flight gate assignment problem. Trans. Res. Procedia **5**, 211–220 (2015)
25. Milano-Malpensa Airport Regulations (2010). http://www.milanomalpensa2.eu/it/download?_fname=/AssetsProtected/contentresources_2/mmo/486/C_2_mmo_1287_file.pdf&_ofname=RegScaloMXP2.1%20mag.%20%2710-%20allegati.pdf
26. Rezaei, J.: Best-worst multi-criteria decision-making method. Omega **53**, 49–57 (2015)
27. Roy, B.: La méthode ELECTRE. Revue d'Informatique et de Recherche Opérationelle (RIRO) **8**, 57–75 (1968)
28. Simon, D.: Biogeography-based optimization. IEEE Trans. Evol. Comput. **12**, 702–713 (2008)
29. Srihari, K., Muthukrishnan, R.: An expert system methodology for aircraft-gate-assignment. Comput. Ind. Eng. **21**, 101–105 (1991)
30. Stadler, W.: A survey of multicriteria optimization, or the vector maximum problem. J. Optim. Theory Appl. **29**, 1–52 (1979)
31. Teodorovic, D., Dell'Orco, M.: Bee colony optimization – a cooperative learning approach to complex transportation problems. In: Advanced OR and AI Methods in Transportation, pp. 51–60 (2005)
32. Teodorovic, D., Lucic, P., Markovic, G., Dell'Orco, M.: Bee colony optimization: principles and applications. In: Neural Network Applications in Electrical Engineering, pp. 151–156 (2006)
33. Xu, J., Bailey, G.: The airport gate assignment problem: mathematical model and a tabu search algorithm. In: Proceedings of the 34th Annual Hawaii International Conference on System Sciences, p. 3032. IEEE (2001)
34. Yan, S., Chang, C.: A network model for gate assignment. J. Adv. Transp. **32**, 176–189 (1998)
35. Yan, S., Huo, C.: Optimization of multiple objective gate assignments. Transp. Res. **35A**, 413–432 (2001)
36. Yan, S., Shieh, C.-Y., Chen, M.: A simulation framework for evaluating airport gate assignments. Transp. Res. **36**, 885–898 (2002)

Shipping Liner Company Stowage Plans: An Optimization Approach

Daniela Ambrosino[1(✉)], Massimo Paolucci[2], and Anna Sciomachen[1]

[1] Department of Economics and Business Studies (DIEC), University of Genoa, Via Vivaldi 5, 16126 Genova, Italy
{daniela.ambrosino,anna.sciomachen}@economia.unige.it
[2] Department of Informatics, Bioengineering, Robotics, and Systems Engineering (DIBRIS), University of Genoa, Via Opera Pia 13, 16145 Genova, Italy
massimo.paolucci@unige.it

Abstract. The problem of determining the stowage plan for each port included in the circular route of a containership is the planning problem daily faced by the shipping line coordinator (SC). The SC has to determine how to stow in each port a given set of containers into bay locations, either on the deck or in the hold, in order to satisfy some structural and operational constraints, all the transport demand and minimize the ship berthing time. The time spent by the ship at the ports is minimized when the SC defines a stowage plan that avoids re-handles, balances the workload of the cranes working in parallel at each port and requires to open and close hatches the least possible when moving containers in the holds. We propose a new MIP formulation, and a MIP heuristics for finding stowage plans for containerships up to 18000 TEUs. Moreover, the new formulation includes new and more accurate stability conditions.

Keywords: Containership stowage planning · Liner shipping optimization · Mixed Integer Linear Programming (MIP) model · Matheuristics

1 Introduction

Among the freight transport modes, maritime transport has the largest share. About the 80% of international traffic uses ships [18]: ships are preferred for transferring cargos over long distances for their cost effectiveness and because they are environmentally friendly [13].

Ports are in close relationship with shippers, exporters, importers, logistics corporation state authorities and others corporations, thus given a great contribution to competitiveness of the national economies; they must be managed in an efficient and sustainable way [19].

Moreover, the major part of the maritime transport is containerized thus particular attention is paid to liner shipping that offer regular services (with a fixed time schedule) on fixed routes. The optimization in container liner shipping has become a very important issued, studied in the recent literature. In [16] the relation between port

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_20

activities and shipping lines is emphasized, while [6] focuses on the improvement of shipping services. In [8] the authors discuss a variety of liner shipping problems arising on the strategic, tactical and operational planning levels.

A survey of optimization techniques required to manage efficiently import/export containerized flows at maritime terminals is presented in [14].

The stowage planning problems concern the operational level liner shipping problems and are influenced by many other tactical and strategic decisions.

The impact of stowage planning and unloading/loading operations on both the efficiency of the terminals and the liner shipping choices has been stressed in [5, 9, 12] among others. In particular, many recent papers are devoted to the optimization of the stowage plans for containerships.

The stowage planning is a problem faced by the shipping line coordinator (SC) and the terminal planner (TP); these two decision makers solve the same problem with a different degree of detail, at different level and interacting during their decision process [3, 7, 15].

The SC has a view of the whole trip of the ship and defines a stowage plan for each port of the ship trip, thus updating the plan accordingly whenever a new *o-d* demand for that ship is received and accepted. The SC stowage plan is sent to the TP before the ship arrival in a port.

The SC stowage plan gives general instructions about how to stow different groups of containers, depending on their destination, type, size and weight. In particular, for each group of containers the SC indicates a portion of the ship in which load it.

Some recent studies deal with this problem, known as Multi Port-MBPP (MP-MBPP). The first attempt to introduce and solve the MP-MBPP is due to [20, 21]. These authors proposed a methodology for generating stowage plans for a containership on a multi-port journey based on a decomposition approach: the decision process was decomposed into two planning sub-processes i.e. a strategic and a tactical one.

In [7] a classification schema for the literature focusing on stowage plans is proposed; this classification is based on four-fields: the decision maker (SC or TP), the type of the output of the decisional problem, the settings of the problem, and the objective function to be minimized. Following the decomposition approach of the decision process proposed in [20], papers focusing only on MP- MBPP at a strategic level are for example [3, 10]. In [10] a linear model with ballast tanks for generating master plans that includes the main stability and stress moments calculations is proposed for handling variable displacements. In [3] the authors proposed a MIP model for solving MP-MBPP for taking into account the hatches, including different type of containers (i.e. standard cargo, reefers and open top) and to guarantee that all containers be loaded on board, even if some re-handles are necessary. Note that re-handles are permitted either for loading some containers in free locations in the hold or for unloading containers from the hold.

In this paper, the MP-MBPP studied in [3] is extended for taking into account another important aspect related to the unloading/loading operations that influences the efficiency of the terminal: we deal with the movements of the hatches. Many shipping companies have the aim of minimizing the number of hatch movements during the unloading and loading phases, thus in the model here described this new aspect is

introduced. Moreover, new and more accurate stability constraints are modelled. To solve the resulting NP-hard model, a MIP heuristics derived from [2] is proposed and an experimental campaign is detailed in order to show the capability of the proposed solution method. The interested reader can refer to [17] for some consideration on the complexity of container stowage problems.

The paper is organized as follows: the problem under investigation together with the new model is presented in details in Sect. 2. Section 3 presents the MIP heuristics able to face this new operative hatch requirement. Computational results based on different loading scenarios for large size containerships up to 18000 TEUs are given in Sect. 4, while Sect. 5 draws the main conclusions and outlines for future works.

2 Problem Under Investigation

Given a ship characterized by its structure, its current cargo, a sequence of ports to visit and a certain demand of transport to satisfy, the problem faced by the SC consists of defining the stowage plan for the containership in such a way to load on board all the containers of the transport demand, satisfy the structural and operative constraints, related to the containers, the ship and the terminals, and minimize the time spent by the ship at the ports of its route for loading/unloading operations.

The transport demand is expressed in number of containers of a given size, type, weight class and related loading and destination port. In this work standard, reefer and open top containers of both 20 and 40 feet are considered. Containers are categorized in three different classes of weight, that are light, medium and heavy; their port of load and destination belong to the circular route of the ship.

From the structure of the ship, described by means of bays, rows, tiers and hatches (see for example [2]), the SC knows the available locations in which it is possible to load the containers of the transport demand. In Fig. 1a is reported the representation of the cells of a bay of a ship; these cells have been divided into 6 subsets thanks to the hatches in the considered bay (Fig. 1b). We will refer to subsets S1, S2, S3 as deck locations, while to subsets S4, S5 and S6 as hold locations. In particular, S1 and S4 are, respectively, the deck and the hold locations of the same hatch. In more details, S1 is the set of cells belonging to bay 22, rows from 08 to 16 and tiers from 72 to 80, while S4 is the set of cells of bay 22 belonging to rows from 08 to 14 and tiers in the hold from 02 to 20.

Thus the whole ship can be split into hold/deck locations in correspondence of each hatch.

The SC has to decide the assignment of the containers to the hatch locations; in particular, the decision regards the number of containers of a given group (having fixed size, type, weight class, origin and destination port) to assign to the locations of a specific hold/deck hatch of the ship. A first attempt to take decisions for groups of containers is the work [1] related to the stowage plan for a single port. The SC assignment must respect the capacity of each deck/hold location and that of all different types of containers (i.e. reefer, open top, etc.) the specific capacity for each type. The SC can assign containers of different groups to the same deck/hold location; however, it is common usage to put in the same deck/hold location containers having the same destination.

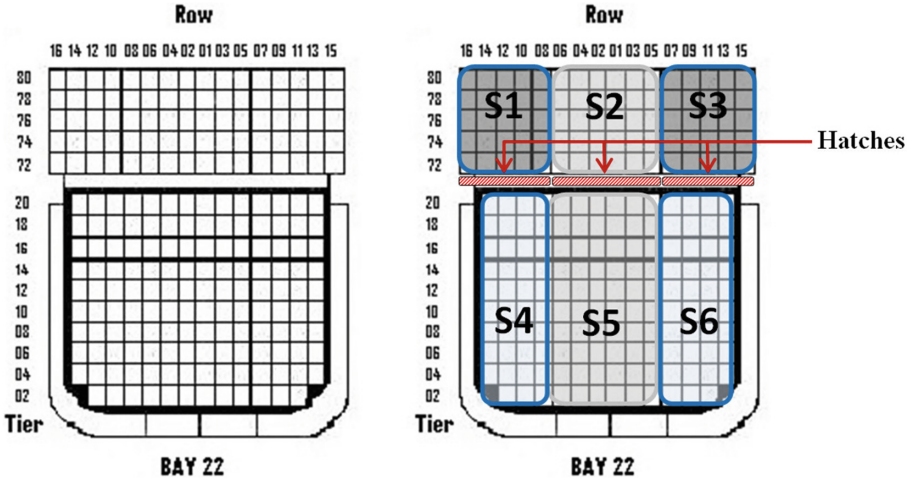


Fig. 1. Deck and hold locations of a bay

The assignment of the containers to the locations must also satisfy some requirements in terms of ship stability. In this work the stability conditions have been derived from [9], that at least to the knowledge of the authors, together with [4], is the only work dealing with stability constraints taking into account the structure of the ship, its draft centre of gravity and trim limits, the maximum and minimum shear limits at particular points of the ship (called stations), and its capacity.

The time spent by the ship at the ports is minimized when the SC is able to define a feasible stowage plan for each port that, during the unloading/loading operations, avoids re-handles (i.e., unproductive movements), balances the workload of the cranes working in parallel at each port (i.e. the loading and unloading operations performed by each crane is balanced), requires to open and close the lowest number of hatches for moving containers in the hold.

In the planning process the SC knows the set of quay cranes serving the ship in parallel at each port and the subset of bays (and thus of hatches) served by each crane at each port.

At a port p re-handles can occur in the locations on the deck above a hatch for both unloading containers bound for port p and stowed in the locations in the hold below the same hatch (re-handles for unloading), and for loading containers departing from the current port p in the free locations in the hold below the same hatch (re-handles for loading); interested readers can refer to [2] for more details. The SC can permit re-handles for loading and thus can move containers from the deck locations, open the hatch and load some containers in the hold free locations; obviously, this last case is not convenient since cause re-handles and hatch movements. Anyway, this choice can be necessary for avoiding to leave containers in the quay.

3 A MIP Model for the MP-MBPP

In this section we propose a mixed integer programming (MIP) model for solving the above described problem; the required notation is the following:

Sets and indexes

P	set of ports (1 = first port, P_{max} = last port);
D	set of *o-d* shipping in the considered circular route of the ship;
D_p	subset of *o-d* shipping passing through port p or having as origin port p;
D_p^-	subset of *o-d* shipping passing through port p;
$A(d)$	subset of destination successive to destination d;
H	set of hatches of the ship;
ST	set of stations of the ship;
H_i	subset of hatches belonging to station i;
Y_p	set of quay cranes serving the ship in parallel at port p;
H_{pc}	subset of hatches served by crane c at port p;
G	set of weight classes (e.g., 1 = light, 2 = medium, 3 = heavy) of the containers;
S	set of size of containers (e.g. 20' and 40');
T	set of type of containers (e.g. 0 = standard, 1 = reefer, 2 = open top);

Constants

$TEUD_h$	capacity in TEU on the deck of hatch h, $h \in H$;
$TEUH_h$	capacity in TEU in the hold of hatch h, $h \in H$;
$TEUD_{ht}$	capacity in TEU for container of type t on the deck of hatch h, $h \in H$, $t \in T$;
$TEUH_{ht}$	capacity in TEU for container of type t in the hold of hatch h, $h \in H$, $t \in T$;
N_{od}^{stg}	number of containers of size s, type t, weight class g to be loaded at origin port o with destination port d, $\forall s \in S$, $t \in T$, $g \in G$, $(o,d) \in D$;
W_g	weight for a container in weight class g, $\forall\ g \in G$;
WD_h	weight limitation for deck locations above hatch h;
WH_h	weight limitation for hold locations under hatch h;
WST_i^+, WST_i^-	weight limitation for locations of station i;
GL^+, GL^-	centre of gravity limits for longitudinal equilibrium
GL_h	components of the centre of gravity for hatch h for longitudinal equilibrium;
WD_h^{max}, WH_h^{max}	upper bound for the total maximum weight of containers in locations on the deck and in the hold associated with hatch h (used for normalization);
GV^+, GV^-	centre of gravity limits for vertical equilibrium
GVD_h, GVH_h	components of the centre of gravity for vertical equilibrium related to locations on deck and in hold associated with hatch h;
PW_f	penalization weight for the components of the objective function, $f = 1,\ldots,3$;

Variables

$xh_{odh}^{stg} \in Z_+, xd_{odh}^{stg} \in Z_+ \; \forall\, s \in S,\, t \in T,\, g \in G, (o,d) \in D,\, h \in H:$ integer non-negative variables indicating the number of containers of size s, type t, weight class g loaded at port o with destination port d, assigned to the locations in the hold xh_{odh}^{stg} (on the deck xd_{odh}^{stg}) below (above) hatch h;

$yh_{pdh} \in \{0,1\}$, $yd_{pdh} \in \{0,1\}$ $\forall p,d \in P$: $p \neq d$, $h \in H$: $yh_{phd} = 1$ ($yd_{phd} = 1$) if at port p the locations in the hold (on the deck) below (above) hatch h are assigned to containers bound for d.

$r_{ph} \in Z_+$ $\forall p \in P$, $h \in H$: integer non-negative variables indicating the number of re-handles executed at port p on locations on the deck above hatch h, either for unloading containers in the hold below h bound for p or for loading containers in free locations in the hold below hatch h;

$f_{ph} \in \{0,1\}$ $\forall p \in P$, $h \in H$: $f_{ph} = 1$ if the containers on the deck over hatch h must be removed at port p to unload/load containers from/in the hold below h.

$mh_{ph} \in \{0,1\}$ $\forall p \in P$, $h \in H$: $mh_{ph} = 1$ if hatch h is removed at port p for unloading or/and loading in the hold below hatch h;

$op_{pc} \in Z_+$ $\forall p \in P$, $c \in Y_p$: integer non-negative variables indicating the total number of loading and unloading operations performed by crane c at port p.

$\Delta_{pce} \in Z_+$ $\forall p$, $\forall c,e \in Y_p$, $c \neq e$: integer non-negative variables indicating the absolute difference between the number of unloading and loading operations performed by cranes c and e at port p.

$\Delta_{max} \in Z_+$: integer non-negative variables indicating the maximum difference of handling operations between each pair of cranes over all the considered ports in P.

The proposed Multi-Port MBPP MIP formulation is then the following:

$$\min PW_1 \cdot \sum_{p \in P} \sum_{h \in H} r_{ph} + PW_2 \cdot \Delta_{max} + PW_3 \cdot \sum_{p \in P} \sum_{h \in H} mh_{ph} \tag{1}$$

subject to

$$\sum_{h \in H} (xh_{odh}^{stg} + xd_{odh}^{stg}) = N_{od}^{stg} \qquad \forall (o,d) \in D, s \in S, t \in T, g \in G \tag{2}$$

$$\sum_{(p,d) \in D_p} \sum_{t \in T} \sum_{g \in G} (xh_{odh}^{20tg} + 2xh_{odh}^{40tg}) \leq TEUH_h \cdot (1 - \sum_{(d > p) \vee (d < p-1)} yd_{p-1dh} + f_{ph}) \qquad \forall p \in P : p \neq 1, p \neq p_{\max}, h \in H \tag{3}$$

$$\sum_{(p,d) \in D} yh_{pdh} \leq 1 \quad \forall p \in P, h \in H \tag{4}$$

$$\sum_{(p,d) \in D} yd_{pdh} \leq 1 \quad \forall p \in P, h \in H \tag{5}$$

$$\sum_{(o,d)\in D_p}\sum_{g\in G}(xh_{odh}^{20tg}+2xh_{odh}^{40tg})\le TEUH_{ht}\cdot yh_{pdh}\quad \forall p\in P, d\in P: d\neq p, t\in T: t>0, h\in H \tag{6}$$

$$\sum_{(o,d)\in D_p}\sum_{t\in T}\sum_{g\in G}(xh_{odh}^{20tg}+2xh_{odh}^{40tg})\le TEUH_{h}\cdot yh_{pdh}\quad \forall p\in P, d\in P: d\neq p, h\in H \tag{7}$$

$$\sum_{(o,d)\in D_p}\sum_{g\in G}(xd_{odh}^{20tg}+2xd_{odh}^{40tg})\le TEUD_{ht}\cdot yd_{pdh}\quad \forall p\in P, d\in P: d\neq p, t\in T: t>0, h\in H \tag{8}$$

$$\sum_{(o,d)\in D_p}\sum_{t\in T}\sum_{g\in G}(xd_{odh}^{20tg}+2xd_{odh}^{40tg})\le TEUD_{h}\cdot yd_{pdh}\quad \forall p\in P, d\in P: d\neq p, h\in H \tag{9}$$

$$yh_{pdh}+yd_{pd'h}\le 1+f_{ph}\ \forall p\in P, d\in P: d\ \neq \mathrm{p},\ d'\in A(d), h\in H \tag{10}$$

$$r_{ph}\ge \sum_{(o,d)\in D_p^-}\sum_{s\in S}\sum_{t\in T}\sum_{g\in G}xd_{odh}^{stg}-TEUD_h\cdot(1-f_{ph})\quad \forall p\in P, h\in H \tag{11}$$

$$\sum_{(o,d)\in D_p}\sum_{s\in S}\sum_{t\in T}\sum_{g\in G}W_g xd_{odh}^{stg}\le WD_h\quad \forall p\in P, h\in H \tag{12}$$

$$\sum_{(o,d)\in D_p}\sum_{s\in S}\sum_{t\in T}\sum_{g\in G}W_g xh_{odh}^{stg}\le WH_h\quad \forall p\in P, h\in H \tag{13}$$

$$GL^-\le \sum_{(o,d)\in D_p}\sum_{h\in H}GL_h\sum_{s\in S}\sum_{t\in T}\sum_{g\in G}\left(\frac{W_g}{WH_h^{\max}}xh_{odh}^{stg}+\frac{W_g}{WD_h^{\max}}xd_{odh}^{stg}\right)\le GL^+\quad \forall p\in P \tag{14}$$

$$GV^-\le \sum_{(o,d)\in D_p}\sum_{h\in H}\sum_{s\in S}\sum_{t\in T}\sum_{g\in G}\left(GVD_h\frac{W_g}{WD_h^{\max}}xd_{odh}^{stg}-GVH_h\frac{W_g}{WH_h^{\max}}xh_{odh}^{stg}\right)\le GV^+\quad \forall p\in P \tag{15}$$

$$WST_i^-\le \sum_{(o,d)\in D_p}\sum_{h\in H_i}\sum_{s\in S}\sum_{t\in T}\sum_{g\in G}W_g(xh_{odh}^{stg}+xd_{odh}^{stg})\le WST_i^+\quad \forall p\in P, i\in ST \tag{16}$$

$$op_{pc}=\sum_{h\in H_{pc}}\left(\sum_{(p,d)\in D}\sum_{s\in S}\sum_{t\in T}\sum_{g\in G}(xh_{odh}^{stg}+xd_{odh}^{stg})+\sum_{(o,p)\in D}\sum_{s\in S}\sum_{t\in T}\sum_{g\in G}(xh_{odh}^{stg}+xd_{odh}^{stg})+2r_{ph}\right)\quad \forall p\in P,\ c\in Y_p \tag{17}$$

$$|op_{pc} - op_{pe}| \leq \Delta_{pce} \quad \forall p \in P, \forall c, e \in Y_p, c \neq e \tag{18}$$

$$\Delta_{max} \geq \Delta_{pce} \quad \forall p \in P, \forall c, e \in Y_p, c \neq e \tag{19}$$

$$\sum_{(p,d)\in D} \sum_{t\in T} \sum_{g\in G} (xh_{odh}^{20tg} + xh_{odh}^{40tg}) + \sum_{(o,p)\in D} \sum_{t\in T} \sum_{g\in G} (xh_{odh}^{20tg} + xh_{odh}^{40tg}) \leq 2TEUH_h \cdot mh_{ph}$$
$$\forall p \in P, h \in H \tag{20}$$

$$\begin{aligned}
&xh_{odh}^{stg} \in Z_+, xd_{odh}^{stg} \in Z_+, \forall s \in S, t \in T, g \in G, (o,d) \in D, h \in H \\
&yh_{pdh} \in \{0,1\}, yd_{pdh} \in \{0,1\}, \forall p \in P, d \in P : d \neq p, h \in H \\
&mh_{ph} \in \{0,1\}, \forall p \in P, h \in H \\
&r_{ph} \in Z_+, \forall p \in P, h \in H \\
&f_{ph} \in \{0,1\}, \forall p \in P, h \in H \\
&op_{pc} \in \boldsymbol{R} \, \forall p \in P, c \in Y_p \\
&\Delta_{pce} \in \boldsymbol{R}, \forall p \in P, c, e \in Y_p : c \neq e \\
&\Delta_{max} \in \boldsymbol{R}
\end{aligned} \tag{21}$$

In the above model the objective function (1) includes as last term the minimization of the number of hatches movements, while the other terms are the same of the model described in [3]; also constraints (2) – (9), (11) and (17) – (19) are the same constraints used in [3]. Constraints (2) satisfy each $o - d$ demand of transport in the route of the ship. Constraints (3) fix variable f_{ph} equal to one when it is necessary to re-handle containers stowed in the locations on the deck above hatch h for loading containers in hold locations of the same hatch: if some containers are stowed at port p below hatch h, the first term of (3) is positive; if in the deck locations of hatch h are stowed containers that have not to be unloaded in p, in the second term of (3) the sum of variables y_d is equal to one, thus f_{ph} is fixed to one. Constraints (4) and (5) define the variables related to the assignment of a destination d to the locations, respectively, in the hold below hatch h and on the deck on hatch h. Constraints (6) and (8) are related to the capacity respectively of hold and of deck locations, for open top containers and reefer ones, while constraints (7) and (9) are related to all the loaded containers.

Constraints (10) define variables f_{ph} in accordance with the loading decisions; more precisely, at each port visited by the ship, is checked if the stowage will cause re-handles at the discharge port; and constraints (11) define the number of re-handles to perform at port p over hatch h.

Equations (12) – (16) represent the new stability constraints: constraints (12) and (13) limit the weight of containers stowed on the hold and in the deck locations of each hatch to the maximum weight allowed for respecting the vessel capacity and draft limits; constraints (14) and (15) are, respectively, the longitudinal and vertical stability constraints. The unbalance in the longitudinal and vertical equilibrium must be within some fixed limits given by the constants of these constraints that reflect trim, GM and draft limits. Constraints (16) are used for checking the share forces: they limit the

weight of containers stowed in the locations of the hatches of each station to be within predefined lower and upper values.

Constraints (17) compute the total number of operations executed by each crane at a given port p and constraints (18) define the absolute difference between the number of unloading and loading operations at port p executed by a pair of different cranes; finally, constraints (19) give the maximum difference of loading and unloading operations between any pair of cranes over all the considered ports.

In (20) the binary variables mh_{ph} used for computing the number of hatches movements are defined. Variable mh_{ph} is equal to one if at port p some containers are either unloaded or loaded on board in the hold of hatch h.

Finally, (21) define the decision variables. Note that those integer variables that are defined as difference or sum of integer ones can be relaxed and defined as real variables.

4 The MIP Heuristics for MP-MBPP

In this section a MIP heuristics based on the MIP model (1) – (21) is briefly described. The proposed MIP heuristics exploits a *relax-and-fix* principle [11] as iteratively solves relaxations of the MIP model progressively fixing subset of binary variables. This approach is derived from the two-level Progressive Random Fixing Procedure (PRFP) proposed in [2]. We will refer to the new heuristics as PRFPHM.

Let us summarize the notation useful for describing the proposed procedure:

M_{PR} linear relaxation of the MIP model (1) – (21) with respect to variables xh^{stg}_{odh} and xd^{stg}_{odh}, mh_{ph} and r_{ph};

S_{PR} solution set of the relaxed model M_{PR};

I_{PR} subset of variables xh^{stg}_{odh} and xd^{stg}_{odh} whose value in solution S_{PR} is integer;

LB_{PR} lower bound obtained by solving M_{PR};

PercFix percentage of the subset I_{PR} of integer variables to fix when solving M_{PF};

M_{PF} partially fixed model is derived from model (1) – (21) by:

- fixing yh_{pdh} and yd_{pdh} to the values in solution S_{PR}
- randomly fixing *PercFix* percent of the variables in I_{PR}

Obj_{PF} objective value obtained by solving M_{PF};

$BestObj_{PF}$ objective value associated with the best so far solution obtained by solving M_{PF};

Iter counter of MIP solver calls.

The main steps of the heuristics are reported in the flow chart in Fig. 2. When model M_{PR} and M_{PF} are solved a time limit on the maximum CPU time is imposed. Note that the resulting M_{PR} model is an easier MIP model.

The MIP heuristics first derives the partially relaxed model M_{PR} from model (1) – (21) and then solves it by means of a MIP solver, allowing a maximum computation time (which is an algorithm parameter). If no feasible solution is found the algorithm terminates; otherwise, the set I_{PR} is obtained from the solution to M_{PR}, and a series of

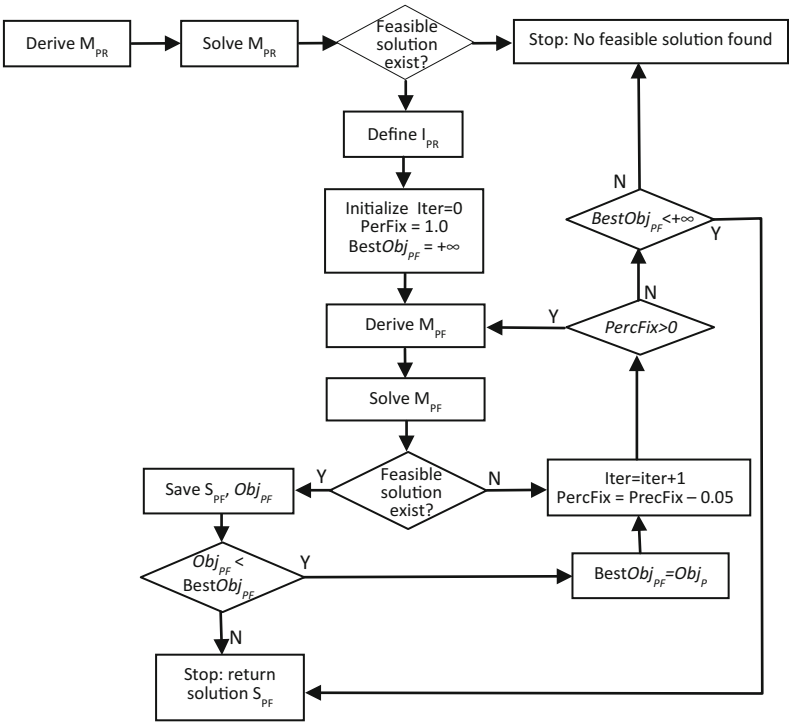


Fig. 2. Flow chart of the MIP heuristics PRFPHM

partially fixed models M_{PF} is solved. In particular, the first M_{PF} is obtained by fixing all the xh_{odh}^{stg} and xd_{odh}^{stg} variables in I_{PR} to the values that they assume in the solution to M_{PR}, being *PercFix* initially set to 1.0. If the MIP solver is able to find a feasible solution to the M_{PF}, then the obtained solution and objective function value are saved and the algorithm checks if the new objective improves the current one; otherwise it terminates returning the current solution. Note that such check always succeeds at the first iteration, since Obj_{PF} is initialized to infinity. Whenever an improved solution is found, the new objective becomes the current best so far and a new iteration is possibly executed by decreasing *PercFix* of 0.05. In case *PercFix* becomes lower than zero, then the algorithm terminates returning two possible output: either a feasible solution, if the initial infinity value of the current best so far objective has been updated, or a signal that no feasible solution has been found. Note that even if in the first iteration for *PercFix* equal to 1 the MIP solver cannot find a feasible solution in the available computation time, a new iteration is performed after having reduced *PercFix*. Thus, this loop continues until a M_{PF} model is solved at an iteration and at the next iteration such solution is not improved or until *PercFix* becomes lower than zero.

5 Computational Results

Model (1) – (21) and PRPFHM heuristics described in the previous section have been implemented in C#, using as MIP solver Cplex 12.6. All the tests ran on a 2.4 GHz Intel Core 2 Duo E6600 computer with 8 GB RAM.

The computational tests are based on randomly generated instances of the MP-MBPP that represent real size scenarios and derived by the experimental campaign used in [3]. The main characteristics of the instances used in the experimental tests here presented are reported in Table 1.

Table 1. Instances characteristics

Instances	Ship characteristics				Cargo composition (%)		
	Tot Cap	R Cap	OT Cap	|H|	ST	R	OT
SS_1	7800	800	1560	117	80	10	10
SS_2					90	5	5
SS_3					90	10	0
SS_4					100	0	0
MS_1	10000	1000	2000	150	80	10	10
MS_2					90	5	5
MS_3					90	10	0
MS_4					100	0	0
LS_1	18032	1680	2576	322	80	10	10
LS_2					90	5	5
LS_3					90	10	0
LS_4					100	0	0

Three ships of different size have been used in these experiments. The total capacity (Tot Cap) is equal to 7800 Teus for the small ships (SS), 10000 Teus for the medium ships (MS) and 18032 Teus in case of large ships (LS). Then, for each containership is specified the capacity for stowing reefer (R Cap) and open top containers (OT Cap). These capacities range from 800 to 1680 for reefer containers and from 1560 to 2576 for open top ones. Finally, in Table 1 for each containership is reported the number of hatches (|H|).

Each ship travels on a circular route with 6 ports and the transport demand is uniformly generated in such a way that for each origin o in the route of a ship, there can be a positive demand for standard (ST), reefer (R) and open top (OT) containers for the three successive ports of the route, in accordance with the four cargo compositions described in the last three columns of Table 1.

The weight classes considered for the containers are the light, medium and heavy class. The weight limit of each class is different for 20' and 40' containers and consists of 20' containers up to 7, 14 and 21 tons and 40' containers up to 10, 20 and 30 tons, respectively.

All instances have been generated in such a way to stress the capability of the solution approach to obtain feasible and effectiveness solutions; thus the occupancy level of the ship (i.e. the ratio between the TEUs on board and the TEUs capacity of the ship) ranges from 80% to 86%, while the required capacity of the ship at each port (i.e. ratio between the TEUs to load and the TEUs capacity of the ship) ranges from 35% to 46.5%.

Model (1) – (21) applied to the considered instances has a number of variables ranging from 95185, in the case of the smallest ship SS, to 261235 for the largest one LS, and a number of constraints ranging from 29050 to 79070, accordingly.

The generated instances are very difficult to be solved up to optimality by model (1) – (21) and, from the executed experimental tests, it has been noted that this difficulty increases with the presence of the term in the objective function related to the minimization of the movements of hatches, as shown in Table 2.

In Table 2 results obtained by model (1) – (21) are compared with those obtained by the model without the hatches movements minimization, that will be denoted (1′) – (21).

When model (1) – (21) is solved, the time limit of three hours is always reached and for instance (LS_3) is not possible to obtain a feasible solution. Moreover, note that instance MS_1 and the LS ones do not present serviceable solutions due to the large number of re-handles.

If the number of movements of hatches is not minimized, thanks to model (1′) – (21) it is possible to find the optimal solution for eight instances of the twelve analyzed, with an average CPU time of 1084 s; optimal solutions are characterized by no re-handles and no more than one cranes unbalance. A feasible solution is also found for the remaining four instances but two of these solutions are not serviceable (LS_1 and LS_2).

It is evident, by results of Table 2, that the larger instances require an unacceptable solution time when model (1) – (21) (or model (1′) – (21′)) is solved by a commercial software like CPLEX; thus, the MIP heuristics briefly described in the previous section is used. The results obtained by the PRPFHM heuristics are shown in Table 3. As before, in Table 3 are compared results obtained for the problem with and without the minimization of the movements of hatches.

During the search phase of the MIP heuristics a time limit of 900 s for each iteration is imposed for the smaller instances, whereas 1800 s has been fixed. for the larger instances. This limit has been further increased to 3600 when the movements of the hatches must be minimized.

Table 3 shows that the solutions obtained for the larger ships are characterized on average by 923 movements of hatches; this means that, on average, 153 hatches among the 322 of the ship are moved at each port visited by the ship.

When the number of movements of hatches is not minimized the solutions obtained are characterized by the fact that, at each port, all the hatches of the ship are moved, for avoiding re-handles and unbalance.

Table 3 also highlights the difference in the average CPU time spent by the heuristics in the two cases. The MIP heuristics is strongly influenced by the difficulty in solving the MIP model.

Table 2. Solutions obtained by model (1) – (21) and (1′) – (21)

	Model (1) – (21)					Model (1′) – (21) No hatch mov.				
Instance	Obj	CPU Time	MovHatches	Rehandles	Unbalance	Obj	CPU Time	MovHatches	Rehandles	Unbalance
SS_1	8996	10800	288	33	2	10	10800	702	0	10
SS_2	7916	10800	261	8	1	1	629	702	0	1
SS_3	8297	10800	270	19	0	1	212	702	0	1
SS_4	7291	10800	243	0	1	1	777	702	0	1
Average	**8125**	**10800**	**265.5**	**15**	**1**	**3.25**	**3104.4**	**702**	**0**	**3.3**
MS_1	374701	10800	635	35493	484	1	631	900	0	1
MS_2	12118	10800	374	83	2	1	3231	900	0	1
MS_3	3012308	10800	373	108	4	1	708	900	0	1
MS_4	10236	10800	338	8	3	1	172	900	0	1
Average	**852340.75**	**10800**	**430**	**8923**	**123.25**	**1**	**1185.5096**	**900**	**0**	**1**
LS_1	542993	10800	1228	50452	1206	506153	10800	1932	50452	1206
LS_2	611367	10800	1250	57194	1481	469081	10800	1932	46718	1455
LS_3	–	10800				4	10800	1932	0	1
LS_4	80052	10800	1108	4536	1079	1	2312	1932	0	1
Average	**411470.67**	**10800**	**1195.33**	**37394**	**1255.33**	**243809.75**	**8677.89**	**1932.00**	**24292.50**	**665.75**

Table 3. Results obtained by the MIP heuristics

	PRPFHM Heuristic					PRPFHM heuristic for solving (1′) – (21) No hatch mov.				
Instance	Obj	CPU Time	MovHatches	Rehandles	Unbalance	Obj	CPU Time	MovHatches	Rehandles	Unbalance
SS_1	6973	396	229	11	0	0	658	702	0	0
SS_2	8572	7532	274	27	82	0	145	702	0	0
SS_3	11744	8735	300	270	2	1	165	702	0	1
SS_4	8730	2122	283	21	30	1	73	702	0	1
Average	**9005**	**4696.13**	**271.48**	**82.125**	**28.5**	**0.5**	**260.1**	**702**	**0**	**0.5**
MS_1	8868	900	292	11	0	0	874	900	0	0
MS_2	13680	3911	388	181	230	2	204	900	0	2
MS_3	11817	6226	387	20	7	2	202	900	0	2
MS_4	11236	4130	349	70	66	2	119	900	0	2
Average	**11400.138**	**3791.66**	**353.98**	**70.5**	**75.75**	**1.5**	**349.7**	**900**	**0**	**1.5**
LS_1	29563	3581	929	169	3	391	3686	1932	39	2
LS_2	34610	5357	959	574	100	1	2281	1932	0	1
LS_3	31089	7982	929	367	175	0	1396	1932	0	0
LS_4	28915	8161	875	240	265	1	330	1932	0	1
Average	**31044**	**6270.21**	**923**	**337.5**	**135.75**	**98.25**	**1923.2**	**1932**	**9.75**	**1.0**

6 Conclusions

This paper focused on the Multi Port Master Bay Plan Problem (MP-MBPP), which is a very crucial optimization problem in maritime logistics, particularly for the efficiency of maritime terminals and port operations. MP-MBPP here analyzed is the problem faced by the ship coordinator.

A MIP model and a MIP heuristics have been proposed aimed at determining stowage plans in circular routes for container ships, where containers of different type, size and destination have to be loaded for being shipped to their destination. The minimization of the berthing time of the ships is the objective function. It includes three important terms like: re-handles, hatches movements and unbalance of the cranes.

From the experimental tests it is evident that due to the minimization of the hatches movements the proposed MIP model is really difficult to be solved. Thus, a strong validation of the weight used in (1) is necessary, for defining weights and for obtaining goods solutions in short CPU times.

Acknowledgements. The authors thank Prof. Paola Gualeni and Prof. Massimo Figari of the Electrical, Electronics and Telecommunication Engineering and Naval Architecture Department (DITEN) of the University of Genova, Italy, for the support given in the definition of the stability ship constraints.

References

1. Ambrosino, D., Anghinolfi, D., Paolucci, M., Sciomachen, A.: A new three-step heuristic for the master bay plan problem. Marit. Econ. Logistics (MEL) **11**(1), 98–120 (2009). Special issue on "OR models in Maritime Transport and Freight Logistics"
2. Ambrosino, D., Paolucci, M., Sciomachen, A.: A MIP heuristic for the multi-port stowage planning. Transp. Res. Procedia **10**, 725–734 (2015)
3. Ambrosino, D., Paolucci, M., Sciomachen, A.: Computational evaluation of a MIP model for multi-port stowage planning problems. Soft. Comput. **21**, 1753–1763 (2017)
4. Delgado-Ortegon, A.: Models and Algorithms for Container Vessel Stowage Optimization. Ph. D. thesis, IT-Universitetet i København, ITU-DS, no. 99 (2013)
5. Imai, A., Nishimura, E., Papadimitriou, S.: Marine container terminal configurations for efficient handling of mega-containerships. Transp. Res. Part E **49**(1), 141–158 (2013)
6. Meng, Q., Wang, S., Andersson, H., Thun, K.: Containership routing and scheduling in liner shipping: overview and future research directions. Transp. Sci. **48**(2), 265–280 (2014)
7. Monaco, M.F., Sammarra, M., Sorrentino, G.: The terminal oriented ship stowage planning problem. Eur. J. Oper. Res. **239**, 256–265 (2014)
8. Mulder, J., Dekker, R.: Optimization in container liner shipping. Econometric Institute Research Papers No. EI2016-05, Erasmus University Rotterdam, Erasmus School of Economics (ESE), Econometric Institute (2016)
9. Pacino, D., Delgado, A., Jensen, R.M., Bebbington, T.: Fast generation of near-optimal plans for eco-efficient stowage of large container vessels. In: Böse, J.W., Hu, H., Jahn, C., Shi, X., Stahlbock, R., Voß, S. (eds.) ICCL 2011. LNCS, vol. 6971, pp. 286–301. Springer, Heidelberg (2011). doi:10.1007/978-3-642-24264-9_22

10. Pacino, D., Delgado, A., Jensen, R.M., Bebbington, T.: An accurate model for seaworthy container vessel stowage planning with ballast tanks. In: Hu, H., Shi, X., Stahlbock, R., Voß, S. (eds.) ICCL 2012. LNCS, vol. 7555, pp. 17–32. Springer, Heidelberg (2012). doi:10.1007/978-3-642-33587-7_2
11. Pochet, Y., Wolsey, L.A.: Production Planning by Mixed Integer Programming. Springer, Berlin (2006)
12. Rashidi, H., Tsang, E.P.K.: Novel constraints satisfaction models for optimization problems in container terminals. Appl. Math. Model. **37**, 3601–3634 (2013)
13. Rodrigue, J.P., Claude, C., Slack, B.: The geography of transport systems, 3rd edn. Routledge, London (2013)
14. Stahlbock, R., Voss, S.: Operations research at container terminal: a literature update. OR Spectr. **30**, 1–52 (2008)
15. Steenken, D., Voss, S., Stahlbock, R.: Container terminal operation and operations research - a classification and literature review. OR Spectr. **26**, 3–49 (2004)
16. Talley, W.K., Ng, M.: Maritime transport chain choice by carriers, ports and shippers. Int. J. Prod. Econ. **142**, 311–316 (2013)
17. Tierney, K., Pacino, D., Jensen, R.M.: On the complexity of container stowage planning problems. Discrete Appl. Math. **169**, 225–230 (2014)
18. UNCTAD. Review of maritime transport, 2014. United Nations Conference on trade and Development (2014)
19. Volkan, Ç.: Sustainable container terminal operations: challenges and enhancements. Karadeniz Araştırmaları, J. Black Sea Stud. **49**, 141–156 (2016)
20. Wilson, I.D., Roach, P.A.: Container stowage planning: a methodology for generating computerised solutions. J. Oper. Res. Soc. **51**(11), 1248–1255 (2000)
21. Wilson, D., Roach, P.A., Ware, J.A.: Container stowage pre-planning: using search to generate solutions, a case study. Knowl. Based Syst. **14**(3–4), 137–145 (2001)

A Multi-depot Periodic Vehicle Routing Model for Petrol Station Replenishment

Pasquale Carotenuto[1(✉)], Stefano Giordani[2], Simone Massari[2], and Fabrizio Vagaggini[2]

[1] Istituto per le Applicazioni del Calcolo "M. Picone", Consiglio Nazionale delle Ricerche, via dei Taurini 19, 00185 Rome, Italy
carotenuto@iac.cnr.it

[2] Dipartimento di Ingegneria dell'Impresa, Università di Roma "Tor Vergata", via del Politecnico 1, 00133 Rome, Italy
stefano.giordani@uniroma2.it

Abstract. The petrol station replenishment problem consists in delivering fuel oils from a set of storage depots to a set of petrol stations during a few days planning horizon. This problem is addressed by an oil company which, for example, has to decide simultaneously the weekly fuel oil replenishment plan for each station, and, for each day of the week, the tank truck (vehicle) routes from depots to stations, in order to deliver the planned fuel oil replenishment amounts to petrol stations. Assuming a fleet of homogeneous tank trucks, the aim is to minimize the total distance travelled by tank trucks during the week, while loading tank trucks possibly near to their capacity in order to maximize the resource utilization. We model the problem as a generalization of the Multi-Depot Periodic Vehicle Routing Problem (MDPVRP) and provide a mathematical formulation. Due to the large size of the real instances which the company has to deal with, we solve the problem heuristically. We propose a hybrid genetic algorithm that successfully address the problem. The algorithm is derived from a known hybrid genetic algorithm for the MDPVRP, and adopts additional techniques and features tailored for the particular fuel oil distribution problem. It is specifically designed to deal with real instances derived from the fuel oil distribution in the European context that are profoundly different from the MDPVRP instances available from the literature. The proposed algorithm is evaluated on a set of real case studies and on a set of randomly generated instances that hold the same characteristics of the former.

Keywords: Fuel oil distribution · Freight transport · Transportation planning · Vehicle routing · Metaheuristics · Genetic algorithm

1 Introduction

In this paper we consider the problem of planning the distribution of petrol products from a set of storage depots to a set of petrol stations, addressed by the logistic department of a major European oil company. This is a very complex problem where many

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_21

variables come into play, such as demand uncertainty and seasonality, agreements for exchange of products with other oil companies, contracts with carriers, etc.

In order to reduce the complexity of the problem, the whole decision process is typically subdivided by the oil company into three phases. In the first (strategic) phase petrol stations are preassigned to one or a few storage depots from which it will be refuelled during the next medium-large term planning horizon. The second (tactical) phase considers a few days planning horizon (e.g., a week) and the subset of petrol stations of a regional area covered by the same few storage depots. In this phase the weekly fuel oil replenishment plan for each station is defined, by determining the days when each petrol station will be replenished, along with the delivery amount of petrol products. Simultaneously, the tank truck (vehicle) routes from depots to stations are determined for each day of the week, in order to deliver the planned fuel oil replenishment amounts to petrol stations. At this level, for simplicity, oil company defines replenishment weekly plans for each petrol station by assuming a single undifferentiated product and a fleet of homogeneous tank trucks. Finally, in the third (operational) phase, the next day final tank truck routes are established by considering also specific operational constraints, including those related to the vehicles characteristics (whose capacity is typically subdivided into compartments), to the distinct products to be delivered (e.g., gasoline and diesel fuel), and fulfilling the actual replenishment demands of petrol stations for the specific day.

In this work, we study the problem addressed in the second phase, where the real scenario faced by the oil company consists of a set of petrol stations that are not all owned by the company. This implies that deciding how much and when to replenish the petrol stations requires to be done in accordance with the petrol stations' owners. Moreover, typically the budget available to a petrol station owner is very limited (specially for a very small petrol station), implying that the petrol order sent to the oil company by a petrol station often covers only its petrol demand for a couple of days. These operational conditions force the company to fulfill the estimated weekly petrol demand of a petrol station (typically for distinct product typologies) with a number of replenishments during the week (i.e., the weekly visit frequency), with the chance to select one out of a set of replenishment plan (or pattern) established in accordance with the petrol station owner, where a replenishment pattern specifies the set of visit days (visit pattern) along with the (possibly distinct) fuel oil amounts (demand pattern) to be delivered in these days, respectively.

We model this tactical petrol stations replenishment problem faced by the oil company as a generalization of the Multi-Depot Periodic Vehicle Routing Problem (MDPVRP) and provide a mathematical formulation. Due to the large size of the real instances which the company has to deal with, we solve the problem heuristically. We propose a hybrid genetic algorithm that successfully addresses the problem. The algorithm is derived from the MDPVRP hybrid genetic algorithm by Vidal *et al.* [27], and adopts additional techniques and features tailored for the particular fuel oil distribution problem. It is specifically designed to deal with real instances derived from the fuel oil distribution in the European context that are profoundly different from the MDPVRP instances available from the literature. Our paper extends the model and algorithm proposed by Carotenuto *et al.* [5] to the case with multiple storage depots.

The proposed algorithm is evaluated on a set of real instances faced by a European oil distribution company, and on a set of randomly generated instances that hold the same characteristics of the former. Such a kind of optimization tool can be used not only in the optimization process of a given oil distribution network, but also for performing scenario analyses and economic assessments simulating variations of existing networks (acquisition or disposal of petrol stations).

The paper is organized as follows. In Sect. 2 we recall the relevant literature, in Sect. 3 we formally define the problem addressed, provide a mathematical formulation of the same, and model it as a generalization of the MDPVRP. The proposed genetic algorithm is detailed in Sect. 4. Section 5 is devoted to the experimental analysis, and finally Sect. 6 gives some conclusions.

2 Literature Review

The problem we study belongs to the class of multi-period petrol station replenishment problems (MPPSRP) (see, e.g., [9]), where the aim is to optimize the delivery of several petrol products to a set of petrol stations over a given planning horizon. They can be viewed as Inventory Routing Problems (IRP) with specific additional constraints such as the use of heterogeneous vehicle with compartments, also known as IRP in fuel delivery (see, e.g., [28]). Malépart *et al.* [15] propose four heuristics for constructing replenishment plans over a horizon of several working days, where some petrol stations manage their own inventories sending their orders to vendor company whenever they want, and for other stations the inventory is managed by the vendor company who decides the replenishment plans. Ng *et al.* [16] study two small petrol distribution networks in Hong Kong, proposing a model for simultaneously assigning trips to tank trucks and stations, assuming stations inventories being managed by the vendor. Cornillier *et al.* [9] propose a heuristic approach to solve the case where the number of stations on any given route is limited to two. Popović *et al.* [20] propose a variable neighborhood search (VNS) heuristic for solving a multi-product multi-period IRP in fuel delivery with multi-compartment homogeneous vehicles, given a distinct deterministic petrol consumption for each fuel type and for each petrol station; they limit the number of stations per route to three. Vidović *et al.* [28] extend this limit to four and propose a mixed integer formulation that can be solved at optimum by commercial solvers only for very small instances (10 petrol stations and 3 days); they also propose some heuristics for solving larger instances up to 50 petrol stations and a period of 5 days.

A special case of the MPPSRP is the periodic version (PPSRP) where a periodic service requirements is specified for each petrol station or frequency of service, i.e., the number of visit days during a week. For the latter problem, Triki [26] proposes a three-phase heuristic method for its solution where first a replenishment plan is selected for each petrol station, then the routing for each day is determined, and, finally, a local improvement technique is applied to further reduce the routing length. Note that our problem is similar to the PPSRP but it generalizes the latter by considering replenishment plans that differ also upon demand patterns.

The single-period case has been also studied (in this case the problem is no longer an IRP but a Vehicle Routing Problem (VRP) with special vehicles). Avella *et al.* [3] and Cornillier *et al.* [8] study this case proposing exact and heuristic algorithms. Cornillier *et al.* [11] propose two heuristics for the more general case with time windows constraints for the petrol delivering to petrol stations; the same authors extend this study by considering also multiple depots [10]. More recently, Benantar *et al.* [4] propose a tabu search algorithm where specific attention is paid to tankers with compartments and customers with different types of demanded products and time windows.

The MDPVRP is a generalization of the multi-depot capacitated VRP (see, e.g.: [14, 24]) that takes into account several planning days with customers that require service on multiple days during the planning period.

After the very early works that propose constructive heuristics for the single-depot case (i.e., the PVRP), Chao *et al.* [6] are the first to provide a two-phase heuristic that allows escaping from poor local optima, enhancing the exploration of the solution space, through the use of deteriorating moves and the relaxation of vehicle capacity. Cordeau *et al.* [7] propose a tabu search heuristic to solve the PVRP and two variants (i.e., the MDPVRP and the periodic travelling salesman problem). Another tabu search heuristic is proposed by Angelelli and Speranza [2] to solve a generalization of PVRP where vehicles can renew their capacity at some intermediate facilities. Drummond *et al.* [12] propose an island-based parallel evolutionary algorithm which evolves individuals representing visit combinations and generates routes with a saving heuristic. Alegre *et al.* [1] propose a scatter search procedure for periodic pick up of raw materials of auto parts, using a two-phase approach that assigns visit combinations and design routes for each period with a neighbourhood-based procedure. Pirkwieser and Raidl [19] introduce an extension of VNS metaheuristics with multilevel refinement strategy particularly suitable for larger PVRP instances. A two-phase GRASP and path relinking metaheuristic is proposed by Pacheco *et al.* [18]. A significant contribution is given by Vidal *et al.* [27] that propose a hybrid genetic algorithm to tackle the MDPVRP and PVRP with heterogeneous capacitated vehicles and constrained route duration. It combines the breadth exploration of evolutionary algorithms, the improvement capabilities of VNS, and an advanced population-diversity management mechanism which allows a broader access to reproduction, preserving at the same time the memory of the characteristics of good solutions represented by the elite individuals. More recently, Nguyen *et al.* [17] develop a genetic algorithm for the variant of the PVRP with time windows, while Rahimi-Vahed *et al.* [22] use a modular heuristic algorithm to solve a particular application of the MDPVRP and the PVRP with the objective of determining the optimal fleet sizing, considering budget constraints in addition to standard constraints on vehicle capacity and route duration.

3 Problem Definition

Let C be the set of n petrol stations (customers). Let D be the set of n_0 available storage depots and let $D_i \subseteq D$ be the subset of (admissible) depots from which station $i \in C$ can

be refurnished. Let $G = (N, A)$ be a direct graph with $N = D \cup C$ being the set of $n_0 + n$ nodes. Let arc $(i, j) \in A$ represents the direct-travel from petrol station (depot) i to petrol station (depot) j with travel cost (length) c_{ij}, and travel time t_{ij}; let τ_i be the service time at station i, which for simplicity is assumed independent from the amount of fuel oil dropped during station replenishment. Let T be the set of p time periods (days) of the planning horizon (week). Each petrol station $i \in C$ presents a visit frequency (visit days) $f_i \leq p$, with at most one visit per day, and a total (weekly) product demand q_i^{tot}. More specifically, let Π_i be the set of available replenishment plans (patterns) of station i, where a replenishment pattern $\pi \in \Pi_i$ is a visit-day combination (i.e., a subset of f_i days of the week) along with the related specific demand combination of total amount q_i^{tot}; let d_{it}^{π} be the demand of station i in time period (day) $t \in T$, if pattern $\pi \in \Pi_i$ is chosen for station i; we assume $\sum_{t \in T} d_{it}^{\pi} = q_i^{tot}$, for each $\pi \in \Pi_i$. Note that with this assumption we generalize the PPSRP that assumes the total demand q_i^{tot} of station i during the planning horizon is equally subdivided among the f_i visit-periods. Fuel oil is delivered with a fleet of homogeneous tank truck, each one with capacity $Q_{\max}$; therefore, the total demand of petrol stations served by a tank truck cannot exceed $Q_{\max}$. Moreover, the duration of each route, that is, the sum of the total route travel time and the total service time of the visited customers, cannot exceed time limit $T_{\max}$.

A feasible solution of the problem consists in selecting one of the available replenishment plans (patterns) and one of the admissible depot, for each station, and routing a fleet of tank trucks, each one from a depot to a subset of petrol stations assigned to that depot while respecting vehicle capacity and route duration constraints, in order to fulfil the selected station replenishment plans. The aim is determine the solution minimizing the total distance (length) travelled by tank trucks during the planning horizon.

Let $y_{io\pi}$ be a binary variable that takes value 1 if petrol station $i \in C$ is refurnished by storage depot $o \in D_i$ with replenishment plan (pattern) $\pi \in \Pi_i$, and 0 otherwise; let x_{ijot} be a binary variable that takes value 1 if arc $(i, j) \in A$ is traversed in time period (day) $t \in T$ by a tank truck departed from storage depot $o \in D$, and 0 otherwise. Let $a_{\pi t}$ be a binary parameter equal to 1 if pattern π requires a visit in time period (day) t and 0 otherwise. The problem can be formulated as follows:

$$\min \ell = \sum_{t \in T} \sum_{o \in D} \sum_{(i,j) \in A} c_{ij} x_{ijot} \tag{1}$$

s.t.

$$\sum_{o \in D_i} \sum_{\pi \in \Pi_i} y_{io\pi} = 1 \quad \forall\, i \in C, \tag{2}$$

$$\sum_{h \in N:\,(h,i) \in A} x_{hiot} + \sum_{j \in N:\,(i,j) \in A} x_{ijot} = 0 \quad \forall\, i \in C, o \in D \backslash D_i, t \in T, \tag{3}$$

$$\sum_{h\in N:(h,i)\in A} x_{hiot} = \sum_{\pi\in\Pi_i} a_{\pi t} y_{io\pi} \quad \forall\, i \in C, o \in D_i, t \in T, \tag{4}$$

$$\sum_{j\in N:(i,j)\in A} x_{ijot} = \sum_{\pi\in\Pi_i} a_{\pi t} y_{io\pi} \quad \forall\, i \in C, o \in D_i, t \in T, \tag{5}$$

$$\sum_{i\in C:\,(i,o)\in A} x_{ioot} = \sum_{j\in C:\,(o,j)\in A} x_{ojot} \quad \forall\, o \in D, t \in T, \tag{6}$$

$$\mathbf{x} \in \mathbf{X}(\mathbf{y}), \tag{7}$$

$$y_{io\pi} \in \{0,1\} \quad \forall\, i \in C, o \in D_i, \pi \in \Pi_i, \tag{8}$$

$$x_{ijot} \in \{0,1\} \quad \forall\, (i,j) \in A, o \in D, t \in T. \tag{9}$$

The objective function (1) is the total length traveled by tank tankers. Constraints (2) assure that exactly one depot and one replenishment plan (pattern) is assigned to each petrol station. Constraints (3) ensure that a tank truck cannot enter and/or leave a petrol station if departed from a non-admissible depot for that station. Constraints (4) and (5) impose that exactly one tanker departed from depot $o \in D_i$ enters and leaves (respectively) each petrol station node $i \in C$ during time period (day) $t \in T$ if station i is refurnished by storage depot $o \in D_i$ and the replenishment plan assigned to that station establishes that the latter must be visited in time period t, and 0 otherwise. Constraints (6) impose that for each time period the number of tank trucks leaving a depot is equal to the number of tank trucks going back to the depot (at the end of its tour). Constraint (7) implicitly represents subtour elimination, tank truck capacity, and route-duration constraints. For example, by introducing a continuous variable u_{iot}, for each $i \in C$, $o \in D$, and $t \in T$, a family of constraints equivalent to (7) without route-duration constraint is:

$$u_{iot} - u_{jot} + Q_{\max}\, x_{ijot} \le Q_{\max} - \sum_{v\in D_j}\sum_{\pi\in\Pi_j} d^{\pi}_{jt} a_{\pi t} y_{jv\pi} \quad \forall i,j \in C{:}(i,j) \in A, o \in D_i \cap D_j, t \in T, \tag{7a}$$

$$\sum_{v\in D_i}\sum_{\pi\in\Pi_i} d^{\pi}_{it} a_{\pi t} y_{iv\pi} \le u_{iot} \le Q_{\max} \qquad \forall\, i \in C, o \in D_i, t \in T. \tag{7b}$$

These constraints generalizes the (MTZ-like) subtour elimination constraints for the capacitated VRP (see, e.g., [24]), ensuring that each tank truck route is connected to a depot and the total demand of the stations visited by a truck cannot exceed its capacity $Q_{\max}$. Similar constraints can be added to include also route duration constraints.

Our petrol station replenishment problem can be modelled as a generalization of the MDPVRP. The MDPVRP extends the multi-depot capacitated VRP adding a time dimension and considering route planning to be performed over a planning horizon of p periods. Each customer i is characterized by a visit frequency f_i, representing the number of required visits during the planning horizon, and a list of possible visit-period

combinations (visit patterns). For each customer i, the demand in each visit-period is assumed to be fixed and equal to q_i with a total demand over the planning horizon equal to $q_i^{tot} = f_i\, q_i$. The MDPVRP aims at selecting a visit-period combination (visit pattern) of f_i periods for each customer i, and finding the feasible routes starting and ending at the same depot (chosen from a set of depots) for each period of the planning horizon in order to fulfil the selected visit patterns, and minimizing the total route travel cost (length) ℓ. The problem is *NP*-hard since it extends the capacitated VRP which is known to be *NP*-hard (see, e.g., [24]).

Our generalization of the MDPVRP considers visit patterns in which the total demand q_i^{tot} of a customer i can be also non-equally subdivided among the f_i visit-periods: in this respect, we generalize the concept of visit pattern accordingly to the definition of replenishment pattern given at the beginning of this section. Moreover, according to our petrol station replenishment problem, we assume that the route visiting customer i must start from depot $o \in D_i \subseteq D$. We also extend the objective function in order to evaluate solutions not only on the basis of the total travelled distance but also with respect to the unused capacity of the vehicles. In particular, following a typical policy of an oil company, it is desirable that tank truck are sufficiently loaded, e.g., with at least $Q_{\min} \leq Q_{\max}$ amount of fuel oil. We model this requirement adding a penalty to the objective function proportional to the waste capacity index $c = \sum_{r \in R} \max\left(0, Q_{\min} - \sum_{i \in C(r)} d_{i,t(r)}^{\pi_i}\right)^2 \Big/ \left(Q_{\max} |R|\right)$, where $C(r)$ is the set of customers visited by route $r \in R$, with R being the set of routes performed by the vehicles during the planning horizon, and $d_{i,t(r)}^{\pi_i}$ being the demand of customer i visited by route r in period (day) $t(r)$, whose amount is established by the selected pattern π_i for that customer. Therefore, the new objective function to be minimized is $z = w_1 \ell + w_2 c$, where w_1 and w_2 are weight parameters.

4 The Proposed Hybrid Genetic Algorithm

The solution approach, called *Multi-Depot Periodic Vehicle Routing Problem Genetic Search* (MDPVRPGS), we propose for solving our generalization of the MDPVRP, is similar to the *Hybrid Genetic Search with Adaptive Diversity Control* (HGSADC) algorithm by Vidal *et al.* [27], which achieves excellent results on MDPVRP instances known in the literature. However, algorithm MDPVRPGS adopts some techniques and features tailored for the studied petrol station replenishment problem. The general scheme of the hybrid genetic algorithm we propose is reported next.

```
Algorithm MDPVRPGS
Initialize subpopulations
while number of iterations without improvement < It_ni and time <
MaxTime do
  Select parent solutions P1 and P2 through binary tournament
  Generate offspring CH from P1 and P2 through crossover
  Improve CH through education procedure with probability p_ed
  if CH is infeasible then
    Insert CH into infeasible subpopulation
    Repair CH and insert CH into feasible subpopulation with
probability p_rep
  if CH is feasible then
    Insert CH into feasible subpopulation
  if maximum subpopulation size is reached then
    Select survivors
  Adjust penalty parameters for violating feasibility conditions
  if best solution not improved for It_div iterations then
    Diversify populations
end while
return best feasible solution
```

Differently from HGSADC, our algorithm is designed to deal with real instances of the considered petrol distribution problem that are profoundly different from the MDPVRP instances available from the literature. First of all, in petrol distribution instances the ratio between the vehicle capacity and the average customer demand is very small (between 3 and 4). This means that feasible solutions are characterized by a large number of routes each one visiting few petrol stations: consequently, some solving techniques for the VRP are not effective for the problem under consideration. Moreover, algorithm HGSADC is designed to work with equal customer demands for all the visit-periods of the selected pattern, while in our problem replenishment plan may contains distinct (daily) demands; furthermore, the route serving a customer cannot start from any depot but only from one of the feasible depots for that customer. Algorithm MDPVRPGS is able to tackle with this more general case of the MDPVRP, that we call MDPVRP+, that adds a further degree of complexity to the routing model and provides greater flexibility to the final decision maker.

Algorithm MDPVRPGS considers a population of feasible and infeasible solutions (individuals) which are kept in two separate subpopulations. The algorithm selects two parent individuals from the entire population and combines them through *crossover*, creating a new solution (*offspring*) which is enhanced through local search procedures (*education* and *repair*). The individual yielded is then included into the correct subpopulation and evaluated on the basis of its fitness function. The reproductive cycle iterates from one generation to the next one, each time selecting the survivors, until a stopping criteria is met.

After having presented its general outline, we describe the proposed algorithm in detail, starting from the representation and evaluation of solutions (Sect. 4.1), and proceeding with the illustration of the basic algorithm components such as parent selection and crossover (Sect. 4.2), education and repair (Sect. 4.3), and population management (Sect. 4.4).

4.1 Solution Representation and Evaluation

The literature on meta-heuristics algorithms shows that a controlled exploration of unfeasible solutions can improve the search performance, facilitating the transition between two structurally different eligible solutions [13].

Let S be the solution space of MDPVRP+ that includes both feasible and unfeasible individuals, the latter obtained by relaxing the constraints on vehicle capacity and on maximum route duration. Let $s \in S$ be a solution, $\pi_i(s) \in \Pi_i$ the related selected pattern for customer i, and $o_i(s) \in D_i$ the depot that will refurnish that customer. Moreover, let $R(s)$ be the set of planned routes (covering the planning horizon) associated to solution s. Each route $r \in R(s)$ is characterized by (vehicle) load $q(r) = \sum_{i \in C(r)} d_{i,t(r)}^{\pi_i(s)}$ and duration $t(r) = \sum_{i \in C(r) \cup \{o_r\}} t_{i,succ(i,r)} + \sum_{i \in C(r)} \tau_i$, where $C(r)$ is the set of customers visited by route r starting from depot node $o_r \in D$ and $succ(i, r)$ is the successive node visited after node $i \in C(r) \cup \{o_r\}$. Clearly, $o_r = o_i(s)$, for each station $i \in C(r)$. Let ω_q and ω_t be the penalties for violating the load capacity constraint and the route duration constraint, respectively. The penalized cost of route r is $\psi(r) = \sum_{i \in C(r) \cup \{o_r\}} c_{i,succ(i,r)} + \omega_q \max\{0, q(r) - Q_{\max}\} + \omega_t \max\{0, t(r) - T_{\max}\}$, that is, the sum of the total distance travelled and the penalty costs for violating vehicle capacity and route duration constraints, respectively.

The penalized cost of solution s is the sum of penalized cost of each route of the solution, and is therefore equal to $\varphi(s) = \sum_{r \in R(s)} \psi(r)$.

Each solution $s \in S$ is characterized by the assignment of exactly one pattern $\pi_i(s) \in \Pi_i$ and exactly one depot $o_i(s) \in D_i$, for each customer i, and by the collection $R(s)$ of generated routes. We recall that a pattern π corresponds to a particular set of time periods (days) when customer i will be visited during the planning horizon, indicating also the amount of goods (fuel oil) to be delivered for each one of the visit days. Similarly to

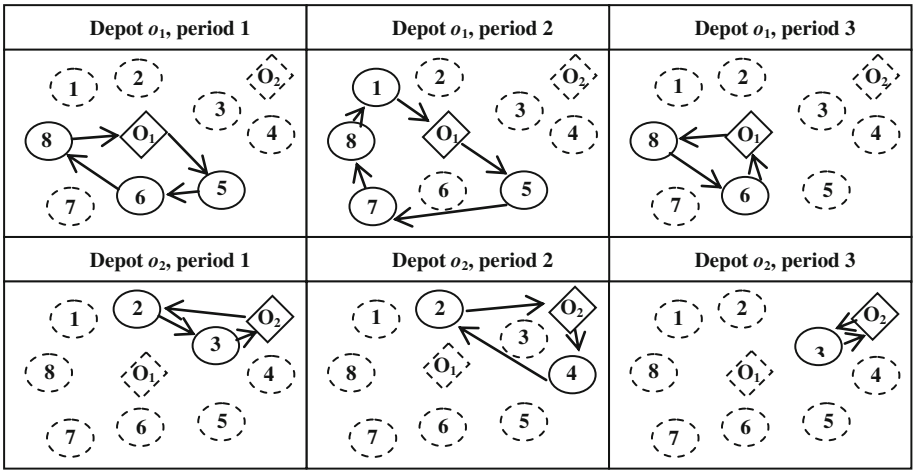


Fig. 1. A solution (individual) of MDPVRP+ and its chromosome representation.

HGSADC algorithm by Vidal *et al.* [27], individual s are represented by 3 distinct chromosomes: (i) the *pattern chromosome* that indicates the pattern $\pi_i(s)$ selected, for each customer $i \in C$; (ii) the *depot chromosome* that indicates the depot $o_i(s)$ selected, for each customer $i \in C$; (iii) the *giant tour chromosome* that contains a single tour (route) without delimiters and depot, for each depot $o \in O$ and period $t \in T$, that is, $n_0 \times p$ sequences $\sigma_{ot}(s)$ of customers to be served. The giant tour is obtained by concatenating in an arbitrary order all the routes $r \in R(s)$ of period t starting from depot o, after having removed the depot from each route. As an example, with $p = 3$ periods and $n_o = 2$ depots, Fig. 1 shows a solution s and the related pattern chromosome, depot chromosome, and giant tour chromosome representations.

Pattern chromosome

Customer *i*	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Periods	{2}	{1, 2}	{1, 3}	{2}	{1, 2}	{1, 3}	{2}	{1,2,3}
Delivery quantities	{5}	{6, 3}	{4, 3}	{6}	{3, 1}	{4, 4}	{3}	{2,1,2}

Depot chromosome

Customer *i*	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Depot	o_1	o_2	o_2	o_2	o_1	o_1	o_1	o_1

Giant tour chromosome

(Depot, period)	**(o_1, 1)**	**(o_1, 2)**	**(o_1, 3)**	**(o_2, 1)**	**(o_2, 2)**	**(o_2, 3)**
Customer sequences	{5,6,8}	{5,7,8,1}	{8,6}	{2,3}	{4,2}	{3}
Delivery quantities	{3,4,2}	{1,3,1,5}	{2,4}	{6,4}	{6,3}	{3}

Let $\theta_{ot}(s)$ be the cost of the giant tour associated to depot o and planning period t of solution s, calculated as the total route distance travelled by a truck having infinite capacity. The "giant tour" cost $\theta(s)$ of solution s is then equal to the sum of all giant tour costs for all the period of the planning horizon, i.e., $\theta(s) = \sum_{o \in D} \sum_{t \in T} \theta_{ot}(s)$.

Representing routes as giant tours allows to perform crossover operation in a simple and efficient way, working on permutations of customers, but requires the use of an efficient algorithm to find the optimal segmentation of a giant tour into routes to get the solution and its cost: our algorithm MDPVRPGS use the Split Algorithm of Prins [21], which reduces the problem of extracting routes from a giant tour to a shortest path problem on an auxiliary acyclic graph.

4.2 Parent Selection and Crossover

The offspring generation scheme starts with the selection of two parents P_1 and P_2 from which individual CH is yielded through crossover using a binary tournament that randomly (with uniform probability) extracts 2 individuals from the entire population and keeps the solution with lower penalized cost with probability 0.8, whereas it selects the worst one with probability 0.2. The binary tournament is performed twice in order to select P_1 and P_2. The chance of combining feasible and unfeasible solutions allows extending the search to the limits of eligibility where we expect to find the best solutions, while the chance of selecting individuals with lower quality is an attempt to get

diversification in the solution generation by exploiting the genetic information of the entire population.

The crossover procedure is composed by four basic steps. The first one defines the rules for the inheritance of the genetic material, through the extraction of two random integers between 0 and the number $n_0 \times p$: let n_1 and n_2 be the two extracted integers, with $0 \leq n_1 \leq n_2 \leq n_0 \times p$. The inheritance rule states that n_1 (depot, period) couples will be entirely inherited at random from P_1 and form the couple set Λ_1, $n_2 - n_1$ (depot, period) couples at random from P_2 to form the couple set Λ_2, while the remaining $n_0 \times p - n_2$ couples will be inherited partly from P_1 and partly from P_2 to form the set Λ_{mix} periods.

In the second step, the genetic material is transmitted from parent P_1 to offspring CH. For each couple $(o, t) \in \Lambda_1$, the customer visit sequence (giant tour) $\sigma_{ot}(P_1)$ is entirely copied from P_1 to CH without modification. For each $(o, t) \in \Lambda_{\text{mix}}$, two random cuts α_{ot} and β_{ot} are done on $\sigma_{ot}(P_1)$: if $\alpha_{ot} < \beta_{ot}$, the subsequence of $\sigma_{ot}(P_1)$ from position α_{ot} to position β_{ot} is transmitted to CH; if $\alpha_{ot} > \beta_{ot}$, CH will inherit from P_1 the subsequences of $\sigma_{ot}(P_1)$ from α_{ot} to the last position and from the first position to β_{ot}; finally, if $\alpha_{ot} = \beta_{ot}$, with probability 0.5 the whole customer sequence $\sigma_{ot}(P_1)$ is transmitted starting from the cut point (i.e., from α_{ot} to the last position and from the first position to α_{ot}).

The third step involves the transmission of genetic material from P_2 to offspring CH. For each $(o, t) \in \Lambda_2 \cup \Lambda_{\text{mix}}$, we consider each customer $i \in \sigma_{ot}(P_2)$ not yet inserted in the current sequence $\sigma_{ot}(CH)$ and copy it at the end of $\sigma_{ot}(CH)$ if i has at least one eligible pattern containing period t. Given P_2, let $q_{it}(P_2)$ be the demand of customer i in period t. If there are eligible patterns $\pi \in \Pi_i$ for which $d_{it}^{\pi} = q_{it}(P_2)$, then one of these pattern will be inherited into the chromosome pattern of CH, otherwise an eligible pattern $\pi' \in \Pi_i$ for which with another demand $d_{it}^{\pi'}$ is transmitted in accordance with the eligible patterns.

Finally, the forth step completes the crossover through transmission to CH of customers with service frequencies not yet satisfied. For each one of these customers (let's say customer i), this is done by calculating the best eligible pattern π_i (with related specific demand combinations) and the best position for each giant tour $\sigma_{ot}(CH)$, with o_i being the depot previously assigned to i during the construction of CH, and $t \in T^{\pi_i} \setminus T_i^{CH}$ where T^{π_i} is the set of visiting periods of pattern π_i and T_i^{CH} the visiting period for customer i yet included in CH. The best position in each giant tour is evaluated through cheapest insertion technique; this may require to modify the amount of demand $q_{it}(P_2)$ inherited from P_2 in order to guarantee the existence of at least one eligible pattern $\pi_i \in \Pi_i$.

4.3 Education and Repair

The education phase consists in a set of local search procedures used to improve the quality of a given solution. If the educated solution is unfeasible a repair procedure is

possibly used to get a feasible solution. Two types of local search are used in the education phase: Route_Improvement (RI) and Pattern_Improvement (PI).

The education phase is applied with a probability p_{ed} to an individual s in the sequence RI-PI-RI. It is worth to notice that the education stage returns an educated solution s' not worse than s, i.e., $\varphi(s') \leq \varphi(s)$.

The RI procedure works on a solution s, trying to improve its routes by moving customers within the routes of each period (i.e., within each giant tour) with a set of moves. Specifically, given a giant tour ρ of a solution and denoting with $C(\rho)$ the related set of visited customers, the following two moves M_1 and M_2 are applied over a customer $u \in C(\rho)$ and one of its neighbour $v \in N(u)$:

- M_1: remove u and place it after v;
- M_2: swap u and v.

The *neighbourhood* $N(u)$ is defined as the set of $h \cdot |C(\rho)|$ customers closest to u in terms of distance among those belonging to the same giant tour ρ of u. Parameter h is the *granularity index* [25], and is a percentage of the number of customers visited by a tour.

Starting from a given solution s, RI applies moves M_1 and M_2 in random order for each period of the planning horizon. For each randomly chosen customer u of the giant tour of a given period and for each randomly chosen neighbour $v \in N(u)$ of u, RI evaluates the obtained neighbour solution s' of s, breaking the search and restarting it from solution s' if the latter is better than s. Solution improvement is evaluated by computing the change in cost of the (one or two) modified routes after the move execution. This process is repeated until all possible moves have been successively evaluated without improvement. Note that at the end of the local search improvement we get possibly a new giant tour for each period; therefore, we apply the Split Algorithm in order to compute the actual cost of the new local optimal solution obtained.

After the second execution of RI procedure, a 2-opt local search on the giant tour of each planning period is applied; finally, an exhaustive search is applied at the single route level if the number of visited customers is very small (e.g., no more than 5): we evaluate all possible sequences of the customers of each single route (all permutations) and select the shortest. To obtain all the possible permutations we use the "plain changes algorithm" (see, e.g., [23]). The basic idea of this algorithm is that all permutations of a given number k of elements can be obtained from the sequence of permutations for $k - 1$ elements by placing the element k into each possible position in each of the permutations of $k - 1$ elements.

The PI procedure is executed with probability p_{PI}. It iterates in random order on each customer i and for each pattern $\pi \in \Pi_i$, looking for the best visit combinations along with the related customer demands. For each customer i, let us consider feasible patterns different from that (let us say π_i) assigned to i in the current solution s. Let $\pi' \in \Pi_i$ be the new pattern considered, the following moves are performed until first improvement is achieved:

- if both π_i and π'_i contain period t and $d_{it}^{\pi_i} = d_{it}^{\pi'_i}$, do nothing;
- if both π_i and π'_i contain period t but $d_{it}^{\pi_i} \neq d_{it}^{\pi'_i}$, consider $d_{it}^{\pi'_i}$ in place of $d_{it}^{\pi_i}$ in the new solution;
- if only π_i contains period t, remove i from the giant tour of depot-period (o_i, t), where o_i is the depot assigned to i;
- if only π'_i contains period t, insert i with quantity $d_{it}^{\pi'_i}$ into the giant tour of depot-period (o_i, t) through cheapest insertion.

This process is repeated until all possible moves have been successively evaluated without improvement.

The educated solution might be unfeasible: in that case the *repair* procedure is applied with probability p_{rep} to make the solution feasible. This procedure consists in temporarily multiplying the penalty parameters by 10 and reapplying the RI-PI-RI sequence. If the individual is still unfeasible, penalty parameters are multiplied by 10 again and the sequence is restarted until a feasible solution is obtained. Finally, the resulted individual is inserted into the feasible subpopulation without deleting the unfeasible one from the unfeasible subpopulation.

4.4 Population Management

Initialization process consists in generating 4μ individuals randomly assigning one pattern to each customer and creating customer visit sequences for each period in random order. Every solution is subjected to education, followed by repair in case of infeasibility (with probability $p_{rep} = 0.5$), and inserted into the proper subpopulation. At the end of initialization, a subpopulation may contain less than μ individuals.

The two (feasible and unfeasible) subpopulations are managed independently. Both subpopulations are programmed to hold a number of individuals between μ and $\mu + \lambda$, which represent the minimum and the maximum size of each subpopulation, respectively. Whenever a new individual is generated, it is added directly to the appropriate subpopulation based on its feasibility. When the maximum size $\mu + \lambda$ is reached in a subpopulation, the *survivor selection* mechanism is activated, that brings the population size back to μ discarding the worst λ individuals.

Penalty parameters are initially set to $\omega_q = \bar{c}/\bar{q}$ and $\omega_t = \bar{v}$, where $\bar{c}$ is the average (cost) distance between two customers, $\bar{q}$ is average amount of demanded product to be delivered, and $\bar{v}$ is the average vehicle speed. These parameters are dynamically adjusted in order to make the generation of naturally feasible solutions easier. Let ξ^{ref} be the reference percentage of naturally feasible solutions and ξ the same percentage evaluated over the last 100 iterations. Then, similarly to Vidal *et al.* [27], the parameters are updated as follows:

- if $\xi \leq \xi^{\text{ref}} - 0.05$, then $\omega_q = \omega_q \cdot 1.2$, and $\omega_t = \omega_t \cdot 1.2$;
- if $\xi \geq \xi^{\text{ref}} + 0.05$, then $\omega_q = \omega_q \cdot 0.85$, and $\omega_t = \omega_t \cdot 0.85$.

The *diversification* method is executed if the best solution is not improved for It_{div} iterations. It eliminates the $2\mu/3$ worst individuals of both the subpopulations. The aim of this step is to reintroduce a significant quantity of new genetic material when the population has lost most of its diversity.

When the stopping criteria is met, before returning the *"best solution"*, a *post optimization* is applied: we apply the same exhaustive search method used in the *education* operator to all routes of all individuals of the feasible population, when the number of visited customers of a route is very small. Then the best solution is updated.

5 Experimentation

Any evolutionary algorithm needs a calibration phase to set the right parameters, enabling the algorithm to perform at its best. In particular we adopt the same parameter's value (see Table 1) resulting from the calibration phase used for the genetic algorithm provided by Carotenuto *et al.* [5].

Table 1. Algorithm parameters.

Parameter	Description	Value
μ	Population dimension	25
λ	Max dimension of a generation	40
p_{ed}	Education probability	1.0
p_{rep}	Repair probability	0.5
p_{PI}	Pattern improvement probability	0.2
h	Granularity index	40%
It_{ni}	Number of iterations without improvement	2500
It_{div}	Number of iterations before diversification	40% It_{ni}
ξ^{ref}	Reference proportion for the penalty adjustment	20%

We experimented the proposed algorithm on three real test cases of an oil company with 49, 194 and 200 petrol stations, respectively, and at most 2 depots. The planning horizon is composed by 6 periods (days), and tank truck capacity $Q_{\max} = 39$ kL. For each one of the three test case we consider two scenarios: Scenario 1 assumes only one open depot, with all stations being replenished from only this depot, while Scenario 2 considers both the two depots open. The aim is to compare solutions of the two scenarios. For the sake of completeness, we also experimented with random realistic instances with $n = 100, 150, 200$ petrol stations and one depot, derived from the real test case with 200 stations. The results are reported in [5].

For brevity, next we report and analyze solutions for the real test case with 194 petrol stations and one or two depots. In particular, the first depot is fixed and represents the effective used storage depot; the second depot represents a possible additional storage depot and is assumed to be located in one out of a set of possible locations indicated by the oil company.

For each scenario, the experimentation aims also at evaluating the tradeoff between the two terms of the objective function z (i.e., the total route length ℓ and the waste capacity index c) with $Q_{\min}$ equal to 35 kL (i.e., approx. 90% of $Q_{\max}$). We run tests to evaluate the goodness of the solutions when changing the value of the weights w_1 and w_2. In particular, we consider $w_1 = (1 - w)/D_{tot}$ and $w_2 = w$, with D_{tot} being the total customer demand in order to make the magnitude of l similar to c, and test with $0 \leq w \leq 1$. We experimented our algorithm with *MaxTime* = 2, 5, 10, 15, 20 min and weight $w = 0, 0.25, 0.5, 0.75, 1$. For each (*MaxTime*, w) combination we executed 10 runs.

Table 2 lists the average results for the real case study with 194 petrol station and one depot (Real test case 2, Scenario 1). The table shows how setting values for w higher than 0.25 leads to "unbalanced" solutions where the increase in distance is not compensated by a significant waste capacity index decrease.

Table 2. Real test case 2: Scenario 1: 194 petrol stations, 1 depot, 6 periods.

MaxTime	Total route length l (Km)					Waste capacity index, c				
	$w = 0$	$w = 0.25$	$w = 0.5$	$w = 0.75$	$w = 1$	$w = 0$	$w = 0.25$	$w = 0.5$	$w = 0.75$	$w = 1$
2	25196	25194	25312	25681	25799	0.67	0.49	0.40	0.45	0.45
5	25151	25082	25277	25377	25772	0.59	0.45	0.42	0.40	0.39
10	24711	24943	25196	25325	25546	0.50	0.39	0.36	0.38	0.35
15	24811	25173	25530	25405	25634	0.54	0.45	0.43	0.37	0.37
20	24849	24926	25202	25699	25550	0.56	0.38	0.37	0.44	0.40

Table 3 lists the average results for the real case study with 194 petrol station and two depots (Real test case 2, Scenario 2). Note that in this case one of the two depots can be assigned to each petrol stations (each station can either be served by depot o_1 or depot o_2). The results, however, show a worsening of the total distance by 20% (and a reduction of the waste capacity of 6%) with respect of the solutions with one depot, assuming to give the same *MaxTime* CPU time to the algorithm for the computation. These results clearly indicate how much the problem complexity increases passing from one to two depots: in particular, we can show that the solution space increases exponentially with the number of stations to be served. Consequently, the algorithm on average needs more computational time to find better results than those with one depot.

Table 3. Real test case, Scenario 2: 194 petrol stations, 2 depots, 6 periods.

MaxTime	Total route length l (Km)					Waste capacity index, c				
	$w = 0$	$w = 0.25$	$w = 0.5$	$w = 0.75$	$w = 1$	$w = 0$	$w = 0.25$	$w = 0.5$	$w = 0.75$	$w = 1$
2	25567	25795	26134	26478	26493	0.73	0.57	0.54	0.51	0.49
5	25723	26013	26117	25665	25772	0.74	0.47	0.51	0.54	0.45
10	25351	25610	26168	25881	26173	0.65	0.52	0.53	0.48	0.40
15	25344	25760	25996	25914	25926	0.62	0.55	0.51	0.46	0.42
20	25234	25650	26027	25934	26005	0.63	0.55	0.53	0.46	0.36

In order to reduce the huge increase of problem complexity with multiple depots, the depots from which each petrol station can be served are partially preassigned on the basis of proximity criteria and oil company's expertise. We experimented with this

additional scenario (called Scenario 3) assuming again the same range of values for MaxTime and weight w, and performed 10 runs for each parameter combinations.

Table 4 lists the average results for Scenario 3 where the addition of the second depot leads to an average reduction of 18% in total distance travelled with respect to the single depot case (Scenario 1) and an average increase of waste capacity of 27%.

Table 4. Real test case 2, Scenario 3: 194 petrol stations, 2 depots, 6 periods.

MaxTime	Total route length l (Km)					Waste capacity index, c				
	$w = 0$	$w = 0.25$	$w = 0.5$	$w = 0.75$	$w = 1$	$w = 0$	$w = 0.25$	$w = 0.5$	$w = 0.75$	$w = 1$
2	20681	20729	20894	20864	21020	0.87	0.59	0.50	0.54	0.50
5	20418	20488	20752	20838	20864	0.72	0.56	0.52	0.51	0.53
10	20589	20555	20646	20993	20975	0.74	0.50	0.48	0.53	0.49
15	20570	20671	20715	20757	20844	0.71	0.58	0.51	0.49	0.46
20	20453	20678	20679	20775	20821	0.67	0.53	0.48	0.48	0.47

Finally, also for the case with multiple depots (with or without partially depot preassignments) we note a similar trend of the results produced by the algorithm, with total route length that increases while waste capacity index tends to decrease by increasing the value of w.

6 Conclusions

In this work, we consider the planning of the final distribution of fuel oil from a set of storage depots to a set of petrol stations faced by a European oil company, which has to decide the weekly replenishment plan for each station, and determine petrol station visiting sequences (vehicle routes) for each day of the week. At tactical level, it is reasonable to assume a fleet of homogeneous vehicles (tank trucks) of given capacity. The problem is modelled as a particular multi-depot periodic vehicle routing problem. We propose a hybrid genetic algorithm to solve this problem. Computational results on real instances show the effectiveness of the proposed approach and the performance achieved when the petrol stations are partially pre-assigned to a depot or not. Further researches may be for example devoted to extend the investigation on the better percentage of petrol stations to assign to depots without decreasing the performance in terms of objective function.

References

1. Alegre, J., Laguna, M., Pacheco, J.: Optimizing the periodic pick-up of raw materials for a manufacturer of auto parts. Eur. J. Oper. Res. **179**(3), 736–746 (2007)
2. Angelelli, E., Speranza, M.: The period vehicle routing problem with intermediate facilities. Eur. J. Oper. Res. **137**(2), 233–247 (2002)
3. Avella, P., Boccia, M., Sforza, A.: Solving a fuel delivery problem by heuristic and exact approaches. Eur. J. Oper. Res. **152**, 170–179 (2004)
4. Benantar, A., Ouafi, R., Boukachour, J.: A petrol station replenishment problem: new variant and formulation. Logistics Res. **9**(6) (2016). doi:10.1007/s12159-016-0133-z

5. Carotenuto, P., Giordani, S., Massari, S., Vagaggini, F.: Periodic capacitated vehicle routing for retail distribution of fuel oils. Transp. Res. Procedia **10**, 735–744 (2015)
6. Chao, I.M., Golden, B.L., Wasil, E.: An improved heuristic for the period vehicle routing problem. Networks **26**(1), 25–44 (1995)
7. Cordeau, J.F., Gendreau, M., Laporte, G.: A tabu search heuristic for multi-depot vehicle routing problems. Networks **30**, 105–119 (1997)
8. Cornillier, F., Boctor, F.F., Laporte, G., Renaud, J.: An exact algorithm for the petrol station replenishment problem. J. Oper. Res. Soc. **59**, 607–615 (2008)
9. Cornillier, F., Boctor, F.F., Laporte, G., Renaud, J.: A heuristic for the multi-period petrol station replenishment problem. Eur. J. Oper. Res. **191**, 295–305 (2008)
10. Cornillier, F., Boctor, F.F., Renaud, J.: Heuristics for the multi-depot petrol station replenishment problem with time windows. Eur. J. Oper. Res. **220**, 361–369 (2012)
11. Cornillier, F., Laporte, G., Boctor, F.F., Renaud, J.: The petrol station replenishment problem with time windows. Comput. Oper. Res. **36**, 919–935 (2009)
12. Drummond, L., Ochi, L., Vianna, D.: An asynchronous parallel metaheuristic for the period vehicle routing problem. Future Gener. Comput. Syst. **17**(4), 379–386 (2001)
13. Glover, F., Hao, J.K.: The case for strategic oscillation. Ann. Oper. Res. **183**, 163–173 (2011)
14. Hoff, A., Andersson, H., Christiansen, M., Hasle, G., Løkketangen, A.: Industrial aspects and literature survey: fleet composition and routing. Comput. Oper. Res. **37**(12), 2041–2061 (2010)
15. Malépart, V., Boctor, F.F., Renaud, J., Labilois, S.: Nouvelles approches pour l'approvisionnement des stations d'essence. Revue Française de Gestion Industrielle **22**, 15–31 (2003)
16. Ng, W., Leung, S., Lam, J., Pan, S.: Petrol delivery tanker assignment and routing: a case study in Hong Kong. J. Oper. Res. Soc. **59**, 1191–1200 (2008)
17. Nguyen, P.K., Crainic, T.G., Toulouse, M.: A hybrid generational genetic algorithm for the periodic vehicle routing problem with time windows. J. Heuristics **20**(4), 383–416 (2014)
18. Pacheco, J., Alvarez, A., Garcia, I., Angel-Bello, F.: Optimizing vehicle routes in a bakery company allowing flexibility in delivery dates. J. Oper. Res. Soc. **63**, 569–581 (2012)
19. Pirkwieser, S., Raidl, G.R.: Multilevel variable neighborhood search for periodic routing problems. In: Cowling, P., Merz, P. (eds.) EvoCOP 2010. LNCS, vol. 6022, pp. 226–238. Springer, Heidelberg (2010). doi:10.1007/978-3-642-12139-5_20
20. Popović, D., Vidović, M., Radivojević, G.: Variable neighborhood search heuristic for the inventory routing problem in fuel delivery. Expert Syst. Appl. **39**, 13390–13398 (2012)
21. Prins, C.: A simple and effective evolutionary algorithm for the vehicle routing problem. Comput. Oper. Res. **31**, 1985–2002 (2004)
22. Rahimi-Vahed, A., Crainic, T.G., Gendreau, M., Rei, W.: Fleet-sizing for multi-depot and periodic vehicle routing problems using a modular heuristic algorithm. Comput. Oper. Res. **53**, 9–23 (2015)
23. Sedgewick, R.: Permutation generation methods. ACM Comput. Surv. **9**(2), 137–164 (1977)
24. Toth, P., Vigo, D.: The Vehicle Routing Problem. SIAM Monographs on Discrete Mathematics and Applications. SIAM, Philadelphia (2002)
25. Toth, P., Vigo, D.: The granular tabu search and its application to the vehicle-routing problem. INFORMS J. Comput. **15**(4), 333–346 (2003)
26. Triki, C.: Solution methods for the periodic petrol station replenishment problem. J. Eng. Res. **10**(2), 69–77 (2013)
27. Vidal, T., Crainic, T.G., Gendreau, M., Lahrichi, N., Rei, W.: A hybrid genetic algorithm for multi-depot and periodic vehicle routing problems. Oper. Res. **60**(3), 611–624 (2012)
28. Vidović, M., Popović, D., Ratković, B.: Mixed integer and heuristics model for the inventory routing problem in fuel delivery. Int. J. Prod. Econ. **147**, 593–604 (2014)

A Multi-criteria Intelligent Control for Traffic Lights Using Reinforcement Learning

Junchen Jin(✉) and Xiaoliang Ma

System Simulation and Control, Department of Transport Science,
KTH Royal Institute of Technology, Teknikringen 10, 10044 Stockholm, Sweden
junchen@kth.se

Abstract. Traffic signal control plays a crucial role in traffic management and operation practices. In the past decade, adaptive signal control systems, capable of adjusting control schemes in response to traffic patterns, have shown the abilities to improve traffic mobility. On the other hand, the negative impacts on environments by increased vehicles attract increased attentions from traffic stakeholders and the general public. Most of the prevalent adaptive signal control systems do not address energy and environmental issues. The present paper proposes an adaptive signal control system capable of taking multi-criteria strategies into account. A general multi-agent framework is introduced for modeling signal control operations. The behavior of each cognitive agent is modeled by a Constrained Markov Decision Process (CMDP). Reinforcement learning algorithms are applied to solve the MDP problem. As a result, the signal controller makes intelligent timing decisions according to a pre-defined policy goal. A case study is carried out for the stage-based control scheme to investigate the effectiveness of the adaptive signal control system from two perspectives, traffic mobility and energy efficiency. The control approach can be further applied to a large network in a decentralized manner.

Keywords: Adaptive traffic light control · Multi-criteria strategies · Intelligent timing decision · Reinforcement Learning

1 Introduction

Traffic lights are commonly deployed at intersections for the purpose of permitting all road users on the conflicting traffic movements to proceed efficiently and safely through space. The efficiency of a traffic signal control system highly depends on how the system is operated. Apart from the traffic efficiency, energy efficiency and emission impacts have attracted increased attentions from traffic stakeholders and the general public. As reported by previous studies e.g. [6], vehicles consume more fuel and emit more pollutant and greenhouse gases when traffic is congested.

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_22

Traffic control systems differ in how they allocate signal phases and how they assign timing for each phase. The way that signal control system allocates phase is called phasing technique. For example, Stage-based phasing is one of the most commonly used phasing techniques in the existing signal control systems. Here, signal stage refers to the compatible traffic movements that are grouped together. In a stage, the associated traffic lights always show same indications. Stage-based phasing follows a round-robin scheduling scheme such that the stage is activated in a specific time span within a signal cycle. Traffic lights always show the same indications in a stage.

Stage-based fixed-time control system (SBFT) and stage-based vehicle actuated control system (SBVA) are two classical types of stage-based signal control systems using different signal timing methods. In terms of SBFT, the duration and order of signal stages remain fixed. Usually, several signal timing plans are selected within a day according to the historical traffic pattern. In practice, SBVA is often superior to SBFT due to its ability in adapting the control scheme to flow fluctuation under varying traffic demand. Specifically, SBVA extends or terminates a green indication on the basis of the presence or absence of vehicles.

Reinforcement learning (RL) has been considered as a well-suited learning method in traffic signal control applications. The study in [9] firstly incorporated an RL algorithm into controlling an isolated signalized intersection. The simulation results showed that RL-based signal control outperformed a fixed time control by reducing average waiting times for all vehicles. Regarding the state representations of a traffic system, RL-based research directions can be divided into two branches. They are intersection-based approach and vehicle-based approach.

In the intersection-based approach, system states are represented by aggregated measures at an intersection. For instance, Abdulhai et al. applied a simple RL technique to an isolated traffic signal, in which states include queue lengths on four approaches and the elapsed times [1]. If value functions are computed with respect to individual vehicle information, such an RL-based signal control approach is named as a vehicle-based approach. The research direction began from Wiering who utilized RL to control traffic lights for the purpose of minimizing the overall waiting time of vehicles [10]. In terms of vehicle-based approaches, the number of states scales acceptably for large networks because it grows linearly with the number of vehicles [3]. However, the requirements are strict for a vehicle-based control system in real-world applications because vehicles are required to send their information to the signal controller in a high frequency.

In summary, there is a still lack of efforts on proposing an adaptive signal control system that considers energy and environment as one of the core design components. To this end, the objective of our study is to also consider energy-efficiency when developing adaptive signal control system in the context of the stage-based phasing technique. This paper is organized as follows. The operation of the proposed signal control is illustrated in the next section. The proposed signal control is enabled by a reinforcement learning approach. The design of the signal controller is presented in Sect. 3. Section 4 elaborates the case study together with analyses and discussions of the results in a simulation environment.

2 Signal Control Approach

2.1 Agent-Based Modeling

Regarding the operation of the stage-based signal control systems, a signal stage can be regarded as an individual cognitive agent. The operation of a stage agent follows a fully observable Markov decision process (MDP) [7]. Since multi-criteria strategy is considered in this paper, constrained Markov decision process (CMDP) is applied in the problem formulation due to its capacity of handling multiple objectives. The CMDP scheme optimizes one cost objective while keeping the other objectives below certain bounds given.

Let us start with the basic notations in the formulation as follows:

- $\mathcal{X}$ is a set of states.
- $\mathcal{U}$ is a set of actions.
- $P : \mathcal{X} \times \mathcal{X} \times \mathcal{U} \rightarrow [0, 1]$ represents the agent dynamics. $P(\mathbf{x}'|\mathbf{x}, \mathbf{u})$ specifies the transition probability of the agent to state $\mathbf{x}'$ from state $\mathbf{x}$ by carrying out the action $\mathbf{u}$.
- $c : \mathcal{X} \times \mathcal{U} \times \mathcal{X} \rightarrow \mathcal{R}$, where $\mathcal{R}$ denotes a set of real number. $c(\mathbf{x}, \mathbf{u}, \mathbf{x}')$ is defined as the immediate cost function by doing action $\mathbf{u}$ when the agent is in the state $\mathbf{x}$ and transmits to the state $\mathbf{x}'$.
- $\pi : \mathcal{X} \rightarrow \mathcal{U}$, control policy that specifies how an agent chooses its actions given the state.
- $\mathbf{d} : \mathcal{X} \times \mathcal{U} \times \mathcal{X} \rightarrow \mathcal{R}$. $\mathbf{d}(\mathbf{x}, \mathbf{u}, \mathbf{x}')$ is a vector of the immediate costs, related to constraints, when doing action $\mathbf{u}$ when the agent is in the state $\mathbf{x}$ transmitting to the state $\mathbf{x}'$.

A stage-based signal control system can be formulated as a multi-agent system. In the multi-agent system, the stage agents share a common environment. The generalization of this signal control system at an isolated intersection can be represented by a tuple:

$$\mathcal{S} =< \mathcal{X}_1, \ldots, \mathcal{X}_n, \mathcal{U}_1, \ldots, \mathcal{U}_n, P_1, \ldots, P_n, c_1, \ldots, c_n, d_1, \ldots, d_n, \pi_1, \ldots, \pi_n >, \tag{1}$$

where n denotes the total number of stages at an isolated intersection. Two main specifications are defined in line with stage-based phasing.

- Only one agent at a time is active.
- All the agents have the same goal and implements the same cost function.

For the purpose of choosing an optimal action sequence, the agent compares different values of cumulative discounted costs. Let's consider $Q_i^{\pi_i}$ as the expectation of cumulative discounted cost for stage agent i applying policy π_i. $T_{i,j}$ denotes the length of the jth active episode for stage i. The agent i starts from an initial state, $\mathbf{x}_{i,1,0}$, and takes an action, $\mathbf{u}_{i,1,0}$. In the case of discounted costs, the future costs are worth less than the current cost. The expected value of cumulative discounted cost is, therefore, written as:

$$
\begin{aligned}
Q_i^{\pi_i} &= \mathbb{E}[\sum_{j=1}^{\infty}\sum_{t=0}^{T_{i,j}}\gamma^t c_{i,j,t}|\mathbf{x}_{i,1,0},\mathbf{u}_{i,1,0}] \\
&= \sum_{j=1}^{\infty}\sum_{t=0}^{T_{i,j}}\sum_{\mathbf{x}_{i,j,t+1}\in\mathcal{X}_i}\gamma^t P_i(\mathbf{x}_{i,j,t+1}|\mathbf{x}_{i,j,t},\mathbf{u}_{i,j,t})\times c_i(\mathbf{x}_{i,j,t},\mathbf{u}_{i,j,t},\mathbf{x}_{i,j,t+1}) \\
&\quad s.t.\ \mathbf{d}_{i,j,t}\leq\mathbf{v}
\end{aligned}
\tag{2}
$$

where $\gamma^t \in [0,1]$ denotes the discount rate; $\mathbf{x}_{i,j,t}$, $\mathbf{u}_{i,j,t}$ and $c_{i,j,t}$, respectively, refer to the state, action and cost for agent i at time step t of the jth active episode. $\mathbf{d}_{i,j,t}$ denotes the constrained costs and $\mathbf{v}$ is a vector of the bounds on other additional objectives.

Equation 2 can be reformulated as a Bellman's equation i.e.

$$
\begin{aligned}
&Q_i^{\pi_i}(\mathbf{x}_{i,j,t},\mathbf{u}_{i,j,t}) = \\
&\sum_{\mathbf{x}_{i,j,t+1}\in\mathcal{X}_i} P_i(\mathbf{x}_{i,j,t+1}|\mathbf{x}_{i,j,t},\mathbf{u}_{i,j,t})[c_i(\mathbf{x}_{i,j,t},\mathbf{u}_{i,j,t},\mathbf{x}_{i,j,t+1})+\gamma Q_i^{\pi_i}(\mathbf{x}_{i,j,t+1},\mathbf{u}_{i,j,t+1})]
\end{aligned}
\tag{3}
$$

The optimal Q-value is basically generated by taking the optimal control policy, represented by

$$
Q^*(\mathbf{x}_{i,j,t},\mathbf{u}_{i,j,t}) = \min_{\pi_i} Q_i^{\pi_i}(\mathbf{x}_{i,j,t},\mathbf{u}_{i,j,t}). \tag{4}
$$

If the system dynamics, including the transition function $P_i, i = 1,\ldots,n$, and the cost function c_i, $\forall i \in n$ are available, Eq. 4 can be solved by optimal control methods. However, the system dynamics are not easy to model due to the uncertainties in the traffic system. To find an optimal action for the active agent in a certain state, the optimal action is given by

$$
\mathbf{u}^*_{i,j,t} = arg\min_{\mathbf{u}'_{i,j,t}\in\mathcal{U}_i} Q^*(\mathbf{x}_{i,j,t},\mathbf{u}'_{i,j,t}) \tag{5}
$$

where $\mathbf{u}^*_{i,j,t}$ denotes the optimal action taken by the active agent i at time t. Therefore, the optimal control problem (in Eq. 4) becomes to minimize the cumulative reward function with respect to the state-action pairs. As only one stage agent is can be activated at a time, the optimal value of cumulative costs of all the stages (Q^*) is simply computed by summing up all the Q values for individual stage agents.

$$
Q^* = \sum_{i\in n} Q_i^* \tag{6}
$$

The task of an agent is to optimize its long-term performance, although only receiving feedback about its immediate cost. Reinforcement learning (RL) algorithms are capable of finding an solutions without completely knowing the information of system dynamics [2]. An RL algorithm runs with a stochastic iterative algorithm using the observations obtained from on-line samples of state-action trails. The RL employed in this study will be presented in the following section.

2.2 Temporal Difference Learning

Temporal Difference (TD) algorithm, a classical type of RL, is applied in the study. A TD algorithm performs an on-line updating procedure by which Q-factor is immediately updated after the state being visited. Q-learning and SARSA(State-Action-Reward-State-Action) are two typical TD learning algorithms. Both of them have been proven to converge to an optimal value if the agent keeps visiting state-action pairs for an infinite number of times [2]. Q-learning is an off-policy learning algorithm meaning that the state-action pair (Q value) directly approximates the optimal Q factors independent of what policy is applied. Here, policy refers to the mapping between a state and an action. On the contrary, SARSA is an on-policy learning algorithm which estimates Q-factor according to a specific behavior policy. The general update process for TD is shown below.

$$Q_{i,j,t+1}(\mathbf{x}_{i,j,t}, \mathbf{u}_{i,j,t}) = Q_{i,j,t}(\mathbf{x}_{i,j,t}, \mathbf{u}_{i,j,t}) + \alpha\delta_{i,j,t}(\mathbf{x}_{i,j,t}, \mathbf{u}_{i,j,t}) \tag{7}$$

where $Q_{i,j,t}(\mathbf{x}_{i,j,t}, \mathbf{u}_{i,j,t})$ refers to the estimated Q-value for stage agent i at time t of the jth active episode, $\alpha \in [0, 1]$ refers to the learning rate. $\delta_{i,j,t}$ denotes the temporal difference between the current estimate and the updated estimate. The TD is given by

$$\delta_{i,j,t}(\mathbf{x}_{i,j,t}, \mathbf{u}_{i,j,t}) = c_{i,j,t+1} + \gamma^t \max_{\mathbf{u}'_{i,j,t+1}} Q_{i,j,t}(\mathbf{x}_{i,j,t+1}, \mathbf{u}'_{i,j,t+1}) - Q_{i,j,t}(\mathbf{x}_{i,j,t}, \mathbf{u}_{i,j,t}) \tag{8}$$

whereas, SARSA is an on-policy TD algorithm that estimates Q-factor concerning a specific behavior policy. TD is computed by

$$\delta_{i,j,t}(\mathbf{x}_{i,j,t}, \mathbf{u}_{i,j,t}) = c_{i,j,t+1} + \gamma^t Q_{i,j,t}(\mathbf{x}_{i,j,t+1}, \pi(\mathbf{x}_{i,j,t+1})) - Q_{i,j,t}(\mathbf{x}_{i,j,t}, \mathbf{u}_{i,j,t}) \tag{9}$$

In this study, softmax is used as the control policy. The probability $Pr_{i,j,t}$ of taking action $\mathbf{u}_{i,j,t+1}$ when the agent i is in state $\mathbf{x}_{i,j,t+1}$ at time t of the jth active episode.

$$Pr_{i,j,t}(\mathbf{u}_{i,j,t+1}|\mathbf{x}_{i,j,t+1}) = \frac{e^{Q(\mathbf{x}_{i,j,t+1}, \mathbf{u}_{i,j,t+1})/\tau}}{\sum_{\mathbf{u}'_{i,j,t+1}} e^{Q(\mathbf{x}_{i,j,t+1}, \mathbf{u}'_{i,j,t+1})/\tau}} \tag{10}$$

where τ is a meta-parameter for softmax policy.

2.3 Simulation-Based Learning Process

The proposed signal control system relies on a learning process to build up the knowledge of signal agents. Traffic simulations facilitate the learning process providing detailed information of the system states and costs. In the signal control system, vehicle characteristics (e.g. vehicle position, speed, and acceleration) in a high resolution are needed to compute the cost functions. SUMO, an open sourced microscopic traffic simulation tool, is applied in the learning

process. SUMO owns an accurate representation of the road network geometry as well as a detailed modeling of individual vehicles' behavior [4]. Specifically, a collision-free car following model is applied in the SUMO simulator to capture vehicle motions (see [5] for more details). Additionally, SUMO allows for on-line interaction via an application programming interface (i.e. TraCI). Thus, at each simulation step, vehicle characteristics data are able to be transmitted from the simulator to the signal control system. Similarly, the change of traffic light indications in the simulator is manipulated according to the actions made by the signal agents. For the sake of avoiding the effects of initial vehicle loadings, signal agents do not start learning until 900 s simulation time.

```
Initialize signal parameters: min_green_time, yellow_time and all_red_time for each stage;
Initialize learning parameters α, γ and τ for reinforcement learning algorithms;

max_greens_time ← genetic_algorithm(va_sig);

for all i ∈ [1, pre_leanrning_num] do
    for all sim_sec1 ∈ [1, sim_period] do
        tl_indications ← va_sig.get_indications();
        u ← ada_sig1.transfer_action(tl_indications);
        x, c ← simulator.get_state_reward();
        ada_sig1.learn_with_action(x, c, u);
        simulator.run_one_sec(tl_indications);
    end for
end for
knowledge = ada_sig1.get_knowledge()

for all j ∈ [1, off_training_num] do
    ada_sig2.set_knowledge(knowledge)
    for all sim_sec2 ∈ [1, sim_period] do
        x, c ← simulator.get_state_reward();
        ada_sig2.learn(x, c);
        ada_sig2.run_one_sec();
        u ← ada_sig2.get_action();
        tl_indications ← ada_sig2.transfer_tl(u);
        simulator.run_one_sec(tl_indications);
    end for
    is_cut_flag = ada_sig2.get_is_cut_flag();
    if is_cut_flag == True then
        off_training_num+ = 1
    else
        knowledge = ada_sig2.get_knowledge()
    end if
end for
```

Fig. 1. Pseudo-code of the simulation-based learning process.

Figure 1 elaborates a two-phase learning process in details with the aid of the pseudo-code. At first, signal parameters and learning parameters are initialized. The first phase is named as 'pre-learning' phase. In the phase, the proposed signal controller is guided by a well-established signal controller, optimized stage-based vehicle actuated signal controller (va_sig). A simulation-based optimization is applied for finding a proper setting of signal parameters with respect to a certain traffic demand (line 3 in the pseudo-code). A genetic algorithm is employed as the optimization algorithm [6]. During the pre-learning phase, traffic light indications are received from va_sig every second (see line 6). Line 7 describes that traffic light indications are then transferred to the action of the adaptive signal control system (ada_sig1). The state ($\mathbf{x}$) and cost (c) are obtained from traffic simulator. Thus, the adaptive signal control system learns knowledge using the information of state, action and cost (line 9). At the end of the pre-learning phase, the knowledge is obtained from the signal controller and is used as an initial knowledge for the next learning phase.

The second learning phase is called off-line training phase. Likewise, the state and cost are obtained from the simulator in the off-line training phase. The state and action are used as the inputs for the proposed adaptive signal control system (adp_sig2). The timing action is acquired from the signal control and then is transferred to traffic light indications. In the off-line training phase, one of the main tasks is to avoid the unacceptable learning behaviors. An indicator, cut_flag, is defined to determine whether the learning process is proper or not. The cut_flag is set to $True$ when the immediate cost value is bigger than a predefined threshold. When cut_flag equals to $True$, the knowledge is not taken into considerations. Otherwise, the knowledge obtained from the current simulation run will be used as another initial knowledge for the next training (see line 25–29).

3 Control Agent Design

In this section, the design of three essential components (i.e. state, action, and cost) is explicitly described. The state of an stage agent consists of three elements: $q_{i,j,t}$, $g_{i,j,t}$ and $q_{i,j,t}^{others}$. Here, $q_{i,j,t}$ and $q_{i,j,t}^{others}$ denote the average queue length (number of vehicles) associated with the lanes that are and are not severed for the stage agent i, respectively. $g_{i,j,t}$ refers to the elapsed green time for the stage agent i. The state vector can be expressed by

$$\mathbf{x}_{i,j,t} = [q_{i,j,t}, q_{i,j,t}^{others}, g_{i,j,t}]^T \tag{11}$$

The action of an agent is either to be ordered to terminate or to extend the green time, i.e.

$$\mathbf{u}_{i,j,t} = g_e \in \{0, 1\} \tag{12}$$

in which $g_e = 0$ means that signal stage agent is to terminate while $g_e = 1$ represents that signal stage agent i extends green time for next time step. The termination action is valid if and only if the value of elapsed green time is higher

than g_i^{low}, where g_i^{low} represents the lower bound of the green time. If the action is ordered to terminate, the stage agent will terminate after accounting for the pre-defined yellow time and inter-green time.

In general, the cost function is defined by decision makers according to their policy goals. In our study, essential criteria are integrated into the multi-agent framework for optimizing the relevant objective functions. In this paper, two cost functions implemented for evaluation include average fuel consumption and average travel delay respectively.

A micro-scale emission model, CMEM, is applied in our study to estimate instantaneous fuel consumption of individual vehicles [8]. CMEM considers the modal operations of a vehicle (e.g. instantaneous speed and acceleration) and estimates fuel consumption based on the detailed vehicle dynamics. Besides, the delay function for vehicle v at time t is defined as

$$d_{v,t} = \begin{cases} t_i - \frac{l_{v,t}}{s_{v,t}^{des}}, & \text{if } s_{v,t} < s_{v,t}^{des} \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

where $d_{v,t}$ refers to the travel delay for vehicle v during the last step of t. $l_{v,t}$ denotes last-step travel distance of vehicle i at time t; t_i is the time step for system to update the next action; $s_{v,t}$ and $s_{v,t}^{des}$ are the speed during the previous time step and desired speed for vehicle v at time t, respectively.

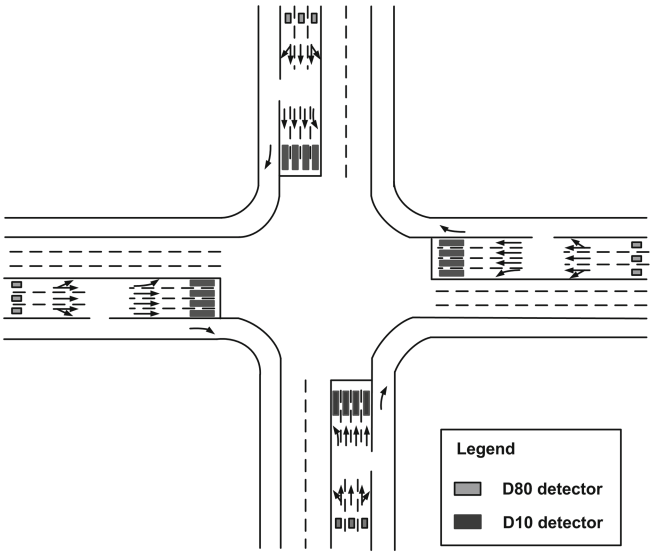
Let us denote $V_{i,j,t}$ as the set of vehicles that are driving on the lanes severed by the signal stage i at time t of jth active episode. The cost function is hence calculated by

$$c_{i,j,t} = \frac{\sum_{v \in V_{i,j,t}} d_{v,t}}{|V_{i,j,t}|} \tag{14}$$

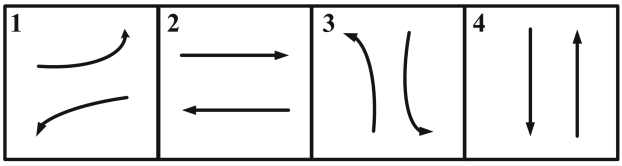
4 CASE STUDY

4.1 Experiment Setup

The proposed adaptive traffic light control is tested on a four-armed isolated intersection. Figure 2 shows the network layout and the pre-defined sequence of four signal stages. Note that traffic lights are not regulated right-turn directions for simplicity purpose. Optimized stage-based vehicle-actuated signal control (OSBVA) is considered as a baseline case. OSBVA extends green time based on the detection information regarding the presence or absence of vehicles. The placement of detectors is shown in 2. A long detector is placed close to the stop line while a short detector is placed 80 m upstream from the stop line on each lane. Generally speaking, the extension time is governed by the time gap between two passing vehicles. If the time gap is less than a pre-determined threshold, the green time is extended by a pre-defined value. In addition, signal stages keep on extending green time if the long detector is occupied. If the specified maximum green time is reached, the signal stage is forced to terminate.



(a) Network layout



(b) Signal stages

Fig. 2. Network layout and the signal stage definition.

Three different traffic demands are implemented in the case studies (see Table 1). In Table 1, the 'Homogeneous' scenario means that traffic demands are identical for all four approaches at the intersection. In the 'Arterial' scenario, traffic flows corresponding to eastbound, and westbound directions are relatively higher than those of the other two directions. In the 'Unbalanced' scenario, traffic flows for eastbound and northbound directions are higher, compared with the other two directions.

4.2 Results and Discussions

The case study firstly generates signal control parameters by a genetic algorithm for OSBVA. The optimization objective is set to be average delay per vehicle and average fuel consumption per kilometer, respectively. Table 2 summarizes the resulted maximum green time and the other signal control parameters for OSBVA. Learning rate and discount rate are two key factors that affect the

Table 1. Traffic volume (vehicles/hour) for each turning movement on the study intersection.

Demand scenario	Eastbound			Westbound			Northbound			Southbound		
	L	T	R	L	T	R	L	T	R	L	T	R
Homogeneous	150	800	150	150	800	150	150	800	150	150	800	150
Arterial	150	800	150	150	800	150	75	400	75	75	400	75
Unbalanced	150	800	150	75	400	75	150	800	150	75	400	75

Note: L, T and R represent the turning rates of left-turn, through and right-turn movements, respectively.

behavior of reinforcement learning algorithm in different ways. For example, the proposed signal control system applying a small learning rate tends to trust the current estimate whereas larger learning rate leads to fast convergence. On the other hand, a small discount rate makes the agents short-sighted by only considering current costs, whereas with a larger rate agents tend to strive for a long-term cost. In this respect, a sensitivity analysis has been carried out to find proper learning parameters. In the analysis, ten simulation runs are directly applied at the off-line training phase. The learning parameters, resulting in the best performance measures, are shown in Table 2. In the table, Adp-QLearning refers to that the adaptive signal control system applies Q-learning as its learning algorithm. Whereas, Adp-SARSA means that the adaptive signal control system applies SARSA.

Table 2. Signal parameters and learning parameters for the control systems ('Homogeneous' demand scenario).

Signal control systems	Signal parameters				Learning parameters	
	g^{max}	g^{min}	y	r	α	γ
OSBVA (delay)	[37, 26, 31, 33]	5	3	2	–	–
OSBVA (fuel)	[32, 29, 37, 38]	5	3	2	–	–
Ada-QLearning (delay)	–	5	3	2	0.4	0.8
Adp-QLearning (fuel)	–	5	3	2	0.4	0.7
Adp-SARSA (delay)	–	5	3	2	0.1	0.9
Adp-SARSA (fuel)	–	5	3	2	0.2	0.7

Note: g^{max}, g^{min} and y represent the minimum green time, yellow time and all-red time, respectively. Delay and fuel in parentheses in the 'Signal control systems' field are corresponding to the optimization function or cost function applied.

Furthermore, the study investigates the influence of the pre-learning process on convergence performance. Signal stage agents follow the behavior of OSBVA in the first five simulation runs if the pre-learning phase is in use. Figure 3 gives an insight of the learning process for Adp-QLearning with and without the

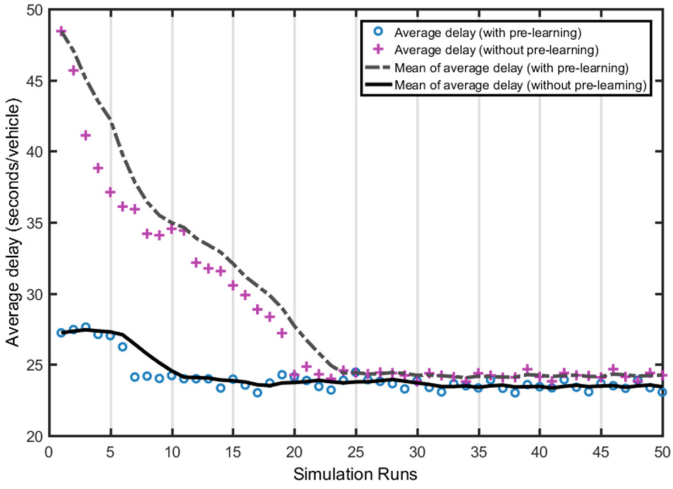


Fig. 3. Convergence performance of Adp-QLearning (delay) signal control system ('Homogeneous' demand scenario)n.

pre-learning phase. In the figure, the "mean of average performance" measure refers to the mean of the performance measure for the previous five simulation runs. Adp-QLeraning applying the pre-learning dramatically decreases the number of simulation runs needed to reach a converged performance. Specifically, Adp-QLearning without the pre-learning phase converges after 25 simulation runs whereas with pre-learning it tends to converge immediately. The converged value of average delay for Adp-QLearning with the pre-learning phase is slightly smaller than the counterpart of Adp-QLearning without the pre-learning phase. This might be caused by the ability of Adp-QLearning with the pre-learning phase to avoid many unacceptable behaviors under the guide of OSBVA. Similar trends are found when average fuel consumption is used as the cost function and also for Adp-SARSA.

Table 3 elaborates the performance measures among different signal control systems in the 'Homogeneous' demand scenario. In general, the adaptive signal control systems give quite promising overall performances. Specifically, fewer average fuel consumptions are achieved by the adaptive traffic light control in all cases, compared to fuel-optimal OSBVA.

SARSA-based signal control achieves lower values of both performance measures compared to the counterpart of Q-learning. The reason behind is as follows. When the signal control system updates the knowledge of an agent, Q-learning simply assumes that an optimal policy is applied during the learning process. The way that Q-learning evolves may lead the signal agents to forget what they have been previously learned. Whereas, SARSA incorporates the action selection policy into its knowledge storing mechanism. The results obtained reveal that SARSA is an approach with superior performance when integrated with the proposed multi-agent traffic light control system.

Table 3. Performance measures among different signal control systems ('Homogeneous' demand scenario).

Signal control systems	Avg. delay (seconds/vehicle)	Avg. fuel (g/km)
OSBVA (delay)	27.20	50.67
OSBVA (fuel)	30.57	49.38
Adp-QLearning (delay)	23.64	48.13
Adp-QLearning (fuel)	24.45	47.39
Adp-SARSA (delay)	23.13	47.62
Adp-SARSA (fuel)	24.07	46.85

Note: Delay and fuel in parentheses in the 'Signal control systems' field are corresponding to the optimization function or cost function applied.

The following experiments try to investigate the adaptiveness of Adp-SARSA when the traffic demand is changing. The 'Arterial' and 'Unbalanced' demand scenarios are implemented. Figure 4 presents the average fuel consumptions generated by Adp-SARSA. Adp-SARSA outperforms OSBVA resulting in fewer fuel consumptions regardless of the setting of traffic demands. In fact, the total traffic demand is reduced by 25% for both 'Arterial' and 'Unbalanced' scenarios. The average fuel consumption is expected to be decreased along with the reduction in traffic flow. The results of Adp-SARSA are in line with the expectation.

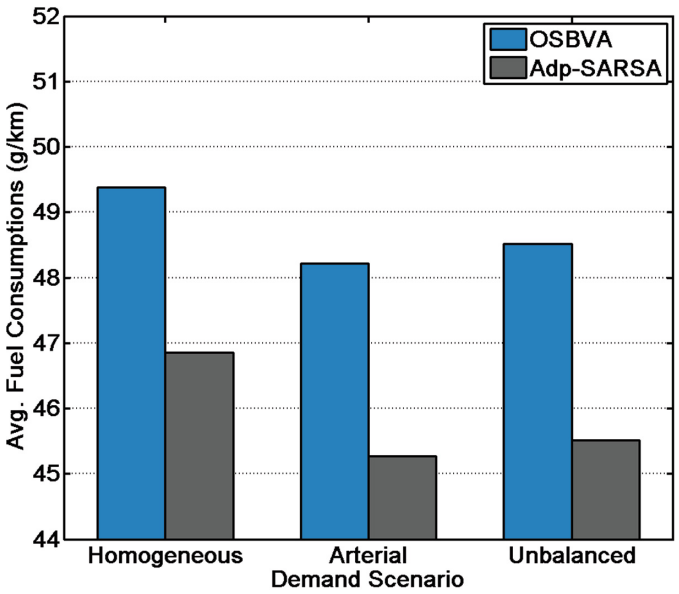


Fig. 4. Average fuel consumption among three demand scenarios for OSBVA and Adp-SARSA.

However, the reduction in average fuel consumption is insignificant when OSBVA is implemented, especially in 'Unbalanced' demand scenario. Consequently, the adaptive signal control can potentially adjust the control schemes of the system when the demand pattern varies.

5 Conclusion

This paper proposes an adaptive signal control system that takes multi-criteria traffic management strategies into consideration. The proposed adaptive signal control system focuses on intelligent timing decisions when stage-based phasing control is applied. The stage-based phasing scheme is modeled as a multi-agent system where a stage is regarded as an individual cognitive agent. The operation of a signal stage agent follows a Markov decision process. The intelligent timing is achieved through utilizing a reinforcement learning algorithm. A two-phase learning process is introduced in the present paper. Particularly, at the first learning phase, the adaptive signal control system is guided by an optimized stage-based vehicle actuated controller for the sake of initializing a reasonable prior knowledge. An open-source microscopic traffic simulator, SUMO, is used in the learning process. A case study is carried out in a simulation environment. A sensitivity analysis is also carried out to understand the relationships between system parameters and performance measures when two different criteria are applied. Finally, three main findings are elaborated in the case study:

- The learning efficiency can be significantly improved by introducing the two-phase learning process;
- The proposed adaptive system shows its superior performance than the optimized stage-based vehicle actuated controller with respect to traffic mobility and energy efficiency;
- When traffic demand varies, the proposed adaptive signal control system has the potential to adjust its control scheme accordingly.

References

1. Abdulhai, B., Pringle, R., Karakoulas, G.J.: Reinforcement learning for true adaptive traffic signal control. J. Transp. Eng. **129**(3), 278–285 (2003)
2. Barto, A.G.: Reinforcement Learning: An Introduction. MIT press, Cambridge (1998)
3. Khamis, M.A., Gomaa, W.: Adaptive multi-objective reinforcement learning with hybrid exploration for traffic signal control based on cooperative multi-agent framework. Eng. Appl. Artif. Intell. **29**, 134–151 (2014)
4. Krajzewicz, D., Erdmann, J., Behrisch, M., Bieker, L.: Recent development and applications of SUMO-simulation of urban mobility. Int. J. Adv. Syst. Meas. **5**(3–4), 128–138 (2012)
5. Krauss, S., Wagner, P., Gawron, C.: Metastable states in a microscopic model of traffic flow. Phys. Rev. E **55**(5), 5597 (1997)
6. Ma, X., Jin, J., Lei, W.: Multi-criteria analysis of optimal signal plans using microscopic traffic models. Transp. Res. Part D Transp. Environ. **32**, 1–14 (2014)

7. Puterman, M.L.: Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley, New York (2014)
8. Scora, G., Barth, M.: Comprehensive modal emissions model (CMEM), version 3.01. User guide. Centre for Environmental Research and Technology, University of California, Riverside (2006)
9. Thorpe, T.L., Anderson, C.W.: Traffic light control using SARSA with three state representations. Technical report, Citeseer (1996)
10. Wiering, M.: Multi-agent reinforcement learning for traffic light control. In: ICML, pp. 1151–1158 (2000)

Analysis of Imprecise Perception in Route Choice Considering Fuzzy Costs

Mario Binetti[1], Marco De Mitri[2], and Michele Ottomanelli[1(✉)]

[1] DICATECh, Technical University of Bari, Bari, Italy
{mario.binetti,michele.ottomanelli}@poliba.it
[2] NIER Ingegneria, Bologna, Italy
m.demitri@niering.it

Abstract. In this work we describe a path choice model in which costs are represented by fuzzy numbers in order to consider uncertainty of users and trying to overcome some limits of well know logit based choice models. Users' estimation costs in order to choice a route over a road network is made in a real uncertain, imprecise and ambiguous environment. As suggested by research in this field, fuzzy numbers is a suitable framework to model such choice context. In this paper, we analyse path costs represented by Fuzzy Numbers. For the fuzzification of the crisp value of the cost we introduce a parameter, that aims at representing the imprecision of users in path cost perception. We assume that users associate a preference index to the paths according to the estimated costs and that the probability of choice of a path is proportional to the preference index. Moreover, we introduce an index for evaluating the overlapping degree of each path with the others. In this way, we overcome the problem due to the Independence from Irrelevant Alternatives (IIA) of Logit based models, which rises for overlapping paths. The method allows to reproduce user choices as compared to other well established ones.

Keywords: Traffic assignment · Route choice · Fuzzy sets · Uncertainty · Overlapping paths

1 Introduction

Traffic assignment models (TAMs) are powerful tools to reproduce users' choices over a transport network. TAMs are based on path choice models; traditional path choice models (deterministic models and stochastic models) are based on the comparison between path costs with crisp numbers. The deterministic choice models assumes that users perceive exactly each path cost, and choose the one having the maximum utility. Thus, each user chooses minimum cost path, and does not consider any other path. This leads to the well-known "All or Nothing" assignment model.

Generally, reproduction of real perceived cost is not an easy task, and the analyst does not exactly know the utility that users associate to each alternative (perceived utility). The *perceived utility* can be different from the average value (i.e. *systematic utility*) [23]. Therefore, some users can choose a path different from the minimum cost

© Springer International Publishing AG 2018
J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8_23

one, and consequently, a choice probability greater than zero occurs for all the possible paths belonging to the users' choice set.

To reproduce this situation the well-established *random utility models* are used (Logit, Gammit, Probit and their further specifications are the most known) (see for ex. [7]).

Several models based on fuzzy set theory for network assignment have been proposed in literature [2, 10, 11, 20, 24–26].

Fuzzy set theory-based route choice model was proposed for the first time by Teodorović and Kikuchi [27]. Their research was followed by many works using fuzzy set theory and based on different approaches.

Wang and Liao [28] considered the traffic assignment problem when the elements of the incidence matrix representing the graph of the transportation network is assumed fuzzy, in the sense that, the arcs chosen into the desired path for travelling from and Origin to a destination is represented as uncertain. Lotan and Koutsopoulos [16] developed a model that uses the concepts of approximate reasoning for route choice in the presence of information by mean of fuzzy inference approach. Peeta and Yu [21] set up a hybrid (fuzzy–probabilistic) model to study the effect of information on travellers' behaviour. Chang and Chen [9] formulated a link-based fuzzy user-optimal route choice problem embedding link interactions using the variational inequality approach.

Henn [14] proposed a route choice model taking account of the imprecision and the uncertainty lying in the dynamic choice process. Okada and Soper [19] focused their work on a shortest path problem on a network in which a fuzzy number is used to represent the length of each arch belonging to the network. Henn and Ottomanelli [15] analysed the role of uncertainty based on the Possibility Theory framework and discussed the pitfalls of some random utility route choice models. They focused on two types of uncertainty: route capacity and free-flow costs. The proposed model, based on possibility theory, shows the different choice behaviour with respect the two uncertainties. Bin et al. [3] proposed a traffic assignment model that is solved by a heuristic-based algorithm which combines the fuzzy shortest path algorithm and the C-Logit method.

Ghatee and Hashemi [12] assumed a fuzzy level of travel demand and found a fuzzy equilibrium flow that satisfies a quasi-logit formula for the network. Murat and Uludag [18] presented route choice model of transportation network based on fuzzy logic model (FLM), logistic regression model (LRM) and survey-based data. The model considered four parameters (traffic safety, travel time, congestion and environmental effects).

Recently, in [17] a method for fusing data relevant both to drivers' experience and provided information about travel time is proposed. The method was based on the Uncertainty-based Information Theory and takes into account the "compatibility" of data originating from different sources, and provides information about acceptability of results.

Our work starts from the work by Akiyama and Kawahara [1] that modelled the route choice behaviour with the values of possibility between fuzzy goal and the fuzzy travel time for the individual routes.

The purpose of this paper is to define a simplified route choice model able to reproduce the imprecision, uncertainty and vagueness typical of the perceived costs. In our hypothesis the user will choice the path according to the comparison between the

estimated costs for all the paths. These values are imprecise, therefore the user must choose on the base of ambiguous quantities and on the base of approximate computing. Such an environment can be represented and handled in the framework of fuzzy numbers. The effective travel time will respect the foresight *in a certain measure*, due to the fuzzy nature of the problem. This measure decreases as much as the effective cost is far from the estimated value.

The paper is developed as it follows. In Sect. 2, we describe in detail the path choice model and some examples on simplified cases. In Sect. 3, we describe the results on a test network with the equilibrium assignment. Finally, in the last section, we focus on the conclusions and possible developments of the research.

2 The Route Choice Model

The proposed model is based on the following basic assumptions:

1. The user "*i*" considers all the alternatives, which define his *choice set* I^i.
2. The user "*i*" associates to each alternative $j \in I^i$ the perceived utility U_j^i
3. The utility U_j^i assigned to each alternative depends on a number of measurable characteristics X_j^i, or *attributes,* that are related to the alternative itself and to the user; then $U_j^i = U^i\left(\mathrm{X}_j^i\right)$.
4. The utility that the user "*i*" associates to alternative *j* is not exactly known by the analyst, and in every case it is imprecise for the users. Such perceived utility can be represented by a fuzzy number.

On the basis of the given assumptions, it is possible to define the probability that the decision-maker (i.e. user) will select the alternative *j* conditional on his choice set I^i proportionally to the preference associated to this alternative, as it will be shown further on.

2.1 The Representation of Path Costs by Fuzzy Number

Path and arc costs are specified by means of Triangular Fuzzy Numbers (TFN) that is represented by the triple (x_L, x, x_R), where x_L, x and x_R are, respectively, the lower (or left) limit, the central value and the upper (or right) limit. The difference $|x_R - x_L|$ is the so-called "support" of the TFN. To obtain the fuzzy number relevant to the crisp value "x" (i.e. fuzzification) [4] we use a parameter $\alpha \in [0, 1]$ and then we determine the relevant TFN by computing the characteristics points given by the lower (x_L) and the upper (x_R) limits of the TFN as follows:

$$x_R = x \cdot (1 + \alpha);\ x_L = x \cdot (1 - \alpha)$$

In Fig. 1 we represent the TFN corresponding to the values 5 and 7, considering $\alpha = 0.200$.

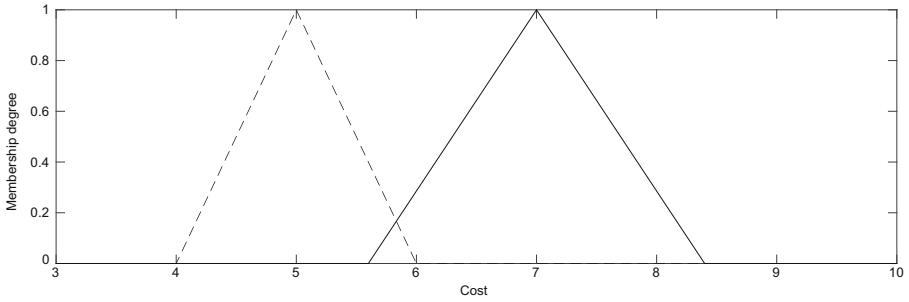


Fig. 1. Representation of costs by Triangular Fuzzy Numbers

Let us consider the sum x + y = z, with x and y *crisp* values. If we add the correspondent TFN (x, y) we obtain "x⊕y = z". In other words, if the value of α is the same for all the TFNs, then the TFN obtained as result of crisp sum is equal to the sum of the TFNs relative to the crisp terms. In this way, our method respects the *independence by the segmentation of the arcs*.

Let consider the following example: let assume two independent paths, the first one made up of a single arc whose costs is 30 and a second one constituted by five arcs whose costs is 6. By applying the proposed method, the two paths have equal crisp cost and equal fuzzy cost: the correspondent TFNs (that represent two paths with the same utility) are identical, as it must be.

The parameter α represents the imprecision of users' prediction (or perception) of path travel time/cost. If this imprecision is null, we assume α = zero. In this case the choice is based on crisp values: the minimum path(s) cost only is (are) chosen. If users will perceive as convenient also other paths (not only the minimum cost one), they will split on them: in this case $\alpha \neq 0$. The value of α is the same for all the paths (only if we have ATIS systems on the network we could have more values for α). Therefore, the support (i.e. base) of the TFN corresponding to the cost "X" (equal to 2αX) is variable, and depends on the crisp value of the cost.

Each user has a different ability to predict (assess) the path travel time. We cannot define a value for each user. Thus, for sake of simplicity, it is advisable to consider an average value of α as representative for all the users belonging to the same class (or category).

Let us analyze the dependence of α on Origin-Destination (O/D) pair. We assume that α is the same for all O/D pairs, i.e. users perceive the state of the network in the same way. If we assign different values of α to different O/D pairs, we would assume that the users of each O/D pair perceive differently the network state. Moreover, we assume that the user is not able to predict the travel time on a certain arc more accurately than other arcs[1]. Thus, we suggest that the value of α will depend from the *urban network*, the *user class* and the *time period*. Naturally, the parameter α have to be

[1] This assumption may fall if in the network there are some VMS. In this situation it might be realistic reduce the α value for the relative arc. In this study, we assume that in the network there aren't VMS.

calibrated. Notwithstanding, we will not discuss about calibration of α since it is out of the scope of this paper. On this topic, the interested reader could refers to the works by Cantarella and Fedele [5, 6], Quattrone and Vitetta [22] or by Hawas [13].

2.2 Definition of the Choice Set I^i

We assume that the fraction of the demand d_{od} on each path is proportional to its convenience. The proposed method is based on the research of "non dominated[2]" paths for the O/D pair. We carry on this research by the comparison of the characteristic points of every TFN. We find the non-dominated paths set by the *dominance check*. A path is considered as non-dominated from the minimum cost one if the lower limit of its TFN is lower than the upper limit of the TFN of the minimum cost path. In Fig. 2 the dashed TFN (20, 25, 30) represent the fuzzy minimum cost path, while the continuous line represents a non-dominated path since its TFN's lower limits (i.e. $x_L = 24$) is lower that the upper limit of the minimum cost TFN (i.e. $x_R = 30$).

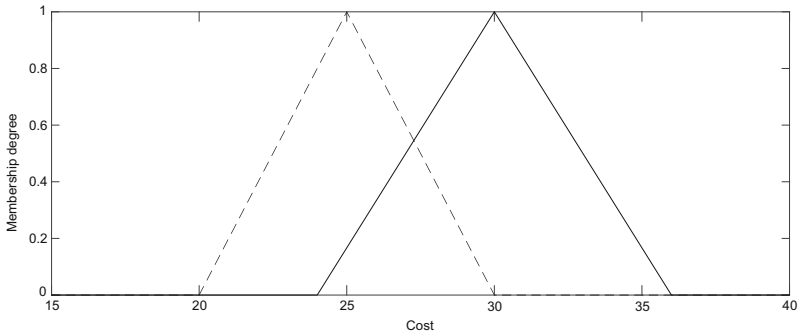


Fig. 2. Example of non-dominated paths

To obtain the non-dominated paths set for each O/D pair without explicit paths enumeration it is possible to use the algorithm proposed in [19] based on the multiple labelling approach, where a number of non-dominated paths can be generated. In addition, the multiple labelling approaches are an exhaustive approach, in which all the possible paths from the source node to the other nodes have to be compared. To this purpose, we modify the algorithm in the step of dominance check according to our method.

The value of α plays an important role in the phase of selection of the paths that will define the choice set. The number of the paths in the choice set depends from the imprecision of costs estimation: the more imprecise are the cost, the higher is the number of elements in the choice set.

[2] The "non dominated" paths are the most convenient paths [19].

For example, let us consider three paths (A, B, C). Let their estimated costs be equal respectively to x≅25, y≅30 and z≅40 min. The relevant TFN can be represented by the following triples:

$TFN_A(x_L, 25, x_R)$, $TFN_B(y_L, 30, y_R)$ and $TFN_B(z_L, 40, z_R)$ where

$$x_R = x \cdot (1+\alpha);\ x_L = x \cdot (1-\alpha)$$

$$y_R = y \cdot (1+\alpha);\ y_L = y \cdot (1-\alpha)$$

$$z_R = z \cdot (1+\alpha);\ z_L = z \cdot (1-\alpha)$$

If we aim at modelling low imprecision, α must assume a low value. Let α be equal to 0.050. We see that paths B and C are dominated from A, as a result only minimum cost path is considered by users in their choice set (Fig. 3).

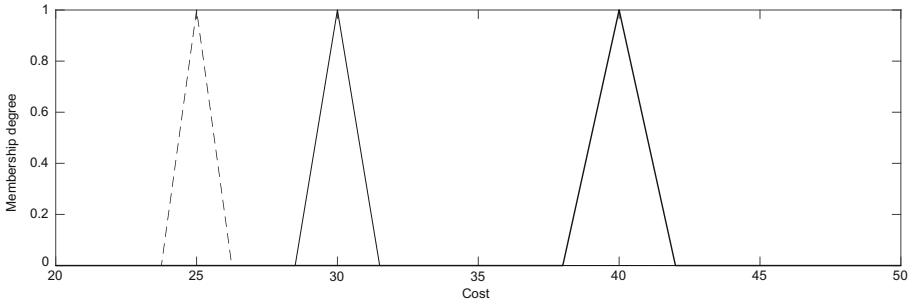


Fig. 3. Example of dominated paths (α = 0.050)

If we want model more imprecision in perceived cost than the previous case, α must assume a higher value. Let α be equal to 0.150 (Fig. 4). In this case the path C only is dominated by A: the choice set is made up of paths A and B.

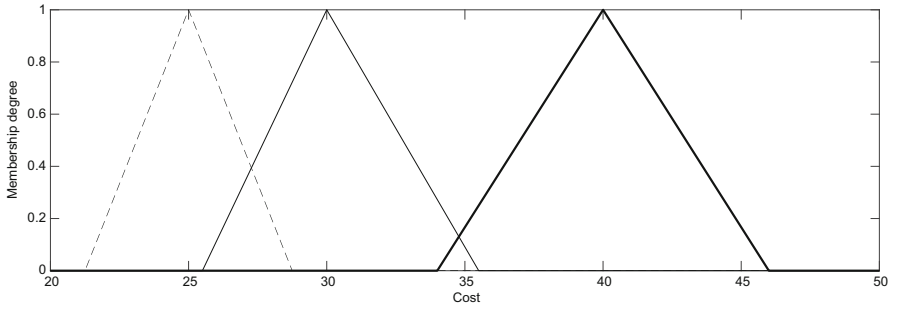


Fig. 4. Example of dominated paths with α = 0.150

As final case, let be $\alpha = 0.300$. In this case there are not dominated paths, so that the choice set is made up of the paths A, B, C all (Fig. 5).

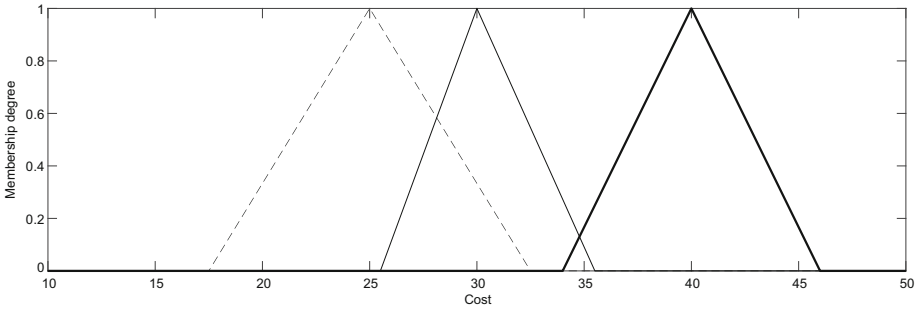


Fig. 5. Example of dominated paths with $\alpha = 0.300$

2.3 Coefficients of Preference

In the previous section, we have shown how to determine (by the dominance check) the choice set for every O/D pair. In this section we are going to show how the demand d_{od} are split among the paths. We assume that the share of demand that will load each path be proportional to the path convenience (utility). In our hypothesis users associate a preference index to each path according to the estimated (perceived) costs.

In this phase, consider the case of independent paths (i.e. they do not have any shared sub-path). In the next section it will be shown how to deal with the case of shared (overlapping) sub-paths.

Then, we associate a numerical value to the convenience of each non-dominated paths. We define for the *i*-th path the coefficient of preference as follows:

$$C_{pref,i} = (\text{min-cost}, R - i\text{-cost}, L) / (i\text{-cost}, R - i\text{-cost}, L) \quad (1)$$

where:

- min-cost is the crisp value of the minimum cost path;
- *i*-cost is the crisp value of the path *i*;
- min-cost,R = min-cost·(1 + α): i.e. the upper limit of the TFN associated to the minimum cost path;
- min-cost,L = min-cost·(1 − α): i.e. the lower limit of the TFN associated to the minimum cost path;
- *i*-cost,R = *i*-cost·(1 + α): i.e. the upper limit of the TFN associated to the path *i*;
- *i*-cost,L = *i*-cost·(1 − α): i.e. the lower limit of the TFN associated to the path *i*;

The coefficient of preference for path *i* is given by the ratio between the shared part of the basis of the two TFNs and the base of the TFN associated to path *i*. The value of $C_{pref,i}$ for every path is between zero and one. In the Fig. 6 we see an example: $\alpha = 0.200$; min-cost = 3; min-cost,L = 2.4; min-cost,R = 3.6; *i*-cost = 4; *i*-cost, L = 3.2; *i*-cost,R = 4.8.

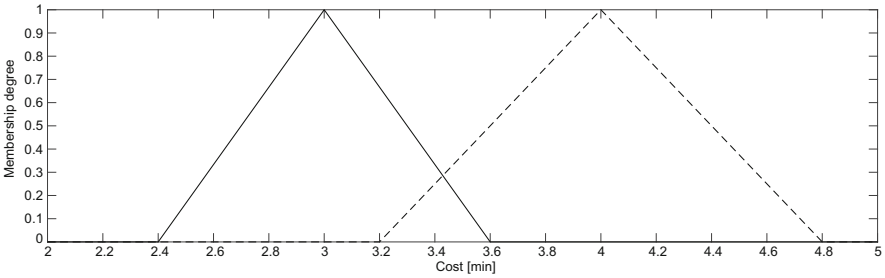


Fig. 6. Representation of path cost membership functions

The coefficient of preference is equal to 1 for the minimum cost path. For the other paths, it is lower as much as higher its crisp cost is. We assume that the probability of choice of a certain path (and consequently the demand fraction that will be assigned to it) is proportional to its coefficient of preference. We calculate the choice probability of the non-dominated paths by the normalisation of the coefficients of preference as follows:

$$p_i' = C_{\text{pref,i}} / \sum C_{\text{pref,i}} \tag{2}$$

Let us focus on the following explaining examples.

Let consider a user whose choice set is made up of two independent paths, namely A and B. Let the cost of B be 10 min greater than the cost of A. What path does the user choose? Under the assumption of *rational decision-maker* user, if his estimation of the path costs were precise, he/she would choose path A that maximize his/her utility. This is true for any value of the cost of paths A and B when B cost is greater than A cost. Therefore, the choice probability of A will be always equal to 100%. In the reality, users do not assume an exact estimation of the path costs [15] because of many source of uncertainty such as approximate reasoning of human beings, different tastes and perception of transportation supply etc. Therefore, choice probabilities depend (differently from the previous case) on the values of path costs and the degree of imprecision/approximation.

For example: let A cost be equal to 2 min and let B cost be equal to 12 min. The user knows that path A will cost *about 2 min* and path B will cost *about 12 min*. It is proper to assume that the choice set will be formed by the A path only. In other words, we expect that the choice probability of A will be equal to 100%. Now suppose that A and B costs are equal respectively to 40 min and 50 min. The user knows that path A will cost *about 40 min* and path B will cost *about 50 min*. Probably he/she will choice also path B, even if path A is clearly more convenient. As a result, we expect that the choice probability of path B will be greater than zero, even if it will be quite smaller than the choice probability of path A. Now suppose that A cost and B cost are equal respectively to 5 h and 20 min and 5 h and 30 min. The user knows that path A will cost *about 5 h 20'* and path B will cost *about 5 h 30'*. Probably the two paths look convenient in the same measure. As a result, the choice probability of the two paths will be very near to 50%.

In the following Table 1 we have summarized the choice probability values obtained by applying the proposed method in the described three examples.

Table 1. Choice probability values ($\alpha = 0.150$)

Path	A	B	A	B	A	B
Cost	2′	12′	40′	50′	5h 20′	5h 30′
Choice probability	1.000	0.000	0.811	0.189	0.531	0.469

Note that the results are realistic for any combination of path cost values, and we obtained these results utilizing only one value of α for modelling the imprecision.

2.4 Overlapping Paths

In general, the paths belonging to the choice set could be partially overlapping [8]. If three or more paths form the choice set we may have unrealistic results (two of them may be overlapping in a certain measure). In this case, by applying the proposed method we may have unrealistic results. To avoid this problem it is necessary to properly define the estimation of choice probabilities.

We introduce an index for evaluating the overlapping degree of each path with the others. We describe the path overlapping in terms of length of shared sub-paths. Therefore, the index will be defined referring to the relevant arc lengths. The *coefficient of overlapping* for the path i is defined as follows:

$$C_{over,i} = [(L_{i,1}/L_i) + (L_{i,2}/L_i) + \ldots + (L_{i,i-1}/L_i) + (L_{i,i+1}/L_i) + \ldots + (L_{i,n}/L_i)] / n_{over} \quad (3)$$

with

- L_i is the length of the path *i;*
- $L_{i,1}(L_{i,2}, L_{i,3}\ldots)$ is the length of the shared sub-paths by the path i and the path 1 (2, 3,…,);
- n_{over} is the number of path partially overlapping in the choice set.

This coefficient is equal to zero for the independent paths. For the overlapping paths, it returns a value that increases proportionally to the length of the shared sub-paths. The value of the coefficient of overlapping is in every case less than one. Now we introduce the *coefficient of independence* as follows:

$$C_{ind,i} = 1 - C_{over,i}$$

It has the opposite meaning of the *coefficient of overlapping*. It is equal to 1 if path i is independent; it is less than 1 if path i has some shared sub-paths. In every case it is

greater than zero. We can now calculate the choice probabilities of the non-dominated paths as follows:

$$p_i = C_{pref,i} \cdot C_{ind,i} / \sum C_{pref,i} \cdot C_{ind,i} \quad (4)$$

If the paths are all independent, all the *coefficients of independence* are equal to one. As a result, the choice probabilities are the same that we will obtain if we applied the normalisation of the coefficients of preference (as we describe in the previous section). Therefore, that case may be considered as a particular case. In this way, we obtain results more realistic in the case (very common in urban networks) of overlapping paths.

As example, let consider the following case. Let the choice set be made up of *n* paths: one independent path and (n−1) practically coincident paths. Let all the paths have the same cost. In this situation, all the coefficients of preference are equal to one. If we calculate the probability without consider the overlapping (i.e. by the Eq. (2)), we obtain that all the paths will have choice probability equal to 1/n. This is actually far from reality. In fact, because of the overlapping, users perceive only two effective alternatives, and will associate to them the same preference as they have the same costs. Therefore, the independent path (first alternative) will have probability of choice equal to 0.50, such as the totally of the remaining paths (second alternative). Each one of these paths has the same cost, so the choice probability of the generic overlapping path will be equal to 0.50/(n−1). If we apply the Eq. 4, we obtain exactly these results.

3 Numerical Analyses

Let us consider the network represented in Fig. 7 with six paths connecting the origin O with the destination D. Let the first one be independent and let its cost be equal to 40. Let the other paths cost be equal to 30. Because of the overlapping, paths 4, 5, 6 are practically the same.

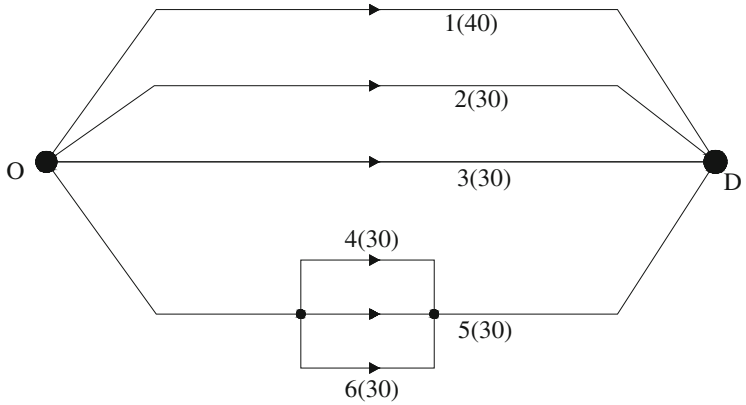


Fig. 7. Test network with overlapping paths

By applying the Eq. 3, the coefficient of overlapping is equal to zero for paths 1, 2, 3; and it is equal to 2/3 = 0,667 for paths 4, 5, 6. The expected path choice probability vector (P_1, P_2, P_3, P_4, P_5, P_6) must have the following characteristics:

1. $P_2 = P_3$ (because path 2 and path 3 have the same cost and are both independent);
5. $P_4 = P_5 = P_6$ (because path 4, path 5 and path 6 have the same cost and are overlapping in the same measure);
6. $P_1 < P_2, P_3$ (because all paths 1, 2, 3 are independent, but path 1 cost is greater than the others);
7. $P_2, P_3 = P_4 + P_5 + P_6$ (because paths 4, 5, 6 form only one alternative, due to their high overlapping, which is as convenient as the alternatives given by paths 2 and 3 - which are independent. Consequently, we expect that users associate the same choice probability to path 2, path 3 and the group of the paths 4, 5, 6).

The following Table 2 shows the results given by the proposed model where we assumed $\alpha = 0.150$. Estimated choice probabilities values reflects all the conditions just exposed, properly reproducing what we expect in route choice behaviour.

Table 2. Simulation results

Path	Lower limit	Crisp cost	Upper limit	$C_{pref,i}$	$\mathbf{P'_i}$	$C_{over,i}$	$C_{ind,i}$	$C_{pref,i} \cdot C_{ind,i}$	**Pi**
1	34.00	40.00	46.00	0.042	**0.008**	0.000	1.000	0.042	**0.012**
2	25.50	30.00	34.50	1.000	**0.198**	0.000	1.000	1.000	**0.329**
3	25.50	30.00	34.50	1.000	**0.198**	0.000	1.000	1.000	**0.329**
4	25.50	30.00	34.50	1.000	**0.198**	0.667	0.333	0.333	**0.110**
5	25.50	30.00	34.50	1.000	**0.198**	0.667	0.333	0.333	**0.110**
6	25.50	30.00	34.50	1.000	**0.198**	0.667	0.333	0.333	**0.110**

4 Equilibrium Assignment

In this case we considered the test network represented in (Fig. 8). To solve the traffic equilibrium problem the MSA (method of successive averages) has been used and then the equilibrium flows on the test network calculated. A sensitivity analysis has been carried out by considering different conditions of length or capacity of certain arcs and travel demand.

The considered network has six paths connecting the O/D pair form origin 1 to destination 9 (Fig. 8). The arc costs are calculated by the well known BPR cost function:

$$C = C_0 \cdot (1 + a \cdot (f_l / Cap_l)^{\gamma})$$

where:

- C_0 is the free flow cost;
- Cap_l is capacity of the arc l;
- f_l is the flow of the arc l;
- a and γ are parameters of the function.

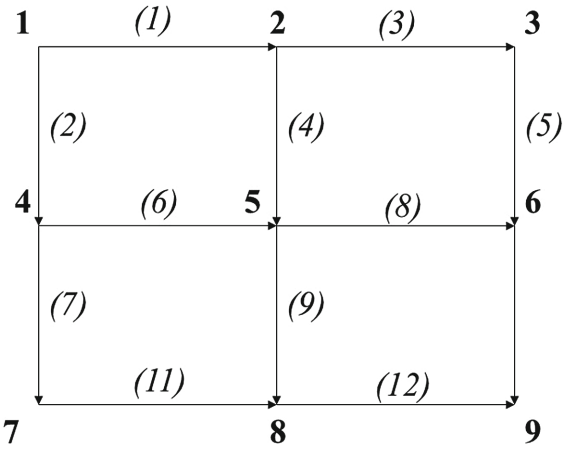


Fig. 8. Test network (Origin 1; Destination 9)

We assumed a = 2, $\gamma = 4$. We also assumed for each arc the capacity be 1000 and free-flow costs equal to 2. We perform several traffic assignments, considering different conditions of the network. For each condition, we applied our model achieving different tests for different values of α. We compared our results with those obtained by applying C-Logit model [7, 8], for different values of his parameters, θ and β_0.

The obtained results will be reported in the next section where the different considered conditions will be described.

4.1 Case 1 - Paths with the Same Free-Flow Cost

Let us consider the following situation: all the arcs have free-flow cost equal to 2, consequently free flow paths cost are all equal to 8. The assignment with C-Logit model [8] shows that the average user cost increases as much as the value of θ increases. The same happens applying the proposed model and increasing the value of α. In fact, θ and α represent (in different ways) the imprecision of the estimated path costs. If imprecision increases, we expect that more users will choice paths that are different from the most convenient one, with the consequence that the average user cost increases. This situation is correctly represented by increasing θ and α both in C-Logit and in our model.

This results were obtained for all the considered assignment cases, that are described as follows.

The vector of the arc flows at the equilibrium obtained by applying our model varies not much by changing the parameter α. We perform the traffic assignments by varying α in the interval $0.100 \div 0.300$. The index RMSE% is always lower than 5% between two probability vectors (obtained by our model with two different values of α, for the same configuration of the network).

Moreover, we saw good correspondence with C-Logit results assuming for the C-Logit parameters the values $\theta = 2.5$ and $\beta_0 = 1$.

Case 1a - Variation of the capacity of certain arcs. *Case 1a1* - Consider the arc 4 (from node 2 to node 5) with capacity equal to 500 (that is the half respect the previous case). We see that the average user cost increases respect to case "1" (we considered the comparison between the results obtained by one model without changing the values of its parameter). This is what we expect. In fact, the reduction of the capacity of one arc will induce an increasing of the network congestion (with the same demand), and obviously of the average user cost.

Case 1a2 - Consider now arc one with doubled capacity (thus equal to 2000). As we expect, we have from the results of the assignments that the average user cost is lower than in case "1". In fact, the increasing of the capacity of one arc will induce a reduction of the network congestion (with the same demand) and of the average user cost. The best correspondence for the results obtained by our model and C-Logit model are obtained when we assumed C-Logit parameter with these values: $\theta = 2.5$ and $\beta_0 = 1$.

Case 1b - Variation of travel demand. Consider the arcs having the same capacity and free flow costs as in case "1", but let the demand be equal to 1800. We see that the average user cost is higher than case "1", as we expect. In fact, the increasing of the demand will induce an increasing of the costs of the relative paths, so the total cost (and obviously the average user cost) will increase.

4.2 Case 2 – Variation of Free-Flow Cost

Case 2a - Let the free-flow cost of the arcs 3 and 5 be $C_0 = 1.5$ instead of 2. We see that the average user cost is lower than the case "1", as we expect. In fact, the reduction of the length of two arcs (and as a result the reduction of the correspondent free-flow cost) will generate a decreasing of the costs of the relative paths. Therefore, we see that the total cost and the average user costs decrease (with the same demand).

Case 2b - Consider the free-flow cost of the arc 4 and 8 very higher than the other arc costs. We expect that users will not consider these arcs, and consequently the relative paths (B [1-2-5-6-9], C [1-2-5-8-9], and D [1-4-5-6-9]).

The other paths (A [1-2-3-6-9], E [1-4-5-8-9], F [1-4-7-8-9]) have the same free flow cost. Moreover, path A is independent; path E and path F have the same degree of overlapping. We expect that the users will not perceive three different alternatives, but only two alternatives having the same length (the first one is path A and the second one is the group of paths E and F). The second alternative may be considered as a path containing a sub-path made of two parallel arcs having the same length. Practically, in this sub-path the capacity is doubled. Consequently, the congestion on them will be lower than on the other arcs. Thus, the second alternative is better than the first one (path A), so we expect that it will be chosen by the majority of the users. We also expect that on the overlapping sub-paths (E and F) the demand will be split in two equal proportions.

At the equilibrium, our model gives the following path choice probabilities (Table 3).

Table 3. Traffic assignment results (α = 0,200)

Path	Lower limit	Crisp cost	Upper limit	$C_{pref,i}$	$C_{over,i}$	$C_{ind,i}$	P_i
A [1-2-3-6-9]	6.840	8.549	10.259	1.000	0.000	1.000	0.430
B [1-2-5-6-9]	35.420	44.275	53.130	–	–	–	0.000
C [1-2-5-8-9]	21.268	26.584	31.901	–	–	–	0.000
D [1-4-5-6-9]	21.268	26.584	31.901	–	–	–	0.000
E [1-4-5-8-9]	7.115	8.894	10.673	0.884	0.250	0.750	0.285
F [1-4-7-8-9]	7.115	8.894	10.673	0.884	0.250	0.750	0.285

Paths B, C, D are dominated. In the Fig. 9, the representation of the TFNs relative to all the paths has been reported. Path E and path F have the same cost, such as path C and path D. As a result, the relative TFN are coincident.

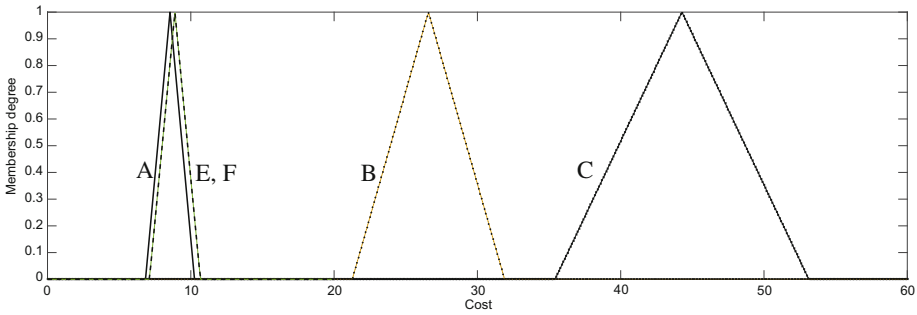


Fig. 9. Membership function

The model gives the expected results: the majority of the users chooses path E and path F, and fractions itself on them in equal parts. The vector of the choice probabilities changes not much by varying the value of α[3] (P_A = 0.439 if α = 0.100; P_A = 0.430 if α = 0.200; P_A = 0.425 if α = 0.300).

5 Conclusion

The aim of this work was to propose a path choice model where fuzzy numbers are utilised to represent the imprecision in the estimation of path costs in road network. The model allow to easily obtain the path choice probabilities overtaking some issue of Logit route choice model and explicitly considering users imprecision in cost perception. By using fuzzy numbers it is possible to consider the non-linearity between the value of path utility and the values of the attributes. The performed applications of the

[3] If α is greater than 0.520 path 3 and path 4 become non-dominated, if α is greater than 0.685 also path 2 become non-dominated (these values of α are not much representatives).

model has given realistic results in all the cases we have examined. It is important to remark that the proposed model is based on one parameter only. Moreover, we need only one value of this parameter to represent correctly users imprecision; both in first phase (selection of non-dominated paths) and in second phase (attribution of choice probabilities). Actually, users may have the problem of choosing between partially overlapping paths. Our model may give unrealistic results due to the IIA property as for Logit based models. To avoid this problem, we propose a correction of the calculation of choice probabilities introducing an index that allows to consider path overlapping. In this way, we obtained realistic results in the load assignment of the network in all the cases we have examined. In the traffic equilibrium assignment, we observe that parameter α is capable to represent the imprecision of the users when they estimate path costs. If the imprecision will increase then α increases. The results of the traffic assignment considering different conditions of the network show that there is a similarity with respect to the C-Logit model that we used as benchmark. The more is the imprecision we want to model, the higher the values of α and θ (for the C-Logit) must be. In this situation, we observe that the average user cost increases applying both our model and C-Logit model, and this is what we expect. Further investigations will be devoted to the analysis of real sized networks.

References

1. Akiyama, T., Kawahara, T.: Traffic assignment model with fuzzy travel time information. In: 9th Mini EURO Conference on Fuzzy Sets in Traffic and Transport Systems, Budva, Yugoslavia, September 1997
2. Arslan, T., Khisty, J.: A rational approach to handling fuzzy perceptions in route choice. Eur. J. Oper. Res. **168**(2), 571–583 (2006)
3. Ban, J., Liu, H., Yang, F., Ran, B.: A traffic assignment model with fuzzy travel time perceptions. In: Proceedings of the 12th World Congress on Intelligent Transport Systems, San Francisco, CA, USA (2005)
4. Bortolan, G., Degani, R.: A review of some methods for ranking fuzzy numbers. Fuzzy Sets Syst. **15**, 1–19 (1985)
5. Cantarella, G.E., Fedele, V.: Fuzzy utility theory for analysing discrete choice behaviour. In: Proceeding of the 4th International Symposium on Uncertainty Modelling and Analysis, pp. 148–154. IEEE Computer Society, Washington (2003)
6. Cantarella, G.E., Fedele, V.: Fuzzy utility models for discrete choice analysis. In: Proceedings of the 3rd International Conference on Traffic and Transport Psychology, Nottingham, UK, 5–9 September 2004
7. Cascetta, E.: Transportation System Analysis: Models and Applications. Springer, New York (2009)
8. Cascetta, E., Nuzzolo, A., Russo, F., Vitetta, V.: A modified logit route choice model overcoming path overlapping problems. Specification and some calibration results for interurban networks. In: XIII International Symposium on Transportation and Traffic Theory, Lyon, France, pp. 697–711, 24 July 1996
9. Chang, M.S., Chen, H.K.: A fuzzy user-optimal route choice problem using a link-based fuzzy variational inequality formulation. Fuzzy Sets Syst. **114**, 339–345 (2000)

10. De Maio, M.L., Vitetta, A.: Route choice on road transport system: a fuzzy approach. J. Intell. Fuzzy Syst. **28**(5), 2015–2027 (2015)
11. Frejinger, E.: Route choice analysis: data, models, algorithms and applications (Doctoral dissertation, EcolePolytechniqueFederale de Lausanne) (2008)
12. Ghatee, M., Hashemi, S.M.: Traffic assignment model with fuzzy level of travel demand: an efficient algorithm based on quasi-Logit formulas. Eur. J. Oper. Res. **194**(2), 432–451 (2009)
13. Hawas, Y.E.: Development and calibration of route choice utility models: factorial experimental design approach. J. Transp. Eng. **130**(2), 159–170 (2004)
14. Henn, V.: Fuzzy route choice model for traffic assignment. Fuzzy Sets Syst. **116**, 77–101 (2000)
15. Henn, V., Ottomanelli, M.: Handling uncertainty in route choice models: from probabilistic to possibilistic approaches. Eur. J. Oper. Res. **175**, 1526–1538 (2006)
16. Kosko, B.: Fuzzy Thinking: The New Science Of Fuzzy Logic. Hyperion, New York (1993)
17. Marinelli, M., Caggiani, L., Ottomanelli, M.: Managing the uncertainty of data fusion from different sources in modelling route choice behaviour. In: Proceedings of the IEEE Conference on Intelligent Transportation Systems, ITSC, Las Palmas, Gran Canaria, pp. 202–207. IEEE Press (2015)
18. Murat, Y.S., Uludag, N.: Route choice modelling in urban transportation networks using fuzzy logic and logistic regression methods. J. Sci. Ind. Res. **67**(1), 19–27 (2008)
19. Okada, S., Soper, T.: A shortest path problem on a network with fuzzy arc lengths. Fuzzy Sets Syst. **109**, 129–140 (2000)
20. Ottomanelli, M., Dell'Orco, M., Sassanelli, D.: Modelling parking choice behaviour using Possibility Theory. Transp. Plann. Technol. **34**(7), 647–667 (2011)
21. Peeta, S., Yu, J.W.: A hybrid model for driver route choice incorporating en-route attributes and real-time information effects. Netw. Spat. Econ. **5**(1), 21–40 (2005)
22. Quattrone, A., Vitetta, A.: Random and fuzzy utility models for road route choice. Transp. Res. Part E **47**(6), 1126–1139 (2011)
23. Ramos, G.M., Daamen, W., Hoogendoorn, S.: A state-of-the-art review: developments in utility theory, prospect theory and regret theory to investigate travellers' behaviour in situations involving travel time uncertainty. Transp. Rev. **34**(1), 46–67 (2014)
24. Ridwan, M.: Fuzzy preference based traffic assignment problem. Transp. Res. Part C Emerg. Technol. **12**(3), 209–233 (2004)
25. Shao, C., Akiyama, T., Sasaki, T.: User equilibrium assignment on network with fuzzy route travel time. In: Proceedings of Sino-Japan Joint Meeting on Fuzzy Sets and Systems, Beijing, China, 15–18 October 1990
26. Teodorovic, D.: Fuzzy logic systems for transportation engineering: the state of the art. Transp. Res. Part A **33**, 337–364 (1999)
27. Teodorovic, D., Kikuchi, S.: Transportation route choice model using fuzzy inference technique. In: Proceedings of the First International Symposium on Uncertainty Modeling and Analysis, pp. 140–145. IEEE (1990)
28. Wang, F., Liao, H.L.: User equilibrium in traffic assignment problem with fuzzy N-A incidence matrix. Fuzzy Sets Syst. **107**, 245–253 (1999)

Author Index

A
Ambrosino, Daniela, 405
Aron, Maurice, 309

B
Bella, Francesco, 355
Bhouri, Neïla, 309, 334
Billot, Romain, 309
Binetti, Mario, 452

C
Canca, David, 198
Carotenuto, Pasquale, 421
Codina, Esteve, 272

D
De Mitri, Marco, 452
Dell'Orco, Mauro, 389
De-Los-Santos, Alicia, 198
Dožić, Slavica, 156

E
El Faouzi, Nour-Eddin, 309

F
Farhi, Nadir, 334

G
Galińska, Barbara, 121
Gastaldi, Massimiliano, 370
Gecchele, Gregorio, 370
Giordani, Stefano, 421
Giovannini, Alessandro, 237

H
Hadas, Yuval, 103
Hadavi, Sheida, 39
Haj-Salem, Habib, 334

I
Ignaccolo, Matteo, 54
Inturri, Giuseppe, 54
Ishiguro, Shogo, 290

J
Jin, Junchen, 438

K
Kalić, Milica, 156
Khanzad, Iran, 183
Kiba-Janiak, Maja, 73

L
Laporte, Gilbert, 198
Le Pira, Michela, 54
Lebacque, Jean Patrique, 334

M
Ma, Xiaoliang, 438
Macharis, Cathy, 39
Malucelli, Federico, 237
Marinelli, Mario, 389
Massari, Simone, 421
Matsumoto, Yukimasa, 290
Mesa, Juan A., 198

N
Nahum, Oren E., 103
Nazemi, Mohsen, 183
Nonato, Maddalena, 237

O
Ottomanelli, Michele, 389, 452

P
Palmisano, Gianvito, 389
Paolucci, Massimo, 405
Pluchino, Alessandro, 54

© Springer International Publishing AG 2018

J. Żak et al. (eds.), *Advanced Concepts, Methodologies and Technologies for Transportation and Logistics*, Advances in Intelligent Systems and Computing 572, DOI 10.1007/978-3-319-57105-8

R
Rosell, Francisca, 272
Rossi, Riccardo, 370
Ruiz, Maurici, 220

S
Sciomachen, Anna, 405
Seguí-Pons, Joana Maria, 220
Seidowsky, Régine, 309
Seyedabrishami, Seyedehsan, 183
Shnaiderman, Matan, 256
Silvestri, Manuel, 355

V
Vagaggini, Fabrizio, 421
Van Raemdonck, Koen, 39

Z
Żak, Jacek, 9, 73, 121
Zarrinmehr, Amirali, 183